



등록특허 10-2590466



**(19) 대한민국특허청(KR)**  
**(12) 등록특허공보(B1)**

(45) 공고일자 2023년10월19일  
(11) 등록번호 10-2590466  
(24) 등록일자 2023년10월12일

- (51) 국제특허분류 (Int. Cl.)  
*C12N 9/22* (2006.01) *A61K 38/00* (2006.01)  
*A61P 35/02* (2006.01) *C07K 14/405* (2006.01)  
*C07K 14/71* (2006.01) *C12N 5/078* (2010.01)  
*C12N 9/14* (2006.01) *C12N 9/88* (2006.01)
- (52) CPC특허분류  
*C12N 9/22* (2013.01)  
*A61K 38/00* (2013.01)
- (21) 출원번호 10-2019-7036681
- (22) 출원일자(국제) 2018년05월25일  
 심사청구일자 2021년04월23일
- (85) 번역문제출일자 2019년12월11일
- (65) 공개번호 10-2020-0011953
- (43) 공개일자 2020년02월04일
- (86) 국제출원번호 PCT/US2018/034726
- (87) 국제공개번호 WO 2018/218194  
 국제공개일자 2018년11월29일
- (30) 우선권주장  
 62/511,194 2017년05월25일 미국(US)  
 62/567,417 2017년10월03일 미국(US)
- (56) 선행기술조사문헌  
 WO2011156430 A2\*  
 WO2017001572 A1\*

\*는 심사관에 의하여 인용된 문헌

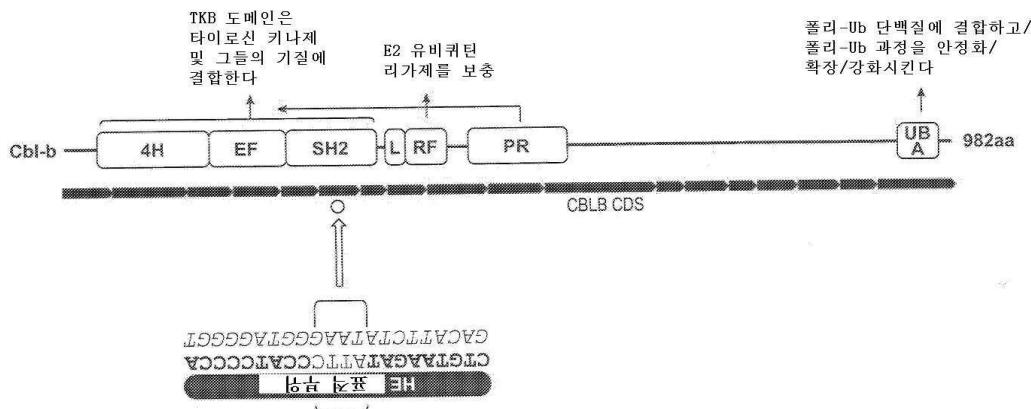
전체 청구항 수 : 총 17 항

심사관 : 이형곤

(54) 발명의 명칭 CBLB 엔도뉴클레아제 변이체, 조성물 및 사용 방법

**(57) 요약**

본 개시내용은 개선된 게놈 편집 조성물 및 CBLB 유전자를 편집하는 방법을 제공한다. 본 개시내용은 추가로 암, 감염성 질환, 자가면역 질환, 염증성 질환 또는 면역결핍의 적어도 하나의 증상의 예방, 치료 또는 개선을 위한 게놈 편집된 세포를 제공한다.

**대 표 도 - 도1**

(52) CPC특허분류

*A61P 35/02* (2018.01)  
*C07K 14/405* (2013.01)  
*C07K 14/71* (2013.01)  
*C12N 5/0634* (2023.05)  
*C12N 9/14* (2013.01)  
*C12N 9/88* (2013.01)  
*C12Y 307/01003* (2013.01)  
*C12Y 406/01* (2013.01)  
*C07K 2319/50* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

서열번호 6과 적어도 90% 동일한 아미노산 서열을 포함하고, 인간 카시타스 B-계통(casitas B-lineage: Cbl) 램프종 원발암유전자 B(CBLB) 유전자 내 표적 부위에 결합하고 이를 절단하는 I-OnuI 호밍 엔도뉴클레아제(homing endonuclease: HE) 변이체를 포함하는 폴리펩타이드로서, 상기 표적 부위가 서열번호 20인, 폴리펩타이드.

#### 청구항 2

제1항에 있어서, I-OnuI HE 변이체가 서열번호 6과 적어도 95% 동일한 아미노산 서열을 포함하는, 폴리펩타이드.

#### 청구항 3

제1항에 있어서, I-OnuI HE 변이체가 서열번호 6과 적어도 99% 동일한 아미노산 서열을 포함하는, 폴리펩타이드.

#### 청구항 4

제1항에 있어서, I-OnuI HE 변이체가 서열번호 6에 제시된 아미노산 서열을 포함하는, 폴리펩타이드.

#### 청구항 5

제1항에 있어서, I-OnuI HE 변이체가 하기 아미노산 치환을 포함하는, 폴리펩타이드:

(a) 서열번호 4의 S24C, L26R, R28D, N32A, K34D, S35L, S36R, V37A, S40R, E42R, G44A, Q46E, T48V, V68T, A70Y, S72A, S78R, K80Q, L138M, T143N, F168L, E178D, C180S, F182V, N184E, I186K, S188R, K189R, K191P, L192V, G193K, Q195G, Q197R, V199R, K207R, Y223R, K225V, K227N, F232H, D236E, 및 V238I;

(b) 서열번호 4의 S24C, L26R, R28D, N32A, K34D, S35L, S36R, V37A, S40R, E42R, G44A, Q46E, T48V, V68T, A70Y, S72A, S78R, K80Q, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S188N, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E, 및 V238I;

(c) 서열번호 4의 S24C, L26R, R28D, N32A, K34D, S35L, S36R, V37A, S40R, E42R, G44S, Q46E, T48S, V68T, A70Y, S72A, S78R, K80Q, D92G, V116L, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S188N, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E, 및 V238I;

(d) 서열번호 4의 S24C, L26R, R28D, R30H, N32A, K34V, S35L, S36R, V37S, S40R, E42R, G44S, Q46E, T48V, V68T, V68K, A70Y, S72A, S78R, K80Q, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S188N, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E, 및 V238I;

(e) 서열번호 4의 S24C, L26G, R28Y, R30H, N32S, K34V, S35L, S36R, V37S, S40R, E42R, G44S, Q46E, T48S, V68T, A70Y, S72A, S78R, K80Q, V116L, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S188N, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E, 및 V238I;

(f) 서열번호 4의 S24C, L26R, R28D, R30H, N32A, K34V, S35L, S36R, V37S, S40R, E42R, G44S, Q46E, T48V, V68T, A70Y, S72A, S78R, K80Q, V116L, L138M, T143N, S159P, F168L, E178D, C180S, F182V, N184E, I186K, S188R, K189R, K191P, L192V, G193K, Q195G, Q197R, V199R, S201G, K207R, Y223R, K225V, K227N, F232H, D236E, 및 V238I; 또는

(g) 서열번호 4의 S24C, L26R, R28D, N32A, K34D, S35L, S36R, V37A, S40R, E42R, G44A, Q46E, T48V, V68T, A70Y, S72A, S78R, K80Q, D92G, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S188N,

S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E, 및 V238I.

#### 청구항 6

제1항에 있어서, I-OnuI HE 변이체가 서열번호 6 내지 12 중 어느 하나에 제시된 아미노산 서열을 포함하는, 폴리펩타이드.

#### 청구항 7

제1항에 있어서, 9.5개의 TALE 반복 단위 내지 15.5개의 TALE 반복 단위를 포함하는 TALE DNA 결합 도메인을 추가로 포함하는, 폴리펩타이드.

#### 청구항 8

제7항에 있어서, TALE DNA 결합 도메인이 서열번호 21에 제시된 폴리뉴클레오타이드 서열에 결합하는, 폴리펩타이드.

#### 청구항 9

제7항에 있어서, 폴리펩타이드가 서열번호 22에 제시된 폴리뉴클레오타이드 서열에 결합하고 이를 절단하는, 폴리펩타이드.

#### 청구항 10

제7항에 있어서, 폴리펩타이드가 서열번호 13 내지 19 중 어느 하나에 제시된 아미노산 서열과 95% 동일한 아미노산 서열을 포함하는, 폴리펩타이드.

#### 청구항 11

제1항 내지 제10항 중 어느 한 항의 폴리펩타이드를 암호화하는 폴리뉴클레오타이드.

#### 청구항 12

제1항 내지 제10항 중 어느 한 항의 폴리펩타이드를 암호화하는 mRNA.

#### 청구항 13

제1항 내지 제10항 중 어느 한 항의 폴리펩타이드를 암호화하는 폴리뉴클레오타이드를 포함하는, 벡터.

#### 청구항 14

제1항 내지 제10항 중 어느 한 항의 폴리펩타이드, 제1항 내지 제10항 중 어느 한 항의 폴리펩타이드를 암호화하는 폴리뉴클레오타이드, 또는 제1항 내지 제10항 중 어느 한 항의 폴리펩타이드를 암호화하는 폴리뉴클레오타이드를 포함하는 벡터를 포함하는, 세포.

#### 청구항 15

제14항에 있어서, 세포가 조작된 항원 수용체를 암호화하는 폴리뉴클레오타이드를 포함하는, 세포.

#### 청구항 16

제15항에 있어서, 조작된 항원 수용체가 조작된 TCR, CAR, Daric 또는 제타카인(zetakine)으로 이루어진 군으로부터 선택되는, 세포.

#### 청구항 17

제14항에 있어서, 세포가 조혈 세포; T 세포; CD3<sup>+</sup>, CD4<sup>+</sup> 및/또는 CD8<sup>+</sup> 세포; 면역 효과기 세포; 세포독성 T 림프구(cytotoxic T lymphocyte: CTL), 종양 침윤성 림프구(tumor infiltrating lymphocyte: TIL) 또는 헬퍼 T 세포; 또는 자연 살해(natural killer: NK) 세포 또는 자연 살해 T(natural killer T: NKT) 세포인, 세포.

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

삭제

청구항 80

삭제

청구항 81

삭제

청구항 82

삭제

청구항 83

삭제

청구항 84

삭제

청구항 85

삭제

청구항 86

삭제

청구항 87

삭제

청구항 88

삭제

청구항 89

삭제

청구항 90

삭제

청구항 91

삭제

청구항 92

삭제

청구항 93

삭제

청구항 94

삭제

청구항 95

삭제

청구항 96

삭제

청구항 97

삭제

청구항 98

삭제

청구항 99

삭제

청구항 100

삭제

청구항 101

삭제

청구항 102

삭제

청구항 103

삭제

청구항 104

삭제

청구항 105

삭제

청구항 106

삭제

청구항 107

삭제

청구항 108

삭제

청구항 109

삭제

청구항 110

삭제

청구항 111

삭제

청구항 112

삭제

청구항 113

삭제

청구항 114

삭제

청구항 115

삭제

청구항 116

삭제

청구항 117

삭제

청구항 118

삭제

청구항 119

삭제

청구항 120

삭제

청구항 121

삭제

청구항 122

삭제

청구항 123

삭제

청구항 124

삭제

청구항 125

삭제

청구항 126

삭제

청구항 127

삭제

청구항 128

삭제

청구항 129

삭제

청구항 130

삭제

청구항 131

삭제

청구항 132

삭제

청구항 133

삭제

청구항 134

삭제

청구항 135

삭제

청구항 136

삭제

청구항 137

삭제

청구항 138

삭제

### 발명의 설명

#### 기술분야

[0001] 관련 출원에 대한 상호 참조

[0002] 본 출원은 미국 특허법(35 U.S.C. § 119(e)) 하에서 2017년 10월 3일자로 출원된 미국 가출원 특허 제 62/567,417호, 및 2017년 5월 25일자로 출원된 미국 가출원 특허 제62/511,194호의 우선권을 주장하며, 이들 각각은 그의 전문이 본 명세서에 참고로 포함된다.

[0003] 서열목록에 관한 언급

[0004] 본 출원과 관련된 서열목록은 종이 사본 대신에 텍스트 형식으로 제공되며, 본원에서 본 명세서 내에 참고로 포함된다. 서열목록을 함유하는 텍스트 파일명은 BLBD\_087\_02WO\_ST25.txt이다. 텍스트 파일은 152KB이며, 2018년 5월 25일자로 생성되었고, 본 명세서의 출원과 동시에 EFS-Web을 통해 전자적으로 제출 중에 있다.

[0005] 기술분야

[0006] 본 개시내용은 개선된 게놈 편집 조성물에 관한 것이다. 더 구체적으로, 본 개시내용은 뉴클레아제 변이체, 조성물, 및 인간 카시타스 B-계통(casitas B-lineage: Cbl) 림프종 원발암유전자 B(lymphoma proto-oncogene B)(CBLB) 유전자를 편집하기 위해 이들을 이용하는 방법에 관한 것이다.

### 배경기술

[0007] 암의 전세계적 부담은 1975 내지 2000에 2배가 되었다. 암은 전세계적으로 이환율 및 사망률의 두 번째 주된 원인이며, 2012년에 대략 1천 4백 십만건의 새로운 사례 및 8백 2십만건의 암관련 사망이 있었다. 가장 통상적인 암은 유방암, 폐 및 기관지암, 전립선암, 결장 및 직장암, 방광암, 피부의 흑색종, 비호지킨 림프종, 갑상선암,

신장 및 신우암, 자궁내막암, 백혈병 및 혀암이다. 새로운 암 사례의 수는 다음 20년 내에 2천 2백만건까지 상승할 것으로 추정된다.

[0008] 면역계는 인간암을 검출하고 이와 싸우는데 중요한 역할을 가진다. 다수의 형질전환된 세포는 면역 보초에 의해 빠르게 검출되며, 물론 발현된 T 세포 수용체(T cell receptor: TCR)를 통한 항원-특이적 T 세포의 활성화를 통해 파괴된다. 따라서, 암은 질환을 지속적으로 억제하고 제거하기 위한 필요한 항-종양 반응을 시작하기 위한 면역계의 부전인 면역학적 장애로 고려될 수 있다. 암과 더 효과적으로 싸우기 위해, 마지막 몇십년에 걸쳐 진행된 특정 면역요법 개입은 T 세포 면역을 향상시키는 데 특별하게 중점을 두었다. 이들 치료는 질환 관해의 산발적 사례만을 수득하였고, 실질적인 전반적 성공을 갖지 않았다.

[0009] 더 최근에, T 세포의 단리, 변형, 확장 및 재주입에 기반한 적응 세포 요법 전략은 초기 임상 시험에서 연구되고 시험되었다. T 세포는 종종 그들의 선택적 인식 및 강력한 효과기 메커니즘에 기인하여 암에 대한 선택의 효과기 세포가 되었다. 이들 치료는 혼합된 비율의 성장을 나타내었지만, 소수의 환자는 T 세포 기반 면역요법에 대한 아직까지 실현되지 않은 잠재력을 강조하는 지속 가능한 관해를 경험하였다.

[0010] 세포용해 T 세포에 의한 종양 세포 관련 항원의 성공적인 인식은 표적화된 종양 용해를 개시하며, 임의의 효과적인 암 면역요법 접근을 뒷받침한다. 종양-침윤성 T 세포(Tumor-infiltrating T cell: TIL)는 TCR과 특별하게 관련된 종양-관련 항원을 발현시키지만; 그러나, TIL의 실질적인 수는 소수의 인간암으로 제한된다. 조작된 T 세포 수용체(TCR) 및 키메라 항원 수용체(chimeric antigen receptor: CAR)는 다수의 암 및 다른 면역 장애에 대한 T 세포-기반 면역요법의 적용 가능성을 잠재적으로 증가시킨다.

[0011] 추가로, 당업계의 조작된 T 세포는 암세포, 면역염증 세포, 기질 세포 및 사이토카인으로 이루어진 복잡한 면역 억제 종양 미세환경에 의해 여전히 조절된다. 이들 성분 중에서, 암세포, 면역염증 세포 및 억제 사이토카인은 T 세포 표현형 및 기능에 유해하게 영향을 미친다. 총괄적으로, 종양 미세환경은 T 세포가 마지막에 고갈된 T 세포로 분화되도록 유도한다.

[0012] T 세포 고갈은 저해 수용체의 증가된 발현 또는 저해 수용체의 증가된 신호전달; 감소된 효과기 사이토카인 생산; 및 암을 지속하고 제거하는 감소된 능력에 의해 나타난 만성 환경에서의 T 세포 기능 장애 상태이다. 고갈된 T 세포는 또한 계층적 방식으로 기능 상실을 나타낸다: 감소된 IL-2 생산 및 생체외 사멸 능력은 고갈의 초기 단계에서 상실되며, TNF- $\alpha$  생산은 중기에 상실되고, IFN- $\gamma$  및 GzmB 생산은 진행된 고갈 단계에서 상실된다. 종양 미세환경에서 대부분의 T 세포는 고갈된 T 세포로 분화되며, 암을 제거하는 능력을 상실하고, 종국적으로 클리어런스된다.

[0013] 지금까지 면역억제 종양 미세환경에 대해 증가된 지속성 및 내성을 갖는 양자 세포 요법의 입증 가능한 임상에는 없었다.

### 발명의 내용

[0014] 본 개시내용을 일반적으로, 부분적으로, 인간 카시타스 B-계통(Cbl) 림프종 원발암유전자 B(CBLB) 유전자 내 표적 부위를 절단하는 호밍 엔도뉴클레아제(homing endonuclease) 변이체 및 메가TAL를 포함하는 조성물 및 이를 이용하는 방법에 관한 것이다.

[0015] 다양한 실시형태에서, 본 개시내용은, 부분적으로, 인간 CBLB 유전자 내 표적 부위를 절단하는 호밍 엔도뉴클레아제(HE) 변이체를 포함하는 폴리펩타이드를 상정한다.

[0016] 특정 실시형태에서, HE 변이체는 LAGLIDADG 호밍 엔도뉴클레아제(LHE) 변이체이다.

[0017] 일부 실시형태에서, 폴리펩타이드는 HE 변이체의 생물학적 활성 단편을 포함한다.

[0018] 소정의 실시형태에서, 생물학적 활성 단편은 대응하는 야생형 HE에 비해 1, 2, 3, 4, 5, 6, 7 또는 8개의 N-말단 아미노산을 결여한다.

[0019] 일부 실시형태에서, 생물학적 활성 단편은 대응하는 야생형 HE에 비해 4개의 N-말단 아미노산을 결여한다.

[0020] 추가 실시형태에서, 생물학적 활성 단편은 대응하는 야생형 HE에 비해 8개의 N-말단 아미노산을 결여한다.

[0021] 특정 실시형태에서, 생물학적 활성 단편은 대응하는 야생형 HE에 비해 1, 2, 3, 4, 5 또는 6개의 C-말단 아미노산을 결여한다.

- [0022] 특정 실시형태에서, 생물학적 활성 단편은 대응하는 야생형 HE에 비해 C-말단 아미노산을 결여한다.
- [0023] 추가적인 실시형태에서, 생물학적 활성 단편은 대응하는 야생형 HE에 비해 2개의 C-말단 아미노산을 결여한다.
- [0024] 특정 실시형태에서, HE 변이체는 I-CreI 및 I-SceI로 이루어진 군으로부터 선택된 LHE의 변이체이다.
- [0025] 소정의 실시형태에서, HE 변이체는 I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-GzeMIII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-NcrI, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI 및 I-Vdi141I로 이루어진 군으로부터 선택된 LHE의 변이체이다.
- [0026] 일부 실시형태에서, HE 변이체는 I-CpaMI, I-HjeMI, I-OnuI, I-PanMI 및 SmaMI로 이루어진 군으로부터 선택된 LHE의 변이체이다.
- [0027] 추가적인 실시형태에서, HE 변이체는 I-OnuI LHE 변이체이다.
- [0028] 추가 실시형태에서, HE 변이체는 서열번호 1 내지 5에 제시된 바와 같은 I-OnuI LHE 아미노산 서열, 또는 이의 생물학적 활성 단편 중의 24, 26, 28, 30, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 68, 70, 72, 75, 76, 78, 80, 82, 180, 182, 184, 186, 188, 189, 190, 191, 192, 193, 195, 197, 199, 201, 203, 223, 225, 227, 229, 231, 232, 234, 236, 238 및 240번으로 이루어진 군으로부터 선택된 아미노산 위치에서의 DNA 인식 계면에서 하나 이상의 아미노산 치환을 포함한다.
- [0029] 소정의 실시형태에서, HE 변이체는 서열번호 1 내지 5에 제시된 바와 같은 I-OnuI LHE 아미노산 서열, 또는 이의 생물학적 활성 단편 중의 24, 26, 28, 30, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 68, 70, 72, 75, 76, 78, 80, 82, 180, 182, 184, 186, 188, 189, 190, 191, 192, 193, 195, 197, 199, 201, 203, 223, 225, 227, 229, 231, 232, 234, 236, 238 및 240번으로 이루어진 군으로부터 선택된 아미노산 위치에서의 DNA 인식 계면에서 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상의 아미노산 치환을 포함한다.
- [0030] 특정 실시형태에서, HE 변이체는 서열번호 1 내지 5에 제시된 바와 같은 I-OnuI LHE 아미노산 서열, 또는 이의 생물학적 활성 단편 중의 19, 24, 26, 28, 30, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 59, 68, 70, 72, 75, 76 77, 78, 80, 82, 168, 180, 182, 184, 186, 188, 189, 190, 191, 192, 193, 195, 197, 199, 201, 203, 223, 225, 227, 229, 231, 232, 234, 236, 238 및 240번으로 이루어진 군으로부터 선택된 아미노산 위치에서 하나 이상의 아미노산 치환을 포함한다.
- [0031] 소정의 실시형태에서, HE 변이체는 서열번호 1 내지 5에 제시된 바와 같은 I-OnuI LHE 아미노산 서열, 또는 이의 생물학적 활성 단편 중의 19, 24, 26, 28, 30, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 59, 68, 70, 72, 75, 76 77, 78, 80, 82, 168, 180, 182, 184, 186, 188, 189, 190, 191, 192, 193, 195, 197, 199, 201, 203, 223, 225, 227, 229, 231, 232, 234, 236, 238 및 240번으로 이루어진 군으로부터 선택된 아미노산 위치에서 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상의 아미노산 치환을 포함한다.
- [0032] 특정 실시형태에서, HE 변이체는 서열번호 1 내지 5 중 어느 하나, 또는 이의 생물학적 활성 단편 중의 24, 26, 28, 30, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 68, 70, 72, 78, 80, 92, 116, 138, 143, 159, 168, 178, 180, 182, 184, 186, 188, 189, 190, 191, 192, 193, 195, 197, 199, 201, 203, 207, 223, 225, 227, 232, 236 및 238번 위치로 이루어진 군의 위치로부터 선택된 적어도 하나의 위치에서 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상의 아미노산을 포함한다.
- [0033] 일부 실시형태에서, HE 변이체는 서열번호 1 내지 5 중 어느 하나, 또는 이의 생물학적 활성 단편 중의 다음의 아미노산 치환: S24C, L26R, L26G, R28D, R28Y, R30H, N32A, N32S, K34D, K34V, S35L, S36R, V37A, V37S, S40R, E42R, G44A, G44S, Q46E, T48V, T48S, V68T, V68K, A70Y, S72A, S78R, K80Q, D92G, V116L, L138M, T143N, S159P, F168L, E178D, C180S, F182V, F182M, N184E, I186K, I186M, S188R, S188N, K189R, S190N, K191P, K191N, L192V, G193K, G193I, Q195G, Q195R, Q197R, V199R, S201G, T203S, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.

- [0034] 추가 실시형태에서, HE 변이체는 서열번호 1 내지 5 중 어느 하나, 또는 이의 생물학적 활성 단편 중의 다음의 아미노산 치환: S24C, L26R, R28D, N32A, K34D, S35L, S36R, V37A, S40R, E42R, G44A, Q46E, T48V, V68T, A70Y, S72A, S78R, K80Q, L138M, T143N, F168L, E178D, C180S, F182V, N184E, I186K, S188R, K189R, K191P, L192V, G193K, Q195G, Q197R, V199R, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.
- [0035] 특정 실시형태에서, HE 변이체는 서열번호 1 내지 5 중 어느 하나, 또는 이의 생물학적 활성 단편 중의 다음의 아미노산 치환: S24C, L26R, R28D, N32A, K34D, S35L, S36R, V37A, S40R, E42R, G44A, Q46E, T48V, V68T, A70Y, S72A, S78R, K80Q, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.
- [0036] 소정의 실시형태에서, HE 변이체는 서열번호 1 내지 5 중 어느 하나, 또는 이의 생물학적 활성 단편 중의 다음의 아미노산 치환: S24C, L26R, R28D, N32A, K34D, S35L, S36R, V37A, S40R, E42R, G44S, Q46E, T48S, V68T, A70Y, S72A, S78R, K80Q, D92G, V116L, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S188N, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.
- [0037] 일부 실시형태에서, HE 변이체는 서열번호 1 내지 5 중 어느 하나, 또는 이의 생물학적 활성 단편 중의 다음의 아미노산 치환: S24C, L26R, R28D, R30H, N32A, K34V, S35L, S36R, V37S, S40R, E42R, G44S, Q46E, T48V, V68T, V68K, A70Y, S72A, S78R, K80Q, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S188N, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.
- [0038] 특정 실시형태에서, HE 변이체는 서열번호 1 내지 5 중 어느 하나, 또는 이의 생물학적 활성 단편 중의 다음의 아미노산 치환: S24C, L26G, R28Y, R30H, N32S, K34V, S35L, S36R, V37S, S40R, E42R, G44S, Q46E, T48S, V68T, A70Y, S72A, S78R, K80Q, V116L, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S188N, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.
- [0039] 소정의 실시형태에서, HE 변이체는 서열번호 1 내지 5 중 어느 하나, 또는 이의 생물학적 활성 단편 중의 다음의 아미노산 치환: S24C, L26R, R28D, R30H, N32A, K34V, S35L, S36R, V37S, S40R, E42R, G44S, Q46E, T48V, V68T, A70Y, S72A, S78R, K80Q, V116L, L138M, T143N, S159P, F168L, E178D, C180S, F182V, N184E, I186K, S188R, K189R, K191P, L192V, G193K, Q195G, Q197R, V199R, S201G, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.
- [0040] 추가 실시형태에서, HE 변이체는 서열번호 1 내지 5 중 어느 하나, 또는 이의 생물학적 활성 단편 중의 다음의 아미노산 치환: S24C, L26R, R28D, N32A, K34D, S35L, S36R, V37A, S40R, E42R, G44A, Q46E, T48V, V68T, A70Y, S72A, S78R, K80Q, D92G, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S188N, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.
- [0041] 특정 실시형태에서, 상기 HE 변이체는 서열번호 6 내지 12 중 임의의 하나의 서열번호에 제시된 아미노산 서열 또는 이의 생물학적 활성 단편에 대해 적어도 80%, 바람직하게는 적어도 85%, 더 바람직하게는 적어도 90%, 또는 더욱더 바람직하게는 적어도 95% 동일한 아미노산 서열을 포함한다.
- [0042] 일정 실시형태에서, HE 변이체는 서열번호 6에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.

- [0043] 추가적인 실시형태에서, HE 변이체는 서열번호 7에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0044] 특정 실시형태에서, HE 변이체는 서열번호 8에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0045] 특정 실시형태에서, HE 변이체는 서열번호 9에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0046] 추가 실시형태에서, HE 변이체는 서열번호 10에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0047] 소정의 실시형태에서, HE 변이체는 서열번호 11에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0048] 소정의 실시형태에서, HE 변이체는 서열번호 12에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0049] 일부 실시형태에서, 폴리펩타이드는 서열번호 20에 제시된 폴리뉴클레오타이드 서열에 결합한다.
- [0050] 특정 실시형태에서, 폴리펩타이드는 DNA 결합 도메인을 추가로 포함한다.
- [0051] 추가적인 실시형태에서, DNA 결합 도메인은 TALE DNA 결합 도메인 및 아연 평거 DNA 결합 도메인으로 이루어진 군으로부터 선택된다.
- [0052] 소정의 실시형태에서, TALE DNA 결합 도메인은 약 9.5개의 TALE 반복 단위 내지 약 15.5개의 TALE 반복 단위를 포함한다.
- [0053] 추가 실시형태에서, TALE DNA 결합 도메인은 CBLB 유전자 내 폴리뉴클레오타이드 서열에 결합한다.
- [0054] 특정 실시형태에서, TALE DNA 결합 도메인은 서열번호 21에 제시된 폴리뉴클레오타이드 서열에 결합한다.
- [0055] 특정 실시형태에서, 폴리뉴클레오타이드는 서열번호 22에 제시된 폴리뉴클레오타이드에 결합하고, 이를 절단한다.
- [0056] 특정 실시형태에서, 아연 평거 DNA 결합 도메인은 2, 3, 4, 5, 6, 7 또는 8개의 아연 평거 모티프를 포함한다.
- [0057] 추가적인 실시형태에서, 폴리펩타이드는 추가로 웨타이드 링커 및 말단-가공 효소 또는 이의 생물학적 활성 단편을 포함한다.
- [0058] 소정의 실시형태에서, 폴리펩타이드는 추가로 바이러스 자기-절단성 2A 웨타이드 웨타이드 및 말단-가공 효소 또는 이의 생물학적 활성 단편을 포함한다.
- [0059] 추가 실시형태에서, 말단-가공 효소 또는 이의 생물학적 활성 단편은 5'-3' 엑소뉴클레아제, 5'-3' 알칼리성 엑소뉴클레아제, 3'-5' 엑소뉴클레아제, 5' 플랩 엔도뉴클레아제, TdT, 헬리카제 또는 주형-독립적 DNA 중합효소 활성을 가진다.
- [0060] 일부 실시형태에서, 말단-가공 효소는 Trex2 또는 이의 생물학적 활성 단편을 포함한다.
- [0061] 추가 실시형태에서, 폴리펩타이드는 서열번호 13 내지 19 중 어느 하나에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0062] 추가적인 실시형태에서, 폴리펩타이드는 서열번호 13에 제시된 아미노산 서열 또는 이의 생물학적 활성 단편을 포함하다.
- [0063] 특정 실시형태에서, 폴리펩타이드는 서열번호 14에 제시된 아미노산 서열 또는 이의 생물학적 활성 단편을 포함하다.
- [0064] 소정의 실시형태에서, 폴리펩타이드는 서열번호 15에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0065] 소정의 실시형태에서, 폴리펩타이드는 서열번호 16에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.

- [0066] 일부 실시형태에서, 폴리펩타이드는 서열번호 17에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0067] 특정 실시형태에서, 폴리펩타이드는 서열번호 18에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0068] 추가 실시형태에서, 폴리펩타이드는 서열번호 19에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0069] 특정 실시형태에서, 폴리펩타이드는 서열번호 20 또는 22에 제시된 폴리뉴클레오타이드 서열에서 인간 CBLB 유전자를 절단한다.
- [0070] 다양한 실시형태에서, 본 개시내용은, 부분적으로, 본 명세서에서 상정된 폴리펩타이드를 암호화하는 폴리뉴클레오타이드를 제공한다.
- [0071] 일부 실시형태에서, 본 개시내용은, 부분적으로, 본 명세서에서 상정된 mRNA를 제공한다.
- [0072] 특정 실시형태에서, 본 개시내용은, 부분적으로, 본 명세서에서 상정된 폴리펩타이드를 암호화하는 cDNA를 제공한다.
- [0073] 다양한 실시형태에서, 본 개시내용은, 부분적으로, 본 명세서에서 상정된 폴리펩타이드를 암호화하는 폴리뉴클레오타이드를 포함하는 백터를 제공한다.
- [0074] 다양한 실시형태에서, 본 개시내용은, 부분적으로, 본 명세서에서 상정된 폴리펩타이드를 포함하는 세포를 제공한다.
- [0075] 추가적인 실시형태에서, 본 개시내용은 부분적으로 본 명세서에서 상정된 폴리펩타이드를 암호화하는 폴리뉴클레오타이드를 포함하는 세포를 제공한다.
- [0076] 추가 실시형태에서, 본 개시내용은, 부분적으로, 본 명세서에서 상정된 백터를 포함하는 세포를 제공한다.
- [0077] 일부 실시형태에서, 본 개시내용은 부분적으로 본 명세서에서 상정된 폴리펩타이드에 의해 도입된 하나 이상의 계놈 변형을 포함하는 세포를 제공한다.
- [0078] 특정 실시형태에서, 세포는 면역효능 인핸서, 면역억제 신호 램퍼, 또는 조작된 항원 수용체 중 하나 이상을 암호화하는 폴리뉴클레오타이드를 포함한다.
- [0079] 추가 실시형태에서, 상기 폴리뉴클레오타이드는 면역효능 인핸서, 면역억제 신호 램퍼 또는 조작된 항원 수용체를 암호화하는 폴리뉴클레오타이드에 작동 가능하게 연결된 RNA 중합효소 II 프로모터를 더 포함한다.
- [0080] 일부 실시형태에서, RNA 중합효소 II 프로모터는 짧은 EF1 $\alpha$  프로모터, 긴 EF1 $\alpha$  프로모터, 인간 ROSA 26 좌위, 유비퀴틴 C(UBC) 프로모터, 포스포글리세레이트 키나제-1(PGK) 프로모터, 거대세포바이러스 인핸서/닭  $\beta$ -액틴(CAG) 프로모터,  $\beta$ -액틴 프로모터 및 콜수증식성 육종 바이러스 인핸서, 결실된 음성 대조 영역(negative control region deleted), dl1587rev 프라이머-결합 부위 치환된(MND) 프로모터로 이루어진 군으로부터 선택된다.
- [0081] 추가적인 실시형태에서, 폴리뉴클레오타이드는 면역효능 인핸서, 면역억제 신호 램퍼, 또는 조작된 항원 수용체에 작동 가능하게 연결되고/되거나, 이들 사이에 개재되고/되거나 측첩되는 하나 이상의 자기-절단 바이러스 웨타이드를 더 암호화한다.
- [0082] 특정 실시형태에서, 자기-절단 바이러스 웨타이드는 2A 웨타이드이다.
- [0083] 소정의 실시형태에서, 폴리뉴클레오타이드는 추가로 이종성 폴리아데닐화 신호를 포함한다.
- [0084] 특정 실시형태에서, 면역억제 신호 램퍼는 면역억제 인자를 상쇄시키는(counteract) 효소 작용을 포함한다.
- [0085] 추가적인 실시형태에서, 면역억제 신호 램퍼는 키뉴레니나제 활성을 포함한다.
- [0086] 특정 실시형태에서, 면역억제 신호 램퍼는 면역억제 인자에 결합하는 액소도메인으로서, 선택적으로 상기 액소도메인은 항체 또는 이의 항원 결합 단편인, 상기 액소도메인; 면역억제 인자에 결합하는 액소도메인 및 막관통도메인; 또는 면역억제 인자에 결합하는 액소도메인, 막관통 도메인 및 세포에 면역억제 신호를 전달할 수 없는 변형된 엔도도메인을 포함한다.

- [0087] 추가 실시형태에서, 면역효능 인핸서는 이중특이성 T 세포 관여자 분자(BiTE), 면역강화 인자 및 플립 수용체 (flip receptor)로 이루어진 군으로부터 선택된다.
- [0088] 일부 실시형태에서, 면역강화 인자는 사이토카인, 케모카인, 사이토톡신, 사이토카인 수용체 및 이들의 변이체로 이루어진 군으로부터 선택된다.
- [0089] 일부 실시형태에서, 플립 수용체는 TGF $\beta$ RII 엑소도메인 및 막관통 도메인; 및 상기 TGF $\beta$ RII 막관통 도메인의 C-말단 단부에 프레임 내 융합된 TLR4, CD28, CD134, CD137, CD278 및/또는 CD3  $\zeta$ 로부터의 엔도도메인을 포함한다.
- [0090] 특정 실시형태에서, 플립 수용체는 TGF $\beta$ RII 엑소도메인; TLR4, CD3, CD4, CD8  $\alpha$ , CD28, CD134 또는 CD137 폴리펩타이드로부터 단리된 막관통 도메인; 및 상기 TGF $\beta$ RII 엑소도메인의 C-말단 단부에 프레임 내 융합된 TLR4, CD28, CD134, CD137, CD278 및/또는 CD3  $\zeta$ 로부터의 엔도도메인을 포함한다.
- [0091] 소정의 실시형태에서, 플립 수용체는 TGF $\beta$ RII 엑소도메인; 및 상기 TGF $\beta$ RII 엑소도메인의 C-말단 단부에 프레임 내 융합된 TLR4, CD3, CD4, CD8  $\alpha$ , CD28, CD134 또는 CD137 폴리펩타이드로부터 단리된 막관통 도메인 및 엔도도메인을 포함한다.
- [0092] 특정 실시형태에서, 조작된 항원 수용체는 조작된 TCR, CAR, Daric 또는 제타카인으로 이루어진 군으로부터 선택된다.
- [0093] 추가적인 실시형태에서, 조작된 수용체는 상기 CBLB 유전자 내로 통합되지 않는다.
- [0094] 추가 실시형태에서, 면역효능 인핸서, 면역억제 신호 댐퍼 또는 조작된 항원 수용체 중 하나 이상을 암호화하는 상기 폴리뉴클레오타이드는 상기 CBLB 유전자 내로 통합된다.
- [0095] 추가 실시형태에서, 면역효능 인핸서, 면역억제 신호 댐퍼, 또는 조작된 항원 수용체 중 하나 이상을 암호화하는 폴리뉴클레오타이드를 포함하는 공여자 수선 주형은 본 명세서에서 상정된 폴리펩타이드에 의해 도입된 DNA 이중 가닥 과손 부위에서 CBLB 유전자 내로 통합된다.
- [0096] 특정 실시형태에서, 세포는 조혈세포이다.
- [0097] 일부 실시형태에서, 세포는 T 세포이다.
- [0098] 특정 실시형태에서, 세포는 CD3+, CD4+ 및/또는 CD8+ 세포이다.
- [0099] 추가적인 실시형태에서, 세포는 면역 효과기 세포이다.
- [0100] 특정 실시형태에서, 세포는 세포독성 T 림프구(cytotoxic T lymphocyte: CTL), 종양 침윤성 림프구(TIL), 또는 헬퍼 T 세포이다.
- [0101] 특정 실시형태에서, 세포는 자연 살해(natural killer: NK) 세포 또는 자연 살해 T(natural killer T: NKT) 세포이다.
- [0102] 바람직한 실시형태에서, 세포는 조작된 항원 수용체를 발현시키도록 유전자 변형된 T 세포이다.
- [0103] 바람직한 실시형태에서, 세포는 키메라 항원 수용체(CAR) 또는 조작된 T 세포 수용체(TCR)를 발현시키기 위해 유전자 조작된 T 세포이다.
- [0104] 추가적인 실시형태에서, 세포의 공급원은 말초혈액 단핵세포, 골수, 림프절 조직, 제대혈, 흉선 조직, 감염 부위로부터의 조직, 복수, 흉수, 비장 조직, 또는 종양이다.
- [0105] 일부 실시형태에서, 세포는 하나 이상의 변형된 CBLB 대립유전자를 포함한다.
- [0106] 추가 실시형태에서, 하나 이상의 변형된 CBLB 대립유전자는 비기능성이거나 또는 실질적으로 감소된 CBLB 기능 및/또는 활성을 가진다.
- [0107] 특정 실시형태에서, 세포는 하나 이상의 변형된 CBLB 대립유전자 내로 도입된 면역효능 인핸서 또는 면역억제 신호 댐퍼를 암호화하는 핵산을 포함하고, 상기 세포는 하나 이상의 변형된 CBLB 대립유전자 내로 도입되지 않는 조작된 항원 수용체를 추가로 포함한다.
- [0108] 일부 실시형태에서, 본 개시내용은, 부분적으로 본 명세서에서 상정된 하나 이상의 세포를 포함하는 복수의 세포를 제공한다.

- [0109] 다양한 실시형태에서, 본 개시내용은, 부분적으로 본 명세서에서 상정된 하나 이상의 세포를 포함하는 조성물을 제공한다.
- [0110] 일부 실시형태에서, 본 개시내용은, 부분적으로 본 명세서에서 상정된 하나 이상의 세포 및 생리적으로 허용 가능한 단체를 포함하는 조성물을 제공한다.
- [0111] 특정 실시형태에서, 본 개시내용은 본 명세서에서 상정된 폴리펩타이드를 암호화하는 폴리뉴클레오타이드를 상기 세포 내로 도입하는 단계를 포함하되, 상기 폴리펩타이드의 발현은 인간 CBLB 유전자 내 표적 부위에서 이중 가닥 파손을 생성하는, 세포에서 인간 CBLB 유전자를 편집하는 방법을 제공한다.
- [0112] 일부 실시형태에서, 본 개시내용은 본 명세서에서 상정된 폴리펩타이드를 암호화하는 폴리뉴클레오타이드를 상기 세포 내로 도입하는 단계를 포함하되, 상기 폴리펩타이드의 발현은 인간 CBLB 유전자 내 표적 부위에서 이중 가닥 파손을 생성하고, 파손은 비상동성 말단 결합(non-homologous end joining: NHEJ)에 의해 수선되는, 세포에서 인간 CBLB 유전자를 편집하는 방법을 제공한다.
- [0113] 특정 실시형태에서, 본 개시내용은 본 명세서에서 상정된 폴리펩타이드를 암호화하는 폴리뉴클레오타이드 및 공여자 수선 주형을 상기 세포 내로 도입하는 단계를 포함하되, 상기 폴리펩타이드의 발현은 인간 CBLB 유전자 내 표적 부위에서 이중 가닥 파손을 생성하고, 상기 공여자 수선 주형은 상기 이중-가닥 파손(double-strand break: DSB) 부위에서 상동 직접 수선(homology directed repair: HDR)에 의해 인간 CBLB 유전자 내로 흔입되는, 인간 CBLB 유전자를 편집하는 방법을 제공한다.
- [0114] 특정 실시형태에서, 세포는 조혈세포이다.
- [0115] 소정의 실시형태에서, 세포는 T 세포이다.
- [0116] 소정의 실시형태에서, 세포는 CD3+, CD4+ 및/또는 CD8+ 세포이다.
- [0117] 일부 실시형태에서, 세포는 면역 효과기 세포이다.
- [0118] 추가적인 실시형태에서, 세포는 세포독성 T 림프구(CTL), 종양 침윤성 림프구(TIL) 또는 헬퍼 T 세포이다.
- [0119] 추가적인 실시형태에서, 세포는 자연 살해(NK) 세포 또는 자연 살해 T(NKT) 세포이다.
- [0120] 특정 실시형태에서, 세포의 공급원은 말초혈액 단핵세포, 골수, 림프절 조직, 제대혈, 흉선 조직, 감염 부위로 부터의 조직, 복수, 흉수, 비장 조직, 또는 종양이다.
- [0121] 특정 실시형태에서, 폴리펩타이드를 암호화하는 폴리뉴클레오타이드는 mRNA이다.
- [0122] 소정의 실시형태에서, 5'-3' 엑소뉴클레아제를 암호화하는 폴리뉴클레오타이드는 상기 세포 내로 도입된다.
- [0123] 추가 실시형태에서, Trex2를 암호화하는 폴리뉴클레오타이드 또는 이의 생물학적 활성 단편은 세포 내로 도입된다.
- [0124] 일부 실시형태에서, 공여자 수선 주형은 상기 야생형 CBLB 유전자에 비해 하나 이상의 돌연변이를 포함하는 CBLB 유전자 또는 이의 일부를 암호화한다.
- [0125] 특정 실시형태에서, 공여자 수선 주형은 면역효능 인핸서, 면역억제 신호 램퍼 또는 조작된 항원 수용체 중 하나 이상을 암호화한다.
- [0126] 일부 실시형태에서, 공여자 수선 주형은 면역효능 인핸서, 면역억제 신호 램퍼 또는 조작된 항원 수용체에 작동 가능하게 연결된 RNA 중합효소 II 프로모터를 추가로 포함한다.
- [0127] 추가 실시형태에서, RNA 중합효소 II 프로모터는 짧은 EF1 $\alpha$  프로모터, 긴 EF1 $\alpha$  프로모터, 인간 ROSA 26 좌위, 유비퀴틴 C(UBC) 프로모터, 포스포글리세레이트 키나제-1(PGK) 프로모터, 거대세포바이러스 인핸서/답  $\beta$ -액틴(CAG) 프로모터,  $\beta$ -액틴 프로모터 및 골수증식성 육종 바이러스 인핸서, 결실된 음성 대조 영역, d1587rev 프라이머-결합 부위 치환된(MND) 프로모터로 이루어진 군으로부터 선택된다.
- [0128] 특정 실시형태에서, 공여자 수선 주형은 면역효능 인핸서, 면역억제 신호 램퍼, 또는 조작된 항원 수용체에 작동 가능하게 연결되고/되거나, 이들 사이에 개재되고/되거나 측접되는 하나 이상의 자기-절단 바이러스 웨بت아이드를 더 암호화한다.
- [0129] 추가적인 실시형태에서, 자기-절단 바이러스 웨بت아이드는 2A 웨بت아이드이다.

- [0130] 소정의 실시형태에서, 공여자 수선 주형은 이중성 폴리아데닐화 신호를 더 포함한다.
- [0131] 추가 실시형태에서, 면역억제 신호 램퍼는 면역억제 인자를 상쇄시키는 효소 작용을 포함한다.
- [0132] 특정 실시형태에서, 면역억제 신호 램퍼는 키뉴레니나제 활성을 포함한다.
- [0133] 일부 실시형태에서, 면역억제 신호 램퍼는 면역억제 인자에 결합하는 엑소도메인으로서, 선택적으로 상기 엑소도메인은 항체 또는 이의 항원 결합 단편인, 상기 엑소도메인; 면역억제 인자에 결합하는 엑소도메인 및 막관통도메인; 또는 면역억제 인자에 결합하는 엑소도메인, 막관통 도메인, 및 세포에 면역억제 신호를 전달할 수 없는 변형된 엔도도메인을 포함한다.
- [0134] 소정의 실시형태에서, 면역억제 신호 램퍼의 엑소도메인 및/또는 막관통 도메인은 TGF $\beta$ RII 엑소도메인 및/또는 막관통 도메인이다.
- [0135] 일부 실시형태에서, 면역효능 인핸서는 이중특이성 T 세포 관여자 분자(BiTE), 면역강화 인자 및 플립 수용체로 이루어진 군으로부터 선택된다.
- [0136] 특정 실시형태에서, 면역강화 인자는 사이토카인, 케모카인, 사이토톡신, 사이토카인 수용체 및 이들의 변이체로 이루어진 군으로부터 선택된다.
- [0137] 특정 실시형태에서, 플립 수용체는 TGF $\beta$ RII 엑소도메인 및 막관통 도메인; 및 상기 TGF $\beta$ RII 엑소도메인의 C-말단 단부에 프레임 내 융합된 TLR4, CD28, CD134, CD137, CD278 및/또는 CD3 $\zeta$ 로부터의 엔도도메인을 포함한다.
- [0138] 특정 실시형태에서, 플립 수용체는 TGF $\beta$ RII 엑소도메인; TLR4, CD3, CD4, CD8 $\alpha$ , CD28, CD134 또는 CD137 폴리펩타이드로부터 단리된 막관통 도메인; 및 상기 TGF $\beta$ RII 엑소도메인의 C-말단 단부에 프레임 내 융합된 TLR4, CD28, CD134, CD137, CD278 및/또는 CD3 $\zeta$ 로부터의 엔도도메인을 포함한다.
- [0139] 추가 실시형태에서, 플립 수용체는 TGF $\beta$ RII 엑소도메인; 및 상기 TGF $\beta$ RII 엑소도메인의 C-말단 단부에 프레임 내 융합된 TLR4, CD3, CD4, CD8 $\alpha$ , CD28, CD134 또는 CD137 폴리펩타이드로부터 단리된 막관통 도메인 및 엔도도메인을 포함한다.
- [0140] 소정의 실시형태에서, 조작된 항원 수용체는 조작된 TCR, CAR, Daric 또는 제타카인으로 이루어진 군으로부터 선택된다.
- [0141] 일부 실시형태에서, 공여자 수선 주형은 상기 DSB의 인간 CBLB 유전자 서열 5'에 대해 상동성인 5' 상동성 아암 및 상기 DSB의 인간 CBLB 유전자 서열 3'에 대해 상동성인 3' 상동성 아암을 포함한다.
- [0142] 추가 실시형태에서, 5' 및 3' 상동성 아암의 상기 길이는 약 100bp 내지 약 2500bp로부터 독립적으로 선택된다.
- [0143] 특정 실시형태에서, 5' 및 3' 상동성 아암의 상기 길이는 약 600bp 내지 약 1500bp로부터 독립적으로 선택된다.
- [0144] 특정 실시형태에서, 5' 상동성 아암은 약 1500bp이고, 상기 3' 상동성 아암은 약 1000bp이다.
- [0145] 일부 실시형태에서, 5' 상동성 아암은 약 600bp이고, 상기 3' 상동성 아암은 약 600bp이다.
- [0146] 소정의 실시형태에서, 바이러스 벡터는 상기 공여자 수선 주형을 상기 세포 내로 도입하기 위해 사용된다.
- [0147] 추가적인 실시형태에서, 바이러스 벡터는 재조합 아데노-연관 바이러스 벡터(rAAV) 또는 레트로바이러스이다.
- [0148] 특정 실시형태에서, rAAV는 AAV2로부터의 하나 이상의 ITR을 가진다.
- [0149] 추가적인 실시형태에서, rAAV는 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9 및 AAV10로 이루어진 군으로부터 선택된 혈청형을 가진다.
- [0150] 추가 실시형태에서, rAAV는 AAV2 또는 AAV6 혈청형을 가진다.
- [0151] 소정의 실시형태에서, 레트로바이러스는 렌티바이러스이다.
- [0152] 일부 실시형태에서, 렌티바이러스는 인테그라제 결핍 렌티바이러스(IDLV)이다.
- [0153] 다양한 실시형태에서, 본 개시내용은 대상체에게 본 명세서에서 상정된 유효량의 조성물을 투여하는 단계를 포함하는, 암, 감염성 질환, 자가면역 질환, 염증성 질환 및 면역결핍 또는 이와 관련된 병태의 적어도 하나의 증상을 치료하거나, 예방하거나 또는 개선시키는 방법을 부분적으로 제공한다.

- [0154] 소정의 실시형태에서, 본 개시내용은 대상체에게 유효량의 본 명세서에서 상정된 조성물을 투여하는 단계를 포함하는 고형암을 치료하는 방법을 제공한다.
- [0155] 특정 실시형태에서, 고형암은 간암, 췌장암, 폐암, 유방암, 난소암, 전립선암, 고환암, 방광암, 뇌암, 육종, 두경부암, 골암, 갑상선암, 신장암 또는 피부암을 포함한다.
- [0156] 일부 실시형태에서, 본 개시내용은 대상체에게 유효량의 본 명세서에서 상정된 조성물을 투여하는 단계를 포함하는 혈액학적 악성 종양을 치료하는 방법을 상정한다.
- [0157] 소정의 실시형태에서, 혈액학적 악성종양은 백혈병, 림프종 또는 다발성 골수종이다.

### 도면의 간단한 설명

- [0158] 도 1은 CBLB 유전자, 및 SH2 도메인을 암호화하는 엑손 6 내 HE 표적 부위(서열번호 20 및 76)를 도시한 도면.
- 도 2는 CBLB HE가 분류의 3회 라운드 내내 키메라 '절반-부위'에 대한 NTD 및 CTD의 조작을 통해 재프로그래밍된 다음, 재프로그래밍된 도메인의 융합 및 완전히 프로그래밍된 HE를 단리시키기 위해 완전한 CBLB 표적 부위에 대한 선별 방법을 도시한 도면.
- 도 3은 염색체 리포터 분석에서 CBLB HE 변이체의 활성을 도시한 도면.
- 도 4는 CBLB.E3 HE 변이체의 효모 표면 친화도 적정을 도시한 도면.
- 도 5는 야생형 I-OnuI 단백질(서열번호 1)에 대한 CBLB HE 변이체(서열번호 77-83)의 정렬을 도시한 도면(동일하지 않은 위치는 강조함).
- 도 6은 CBLB.E3 메가TAL를 생성하기 위해 CBLB.E3 HE 변이체에 융합된 TAL RVD에 대한 결합 부위(서열번호 22 및 84)를 도시한 도면.
- 도 7은 CBLB 메가TAL를 이용하는 CAR T 세포에서 일치되는 CBLB 유전자 편집률을 도시한 도면.
- 도 8은 CBLB 메가TAL로 처리된 T 세포에서의 감소된 세포내 CBLB 단백질 발현을 도시한 도면.
- 도 9는 CBL-편집된 T 세포가 항-CD3 항체에 의해, 또는 CAR 특이적 항원에 의해 자극될 때 더 많은 IFN $\gamma$  및 TNF $\alpha$ 를 생산한다는 것을 도시한 도면.
- 도 10은 다양한 CBLB 메가TAL 및 TREX2를 암호화하는 mRNA의 공동전달에 의해 편집된 T 세포의 TIDE 분석을 도시한 도면. 표적 좌위에서의 편집률은 61 내지 93%였다.
- 도 11은 CBLB 메가TAL 및 AAV 공여자를 이용하여 MND-GFP 발현 카세트를 CBLB 표적 부위에 삽입하는 HDR 전략의 다이어그램을 도시한 도면.
- 도 12는 모의(mock) 편집된 T 세포에 비해 도 11에 도시된 HDR 전략을 이용하여 편집된 T 세포에서의 GFP 발현의 안정한 증가를 도시한 도면.
- 도 13은 생체내 마우스 연구 결과를 도시한 도면. A549 종양을 갖는 마우스를 비히를 처리된 T 세포(상부 좌측 패널), CBLB 메가TAL로 편집된 T 세포(상부 우측 패널), 항-EGFR CAR T 세포(하부 좌측 패널) 및 CBLB 메가TAL로 편집된 항-EGFR CAR T 세포(하부 우측 패널)로 처리하였다. 시간에 따라 평균 종양 용적을 측정하였다.

### 발명을 실시하기 위한 구체적인 내용

- [0159] 서열 식별자의 간단한 설명
- [0160] 서열번호 1은 야생형 I-OnuI LAGLIDADG 호밍 엔도뉴클레아제(LHE)의 아미노산 서열이다.
- [0161] 서열번호 2는 야생형 I-OnuI LHE의 아미노산 서열이다.
- [0162] 서열번호 3은 야생형 I-OnuI LHE의 생물학적 활성 단편의 아미노산 서열이다.
- [0163] 서열번호 4은 야생형 I-OnuI LHE의 생물학적 활성 단편의 아미노산 서열이다.
- [0164] 서열번호 5는 야생형 I-OnuI LHE의 생물학적 활성 단편의 아미노산 서열이다.
- [0165] 서열번호 6 내지 12는 인간 CBLB 유전자에서 표적 부위에 결합하고 절단하도록 재프로그래밍된 I-OnuI LHE 변이

체의 아미노산 서열을 제시한다.

[0166] 서열번호 13 내지 19는 인간 CBLB 유전자에서 표적 부위에 결합하고 절단하도록 재프로그래밍된 I-OnuI LHE 변이체의 아미노산 서열을 제시한다.

[0167] 서열번호 20은 인간 CBLB 유전자의 엑손 6에서 I-OnuI LHE 변이체 표적 부위이다.

[0168] 서열번호 21은 인간 CBLB 유전자의 엑손 6에서 TALE DNA 결합 도메인 표적 부위이다.

[0169] 서열번호 22는 인간 CBLB 유전자의 엑손 6에서 메가TAL 표적 부위이다.

[0170] 서열번호 23, 25, 27 및 29는 인간 CBLB 유전자의 엑손 6에서 I-OnuI LHE 변이체 N-말단 도메인 표적 부위를 제시한다.

[0171] 서열번호 24, 26 및 28은 인간 CBLB 유전자의 엑손 6에서 I-OnuI LHE 변이체 C-말단의 도메인 표적 부위를 제시한다.

[0172] 서열번호 30은 CBLB.E3 표면 디스플레이 플라스미드의 폴리뉴클레오타이드 서열이다.

[0173] 서열번호 31 내지 36은 CBLB 메가TAL을 암호화하는 mRNA 서열을 제시한다.

[0174] 서열번호 37은 뮤린 Trex2를 암호화하는 mRNA 서열이다.

[0175] 서열번호 38은 뮤린 Trex2를 암호화하는 아미노산 서열이다.

[0176] 서열번호 39 내지 49는 다양한 링커의 아미노산 서열을 제시한다.

[0177] 서열번호 50 내지 74는 프로테아제 절단 부위 및 자기-절단성 폴리펩타이드 절단 부위의 아미노산 서열을 제시한다. 앞서 언급한 서열에서, X는 존재한다면, 임의의 아미노산 또는 아미노산의 부재를 지칭한다.

#### A. 개요

[0179] 본 개시내용은 일반적으로, 부분적으로, 개선된 게놈 편집 조성을 및 이의 사용 방법에 관한 것이다. 임의의 특정 이론에 의해 결합되는 일 없이, 다양한 실시형태에서 상정된 게놈 편집 조성을 암, 감염성 질환, 자가면역 질환, 염증성 질환 및 면역결핍, 또는 이들과 관련된 병태를 예방하거나, 또는 이들 중 적어도 하나의 증상을 개선시키기 위해 사용될 수 있다. 존재하는 적응 세포 요법을 방해하는 한 가지 제한 또는 문제는 종양 미세환경에 의해 매개되는 고갈에 기인하여 면역 효과기 세포의 저반응성(hyporesponsiveness)이다. 고갈된 T 세포는 천연, 효과기 또는 기억 T 세포와 현저히 별개인 독특한 분자 서명을 가진다. 고갈된 T 세포는 감소된 사이토카인 발현 및 효과기 기능을 갖는 T 세포이다.

[0180] 카시타스 B-계통(Cbl) 림프종 원발암유전자 B(CBLB)는 RING-핑거 패밀리 또는 E3 유비퀴틴 계통의 구성원이고, 조혈 계통 세포를 포함하는 다양한 조직 및 세포 유형에서 발현된다. CBLB는 E2-유비퀴틴 접합 효소의 그의 RING 핑거 도메인에 대한 보충을 통해 표적 기질 단백질의 유비퀴틴화를 용이하게 한다. CBLB 기질 단백질의 유비퀴틴화는 단백질 분해를 촉진시키거나 또는 단백질-단백질 상호작용을 방해할 수 있다. CBLB는 그의 SH2-함유 타이로신 키나제 결합 단백질, SH3 도메인-함유 단백질과 상호작용하는 그의 단백질-풍부 서열, 또는 유비퀴틴-태그된 단백질과 상호작용하는 그의 유비퀴틴-연관 도메인을 통해 기질 단백질에 결합한다.

[0181] CBLB는 또한 효과기 T 세포 활성 및 지속의 음성 조절에 연루된다. CBLB 네아웃 마우스 T 세포는 과증식성이고, 항원 자극에 반응하여 IL2 및 IFN $\gamma$ 의 강조된 수준을 생성하며, TGF $\beta$ -매개된 억제에 내성이 있고, 더 낮은 활성화 역치를 갖는데, 이는 T 세포 활성화를 음성 조절하는 데 어떤 역할을 한다는 것을 나타낸다. 임의의 특정 이론에 의해 구속되는 일 없이, 조작된 뉴클레아제를 이용하는 T 세포에서의 CBLB 유전자 붕괴는 더 효능있고 지속적인 양자 세포 면역요법(adoptive cellular immunotherapy)을 초래할 것임이 상정된다.

[0182] 특정 실시형태에서, 본 명세서에서 상정된 게놈 편집된 면역 효과기 세포는 CBLB 발현, CBLB 활성 및/또는 CBLB 를 통한 신호전달을 제거하거나, 감소시키거나 또는 감소시킴으로써 고갈에 대해 더 내성으로 만들어진다.

[0183] 다양한 실시형태에서 상정된 게놈 편집 조성을 및 방법은 카시타스 B-계통(Cbl) 림프종 원발암유전자 B(CBLB) 유전자 내 표적 부위에 결합하고 이를 절단하도록 설계된 뉴클레아제 변이체를 포함한다. 특정 실시형태에서 상정된 뉴클레아제 변이체는 폴리뉴클레오타이드 주형, 예를 들어, 공여자 수선 주형의 부재 하에 비상동성 말단 결합(NHEJ)에 의해, 또는 공여자 수선 주형의 존재 하에 상동 직접 수선(HDR), 즉, 상동성 재조합에 의해 수선 될 수 있는 표적 폴리뉴클레오타이드 서열에서 이중-가닥 파손을 도입하기 위해 사용될 수 있다. 특정 실시형태

에서 상정된 뉴클레아제 변이체는 공여자 수선 주형의 존재 하에 세포의 염기-절단-수선(base-excision-repair: BER) 기작 또는 상동성 재조합을 이용하여 수선될 수 있는 단일-가닥 DNA 파손을 생성하는 틈내기효소로서 설계될 수 있다. NHEJ는 유전자 기능을 파괴하는 작은 삽입 및 결실의 제형을 빈번하게 초래하는 오류-유발 과정이다. 상동성 재조합은 수선을 위한 주형으로서 상동성 DNA를 필요로 하며, 표적 부위에 측접하는 영역에 대해 상동성을 보유하는 서열에 의해 측면 중 하나 상에 측접된, 표적 부위에서 목적으로 하는 서열을 함유하는 공여자 DNA의 도입에 의해 구체화되는 제한 없는 다양한 변형을 생성하도록 영향력을 가질 수 있다.

- [0184] 하나의 바람직한 실시형태에서, 본 명세서에서 상정된 계놈 편집 조성물은 인간 CBLB 유전자를 표적화하는 호밍 엔도뉴클레아제 변이체 또는 메가TAL를 포함한다.
- [0185] 하나의 바람직한 실시형태에서, 본 명세서에서 상정된 계놈 편집 조성물은 호밍 엔도뉴클레아제 변이체 또는 메가TAL 및 말단-가공 효소, 예를 들어, Trex2를 포함한다.
- [0186] 다양한 실시형태에서, 계놈 편집된 세포가 상정된다. 계놈 편집된 세포는 편집된 CBLB 유전자를 포함하되, 편집 전략은 CBLB 발현을 감소 또는 제거하도록 설계된다. 특정 실시형태에서, CAR T 세포, 조작된 TCR T 세포 또는 DARIC T 세포는 편집된 CBLB 유전자를 포함한다.
- [0187] 다양한 실시형태에서, DNA 파손은 T 세포, 예를 들어, 면역 효과기 세포 내 CBLB 유전자의 표적 부위에서 생성되고, 절단된 계놈 서열 말단의 NHEJ는 CBLB 발현이 거의 없거나 또는 전혀 없는 세포, 및 바람직하게는 작용성 CBLB 발현 및/또는 신호전달이 결여되거나 또는 실질적으로 결여된, 예를 들어, T 세포 고갈을 증가시키는 능력이 결여된 T 세포를 초래할 수 있다. 임의의 특정 이론에 의해 결합되는 일 없이, 작용성 CBLB 발현이 결여된 T 세포는 면역억제 및 T 세포 고갈에 대해 더 내성이 있고, 따라서, 더 지속적이고 치료적으로 효능이 있다.
- [0188] 다양한 다른 실시형태에서, 절단된 CBLB 계놈 서열의 수선을 위한 공여자 주형이 제공된다. CBLB 유전자는 DNA 파손-부위에서 상동성 재조합에 의해 주형의 서열로 수선된다. 특정 실시형태에서, 수선 주형은 작용성 CBLB 발현을 파괴하고, 바람직하게는 실질적으로 감소시키거나 또는 제거하는 폴리뉴클레오타이드 서열을 포함한다.
- [0189] 특정 실시형태에서, CBLB 유전자는 면역효능 인핸서, 면역억제 신호 램퍼 또는 조작된 항원 수용체를 암호화하는 폴리뉴클레오타이드로 수선된다.
- [0190] 특정 실시형태에서, CBLB 유전자는 면역효능 인핸서, 면역억제 신호 램퍼 또는 조작된 항원 수용체를 암호화하는 폴리뉴클레오타이드로 수선되고, 면역효능 인핸서, 면역억제 신호 램퍼 또는 조작된 항원 수용체의 발현을 전사적으로 제어하도록 내인성 CBLB 프로모터를 채택하기 위해 CBLB 유전자 내로 도입된다.
- [0191] 바람직한 실시형태에서, 본 명세서에서 상정된 계놈 편집 조성물 및 방법은 인간 CBLB 유전자를 편집하기 위해 사용된다.
- [0192] 따라서, 본 명세서에서 상정된 방법 및 조성물은 존재하는 적응 세포 요법에 비해 다수의 개선을 나타낸다.
- [0193] 재조합(즉, 조작된) DNA에 대한 기법, 웨타이드 및 올리고뉴클레오타이드 합성, 면역분석, 조직 배양, 형질전환(예를 들어, 전기천공법, 리포택션), 효소 반응, 정제 및 관련된 기법 및 절차는 본 명세서 전체적으로 인용되고 논의된 바와 같은 미생물, 분자 생물학, 생화학, 분자 유전학, 세포 생물학, 바이러스학 및 면역학에서 다양한 일반적이고 더 구체적인 참고문헌에 기재되는 바와 같이 일반적으로 수행될 수 있다. 예를 들어, 문헌 [Sambrook et al., Molecular Cloning: A Laboratory Manual, 3d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; Current Protocols in Molecular Biology (John Wiley and Sons, updated July 2008); Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology, Greene Pub. Associates and Wiley-Interscience; Glover, DNA Cloning: A Practical Approach, vol. I & II (IRL Press, Oxford Univ. Press USA, 1985); Current Protocols in Immunology (Edited by: John E. Coligan, Ada M. Kruisbeek, David H. Margulies, Ethan M. Shevach, Warren Strober 2001 John Wiley & Sons, NY, NY); Real-Time PCR: Current Technology and Applications, Edited by Julie Logan, Kirstin Edwards and Nick Saunders, 2009, Caister Academic Press, Norfolk, UK; Anand, Techniques for the Analysis of Complex Genomes, (Academic Press, New York, 1992); Guthrie and Fink, Guide to Yeast Genetics and Molecular Biology (Academic Press, New York, 1991); Oligonucleotide Synthesis (N. Gait, Ed., 1984); Nucleic Acid The Hybridization (B. Hames & S. Higgins, Eds., 1985); Transcription and Translation (B. Hames & S. Higgins, Eds., 1984); Animal Cell Culture (R. Freshney, Ed., 1986); Perbal, A Practical Guide to Molecular Cloning (1984); Next-Generation Genome Sequencing (Janitz, 2008 Wiley-VCH); PCR Protocols (Methods in Molecular Biology) (Park, Ed., 3rd Edition, 2010 Humana Press);

Immobilized Cells And Enzymes (IRL Press, 1986); the treatise, Methods In Enzymology (Academic Press, Inc., N.Y.); Gene Transfer Vectors For Mammalian Cells (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); Harlow and Lane, Antibodies, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1998); Immunochemical Methods In Cell And Molecular Biology (Mayer and Walker, eds., Academic Press, London, 1987); Handbook Of Experimental Immunology, Volumes I-IV (D. M. Weir and CC Blackwell, eds., 1986); Roitt, Essential Immunology, 6th Edition, (Blackwell Scientific Publications, Oxford, 1988); Current Protocols in Immunology (Q. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach and W. Strober, eds., 1991); Annual Review of Immunology];뿐만 아니라 [Advances in Immunology]과 같은 저널 내의 모노그래프를 참조.

[0194]

B. 정의

[0195]

본 개시내용을 더 상세하게 제시하기 전에, 본 명세서에서 사용될 소정의 용어의 정의를 제공하는 것은 본 개시 내용의 이해에 도움을 줄 수 있다.

[0196]

달리 정의되지 않는 한, 본 명세서에서 사용되는 모든 기술적 및 과학적 용어는 본 발명이 속하는 기술분야의 당업자에 의해 통상적으로 이해되는 것과 동일한 의미를 가진다. 본 명세서에 기재된 것과 유사한 또는 동등한 임의의 방법 및 물질이 특정 실시형태의 실행 또는 시험에서 사용될 수 있지만, 조성물, 방법 및 물질의 바람직한 실시형태가 본 명세서에 기재된다. 본 개시내용의 목적을 위해, 다음의 용어를 이하에 정의한다. 추가적인 정의를 본 개시내용 전체적으로 제시한다.

[0197]

단수의 항목은 본 명세서에서 항목의 문법적 대상의 하나 이상(즉, 적어도 하나, 또는 하나 이상)을 지칭하기 위해 사용된다. 예로서, "요소"는 하나의 요소 또는 하나 이상의 요소를 의미한다.

[0198]

대안의(예를 들어, "또는")의 사용은 대안 중 하나, 둘 다, 또는 이들의 임의의 조합을 의미하는 것으로 이해되어야 한다.

[0199]

용어 "및/또는"은 대안 중 하나, 또는 둘 다를 의미하는 것으로 이해되어야 한다.

[0200]

본 명세서에서 사용되는 용어 "약" 또는 "대략"은 기준 양, 수준, 값, 숫자, 빈도, 백분율, 치수, 크기, 양, 중량 또는 길이에 대해 15%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2% 또는 1%만큼 변하는 기준 양, 수준, 값, 숫자, 빈도, 백분율, 치수, 크기, 양, 중량 또는 길이를 지칭한다. 일 실시형태에서, 용어 "약" 또는 "대략"은 기준 양, 수준, 값, 숫자, 빈도, 백분율, 치수, 크기, 양, 중량 또는 길이에 대해 ± 15%, ± 10%, ± 9%, ± 8%, ± 7%, ± 6%, ± 5%, ± 4%, ± 3%, ± 2% 또는 ± 1%의 기준 양, 수준, 값, 숫자, 빈도, 백분율, 치수, 크기, 양, 중량 또는 길이의 범위를 지칭한다.

[0201]

일 실시형태에서, 범위, 예를 들어, 1 내지 5, 약 1 내지 5, 또는 약 1 내지 약 5는 범위에 의해 포함되는 각각의 수치적 값을 지칭한다. 예를 들어, 하나의 비제한적 그리고 단지 예시적인 실시형태에서, 범위 "1 내지 5"는 표현 1, 2, 3, 4, 5; 또는 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 또는 5.0; 또는 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9 또는 5.0과 동등하다.

[0202]

본 명세서에서 사용되는 용어 "실질적으로"는 기준 양, 수준, 값, 숫자, 빈도, 백분율, 치수, 크기, 양, 중량 또는 길이에 비해 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 이상인 기준 양, 수준, 값, 숫자, 빈도, 백분율, 치수, 크기, 양, 중량 또는 길이를 지칭한다. 일 실시형태에서, "실질적으로 동일한"은 기준 양, 수준, 값, 숫자, 빈도, 백분율, 치수, 크기, 양, 중량 또는 길이와 거의 동일한 효과, 예를 들어, 생리적 효과를 생산하는 기준 양, 수준, 값, 숫자, 빈도, 백분율, 치수, 크기, 양, 중량 또는 길이를 지칭한다.

[0203]

본 명세서 전체적으로, 문맥에서 달리 요구되지 않는 한, 단어 "포함하다(comprise)", "포함하다(comprises)" 및 "포함하는(comprising)"은 단계를 또는 요소들의 언급된 단계 또는 요소 또는 그룹의 포함하지만, 임의의 다른 단계 또는 요소 또는 단계들의 제외는 아니라는 것을 나타내는 것으로 이해될 것이다. "이루어진"은 어구 "이루어진"에 따르는 것을 포함하지만, 이것으로 제한되지 않음을 의미한다. 따라서, 어구 "이루어진"은 열거된 요소가 필요하거나 또는 필수라는 것, 다른 요소가 존재하지 않을 수도 있다는 것을 나타낸다. "본질적으로 이루어진"은 어구 다음에 열거된 임의의 요소를 포함하며, 열거된 요소에 대해 본 개시내용에서 구체화된 활성 또는 작용을 방해하지 않거나 또는 기여하지 않는 다른 요소로 제한되지 않음을 의미한다. 따라서, 어구 "본질적으로 이루어진"은 열거된 요소가 필요하거나 또는 필수이지만, 열거된 요소의 활성 또는 작용에 실질적으로 영

향을 미치는 다른 요소가 존재하지 않는다는 것을 나타낸다.

- [0204] 본 명세서 전체적으로 "일 실시형태", "실시형태", "특정 실시형태", "관련된 실시형태", "특정 실시형태", "추가적인 실시형태" 또는 "추가 실시형태" 또는 이들의 조합에 대한 언급은 실시형태와 관련하여 기재된 특정 특성, 구조 또는 특징이 적어도 일 실시형태에 포함된다는 것을 의미한다. 따라서, 본 명세서 전체적으로 다양한 곳에서 앞서 언급한 어구의 출현은 반드시 모두 동일한 실시형태에 관한 것은 아니다. 더 나아가, 특정 특성, 구조 또는 특징은 하나 이상의 실시형태에서 임의의 적합한 방식으로 조합될 수 있다. 또한 일 실시형태에서 특정의 긍정적 인용은 특정 실시형태에서의 특성을 제외하기 위한 기초로서 작용한다는 것이 이해된다.
- [0205] 용어 "생체외"는 일반적으로 유기체 외부에서 일어나는 활동, 예컨대 유기체 외부의 인공 환경 내 살아있는 조직에서 또는 이런 조직 상에서(바람직하게는 천연 조건의 최소 변경) 행해지는 실험 또는 측정을 지칭한다. 특정 실시형태에서, "생체외" 절차는 유기체로부터 취한 그리고 실험실 장치에서, 보통 멸균 조건 하에, 전형적으로 환경에 따라서 몇 시간 또는 약 24시간까지(48시간 또는 72시간을 포함) 배양 또는 조절한 살아있는 세포 또는 조직을 수반한다. 특정 실시형태에서, 이러한 조직 또는 세포는 수집되고 냉동될 수 있으며, 이후에 생체외 처리를 위해 해동될 수 있다. 살아있는 세포 또는 조직을 이용하여 며칠 이상 지속되는 조직 배양 실험 또는 절차는 전형적으로 "시험판내"(특정 실시형태에서, 이 용어가 생체외와 상호 호환적으로 사용될 수 있지만)가 되는 것으로 고려된다.
- [0206] 용어 "생체내"는 일반적으로 유기체 내에서 생기는 활성을 지칭한다. 일 실시형태에서, 세포 제놈은 생체내에서 조작되거나, 편집되거나 또는 변형된다.
- [0207] "향상시키다" 또는 "촉진시키다" 또는 "증가시키다" 또는 "확장하다" 또는 "강화하다"는 일반적으로 비히클 또는 대조군 중 하나에 의해 야기되는 반응에 비해 더 큰 반응(즉, 생리적 반응)을 생산하거나, 유발하거나 또는 야기하기 위한 본 명세서에서 상정된 뉴클레아제 변이체, 게놈 편집 조성물 또는 게놈 편집 세포의 능력을 지칭한다. 측정 가능한 반응은, 특히 당업계 및 본 명세서의 설명에서의 이해로부터 분명한 촉매적 활성, 결합 친화도, 결합 부위 특이성, 결합 부위 선택성, 지속성, 세포용해 활성의 증가, 및/또는 전염증 사이토카인의 증가를 포함할 수 있다. "증가된" 또는 "향상된" 양은 전형적으로 "통계적으로 유의한" 양이고, 비히클 또는 대조군에 의해 생산되는 반응의 1.1, 1.2, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30배 이상(예를 들어, 500, 1000배)(그 사이의 그리고 1 초과의 모든 정수 및 소수점을 포함, 예를 들어, 1.5, 1.6, 1.7, 1.8 등)인 증가를 포함할 수 있다.
- [0208] "감소시키다" 또는 "낮추다" 또는 "줄이다" 또는 "줄여들게 하다" 또는 "없애다" 또는 "제거하다" 또는 "저해하다" 또는 "감쇄하다"는 일반적으로 비히클 또는 대조군 중 하나에 의해 야기되는 반응에 비해 더 적은 반응(즉, 생리적 반응)을 생산하거나, 유발하거나 또는 야기하기 위한 본 명세서에서 상정된 뉴클레아제 변이체, 게놈 편집 조성물 또는 게놈 편집 세포의 능력을 지칭한다. 측정 가능한 반응은 비표적(off-target) 결합 친화도, 비표적 절단 특이성, T 세포 고갈 등의 감소를 포함할 수 있다. "감소된" 또는 "줄어든" 양은 전형적으로 "통계적으로 유의한" 양이며, 비히클 또는 대조군에 의해 생산된 반응(기준 반응)의 1.1, 1.2, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30배 이상(예를 들어, 500, 1000배)(그 사이의 그리고 1 초과의 모든 정수 및 소수점을 포함, 예를 들어, 1.5, 1.6, 1.7, 1.8 등)인 감소를 포함할 수 있다.
- [0209] "유지하다" 또는 "보존하다" 또는 "유지" 또는 "변화 없음" 또는 "실질적인 변화 없음" 또는 "실질적인 감소 없음"은 일반적으로 비히클 또는 대조군에 의해 야기되는 반응에 비해 실질적으로 유사한 또는 비슷한 생리적 반응(즉, 하류 효과)를 생산하거나, 유발하거나 또는 야기하기 위한 본 명세서에서 상정된 뉴클레아제 변이체, 게놈 편집 조성물 또는 게놈 편집 세포의 능력을 지칭한다. 비슷한 반응은 기준 반응과 상당히 상이하지 않은 또는 측정 가능하게 상이하지 않은 것이다.
- [0210] 본 명세서에서 사용되는 용어 "특이적 결합 친화도" 또는 "특이적으로 결합하다" 또는 "특이적으로 결합된" 또는 "특이적 결합" 또는 "특이적으로 표적화하다"는 배경 결합보다 더 큰 결합 친화도로 하나의 분자의 다른 분자에 대한 결합, 예를 들어, DNA에 대한 폴리펩타이드 결합의 DNA 결합 도메인을 기재한다. 결합 도메인이, 예를 들어, 약  $10^5 \text{ M}^{-1}$  이상의 친화도 또는  $K_a$ (즉, 1/M의 단위로 특정 결합 상호작용의 평형 결합 상수)로 표적 부위에 결합하거나 또는 회합된다면, 결합 도메인은 표적 부위에 "특이적으로 결합한다". 특정 실시형태에서, 결합 도메인은 약  $10^6 \text{ M}^{-1}$ ,  $10^7 \text{ M}^{-1}$ ,  $10^8 \text{ M}^{-1}$ ,  $10^9 \text{ M}^{-1}$ ,  $10^{10} \text{ M}^{-1}$ ,  $10^{11} \text{ M}^{-1}$ ,  $10^{12} \text{ M}^{-1}$  또는  $10^{13} \text{ M}^{-1}$  이상의  $K_a$ 로 표적 부위에 결합한다. "고친화도" 결합 도메인은 적어도  $10^7 \text{ M}^{-1}$ , 적어도  $10^8 \text{ M}^{-1}$ , 적어도  $10^9 \text{ M}^{-1}$ , 적어도  $10^{10} \text{ M}^{-1}$ , 적

어도  $10^{11} \text{ M}^{-1}$ , 적어도  $10^{12} \text{ M}^{-1}$ , 적어도  $10^{13} \text{ M}^{-1}$  이상의  $K_a$ 를 갖는 해당 결합 도메인을 지칭한다.

[0211] 대안적으로, 친화도는 특정 실시형태에서 M 단위(예를 들어,  $10^{-5} \text{ M}$  내지  $10^{-13} \text{ M}$  이하)를 갖는 특정 결합 상호작용의 평형 해리 상수( $K_d$ )로서 정의될 수 있다. 특정 실시형태에서 상정된 DNA 표적 부위에 대한 하나 이상의 DNA 결합 도메인을 포함하는 뉴클레아제 변이체의 친화도는 통상적인 기법, 예를 들어, 효모 세포 표면 디스플레이를 이용하거나, 또는 결합 회합 또는 표지된 리간드를 이용하는 변위 분석에 의해 용이하게 결정될 수 있다.

[0212] 일 실시형태에서, 특이적 결합의 친화도는 배경 결합보다 약 2배 초과, 배경 결합보다 약 5배 초과, 배경 결합보다 약 10배 초과, 배경 결합보다 약 20배 초과, 배경 결합보다 약 50배 초과, 배경 결합보다 약 100배 초과 또는 배경 결합보다 약 1000배 초과 또는 그 이상이다.

[0213] 용어 "선택적으로 결합한다" 또는 "선택적으로 결합된" 또는 "선택적으로 결합하는" 또는 "선택적으로 표적화한다"는 복수의 비표적 분자의 존재 하의 표적 분자(표적 상의 결합)에 대한 하나의 분자의 우선적인 결합을 설명한다. 특정 실시형태에서, HE 또는 메가TAL은 HE보다 약 5, 10, 15, 20, 25, 50, 100 또는 1000배 더 빈번하게 표적 상의 DNA 결합 부위에 선택적으로 결합하거나 또는 메가TAL은 비표적 DNA 표적 결합 부위에 결합한다.

[0214] "표적 상의"는 표적 부위 서열을 지칭한다.

[0215] "비표적"은 표적 부위 서열과 유사하지만 동일하지 않은 서열을 지칭한다.

[0216] "표적 부위" 또는 "표적 서열"은 결합 및/또는 절단이 존재하는 충분한 조건 하에서, 결합 분자에 결합하고/하거나 절단하는 핵산의 일부를 정하는 염색체 또는 염색체와 핵산 서열이다. 표적 부위 또는 표적 서열의 한 가닥만을 언급하는 폴리뉴클레오타이드 서열 또는 서열번호를 지칭할 때, 뉴클레아제 변이체에 의해 결합되고/되거나 절단되는 표적 부위 또는 표적 서열이 이중 가닥이고 기준 서열 및 그의 상보체를 포함한다는 것이 이해될 것이다. 바람직한 실시형태에서, 표적 부위는 인간 CBLB 유전자 내 서열이다.

[0217] "재조합"은 비상동성 말단 결합(NHEJ) 및 상동성 재조합에 의한 공여자 포획을 포함하지만, 이들로 제한되지 않는, 두 폴리뉴클레오타이드 사이의 유전자 정보 교환 과정을 지칭한다. 본 개시내용의 목적을 위해, "상동성 재조합(HR)"은, 예를 들어, 상동 직접 수선(HDR) 메커니즘을 통해 세포에서 이중-가닥 파손의 수선 동안 일어나는 이러한 교환의 구체화된 형태를 지칭한다. 이 과정은 뉴클레오타이드 서열 상동성을 필요로 하며, "표적" 분자(즉, 이중-가닥 파손을 경험한 것)를 수선하기 위한 "공여자" 분자를 사용하고, "비교차형(non-crossover) 유전자 전환" 또는 "짧은 관 유전자 전환"으로서 다양하게 알려져 있는데, 이것이 공여자로부터 표적까지 유전자 정보의 전달을 야기하기 때문이다. 임의의 특정 이론에 의해 결합되는 일 없이, 이러한 전달은 파손된 표적과 공여자 사이에 형성되는 이형이중가닥 DNA의 미스매치 수정, 및/또는 "합성-의존적 가닥 어닐링"을 수반할 수 있으며, 이때 공여자는 표적의 부분, 및/또는 관련된 과정의 부분이 되는 유전자 정보를 재합성하기 위해 사용된다. 공여자 폴리뉴클레오타이드의 부분 또는 모두가 표적 폴리뉴클레오타이드 내로 흔입되도록 이러한 전문화된 HR은 종종 표적 분자 서열의 변경을 초래한다.

[0218] "NHEJ" 또는 "비상동성 말단 결합"은 공여자 수선 주형 또는 상동성 서열의 부재 하에서 이중-가닥 파손의 해결을 지칭한다. NHEJ는 파손 부위에서의 삽입 및 결실을 초래할 수 있다. NHEJ는 몇몇 하위 경로에 의해 매개되는데, 이들 각각은 별개의 돌연변이 결과를 가진다. 고전적 NHEJ 경로(cNHEJ)는 KU/DNA-PKcs/Lig4/XRCC4 복합체를 필요로 하며, 최소 가공과 함께 다시 말단을 결찰시키고, 종종 파손의 정확한 수선을 야기한다. 대안의 NHEJ 경로(altNHEJ)는 또한 dsDNA 파손을 해결하는 데 활성이지만, 이들 경로는 상당히 더 돌연변이 유발성이며, 종종 삽입 및 결실에 의해 표시되는 파손의 부정확한 수선을 초래한다. 임의의 특정 이론에 구속되는 일 없이, 말단-가공 효소, 예를 들어, 엑소뉴클레아제, 예를 들어, Trex2에 의한 dsDNA 파손의 변형은 부정확한 수선 가능성을 증가시킬 수 있다는 것이 상정된다.

[0219] "절단"은 DNA 분자의 공여 골격의 파손을 지칭한다. 절단은 포스포다이에스터 결합의 효소적 또는 화학적 가수분해를 포함하지만, 이들로 제한되지 않는 다양한 방법에 의해 개시될 수 있다. 단일-가닥 절단과 이중-가닥 절단이 둘 다 가능하다. 이중-가닥 절단은 2개의 별개의 단일-가닥 절단 사건의 결과로서 일어날 수 있다. DNA 절단은 평활말단 또는 엇갈린 말단 중 하나의 생산을 초래할 수 있다. 특정 실시형태에서, 본 명세서에서 상정된 폴리펩타이드 및 뉴클레아제 변이체, 예를 들어, 호밍 엔도뉴클레아제 변이체, 메가TAL 등은 표적화된 이중-가닥 DNA 절단을 위해 사용된다. 엔도뉴클레아제 절단 인식 부위는 DNA 가닥 중 하나 상에 있을 수 있다.

- [0220] "외인성" 문자는 세포에 정상적으로 존재하지 않지만, 하나 이상의 유전적, 생화학적 또는 다른 방법에 의해 세포 내로 도입되는 문자이다. 예시적인 외인성 문자는 소유기 문자, 단백질, 핵산, 탄수화물, 지질, 당단백질, 지방단백질, 다당류, 상기 문자의 임의의 변형된 유도체, 또는 상기 문자 중 하나 이상을 포함하는 임의의 복합체를 포함하지만, 이들로 제한되지 않는다. 세포 내로 외인성 문자의 도입을 위한 방법은 당업자에게 공지되어 있으며, 지질-매개 전달(즉, 중성 및 양이온성 지질을 포함하는 리포좀), 전기천공법, 직접 주사, 세포 융합, 입자총, 생체고분자 나노입자, 인산칼슘 공동침전, DEAE-덱스트란-매개 전달 및 바이러스 벡터-매개 전달을 포함하지만, 이들로 제한되지 않는다.
- [0221] "내인성" 문자는 특정 환경 조건 하에서 특정 발생 단계에서 특정 세포에 정상적으로 존재하는 것이다. 추가적인 내인성 문자는 단백질을 포함할 수 있다.
- [0222] "유전자"는, 이러한 조절 서열이 암호 및/또는 전사된 서열에 인접하든 아니든 유전자 산물을 암호화하는 DNA 영역뿐만 아니라 유전자 산물의 생산을 조절하는 모든 DNA 영역을 지칭한다. 유전자는 프로모터 서열, 인핸서, 사일런서, 절연체, 경계 요소, 종결자, 폴리아데닐화 서열, 전사후 반응 요소, 번역 조절 서열, 예컨대 리보솜 결합 부위 및 내부 리보솜 유입 부위, 복제기점, 기질 부착 부위 및 좌우 제어 영역을 포함하지만, 이들로 제한되지 않는다.
- [0223] "유전자 발현"은 유전자에 함유된 정보의 유전자 산물로의 전환을 지칭한다. 유전자 산물은 유전자의 직접 전사 산물(예를 들어, mRNA, tRNA, rRNA, 안티센스 RNA, 리보자임, 구조적 RNA 또는 임의의 다른 유형의 RNA) 또는 mRNA의 번역에 의해 생산되는 단백질일 수 있다. 유전자 산물은 또한 캡핑, 폴리아데닐화, 메틸화 및 편집과 같은 과정에 의해 변형되는 RNA 및, 예를 들어, 메틸화, 아세틸화, 포스포릴화, 유비퀴틴화, ADP-리보실화, 미리스틸화 및 글리코실화에 의해 변형된 단백질을 포함한다.
- [0224] 본 명세서에서 사용되는 용어 "유전자 조작된" 또는 "유전자 변형된"은 DNA 또는 RNA 형태의 유전자와 물질의 세포 내 총 유전자 물질에 대한 염색체 또는 염색체외 첨가를 지칭한다. 유전자 변형은 세포 계놈에서 특정 부위로 표적화 또는 비표적화될 수 있다. 일 실시형태에서, 유전자 변형은 부위 특이적이다. 일 실시형태에서, 유전자 변형은 부위 특이적이지 않다.
- [0225] 본 명세서에서 사용되는 용어 "계놈 편집"은 유전자 또는 유전자 산물의 발현을 회복시키고/시키거나, 수정하고/하거나 파괴하고/하거나 변형시키는 세포 계놈 내 표적 부위에서의 유전자 물질의 치환, 결실 및/또는 도입을 지칭한다. 특정 실시형태에서 상정된 계놈 편집은 세포 계놈에서, 선택적으로 공여자 수선 주형의 존재 하에 표적 부위에서 또는 표적 부위 근위에서 DNA 손상을 생성하기 위해 세포 내로 하나 이상의 뉴클레아제 변이체를 도입하는 것을 포함한다.
- [0226] 본 명세서에서 사용되는 용어 "유전자 요법"은 유전자 또는 유전자 산물의 발현을 회복하거나, 수정하거나 또는 변형시키는 또는 치료적 폴리펩타이드를 발현시키는 목적을 위한 세포 내 총 유전자 물질 내로의 유전자와 물질의 도입을 지칭한다. 특정 실시형태에서, 유전자 또는 유전자 산물의 발현을 회복시키거나, 수정하거나, 파괴하거나 또는 변형시키는, 또는 치료적 폴리펩타이드를 발현시키는 목적을 위해 계놈 편집에 의한 세포 계놈 내로의 유전자 물질의 도입이 유전자 요법으로 고려된다.
- [0227] "면역 장애"는 면역계로부터 반응을 유발하는 질환을 지칭한다. 특정 실시형태에서, 용어 "면역 장애"는 암, 이식편대숙주병, 자가면역 질환 또는 면역결핍을 지칭한다. 일 실시형태에서, 면역 장애는 감염성 질환을 포함한다.
- [0228] 본 명세서에서 사용되는 용어 "암"은 일반적으로 비정상 세포가 제어 없이 분할되고 조직 근처에서 침윤할 수 있는 질환 또는 병태의 부류에 관한 것이다.
- [0229] 본 명세서에서 사용되는 용어 "악성종양"은 종양 세포의 그룹이 제어되지 않는 성장(즉, 정상 한계 이상의 분할), 침윤(즉, 인접한 조직 상의 침범 및 파괴) 및 전이(즉, 럼프 또는 혈액을 통한 신체 내 다른 위치로의 확산) 중 하나 이상을 나타내는 암을 지칭한다.
- [0230] 본 명세서에서 사용되는 용어 "전이하다"는 신체의 한 부분으로부터 다른 부분까지의 암의 확산을 지칭한다. 확산을 갖는 세포에 의해 형성된 종양은 "전이성 종양" 또는 "전이"로 불린다. 전이성 종양은 본래의(원발성) 종양에서와 유사한 세포를 함유한다.
- [0231] 본 명세서에서 사용되는 용어 "양성" 또는 "비악성"은 더 크게 성장할 수 있지만, 신체의 다른 부분으로 확산되지 않는 종양을 지칭한다. 양성 종양은 자기 제한적이며, 전형적으로 침윤하거나 또는 전이하지 않는다.

- [0232] "암세포" 또는 "종양 세포"는 암성 성장의 개개 세포 또는 조직을 지칭한다. 종양은 일반적으로 양성, 전악성 또는 악성일 수 있는 세포의 비정상적 성장에 의해 형성되는 팽윤 또는 병변을 지칭한다. 대부분의 암은 종양을 형성하지만, 일부, 예를 들어, 백혈병은 반드시 종양을 형성하지는 않는다. 종양을 형성하는 해당 암에 대해, 용어 암(세포)과 종양(세포)은 상호 호환적으로 사용된다. 개체에서 종양의 양은 종양의 수, 용적 또는 중량으로서 측정될 수 있는 "종양 부담"이다.
- [0233] "이식편대 숙주질환" 또는 "GVHD"는 세포, 조직 또는 고형 기관 이식 후에 일어날 수 있는 합병증을 지칭한다. 이식된 공여자 세포는 이식 수용자 신체를 공격하는 GVHD는 줄기 세포 또는 골수 이식 후에 일어날 수 있다. 인간에서 급성 GVHD는 이식후 약 60일 내에 일어나며, 세포용해 림프구의 작용에 의해 피부, 간 및 장에 대한 손상을 초래한다. 만성 GVHD는 이후에 일어나며, 주로 피부에 영향을 미치는 전신 자가면역 질환이고, B 세포의 다클론성 활성화 및 Ig 및 자가항체의 과다생성을 초래한다. 고형-기관 이식 이식편대숙주병(SOT-GVHD)은 두 가지 형태로 일어난다. 더 통상적인 유형은 항체 매개되며, O형을 갖는 공여자로부터의 항체는 A, B 또는 AB 혈액형을 갖는 수용자에서 수용자의 적혈구를 공격하여, 가벼운 일시적, 용혈성 빈혈을 야기한다. SOT-GVHD의 두 번째 형태는 높은 사망률과 관련된 세포 유형이 되며, 공여자-유래 T 세포는 피부, 간, 위장관 및 골수에서 가장 흔한 면역학적으로 이질적인 숙주 조직에 대해 면역학적 공격을 생산하여, 이들 기관에서 합병증을 야기한다.
- [0234] "이식편대백혈병" 또는 "GVL"은 공여자의 이식된 조직, 예컨대 골수 또는 말초 혈액에 존재하는 면역 세포에 의한 사람의 백혈병 세포에 대한 면역 반응을 지칭한다.
- [0235] "자가면역 질환"은 신체가 그 자신의 조직의 일부 구성성분에 대해 면역원성(즉, 면역계) 반응을 생산하는 질환을 지칭한다. 다시 말해서, 면역계는 "자기" 및 표적으로서 신체 내의 일부 조직 또는 시스템을 인식하는 그의 능력을 상실하며, 그것이 외래 물질인 것처럼 공격한다. 자가면역 질환의 예시적 예는 관절염, 염증성 장 질환, 하시모토 갑상선염, 그레이브스병, 루푸스, 다발성 경화증, 류마티스 관절염, 용혈성 빈혈, 항면역 갑상선염, 전신 홍반성 루푸스, 셀리악병, 크론병, 결장염, 당뇨병, 경피증, 건선 등을 포함하지만, 이들로 제한되지 않는다.
- [0236] "면역결핍"은 질환에 의해 또는 화학물질의 투여에 의해 면역계가 손상된 환자 상태를 의미한다. 이 병태는 시스템이 외래 물질을 방어하는 데 필요한 혈액 세포의 수 및 유형이 결여되거나 한다. 면역결핍증 병태 또는 질환은 당업계에 공지되어 있고, 예를 들어, AIDS(후천성 면역결핍 증후군), SCID(중증 합병 면역결핍 질환), 선택적 IgA 결핍증, 공통 가변 면역결핍증, X-연관 무감마글로불린혈증, 만성 육아종성 질환, 과-IgM 증후군, 위스콧-알드리치 증후군(WAS) 및 당뇨병을 포함한다.
- [0237] "감염성 질환"은 사람 간에 또는 유기체 간에 전염될 수 있고, 미생물 또는 바이러스 제제(예를 들어, 보통 감기)에 의해 야기되는 질환을 지칭한다. 감염성 질환은 당업계에 공지되어 있으며, 예를 들어, 간염, 성적으로 전염되는 질환(예를 들어, 클라미디아, 임질), 결핵, HIV/AIDS, 디프테리아, 간염 B, 간염 C, 콜레라 및 인플루엔자를 포함한다.
- [0238] 본 명세서에서 사용되는 용어 "개체" 및 "대상체"는 종종 상호 호환적으로 사용되며, 뉴클레아제 변이체, 게놈 편집 조성물, 유전자 요법 백터, 게놈 편집 백터, 게놈 편집된 세포, 및 본 명세서의 다른 곳에 상정된 방법으로 치료될 수 있는 면역 장애의 증상을 나타내는 임의의 동물을 지칭한다. 적합한 대상체(예를 들어, 환자)는 실험실 동물(예컨대, 마우스, 래트, 토끼 또는 기니피), 농장 동물 및 가축 동물 또는 반려동물(예컨대, 고양이 또는 개)을 포함한다. 비인간 영장류 및 바람직하게는 인간 대상체가 포함된다. 전형적인 대상체는 면역 장애로 진단되거나 또는 면역 장애를 가질 위험에 있는 인간 환자를 포함한다.
- [0239] 본 명세서에서 사용되는 용어 "환자"는 뉴클레아제 변이체, 게놈 편집 조성물, 유전자 요법 백터, 게놈 편집 백터, 게놈 편집된 세포, 및 본 명세서의 다른 곳에 상정된 방법으로 치료될 수 있는 면역 장애로 진단된 대상체를 지칭한다.
- [0240] 본 명세서에서 사용되는 "치료" 또는 "치료하는"은 질환 또는 병리적 병태의 증상 또는 병리에 대한 임의의 유리한 또는 바람직한 효과를 포함하고, 치료 중인 질환 또는 병태, 예를 들어, 암, GVHD, 감염성 질환, 자가면역 질환, 염증성 질환 및 면역결핍의 하나 이상의 측정 가능한 마커의 심지어 최소 감소를 포함할 수 있다. 치료는 선택적으로 질환 또는 병태 진행의 지연을 수반할 수 있다. "치료"는 질환 또는 병태, 또는 이의 관련된 증상의 완전한 근절 또는 치유를 반드시 나타내지는 않는다.
- [0241] 본 명세서에서 사용하는 바와 같은, "예방하다" 및 유사한 단어, 예컨대 "예방", "예방되는", "예방하는" 등은 질환 또는 병태, 예를 들어, 암, GVHD, 감염성 질환, 자가면역 질환, 염증성 질환 및 면역결핍의 발생 또는 재

발 가능성을 예방하거나, 저해하거나 또는 감소시키기 위한 접근을 나타낸다. 또한 질환 또는 병태의 개시 또는 재발을 지연시키거나 또는 질환 또는 병태의 발생 또는 재발을 지연시키는 것을 지칭한다. 본 명세서에서 사용하는 바와 같은 "예방" 및 유사한 단어는 또한 질환 또는 병태의 개시 또는 재발 전에 질환 또는 병태의 강도, 효과, 증상 및/또는 부담을 감소시키는 것을 포함한다.

[0242] 본 명세서에서 사용되는 어구 "의 적어도 하나의 증상을 개선시키는"은 대상체에 대해 치료 중인 질환 또는 병태, 예를 들어, 암, GVHD, 감염성 질환, 자가면역 질환, 염증성 질환 및 면역결핍의 하나 이상의 증상을 감소시키는 것을 지칭한다. 특정 실시형태에서, 치료 중인 질환 또는 병태는 암이되, 개선될 하나 이상의 증상은 쇠약, 피로, 숨가쁨, 잦은 명 및 출혈, 빈번한 감염, 림프절 비대, 복부 팽창 또는 복부 통증(팽장된 복부 기관에 기인), 뼈 또는 관절 통증, 골절, 계획하지 않은 체중 감소, 식욕 부진, 도한, 지속적 미열 및 감소된 배뇨작용(손상된 신장 기능에 기인)을 포함하지만, 이들로 제한되지 않는다.

[0243] 본 명세서에서 사용되는 용어 "양"은 임상 결과를 포함하는 유리한 또는 목적으로 하는 예방적 또는 치료적 결과를 달성하는 데 충분한 뉴클레아제 변이체, 게놈 편집 조성물, 또는 게놈 편집 세포의 "효과적인 양" 또는 "유효량"을 지칭한다.

[0244] "예방적 유효량"은 목적으로 하는 예방적 결과를 달성하는 데 충분한 뉴클레아제 변이체, 게놈 편집 조성물 또는 게놈 편집 세포의 양을 지칭한다. 전형적으로, 반드시는 아니지만, 예방적 용량은 질환 전에 또는 초기에 대상체에서 사용되기 때문에, 예방적 유효량은 치료적 유효량보다 적다.

[0245] 뉴클레아제 변이체, 게놈 편집 조성물 또는 게놈 편집 세포의 "치료적 유효량"은 개체의 질환 상태, 연령, 성별 및 체중 및 개체에서 목적으로 하는 반응을 유발하는 능력과 같은 인자에 따라 다를 수 있다. 치료적 유효량은 또한 임의의 독성 또는 유해한 효과가 치료적으로 유리한 효과보다 더 큰 것이다. 용어 "치료적 유효량"은 대상체(예를 들어, 환자)를 "치료"하는 데 효과적인 양을 포함한다. 치료적 양이 표시될 때, 특정 실시형태에서 상정되는 투여될 조성물의 정확한 양은 본 명세서를 고려하여 그리고 환자(대상체)의 연령, 체중, 종양 크기, 감염 또는 전이 정도 및 병태에서 개개 차이를 고려하여 결정될 수 있다.

#### C. 뉴클레아제 변이체

[0247] 본 명세서의 특정 실시형태에서 상정되는 뉴클레아제 변이체는 CBLB 유전자 내 표적 부위를 게놈 편집하는 데 적합하고, 하나 이상의 DNA 결합 도메인 및 하나 이상의 DNA 절단 도메인(예를 들어, 하나 이상의 엔도뉴클레아제 및/또는 엑소뉴클레아제 도메인), 및 선택적으로 본 명세서에서 상정된 하나 이상의 링커를 포함한다. 용어 "재프로그래밍된 뉴클레아제", "조작된 뉴클레아제" 또는 "뉴클레아제 변이체"는 상호 호환적으로 사용되며, 하나 이상의 DNA 결합 도메인 및 하나 이상의 DNA 절단 도메인을 포함하는 뉴클레아제를 지칭하되, 상기 뉴클레아제는 CBLB 유전자 내 이중 가닥 DNA 표적 서열에 결합하고 절단하도록 모 또는 천연 유래 뉴클레아제로부터 설계되고/되거나 변형되었다.

[0248] 특정 실시형태에서, 뉴클레아제 변이체는 CBLB 유전자의 엑손 6에서, 바람직하게는 CBLB 유전자의 엑손 6에서 서열번호 20, 그리고 더 바람직하게는 CBLB 유전자의 엑손 6에서 서열번호 20에서의 서열 "ATTC"에서 표적 서열에 결합하고, 절단한다.

[0249] 뉴클레아제 변이체는 천연 유래 뉴클레아제로부터 또는 이전의 뉴클레아제 변이체로부터 설계되고/되거나 변형될 수 있다. 특정 실시형태에서 상정된 뉴클레아제 변이체는 하나 이상의 추가적인 작용성 도메인, 예를 들어, 5'-3' 엑소뉴클레아제, 5'-3' 알칼리성 엑소뉴클레아제, 3'-5' 엑소뉴클레아제(예를 들어, Trex2), 5' 플랩 엔도뉴클레아제, 헬리카제, 주형-의존적 DNA 중합효소 또는 주형-독립적 DNA 중합효소 활성을 나타내는 말단-가공효소의 말단-가공 효소 도메인을 추가로 포함할 수 있다.

[0250] CBLB 유전자 내 표적 서열에 결합하고 절단하는 뉴클레아제 변이체의 예시적 예는 호밍 엔도뉴클레아제(메가뉴클레아제) 변이체 및 메가TAL를 포함하지만, 이들로 제한되지 않는다.

#### 1. 호밍 엔도뉴클레아제(메가뉴클레아제) 변이체

[0252] 다양한 실시형태에서, 호밍 엔도뉴클레아제 또는 메가뉴클레아제는 CBLB 유전자 내 표적 부위에서 이중-가닥 파손(DSB)을 도입하도록 재프로그래밍된다. 특정 실시형태에서, 호밍 엔도뉴클레아제 변이체는 CBLB 유전자의 엑손 6에서, 바람직하게는 CBLB 유전자의 엑손 6에서 서열번호 20에서, 그리고 더 바람직하게는 CBLB 유전자의 엑손 6에서 서열번호 20에서의 서열 "ATTC"에서 이중 가닥 파손을 도입한다.

[0253] "호밍 엔도뉴클레아제" 및 "메가뉴클레아제"는 상호 호환적으로 사용되며, 12 내지 45개의 염기쌍 절단 부위를

인식하고, 서열 및 구조 모티프에 기반하여 5개의 패밀리(LAGLIDADG, GIY-YIG, HNH, His-Cys 박스 및 PD-(D/E)XK)로 통상적으로 그룹화되는 천연 유래 호밍 엔도뉴클레아제를 지칭한다.

[0254] "기준 호밍 엔도뉴클레아제" 또는 "기준 메가뉴클레아제"는 천연에서 발견되는 야생형 호밍 엔도뉴클레아제 또는 호밍 엔도뉴클레아제를 지칭한다. 일 실시형태에서, "기준 호밍 엔도뉴클레아제"는 기본적 활성을 증가시키도록 변형된 야생형 호밍 엔도뉴클레아제를 지칭한다.

[0255] "조작된 호밍 엔도뉴클레아제", "재프로그래밍된 호밍 엔도뉴클레아제", "호밍 엔도뉴클레아제 변이체", "조작된 메가뉴클레아제", "재프로그래밍된 메가뉴클레아제" 또는 "메가뉴클레아제 변이체"는 하나 이상의 DNA 결합 도메인 및 하나 이상의 DNA 절단 도메인을 포함하는 호밍 엔도뉴클레아제를 지칭하되, 호밍 엔도뉴클레아제는 CBLB 유전자 내 DNA 표적 서열에 결합하고 이를 절단하도록 모 또는 천연 유래 호밍 엔도뉴클레아제로부터 설계되고/되거나 변형되었다. 호밍 엔도뉴클레아제 변이체는 천연 유래 호밍 엔도뉴클레아제로부터 또는 다른 호밍 엔도뉴클레아제 변이체로부터 설계되고/되거나 변형될 수 있다. 특정 실시형태에서 상정된 호밍 엔도뉴클레아제 변이체는 하나 이상의 추가적인 작용성 도메인, 예를 들어, 5'-3' 엑소뉴클레아제, 5'-3' 알칼리성 엑소뉴클레아제, 3'-5' 엑소뉴클레아제(예를 들어, Trex2), 5' 플랩 엔도뉴클레아제, 헬리카제, 주형-의존적 DNA 중합효소 또는 주형-독립적 DNA 중합효소 활성을 나타내는 말단-가공 효소의 말단-가공 효소 도메인을 추가로 포함할 수 있다.

[0256] 호밍 엔도뉴클레아제(HE) 변이체는 천연에서 존재하지 않으며, 재조합 DNA 기법에 의해 또는 무작위 돌연변이 유발에 의해 얻을 수 있다. HE 변이체는 하나 이상의 아미노산 변경을 만들고 써, 예를 들어, 천연 유래 HE 또는 HE 변이체에서 하나 이상의 아미노산을 돌연변이, 치환, 침가 또는 결실시킴으로써 얻을 수 있다. 특정 실시형태에서, HE 변이체는 DNA 인식 계면에 대한 하나 이상의 아미노산 변경을 포함한다.

[0257] 특정 실시형태에서 상정된 HE 변이체는 하나 이상의 추가적인 링커 및/또는 추가적인 작용성 도메인, 예를 들어, 5'-3' 엑소뉴클레아제, 5'-3' 알칼리성 엑소뉴클레아제, 3'-5' 엑소뉴클레아제(예를 들어, Trex2), 5' 플랩 엔도뉴클레아제, 헬리카제, 주형-의존적 DNA 중합효소 또는 주형-독립적 DNA 중합효소 활성을 나타내는 말단-가공 효소의 말단-가공 효소 도메인을 추가로 포함할 수 있다. 특정 실시형태에서, HE 변이체는 5'-3' 엑소뉴클레아제, 5'-3' 알칼리성 엑소뉴클레아제, 3'-5' 엑소뉴클레아제(예를 들어, Trex2), 5' 플랩 엔도뉴클레아제, 헬리카제, 주형-의존적 DNA 중합효소 또는 주형-독립적 DNA 중합효소 활성을 나타내는 말단-가공 효소에 의해 T 세포 내로 도입된다. HE 변이체 및 3'가공 효소는 별개로, 예를 들어, 상이한 백터 또는 별개의 mRNA에서, 또는 함께, 예를 들어, 융합 단백질로서, 또는 바이러스 자기-절단성 웹타이드 또는 IRES 요소에 의해 분리된 다시스 트론성 작제물에서 도입될 수 있다.

[0258] "DNA 인식 계면"은 핵산 표적 염기와 상호작용하는 HE 아미노산 잔기뿐만 아니라 인접한 해당 잔기를 지칭한다. 각각의 HE에 대해, DNA 인식 계면은 측쇄 대 측쇄 및 측쇄 대 DNA 접촉의 광대한 그물망을 포함하며, 이 중 대부분은 특정 핵산 표적 서열을 인식하기에 충분히 독특하다. 따라서, 특정 핵산 서열에 대응하는 DNA 인식 계면의 아미노산 서열은 상당히 다르며, 임의의 천연 또는 HE 변이체의 특징이다. 비제한적 예로서, 특정 실시형태에 상정된 HE 변이체는 HE 변이체의 라이브러리를 작제함으로써 유도될 수 있으며, 이때 천연 HE(또는 이전에 생성된 HE 변이체)의 DNA 인식 계면에서 국소화된 하나 이상의 아미노산 잔기가 변한다. 라이브러리는 절단 분석을 이용하여 각각의 예측된 CBLB 표적 부위에 대한 표적 절단 활성을 대해 선별될 수 있다(예를 들어, 문헌 [Jarjour et al., 2009. Nuc. Acids Res. 37(20): 6871-6880] 참조).

[0259] LAGLIDADG 호밍 엔도뉴클레아제(LHE)는 호밍 엔도뉴클레아제의 가장 잘 연구된 패밀리이며, 고세균에서 그리고 녹조류 및 진균에서의 세포소기관 DNA에서 주로 암호화되고, 가장 높은 전반적인 DNA 인식 특이성을 나타낸다. LHE는 단백질 쇄 당 1 또는 2개의 LAGLIDADG 촉매적 모티프를 포함하며, 동종이량체 또는 단일쇄 단량체로서 각각 작용한다. LAGLIDADG 단백질의 구조적 연구는  $\alpha\beta\beta\alpha\beta\beta\alpha$  폴드에 의해 특성규명되는 고도로 보존된 코어 구조를 동정하였고(Stoddard 2005), LAGLIDADG 모티프는 이 폴드의 첫 번째 나선에 속한다. LHE의 고도로 효율적인 그리고 특이적 절단은 신규한, 고도로 특이적인 엔도뉴클레아제를 유도하기 위한 단백질 스캐폴드를 나타낸다. 그러나, 비천연 또는 비정규 표적 부위를 결합하고 절단하기 위한 LHE의 조작은 표적 부위에서 염기쌍 위치의 2/3까지에서, 적절한 LHE 스캐폴드의 선택, 표적 좌위의 시험, 추정적 표적 부위의 선택, 및 DNA 접점 및 절단 특이성을 변경하기 위한 LHE의 광대한 변경을 필요로 한다.

[0260] 일 실시형태에서, 재프로그래밍된 LHE 또는 LHE 변이체가 설계될 수 있는 LHE는 I-CreI 및 I-SceI를 포함하지만, 이들로 제한되지 않는다.

- [0261] 재프로그래밍된 LHE 또는 LHE 변이체로부터 설계될 수 있는 LHE의 예시적인 예는 I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-NcrI, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI 및 I-Vdi141I를 포함하지만, 이들로 제한되지 않는다.
- [0262] 일 실시형태에서, 재프로그래밍된 LHE 또는 LHE 변이체는 I-CpaMI 변이체, I-HjeMI 변이체, I-OnuI 변이체, I-PanMI 변이체 및 I-SmaMI 변이체로 이루어진 군으로부터 선택된다.
- [0263] 일 실시형태에서, 재프로그래밍된 LHE 또는 LHE 변이체는 I-OnuI 변이체이다. 예를 들어, 서열번호 6 내지 12 참조.
- [0264] 일 실시형태에서, CBLB 유전자를 표적화하는 재프로그래밍된 I-OnuI LHE 또는 I-OnuI 변이체는 천연 I-OnuI 또는 이의 생물학적 활성 단편(서열번호 1-5)으로부터 생성되었다. 바람직한 실시형태에서, 인간 CBLB 유전자를 표적화하는 재프로그래밍된 I-OnuI LHE 또는 I-OnuI 변이체는 존재하는 I-OnuI 변이체로부터 생성되었다. 일 실시형태에서, 재프로그래밍된 I-OnuI LHE는 서열번호 20에 제시된 인간 CBLB 유전자 표적 부위에 대해 생성되었다.
- [0265] 특정 실시형태에서, 인간 CBLB 유전자에 결합하고, 이를 절단하는 재프로그래밍된 I-OnuI LHE 또는 I-OnuI 변이체는 DNA 인식 계면에서 하나 이상의 아미노산 치환을 포함한다. 특정 실시형태에서, 인간 CBLB 유전자에 결합하고 이를 절단하는 I-OnuI LHE는 I-OnuI(Taekuchi et al. 2011. Proc Natl Acad Sci U. S. A. 2011 Aug 9; 108(32): 13077-13082), 또는 서열번호 6 내지 12 중 임의의 하나에 제시된 바와 같은 I-OnuI LHE 변이체, 이들의 생물학적 활성 단편, 및/또는 이들의 추가적인 변이체의 DNA 인식 계면과 적어도 70%, 적어도 71%, 적어도 72%, 적어도 73%, 적어도 74%, 적어도 75%, 적어도 76%, 적어도 77%, 적어도 78%, 적어도 79%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98% 또는 적어도 99%의 서열 동일성을 포함한다.
- [0266] 일 실시형태에서, 인간 CBLB 유전자에 결합하고, 이를 절단하는 I-OnuI LHE는 I-OnuI(Taekuchi et al. 2011. Proc Natl Acad Sci U. S. A. 2011 Aug 9; 108(32): 13077-13082) 또는 서열번호 6 내지 12 중 임의의 하나에 제시된 바와 같은 I-OnuI LHE 변이체, 이들의 생물학적 활성 단편 및/또는 이들의 추가적인 변이체의 DNA 인식 계면과 적어도 70%, 더 바람직하게는 적어도 80%, 더 바람직하게는 적어도 85%, 더 바람직하게는 적어도 90%, 더 바람직하게는 적어도 95%, 더 바람직하게는 적어도 97%, 더 바람직하게는 적어도 99%의 서열 동일성을 포함한다.
- [0267] 특정 실시형태에서, 인간 CBLB 유전자에 결합하고 이를 절단하는 I-OnuI LHE 변이체는 서열번호 1 내지 12 중 임의의 하나에 제시된 바와 같은 I-OnuI, 이들의 생물학적 활성 단편 및/또는 이들의 추가적인 변이체의 DNA 인식 계면에서의 하나 이상의 아미노산 치환 또는 변형을 포함한다.
- [0268] 특정 실시형태에서, 인간 CBLB 유전자에 결합하고, 이를 절단하는 I-OnuI LHE 변이체는 DNA 인식 계면에서, 특히 I-OnuI(서열번호 1 내지 5) 또는 서열번호 6 내지 12 중 임의의 하나에 제시된 I-OnuI 변이체, 이의 생물학적 활성 단편, 및/또는 이의 추가적인 변이체의 위치 24 내지 50, 68 내지 82, 180 내지 203 및 223 내지 240번에 위치된 서브도메인에서의 하나 이상의 아미노산 치환 또는 변형을 포함한다.
- [0269] 특정 실시형태에서, HE 변이체는 I-OnuI(서열번호 1 내지 5), 또는 서열번호 6 내지 12 중 어느 하나에 제시된 바와 같은 I-OnuI 변이체 및/또는 이의 추가적인 변이체로 이루어진 군으로부터 선택된 아미노산 위치: 24, 26, 28, 30, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 68, 70, 72, 75, 76, 78, 80, 82, 180, 182, 184, 186, 188, 189, 190, 191, 192, 193, 195, 197, 199, 201, 203, 223, 225, 227, 229, 231, 232, 234, 236, 238 및 240번에서의 DNA 인식 계면에서 하나 이상의 아미노산 치환을 포함한다.
- [0270] 특정 실시형태에서, HE 변이체는 I-OnuI(서열번호 1 내지 5) 또는 서열번호 6 내지 12 중 어느 하나에 제시된 I-OnuI 변이체, 이들의 생물학적 활성 단편, 및/또는 이들의 추가적인 변이체로 이루어진 군으로부터 선택된 아미노산 위치: 19, 24, 26, 28, 30, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 59, 68, 70, 72, 75, 76, 77, 78, 80, 82, 168, 180, 182, 184, 186, 188, 189, 190, 191, 192, 193, 195, 197, 199, 201, 203, 223, 225, 227, 229, 231, 232, 234, 236, 238 및 240번에서의 하나 이상의 아미노산 치환을 포함한다.

- [0271] 특정 실시형태에서, 인간 CBLB 유전자에 결합하고, 이를 절단하는 I-OnuI LHE 변이체는 DNA 인식 계면에서, 특히 I-OnuI(서열번호 1 내지 5) 또는 서열번호 6 내지 12 중 임의의 하나에 제시된 I-OnuI 변이체, 이의 생물학적 활성 단편, 및/또는 이의 추가적인 변이체의 위치 24 내지 50, 68 내지 82, 180 내지 203 및 223 내지 240번에 위치된 서브도메인에서의 5, 10, 15, 20, 25, 30, 35 또는 40개 이상의 아미노산 치환 또는 변형을 포함한다.
- [0272] 특정 실시형태에서, 인간 CBLB 유전자에 결합하고, 이를 절단하는 I-OnuI LHE 변이체는 I-OnuI(서열번호 1 내지 5) 또는 서열번호 6 내지 12 중 어느 하나에 제시된 바와 같은 I-OnuI 변이체, 이의 생물학적 활성 단편, 및/또는 이의 추가적인 변이체의 24, 26, 28, 30, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 59, 68, 70, 72, 75, 76, 78, 80, 82, 168, 180, 182, 184, 186, 188, 189, 190, 191, 192, 193, 195, 197, 199, 201, 203, 223, 225, 227, 229, 231, 232, 234, 236, 238 및 240번으로 이루어진 군으로부터 선택된 아미노산 위치에서의 5, 10, 15, 20, 25, 30, 35 또는 40개 이상의 아미노산 치환 또는 변형을 포함한다.
- [0273] 특정 실시형태에서, 인간 CBLB 유전자에 결합하고 절단하는 I-OnuI LHE 변이체는 I-OnuI(서열번호 1 내지 5)의 19, 24, 26, 28, 30, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 59, 68, 70, 72, 75, 76 77, 78, 80, 82, 168, 180, 182, 184, 186, 188, 189, 190, 191, 192, 193, 195, 197, 199, 201, 203, 223, 225, 227, 229, 231, 232, 234, 236, 238 및 240번으로 이루어진 군으로부터 선택된 아미노산 위치에서의 5, 10, 15, 20, 25, 30, 35 또는 40개 이상의 아미노산 치환 또는 변형 또는 서열번호 6 내지 12 중 어느 하나에 제시된 바와 같은 I-OnuI 변이체, 이의 생물학적 활성 단편, 및/또는 이의 추가적인 변이체를 포함한다.
- [0274] 일 실시형태에서, 인간 CBLB 유전자에 결합하고 이를 절단하는 I-OnuI LHE 변이체는 전체 I-OnuI 서열 내의 아무데나 위치된 추가적인 위치에서 하나 이상의 아미노산 치환 또는 변형을 포함한다. 치환되고/되거나 변형될 수 있는 잔기는 직접적으로 또는 물 분자를 통해, 핵산 표적과 접촉하거나 또는 핵산 골격과 또는 뉴클레오타이드 염기와 상호작용하는 아미노산을 포함한다. 하나의 비제한적 예에서, 인간 CBLB 유전자에 결합하고, 이를 절단하는 본 명세서에서 상정된 I-OnuI LHE 변이체는 I-OnuI(서열번호 1 내지 5) 또는 서열번호 6 내지 12 중 임의의 하나에 제시된 바와 같은 I-OnuI 변이체, 이의 생물학적 활성 단편, 및/또는 이의 추가적인 변이체의 위치: 24, 26, 28, 30, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 68, 70, 72, 78, 80, 92, 116, 138, 143, 159, 168, 178, 180, 182, 184, 186, 188, 189, 190, 191, 192, 193, 195, 197, 199, 201, 203, 207, 223, 225, 227, 232, 236 및 238번으로 이루어진 위치군으로부터 선택되는 적어도 하나의 위치에서 하나 이상의 치환 및/또는 변형, 바람직하게는 적어도 5, 바람직하게는 적어도 10, 바람직하게는 적어도 15, 바람직하게는 적어도 20, 더 바람직하게는 적어도 25, 더 바람직하게는 적어도 30, 더욱더 바람직하게는 적어도 35, 또는 더욱더 바람직하게는 적어도 40개 이상의 아미노산을 포함한다.
- [0275] 소정의 실시형태에서, HE 변이체는 CBLB 엑손 6 표적 부위를 절단하며, I-OnuI(서열번호 1 내지 5) 또는 서열번호 6 내지 12 중 어느 하나에 제시된 바와 같은 I-OnuI 변이체, 이들의 생물학적 활성 단편, 및/또는 이들의 추가적인 변이체의 다음의 아미노산 치환: S24C, L26R, L26G, R28D, R28Y, R30H, N32A, N32S, K34D, K34V, S35L, S36R, V37A, V37S, S40R, E42R, G44A, G44S, Q46E, T48V, T48S, V68T, V68K, A70Y, S72A, S78R, K80Q, D92G, V116L, L138M, T143N, S159P, F168L, E178D, C180S, F182V, F182M, N184E, I186K, I186M, S188R, S188N, K189R, S190N, K191P, K191N, L192V, G193K, G193I, Q195G, Q195R, Q197R, V199R, S201G, T203S, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.
- [0276] 일부 실시형태에서, HE 변이체는 CBLB 표적 부위를 절단하며, I-OnuI(서열번호 1 내지 5) 또는 서열번호 6 내지 12 중 어느 하나에 제시된 바와 같은 I-OnuI 변이체, 이들의 생물학적 활성 단편, 및/또는 이들의 추가적인 변이체의 다음의 아미노산 치환: S24C, L26R, R28D, N32A, K34D, S35L, S36R, V37A, S40R, E42R, G44A, Q46E, T48V, V68T, A70Y, S72A, S78R, K80Q, L138M, T143N, F168L, E178D, C180S, F182V, N184E, I186K, S188R, K189R, K191P, L192V, G193K, Q195G, Q197R, V199R, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.
- [0277] 특정 실시형태에서, HE 변이체는 CBLB 표적 부위를 절단하며, I-OnuI(서열번호 1 내지 5) 또는 서열번호 6 내지 12 중 어느 하나에 제시된 바와 같은 I-OnuI 변이체, 이들의 생물학적 활성 단편, 및/또는 이들의 추가적인 변이체의 다음의 아미노산 치환: S24C, L26R, R28D, N32A, K34D, S35L, S36R, V37A, S40R, E42R, G44A, Q46E, T48V, V68T, A70Y, S72A, S78R, K80Q, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M,

S188N, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.

[0278] 특정 실시형태에서, HE 변이체는 CBLB 표적 부위를 절단하며, I-OnuI(서열번호 1 내지 5) 또는 서열번호 6 내지 12 중 어느 하나에 제시된 바와 같은 I-OnuI 변이체, 이들의 생물학적 활성 단편, 및/또는 이들의 추가적인 변이체의 다음의 아미노산 치환: S24C, L26R, R28D, N32A, K34D, S35L, S36R, V37A, S40R, E42R, G44S, Q46E, T48S, V68T, A70Y, S72A, S78R, K80Q, D92G, V116L, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S188N, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.

[0279] 특정 실시형태에서, HE 변이체는 CBLB 표적 부위를 절단하며, I-OnuI(서열번호 1 내지 5) 또는 서열번호 6 내지 12 중 어느 하나에 제시된 바와 같은 I-OnuI 변이체, 이들의 생물학적 활성 단편, 및/또는 이들의 추가적인 변이체의 다음의 아미노산 치환: S24C, L26R, R28D, R30H, N32A, K34V, S35L, S36R, V37S, S40R, E42R, G44S, Q46E, T48V, V68T, V68K, A70Y, S72A, S78R, K80Q, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S188N, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.

[0280] 특정 실시형태에서, HE 변이체는 CBLB 표적 부위를 절단하며, I-OnuI(서열번호 1 내지 5) 또는 서열번호 6 내지 12 중 어느 하나에 제시된 바와 같은 I-OnuI 변이체, 이들의 생물학적 활성 단편, 및/또는 이들의 추가적인 변이체의 다음의 아미노산 치환: S24C, L26G, R28Y, R30H, N32S, K34V, S35L, S36R, V37S, S40R, E42R, G44S, Q46E, T48S, V68T, A70Y, S72A, S78R, K80Q, V116L, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S188N, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.

[0281] 특정 실시형태에서, HE 변이체는 CBLB 표적 부위를 절단하며, I-OnuI(서열번호 1 내지 5) 또는 서열번호 6 내지 12 중 어느 하나에 제시된 바와 같은 I-OnuI 변이체, 이들의 생물학적 활성 단편, 및/또는 이들의 추가적인 변이체의 다음의 아미노산 치환: S24C, L26R, R28D, R30H, N32A, K34V, S35L, S36R, V37S, S40R, E42R, G44S, Q46E, T48V, V68T, A70Y, S72A, S78R, K80Q, V116L, L138M, T143N, S159P, F168L, E178D, C180S, F182V, N184E, I186K, S188R, K189R, K191P, L192V, G193K, Q195G, Q197R, V199R, S201G, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.

[0282] 특정 실시형태에서, HE 변이체는 CBLB 표적 부위를 절단하며, I-OnuI(서열번호 1 내지 5) 또는 서열번호 6 내지 12 중 어느 하나에 제시된 바와 같은 I-OnuI 변이체, 이들의 생물학적 활성 단편, 및/또는 이들의 추가적인 변이체의 다음의 아미노산 치환: S24C, L26R, R28D, N32A, K34D, S35L, S36R, V37A, S40R, E42R, G44A, Q46E, T48V, V68T, A70Y, S72A, S78R, K80Q, D92G, L138M, T143N, S159P, F168L, E178D, C180S, F182M, N184E, I186M, S188N, S190N, K191N, L192V, G193I, Q195R, Q197R, V199R, T203S, K207R, Y223R, K225V, K227N, F232H, D236E 및 V238I 중 적어도 5, 적어도 15, 바람직하게는 적어도 25, 더 바람직하게는 적어도 35, 또는 훨씬 더 바람직하게는 적어도 40개 이상을 포함한다.

[0283] 특정 실시형태에서, 인간 CBLB 유전자에 결합하고, 이를 절단하는 I-OnuI LHE 변이체는 서열번호 6 내지 12의 임의의 하나에 제시된 아미노산 서열에 대해 적어도 80%, 바람직하게는 적어도 85%, 더 바람직하게는 적어도 90%, 또는 더욱더 바람직하게는 적어도 95% 동일한 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.

[0284] 특정 실시형태에서, I-OnuI LHE 변이체는 서열번호 6 내지 12 중 임의의 하나에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.

[0285] 특정 실시형태에서, I-OnuI LHE 변이체는 서열번호 6에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.

[0286] 특정 실시형태에서, I-OnuI LHE 변이체는 서열번호 7에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.

- [0287] 특정 실시형태에서, I-OnuI LHE 변이체는 서열번호 8에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0288] 특정 실시형태에서, I-OnuI LHE 변이체는 서열번호 9에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0289] 특정 실시형태에서, I-OnuI LHE 변이체는 서열번호 10에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0290] 특정 실시형태에서, I-OnuI LHE 변이체는 서열번호 11에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0291] 특정 실시형태에서, I-OnuI LHE 변이체는 서열번호 12에 제시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함한다.
- [0292] 2. 메가TAL
- [0293] 다양한 실시형태에서, 호밍 엔도뉴클레아제 변이체를 포함하는 메가TLA은 CBLB 유전자 내 표적 부위에서 이중-가닥 파손(DSB)을 도입하도록 재프로그래밍된다. 특정 실시형태에서, 메가TAL은 CBLB 유전자의 엑손 6에서, 바람직하게는 CBLB 유전자의 엑손 6 내 서열번호 20에서, 그리고 더 바람직하게는 CBLB 유전자의 엑손 6 내 서열번호 20에서의 서열 "ATTC"에서 DSB를 도입한다.
- [0294] "메가TAL"은 CBLB 유전자 내 DNA 표적 서열에 결합하고, 이를 절단하며, 선택적으로 하나 이상의 링커 및/또는 추가적인 작용성 도메인, 예를 들어, 5'-3' 엑소뉴클레아제, 5'-3' 알칼리성 엑소뉴클레아제, 3'-5' 엑소뉴클레아제(예를 들어, Trex2), 5' 플랩 엔도뉴클레아제, 헬리카제 또는 주형-독립적 DNA 중합효소 활성을 나타내는 말단-가공 효소의 말단-가공 효소 도메인을 포함하는, TALE DNA 결합 도메인 및 호밍 엔도뉴클레아제 변이체를 포함하는 폴리펩타이드를 지칭한다.
- [0295] 특정 실시형태에서, 메가TAL은 5'-3' 엑소뉴클레아제, 5'-3' 알칼리성 엑소뉴클레아제, 3'-5' 엑소뉴클레아제(예를 들어, Trex2), 5' 플랩 엔도뉴클레아제, 헬리카제, 주형-의존적 DNA 중합효소 또는 주형-독립적 DNA 중합효소 활성을 나타내는 말단-가공 효소를 따라서 세포 내로 도입된다. 메가TAL 및 3'가공 효소는 별개로, 예를 들어, 상이한 벡터 또는 별개의 mRNA에서, 또는 함께, 예를 들어, 융합 단백질로서, 또는 바이러스 자기-절단성 펩타이드 또는 IRES 요소에 의해 분리된 다이스트론성 작제물에서 도입될 수 있다.
- [0296] "TALE DNA 결합 도메인"은 식물 전사체를 조작하기 위해 식물 전사 활성체를 모방하는 전사 활성체-유사 효과기 (TALE 또는 TAL-효과기)의 DNA 결합 부분이다(예를 들어, 문헌[Kay et al., 2007. Science 318:648-651] 참조). 특정 실시형태에서 상정된 TALE DNA 결합 도메인은 드노보(de novo)로 또는 천연 유래 TALE로부터, 예를 들어, 세균점무늬병(Xanthomonas campestris pv. vesicatoria), 잔토모나스 가드네리(Xanthomonas gardneri), 잔토모나스 트랜스루센스(Xanthomonas translucens), 잔토모나스 악소노포디스(Xanthomonas axonopodis), 잔토모나스 퍼포란스(Xanthomonas perforans), 잔토모나스 알팔파(Xanthomonas alfalfa), 잔토모나스 시트리(Xanthomonas citri), 잔토모나스 유베시카토리아(Xanthomonas euvesicatoria) 및 잔토모나스 오리자에(Xanthomonas oryzae)로부터의 AvrBs3, 및 랄스토니아 솔라나세아룸(Ralstonia solanacearum)으로부터의 brg11 및 hpx17로부터 조작된다. DNA 결합 도메인을 유도 및 설계하기 위한 TALE 단백질의 예시적인 예는 미국 특허 제9,017,967호 및 이의 참고문헌에 개시되어 있으며, 이를 모두는 본 명세서에 그들의 전문이 참고로 포함된다.
- [0297] 특정 실시형태에서, 메가TAL은 TALE DNA 결합 도메인의 그의 대응하는 표적 DNA 서열에 대한 결합에 연루된 하나 이상의 반복 단위를 포함하는 TALE DNA 결합 도메인을 포함한다. 단일 "반복 단위"(또한 "반복부"로서 지칭됨)는 전형적으로 길이가 33 내지 35개의 아미노산이다. 각각의 TALE DNA 결합 도메인 반복부 단위는 전형적으로 반복부의 위치 12 및/또는 13에서 반복부 가변 2-잔기(RVD)를 구성하는 1 또는 2개의 DNA-결합 잔기를 포함한다. 이들 TALE DNA 결합 도메인의 DNA 인식을 위한 천연(정규) 암호는 위치 12 및 13에서 HD 서열이 사이토신(C)에 대한 결합을 야기하고, NG가 T, NI 내지 A에 결합하며, NN은 G 또는 A에 결합하고, NG가 T에 결합하도록 결정되었다. 특정 실시형태에서, 비-정규(비전형적) RVD가 상정된다.
- [0298] 특정 실시형태에서 상정된 특정 메가TAL에서 사용하기에 적합한 비정규 RVD의 예시적 예는 구아닌(G)의 인식을 위해 HH, KH, NH, NK, NQ, RH, RN, SS, NN, SN, KN; 아데닌(A)의 인식을 위해 NI, KI, RI, HI, SI; 티민(T) NG, HG, KG, RG; 사이토신(C)의 인식을 위해 RD, SD, HD, ND, KD, YG; A 또는 G의 인식을 위해 NV, HN; 및 A 또는 T 또는 G 또는 C의 인식을 위해 H\*, HA, KA, N\*, NA, NC, NS, RA, S\*을 포함하지만, 이들로 제한되지 않

되, (\*)는 위치 13에서 아미노산이 없다는 것을 의미한다. 특정 실시형태에서 상정된 특정 메가TAL에서 사용하기에 적합한 RVD의 추가적인 예시적 예는 본 명세서에 전문이 참고로 포함된 미국 특허 제8,614,092호에 개시된 것을 추가로 포함한다.

[0299] 특정 실시형태에서, 본 명세서에서 상정된 메가TAL은 3 내지 30개의 반복 단위를 포함하는 TALE DNA 결합 도메인을 포함한다. 특정 실시형태에서, 메가TAL은 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 또는 30개의 TALE DNA 결합 도메인 반복 단위를 포함한다. 바람직한 실시형태에서, 본 명세서에서 상정된 메가TAL은 5 내지 15개의 반복 단위, 더 바람직하게는 7 내지 15개의 반복 단위, 더 바람직하게는 9 내지 15개의 반복 단위, 및 더 바람직하게는 9, 10, 11, 12, 13, 14 또는 15개의 반복 단위를 포함하는 TALE DNA 결합 도메인을 포함한다.

[0300] 특정 실시형태에서, 본 명세서에서 상정된 메가TAL은 3 내지 30개의 반복 단위를 포함하는 TALE DNA 결합 도메인 및 한 세트의 TALE 반복 단위의 C-말단에 위치된 20개의 아미노산을 포함하는 추가적인 절단된 TALE 반복부 단위, 즉, 추가적인 C-말단 절반-TELE DNA 결합 도메인 반복부 단위(본 명세서의 다른 곳에 개시된 C-캡의 아미노산 -20 내지 -1, 이하 참조)를 포함한다. 따라서, 특정 실시형태에서, 본 명세서에서 상정된 메가TAL은 3.5 내지 30.5개의 반복 단위를 포함하는 TALE DNA 결합 도메인을 포함한다. 특정 실시형태에서, 메가TAL은 3.5, 4.5, 5.5, 6.5, 7.5, 8.5, 9.5, 10.5, 11.5, 12.5, 13.5, 14.5, 15.5, 16.5, 17.5, 18.5, 19.5, 20.5, 21.5, 22.5, 23.5, 24.5, 25.5, 26.5, 27.5, 28.5, 29.5 또는 30.5개의 TALE DNA 결합 도메인 반복 단위를 포함한다. 바람직한 실시형태에서, 본 명세서에서 상정된 메가TAL은 5.5 내지 15.5개의 반복 단위, 더 바람직하게는 7.5 내지 15.5개의 반복 단위, 더 바람직하게는 9.5 내지 15.5개의 반복 단위, 및 더 바람직하게는 9.5, 10.5, 11.5, 12.5, 13.5, 14.5 또는 15.5개의 반복 단위를 포함하는 TALE DNA 결합 도메인을 포함한다.

[0301] 특정 실시형태에서, 메가TAL은 "N-말단의 도메인(NTD)" 폴리펩타이드, 하나 이상의 TALE 반복부 도메인/단위, "C-말단의 도메인(CTD)" 폴리펩타이드 및 호밍 엔도뉴클레아제 변이체를 포함하는 TLA 효과기 구조를 포함한다. 일부 실시형태에서, NTD, TALE 반복부 및/또는 CTD 도메인은 동일한 종으로부터 유래된다. 다른 실시형태에서, NTD, TALE 반복부 및/또는 CTD 도메인 중 하나 이상은 상이한 종으로부터 유래된다.

[0302] 본 명세서에서 사용되는 용어 "N-말단의 도메인(NTD)" 폴리펩타이드는 천연 유래 TALE DNA 결합 도메인의 N-말단 부분 또는 단편에 측접하는 서열을 지칭한다. NTD 서열은, 존재한다면, TALE DNA 결합 도메인 반복 단위가 DNA에 결합하는 능력을 보유하는 한, 임의의 길이를 가질 수 있다. 특정 실시형태에서, NTD 폴리펩타이드는 TALE DNA 결합 도메인에 대해 N-말단의 적어도 120개 내지 적어도 140개 이상의 아미노산을 포함한다(0은 가장 N-말단의 반복부 단위의 아미노산 1임). 특정 실시형태에서, NTD 폴리펩타이드는 TALE DNA 결합 도메인에 대해 N-말단의 적어도 약 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139 또는 적어도 140개의 아미노산을 포함한다. 일 실시형태에서, 본 명세서에서 상정된 메가TAL은 잔토모나스 TALE 단백질의 적어도 약 아미노산 +1 내지 +122 내지 적어도 약 +1 내지 +137의 NTD 폴리펩타이드를 포함한다(0은 가장 N-말단의 반복부 단위의 아미노산 1이다). 특정 실시형태에서, NTD 폴리펩타이드는 잔토모나스 TALE 단백질의 TALE DNA 결합 도메인에 대해 N-말단의 적어도 약 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136 또는 137개의 아미노산을 포함한다. 일 실시형태에서, 본 명세서에서 상정된 메가TAL은 랄스토니아 TALE 단백질의 적어도 아미노산 +1 내지 +121의 NTD 폴리펩타이드를 포함한다(0은 가장 N-말단의 반복부 단위의 아미노산 1이다). 특정 실시형태에서, NTD 폴리펩타이드는 랄스토니아 TALE 단백질의 TALE DNA 결합 도메인에 대해 N-말단의 적어도 약 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136 또는 137개의 아미노산을 포함한다.

[0303] 본 명세서에서 사용되는 용어 "C-말단의 도메인(CTD)" 폴리펩타이드는 천연 유래 TALE DNA 결합 도메인의 C-말단 부분 또는 단편에 측접하는 서열을 지칭한다. CTD 서열은, 존재한다면, TALE DNA 결합 도메인 반복 단위가 DNA에 결합하는 능력을 보유하는 한, 임의의 길이를 가질 수 있다. 특정 실시형태에서, CTD 폴리펩타이드는 TALE DNA 결합 도메인의 마지막 완전한 반복부에 대한 C-말단의 적어도 20 내지 적어도 85개의 아미노산을 포함한다(처음 20개의 아미노산은 마지막 C-말단의 전체 반복부 단위에 대한 절반-반복부 단위 C-말단이다). 특정 실시형태에서, CTD 폴리펩타이드는 TALE DNA 결합 도메인의 마지막 완전한 반복부에 대해 C-말단의 적어도 약 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 443, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84 또는 적어도 85개의 아미노산을 포함한다. 일 실시형태에서, 본 명세서에서 상정된 메가TAL은 잔토모나스 TALE 단백질의 적어도 약 아미노산 -20 내지 -1의 CTD 폴리펩타이드를 포함한다(-20은 마지막 C-말단의 전체 반복부 단위에 대한 절반-반복부 단위 C-말단의 아미

노산 1이다). 특정 실시형태에서, CTD 폴리펩타이드는 잔토모나스 TALE 단백질의 TALE DNA 결합 도메인의 마지막 전체 반복부에 대해 C-말단의 적어도 약 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1개의 아미노산을 포함한다. 일 실시형태에서, 본 명세서에서 상정된 메가TAL은 랄스토니아 TALE 단백질의 적어도 약 아미노산 -20 내지 -1의 CTD 폴리펩타이드를 포함한다(-20은 마지막 C-말단의 전체 반복부 단위에 대한 절반-반복부 단위 C-말단의 아미노산 1이다). 특정 실시형태에서, CTD 폴리펩타이드는 랄스토니아 TALE 단백질의 TALE DNA 결합 도메인의 마지막 전체 반복부에 대해 C-말단의 적어도 약 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1개의 아미노산을 포함한다.

[0304] 특정 실시형태에서, 본 명세서에서 상정된 메가TAL은 선택적으로 본 명세서의 다른 곳에 상정된 하나 이상의 링커 폴리펩타이드와 서로 결합된 표적 서열에 결합하도록 조작된 TALE DNA 결합 도메인, 표적 서열에 결합하고 절단하도록 재프로그래밍된 호밍 엔도뉴클레아제, 및 선택적으로 NTD 및/또는 CTD 폴리펩타이드를 포함하는 융합 폴리펩타이드를 포함한다. 임의의 특정 이론에 의해 결합되는 일 없이, TALE DNA 결합 도메인, 및 선택적으로 NTD 및/또는 CTD 폴리펩타이드를 포함하는 메가TAL은 호밍 엔도뉴클레아제 변이체에 추가로 융합된 링커 폴리펩타이드에 융합된다는 것이 상정된다. 따라서, TALE DNA 결합 도메인은 호밍 엔도뉴클레아제 변이체의 DNA 결합 도메인에 의해 결합된 표적 서열로부터 떨어져서 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 또는 15개의 뉴클레오타이드 내인 DNA 표적 서열에 결합한다. 이런 방법으로, 본 명세서에서 상정된 메가TAL은 게놈 편집의 특이성 및 효율을 증가시킨다.

[0305] 일 실시형태에서, 메가TAL은 재프로그래밍된 호밍 엔도뉴클레아제의 결합 부위 상류의 약 4, 5 또는 6개의 뉴클레오타이드, 바람직하게는 5 또는 6개의 뉴클레오타이드 내인 뉴클레오타이드 서열에 결합하는 호밍 엔도뉴클레아제 변이체 및 TALE DNA 결합 도메인을 포함한다.

[0306] 일 실시형태에서, 메가TAL은 호밍 엔도뉴클레아제 변이체에 의해 결합되고 절단되는 뉴클레오타이드 서열(서열 번호 20) 상류의 5개 뉴클레오타이드인(즉, TALE 결합 부위와 HE 결합 사이에 4개의 뉴클레오타이드가 있음) 서열번호 21에 제시된 뉴클레오타이드 서열에 결합하는 호밍 엔도뉴클레아제 변이체 및 TALE DNA 결합 도메인을 포함한다. 바람직한 실시형태에서, 메가TAL 표적 서열은 서열번호 22이다.

[0307] 특정 실시형태에서, 본 명세서에서 상정된 메가TAL은 하나 이상의 TALE DNA 결합 반복 단위 및 I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-NcrI, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, I-Vdi141I로 이루어진 군으로부터 선택되는 LHE 및 이들의 변이체, 또는 바람직하게는 I-CpaMI, I-HjeMI, I-OnuI, I-PanMI, SmaMI 및 이들의 변이체, 또는 더 바람직하게는 I-OnuI 및 이의 변이체로부터 설계되거나 또는 재프로그래밍된 LHE 변이체를 포함한다.

[0308] 특정 실시형태에서, 본 명세서에서 상정된 메가TAL은 NTD, 하나 이상의 TALE DNA 결합 반복 단위, CTD, 및 I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-NcrI, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, I-Vdi141I 및 이들의 변이체, 또는 바람직하게는 I-CpaMI, I-HjeMI, I-OnuI, I-PanMI, SmaMI 및 이들의 변이체, 또는 더 바람직하게는 I-OnuI 및 이들의 변이체로 이루어진 군으로부터 선택되는 LHE 변이체를 포함한다.

[0309] 특정 실시형태에서, 본 명세서에서 상정된 메가TAL은 NTD, 약 9.5 내지 약 15.5개의 TALE DNA 결합 반복 단위, 및 I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-NcrI, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, I-Vdi141I 및 이들의 변이체, 또는 바람직하게는 I-CpaMI, I-HjeMI, I-OnuI, I-PanMI, SmaMI 및 이들의 변이체, 또는 더 바람직하게는 I-OnuI 및 이들의 변이체로 이루어진 군으로부터 선택된 LHE 변이체를 포함한다.

[0310] 특정 실시형태에서, 본 명세서에서 상정된 메가TAL은 약 122개의 아미노산 내지 137개의 아미노산, 약 9.5, 약 10.5, 약 11.5, 약 12.5, 약 13.5, 약 14.5 또는 약 15.5개의 결합 반복 단위의 NTD, 약 20개의 아미노산 내지 약 85개의 아미노산의 CTD, 및 I-OnuI LHE 변이체를 포함한다. 특정 실시형태에서, NTD, DNA 결합 도메인 및

CTD 중 임의의 하나, 둘 또는 모두는 임의의 적합한 조합에서 동일한 종 또는 상이한 종으로부터 설계될 수 있다.

- [0311] 특정 실시형태에서, 본 명세서에서 상정된 메가TAL은 서열번호 13 내지 19 중 임의의 하나에 제시된 아미노산 서열을 포함한다.
- [0312] 특정 실시형태에서, 본 명세서에서 상정된 메가TAL-Trex2 융합 단백질은 서열번호 13 내지 19 및 38에 제시된 아미노산 서열을 포함한다.
- [0313] 특정 실시형태에서, 메가TAL은 TALE DNA 결합 도메인을 포함하고, I-OnuI LHE 변이체는 서열번호 22에 제시된 뉴클레오타이드 서열에 결합하고, 이를 절단한다.
- [0314] 3. 말단-가공 효소
- [0315] 특정 실시형태에 상정된 게놈 편집 조성물 및 방법은 뉴클레아제 변이체 및 말단-가공 효소의 하나 이상의 복제물을 이용하여 세포 게놈을 편집하는 것을 포함한다. 특정 실시형태에서, 단일 폴리뉴클레오타이드는 링커, 자기-절단성 웨타이드 서열, 예를 들어, 2A 서열에 의해, 또는 IRES 서열에 의해 분리되는 호밍 엔도뉴클레아제 변이체 및 말단-가공 효소를 암호화한다. 특정 실시형태에서, 게놈 편집 조성물은 뉴클레아제 변이체를 암호화하는 폴리뉴클레오타이드 및 말단-가공 효소를 암호화하는 별개의 폴리뉴클레오타이드를 포함한다. 특정 실시형태에서, 게놈 편집 조성물은 단일 융합 폴리웨타이드에서 호밍 엔도뉴클레아제 변이체 및 말단-가공 효소를 암호화하는 폴리뉴클레오타이드를 포함한다. 일 실시형태에서, 융합 폴리웨타이드는 메가TAL, 및 각각 자기-절단성 웨타이드에 의해 분리되는 말단-가공 효소의 하나 이상의 복제물을 포함한다.
- [0316] 용어 "말단-가공 효소"는 폴리뉴클레오타이드 쇄의 노출된 말단을 변형시키는 효소를 지칭한다. 폴리뉴클레오타이드는 이중 가닥 DNA(dsDNA), 단일 가닥 DNA(ssDNA), RNA, DNA 및 RNA의 이중 가닥 혼성체, 및 합성 DNA(예를 들어, A, C, G 및 T 이외의 염기를 함유)일 수 있다. 말단-가공 효소는 하나 이상의 뉴클레오타이드를 첨가하고/하거나 하나 이상의 뉴클레오타이드를 제거하고/하거나 인산염기를 제거 또는 변형시키고/시키거나 하이드록실기를 제거 또는 변형시킴으로써 노출된 폴리뉴클레오타이드 쇄를 변형시킬 수 있다. 말단-가공 효소는 엔도뉴클레아제 절단 부위에서, 또는 다른 화학적 또는 기계적 수단, 예컨대 전단(예를 들어, 미세한 게이지의 바늘을 통과, 가열, 음파처리, 미니 비드 텁블링, 및 네뷸라이징에 의해), 이온화 방사선, 자외선 방사, 산소 라디칼, 화학적 가수분해 및 화학요법제에 의해 생성되는 말단에서 말단을 변형시킬 수 있다.
- [0317] 특정 실시형태에서, 특정 실시형태에서 상정된 게놈 편집 조성물 및 방법은 호밍 엔도뉴클레아제 변이체 또는 메가TAL 및 DNA 말단-가공 효소를 이용하여 세포 게놈을 편집하는 것을 포함한다.
- [0318] 용어 "DNA 말단-가공 효소"는 DNA의 노출된 말단을 변형시키는 효소를 지칭한다. DNA 말단-가공 효소는 평활 말단 또는 엇갈린 말단(5' 또는 3' 돌출부를 갖는 말단)을 변형시킬 수 있다. DNA 말단-가공 효소는 단일 가닥 또는 이중 가닥 DNA를 변형시킬 수 있다. DNA 말단-가공 효소는 엔도뉴클레아제 절단 부위에서 또는 다른 화학적 또는 기계적 수단, 예컨대 전단(예를 들어, 미세한 게이지의 바늘을 통과, 가열, 음파처리, 미니 비드 텁블링, 및 네뷸라이징에 의해), 이온화 방사선, 자외선 방사, 산소 라디칼, 화학적 가수분해 및 화학요법제에 의해 생성되는 말단에서 말단을 변형시킬 수 있다. DNA 말단-가공 효소는 하나 이상의 뉴클레오타이드를 첨가하고/하거나 하나 이상의 뉴클레오타이드를 제거하고/하거나 인산염기를 제거 또는 변형시키고/시키거나 하이드록실기를 제거 또는 변형시킴으로써 노출된 DNA 말단을 변형시킬 수 있다.
- [0319] 본 명세서에서 상정된 특정 실시형태에서 사용하기에 적합한 DNA 말단-가공 효소의 예시적인 예는 5'-3' 엑소뉴클레아제, 5'-3' 알칼리성 엑소뉴클레아제, 3'-5' 엑소뉴클레아제, 5' 플랩 엔도뉴클레아제, 헬리카제, 포스파타제, 가수분해효소 및 주형-독립적 DNA 중합효소를 포함하지만, 이들로 제한되지 않는다.
- [0320] 본 명세서에서 상정된 특정 실시형태에서 사용하기에 적합한 DNA 말단-가공 효소의 추가적인 예시적인 예는 Trex2, Trex1, 막관통 도메인이 없는 Trex1, Apollo, Artemis, DNA2, Exo1, ExoT, ExoIII, Fen1, Fan1, MreII, Rad2, Rad9, TdT(말단의 데옥시뉴클레오티딜 트랜스퍼라제), PNKP, RecE, RecJ, RecQ, 람다 엑소뉴클레아제, Sox, 백시니아 DNA 중합효소, 엑소뉴클레아제 I, 엑소뉴클레아제 III, 엑소뉴클레아제 VII, NDK1, NDK5, NDK7, NDK8, WRN, T7-엑소뉴클레아제 유전자 6, 조류 골수아세포종 바이러스 통합 단백질(IN), Bloom, Antarctic 포스파타제, 알칼리성 포스파타제, 폴리뉴클레오타이드 키나제(PNK), Ape1, Mung Bean 뉴클레아제, Hex1, TTRAP (TDP2), Sgs1, Sae2, CUP, Pol mu, Pol Lambda, MUS81, EME1, EME2, SLX1, SLX4 및 UL-12를 포함하지만, 이들로 제한되지 않는다.

- [0321] 특정 실시형태에서, 본 명세서에서 상정된 세포 계놈을 편집하기 위한 계놈 편집 조성물 및 방법은 호밍 엔도뉴 클레아제 변이체 또는 메가TAL 및 엑소뉴클레아제를 포함하는 폴리펩타이드를 포함한다. 용어 "엑소뉴클레아제"는 3' 또는 5' 말단 중 하나에서 포스포다이에스터 결합을 과손하는 가수분해 반응을 통해 폴리뉴클레오타이드 쇄의 말단에서 포스포다이에스터 결합을 절단하는 효소를 지칭한다.
- [0322] 본 명세서에서 상정된 특정 실시형태에서 사용하기에 적합한 엑소뉴클레아제의 예시적 예는 hExoI, 효모 ExoI, 이콜라이(*E. coli*) ExoI, hTREX2, 마우스 TREX2, 래트 TREX2, hTREX1, 마우스 TREX1 및 래트 TREX1을 포함하지만, 이들로 제한되지 않는다.
- [0323] 특정 실시형태에서, DNA 말단-가공 효소는 3' 또는 5' 엑소뉴클레아제, 바람직하게는 Trex 1 또는 Trex2, 더 바람직하게는 Trex2, 및 더욱더 바람직하게는 인간 또는 마우스 Trex2이다.
- [0324] D. 표적 부위
- [0325] 특정 실시형태에서 상정된 뉴클레아제 변이체는 임의의 적합한 표적 서열에 결합하도록 설계될 수 있고, 천연 유래 뉴클레아제에 비해 신규한 결합 특이성을 가질 수 있다. 특정 실시형태에서, 표적 부위는 프로모터, 인핸서, 리프레서 요소 등을 포함하지만, 이들로 제한되지 않는, 유전자의 조절 영역이다. 특정 실시형태에서, 표적 부위는 유전자 또는 스플라이스 부위의 암호 영역이다. 특정 실시형태에서, 뉴클레아제 변이체는 유전자의 발현을 하향조절하거나 또는 감소시키도록 설계된다. 특정 실시형태에서, 뉴클레아제 변이체 및 공여자 수선 주형은 목적으로 하는 표적 서열을 수선하거나 또는 결실시키도록 설계될 수 있다.
- [0326] 다양한 실시형태에서, 뉴클레아제 변이체는 인간 카시타스 B-계통(Cbl) 림프종 원발암유전자 B(CBLB) 유전자에서 표적 서열에 결합하고 이를 절단한다. CBL은 또한 CBL2; 소아 콜수단핵구 백혈병(NSLL)과 함께 또는 이것이 없는 누난 증후군-유사 장애; C-CBL; RING 평거 단백질 55(RNF55); 취약 부위, 엽산 유형, 회귀, fra (11)(q23.3) (FRA11B); E3 유비퀴틴-단백질 리가제 CBL; Cas-Br-M(뮤린) 제한숙주역(ecotropic) 레트로바이러스 형질전환 서열; Cbl 원발암유전자, E3 유비퀴틴 단백질 리가제; RING-type E3 유비퀴틴 트랜스퍼라제 CBL; 카시타스 B-계통 림프종 원발암유전자; 종양유전자 CBL2; 원발암유전자 c-Cbl; 및 신호전달 단백질 CBL로서 지칭된다.
- [0327] 이 유전자는 RING 평거 E3 유비퀴틴 리가제를 암호화하는 원발암유전자이다. 암호화된 단백질은 프로테아좀에 의한 분해를 위한 기질을 표적화하는 데 필요한 효소 중 하나이다. 이 단백질은 유비퀴틴 접합 효소(E2)로부터의 유비퀴틴의 특정 기질에 대한 전달을 매개한다. 이 단백질은 또한 수많은 타이로신-인산화된 기질과 상호작용하도록 허용하고 프로테아좀 분해를 위해 그들을 표적화하는 N-말단의 포스포타이로신 결합 도메인을 함유한다. CBLB는 T 세포 활성화 및 지속을 비롯한 다수의 신호전달 경로의 음성 조절자로서 작용한다.
- [0328] 특정 실시형태에서, 호밍 엔도뉴클레아제 변이체 또는 메가TAL은 CBLB 유전자 내 표적 부위에서 이중-가닥 파손(DSB)을 도입한다. 특정 실시형태에서, 호밍 엔도뉴클레아제 변이체 또는 메가TAL은 CBLB 유전자의 엑소 6에서, 바람직하게는 CBLB 유전자의 엑손 6에서 서열번호 20, 더 바람직하게는 CBLB 유전자의 엑손 6에서 서열번호 20에서 서열 "ATTC"에서 DSB를 도입한다.
- [0329] 바람직한 실시형태에서, 호밍 엔도뉴클레아제 변이체 또는 메가TAL은 이중 가닥 DNA를 절단하고, 서열번호 20 또는 22에 제시된 폴리뉴클레오타이드 서열 내로 DSB를 도입한다.
- [0330] 바람직한 실시형태에서, CBLB 유전자는 인간 CBLB 유전자이다.
- [0331] E. 공여자 수선 주형
- [0332] 뉴클레아제 변이체는 표적 서열에서 DSB를 도입하기 위해 사용될 수 있고; DSB는 하나 이상의 공여자 수선 주형의 존재 하에 상동 직접 수선(HDR) 메커니즘을 통해 수선될 수 있다.
- [0333] 다양한 실시형태에서, 공여자 수선 주형은 면역효능 인핸서, 면역억제 신호 램퍼 또는 조작된 항원 수용체를 암호화하는 하나 이상의 폴리뉴클레오타이드를 포함한다.
- [0334] 다양한 실시형태에서, 상이한 면역억제 경로를 표적화하는 면역효능 인핸서 및/또는 면역억제 신호 램퍼를 독립적으로 암호화하는 복수의 공여자 수선 주형의 존재 하에 조작된 뉴클레아제의 세포를 제공하는 것은 치료 효능 및 지속성이 증가된 계놈 편집된 T 세포를 수득한다는 것이 상정된다. 예를 들어, PD-1, LAG-3, CTLA-4, TIM3, IL-10R, TIGIT 및 TGF $\beta$ RII 경로의 조합을 표적화하는 면역효능 인핸서 또는 면역억제 신호는 특정 실시형태에서 바람직할 수 있다.

- [0335] 특정 실시형태에서, 공여자 수선 주형은 서열을 계놈 내로 삽입하는 데 사용된다. 특정 바람직한 실시형태에서, 공여자 수선 주형은 계놈 내 서열을 수선하거나 또는 변형시키기 위해 사용된다.
- [0336] 다양한 실시형태에서, 공여자 수선 주형은 공여자 수선 주형을 포함하는 아데노-연관 바이러스(AAV), 레트로바이러스, 예를 들어, 렌티바이러스, IDLV 등, 단순 포진 바이러스, 아데노바이러스 또는 백시니아 바이러스 벡터를 이용하여 세포를 형질도입함으로써, 조혈세포, 예를 들어, T 세포 내로 도입된다.
- [0337] 특정 실시형태에서, 공여자 수선 주형은 DSB 부위에 측접하는 하나 이상의 상동성 아암을 포함한다.
- [0338] 본 명세서에서 사용되는 용어 "상동성 아암"은 표적 부위에서 뉴클레아제에 의해 도입되는 DNA 파손에 측접하는 DNA 서열과 동일하거나 또는 거의 동일한 공여자 수선 주형 내 핵산 서열을 지칭한다. 일 실시형태에서, 공여자 수선 주형은 DNA 파손 부위의 DNA 서열 5'과 동일하거나 또는 거의 동일한 핵산 서열을 포함하는 5' 상동성 아암을 포함한다. 일 실시형태에서, 공여자 수선 주형은 DNA 파손 부위의 DNA 서열 3'과 동일하거나 또는 거의 동일한 핵산 서열을 포함하는 3' 상동성 아암을 포함한다. 바람직한 실시형태에서, 공여자 수선 주형은 5'상동성 아암 및 3'상동성 아암을 포함한다. 공여자 수선 주형은 DSB 부위에 바로 인접한 계놈 서열에 대해 상동성, 또는 DSB 부위로부터 다수의 염기쌍 내에서 계놈 서열에 대한 상동성을 포함할 수 있다. 특정 실시형태에서, 상동성 아암의 쌍은 표적 부위의 재절단을 최소화하기 위해 표적 부위 내 돌연변이에 의한 이중 가닥 파손을 위한 표적 부위를 포함하는 폴리뉴클레오타이드 서열을 포함하는 상동성 아암을 포함한다. 일 실시형태에서, 공여자 수선 주형은 약 5bp, 약 10bp, 약 25bp, 약 50bp, 약 100bp, 약 200bp, 약 300bp, 약 400bp, 약 500bp, 약 600bp, 약 700bp, 약 800bp, 약 900bp, 약 1000bp, 약 1100bp, 약 1200bp, 약 1300bp, 약 1400bp, 약 1500bp, 약 1600bp, 약 1700bp, 약 1800bp, 약 1900bp, 약 2000bp, 약 2100bp, 약 2200bp, 약 2300bp, 약 2400bp, 약 2500bp, 약 2600bp, 약 2700bp, 약 2800bp, 약 2900bp, 또는 약 3000bp 이상(상동성 서열의 임의의 개재된 길이를 포함)의 계놈 서열에 대해 상동성인 핵산 서열 또는 상동성 아암을 포함한다.
- [0339] 특정 실시형태에 상정된 상동성 아암의 적합한 길이의 예시적 예는 독립적으로 선택될 수 있으며, 5 bp, 약 10 bp, 약 25 bp, 약 50 bp, 약 100bp, 약 200bp, 약 300bp, 약 400bp, 약 500bp, 약 600bp, 약 700bp, 약 800bp, 약 900bp, 약 1000bp, 약 1100bp, 약 1200bp, 약 1300bp, 약 1400bp, 약 1500bp, 약 1600bp, 약 1700bp, 약 1800bp, 약 1900bp, 약 2000bp, 약 2100bp, 약 2200bp, 약 2300bp, 약 2400bp, 약 2500bp, 약 2600bp, 약 2700bp, 약 2800bp, 약 2900bp, 또는 약 3000bp 이상의 상동성 아암(상동성 아암의 모든 개재 길이를 포함)을 포함하지만, 이들로 제한되지 않는다.
- [0340] 적합한 상동성 아암 길이의 추가적인 예시적 예는 약 100bp 내지 약 3000bp, 약 200bp 내지 약 3000bp, 약 300bp 내지 약 3000bp, 약 400bp 내지 약 3000bp, 약 500bp 내지 약 3000bp, 약 500bp 내지 약 2500bp, 약 500bp 내지 약 2000bp, 약 750bp 내지 약 2000bp, 약 750bp 내지 약 1500bp, 또는 약 1000bp 내지 약 1500bp (상동성 아암의 모든 개재 길이를 포함)를 포함하지만, 이들로 제한되지 않는다.
- [0341] 특정 실시형태에서, 5' 및 3'상동성 아암의 길이는 약 500bp 내지 약 1500bp로부터 독립적으로 선택된다. 일 실시형태에서, 5' 상동성 아암은 약 1500bp이고, 상기 3'상동성 아암은 약 1000bp이다. 일 실시형태에서, 5'상동성 아암은 약 200bp 내지 약 600bp이고, 3'상동성 아암은 약 200bp 내지 약 600bp이다. 일 실시형태에서, 5' 상동성 아암은 약 200bp이고, 상기 3'상동성 아암은 약 200bp이다. 일 실시형태에서, 5' 상동성 아암은 약 300bp이고, 상기 3'상동성 아암은 약 300bp이다. 일 실시형태에서, 5' 상동성 아암은 약 400bp이고, 상기 3'상동성 아암은 약 400bp이다. 일 실시형태에서, 5' 상동성 아암은 약 500bp이고, 상기 3'상동성 아암은 약 500bp이다. 일 실시형태에서, 5' 상동성 아암은 약 600bp이고, 상기 3'상동성 아암은 약 600bp이다.
- [0342] 공여자 수선 주형은 특정 실시형태에서 하나 이상의 발현 카세트를 포함한다. 소정의 실시형태에서, 공여자 수선 주형은 프로모터 및/또는 인핸서, 비번역 영역(UTR), 코작 서열, 폴리아데닐화 신호, 추가적인 제한 효소 부위, 다중 클로닝 부위, 내부 리보솜 유입 부위(IRES), 재조합효소 인식 부위(예를 들어, LoxP, FRT 및 Att 부위), 종결 코돈, 전사 종결 신호 및 폴리뉴클레오타이드 암호화 자기-절단성 폴리펩타이드 및 에피토프 태그를 포함하지만, 이들로 제한되지 않는 하나 이상의 폴리뉴클레오타이드 및 하나 이상의 상동성 아암을 포함할 수 있다.
- [0343] 다양한 실시형태에서, 공여자 수선 주형은 5'상동성 아암, RNA 중합효소 II 프로모터, 하나 이상의 폴리뉴클레오타이드(면역효능 인핸서, 면역억제 신호 뎁퍼 또는 조작된 항원 수용체를 암호화) 및 3'상동성 아암을 포함한다.
- [0344] 다양한 실시형태에서, 표적 부위는 5'상동성 아암, 하나 이상의 폴리뉴클레오타이드(자기-절단 바이러스 웨타이드, 예를 들어, T2A, 면역효능 인핸서, 면역억제 신호 뎁퍼 또는 조작된 항원 수용체를 암호화), 선택적으로 폴리(A) 신호 또는 자기-절단성 웨타이드 및 3'상동성 아암을 포함하는 공여자 수선 주형에 의해 변형되어, 하나

이상의 폴리뉴클레오타이드의 발현은 내인성 CBLB 프로모터에 의해 지배된다.

[0345] 1. 면역효능 인핸서

특정 실시형태에서, 본 명세서에서 상정된 게놈 편집된 면역 효과기 세포는 면역효능 인핸서를 암호화하는 공여자 수선 주형의 존재 하에 CBLB 유전자에서 DSB를 도입함으로써 면역억제 인자에 대해 더 강하고/강하거나 내성으로 만들어진다. 본 명세서에서 사용되는 용어 "면역효능 인핸서"는 T 세포 활성화를 자극하고/하거나 강력하게 하고/하거나, 면역강화 인자, 및 종양 미세환경으로부터의 면역억제 신호를 T 세포 또는 다른 면역 세포에서 면역조절 신호로 전환시키는 비천연유래 폴리펩타이드로 작용하는 비천연 유래 분자를 지칭한다.

특정 실시형태에서, 면역효능 인핸서는 이중특이성 T 세포 관여자(BiTE) 분자; 사이토카인, 케모카인, 사이토톡신 및/또는 사이토카인 수용체를 포함하지만, 이들로 제한되지 않는, 면역강화 인자; 및 플립 수용체로 이루어진 군으로부터 선택된다.

일부 실시형태에서, 면역효능 인핸서, 면역강화 인자 또는 플립 수용체는 단백질 탈안정화 도메인을 포함하는 융합 폴리펩타이드이다.

[0349] a. 이중특이성 T 세포 관여자(BiTE) 분자

특정 실시형태에서, 본 명세서에서 상정된 게놈 편집된 면역 효과기 세포는 이중특이성 T 세포 관여자(BiTE) 분자를 암호화하는 폴리뉴클레오타이드를 포함하는 공여자 수선 주형의 존재 하에 CBLB 유전자에서 DSB를 도입함으로써 더 강하게 된다. BiTE 분자는 본 명세서의 다른 곳에 상정된 바와 같은 표적 항원, 링커 또는 스페이서에 결합하는 제1 결합 도메인, 및 면역 효과기 세포 상에서 자극 또는 공자극 분자에 결합하는 제2 결합 도메인을 포함하는 이분(bipartite) 분자이다. 제1 및 제2 결합 도메인은 리간드, 수용체, 항체 또는 항원 결합 단편, 렉틴 및 탄수화물로부터 독립적으로 선택될 수 있다.

특정 실시형태에서, 제1 및 제2 결합 도메인은 항원 결합 도메인이다.

특정 실시형태에서, 제1 및 제2 결합 도메인은 항체 또는 이들의 항원 결합 단편이다. 일 실시형태에서, 제1 및 제2 결합 도메인은 단일쇄 가변 단편(scFv)이다.

특정 실시형태에서 제1 결합 도메인에 의해 인식되고 결합될 수 있는 예시적인 예는 알파 엽산 수용체, 5T4,  $\alpha$   $\beta$  6 인테그린, BCMA, B7-H3, B7-H6, CAIX, CD16, CD19, CD20, CD22, CD30, CD33, CD44, CD44v6, CD44v7/8, CD70, CD79a, CD79b, CD123, CD138, CD171, CEA, CSPG4, EGFR, ErbB2(HER2)를 포함하는 EGFR 패밀리, EGFRvIII, EGP2, EGP40, EPCAM, EphA2, EpCAM, FAP, 태아 AchR, GD2, GD3, 글리피칸-3(GPC3), HLA-A1+MAGE1, HLA-A2+MAGE1, HLA-A3+MAGE1, HLA-A1+NY-ESO-1, HLA-A2+NY-ESO-1, HLA-A3+NY-ESO-1, IL-11R  $\alpha$ , IL-13R  $\alpha$  2, 람다, 루이스-Y, 카파, 메소텔린, Muc1, Muc16, NCAM, NKG2D 리간드, NY-ESO-1, PRAME, PSCA, PSMA, ROR1, SSX, 서비빈, TAG72, TEMs, VEGFR2 및 WT-1을 포함하지만, 이들로 제한되지 않는다.

표적 항원의 다른 예시적 실시형태는 MHC-펩타이드 복합체를 포함하되, 선택적으로 웨პ타이드는 알파 엽산 수용체, 5T4,  $\alpha$   $\beta$  6 인테그린, BCMA, B7-H3, B7-H6, CAIX, CD16, CD19, CD20, CD22, CD30, CD33, CD44, CD44v6, CD44v7/8, CD70, CD79a, CD79b, CD123, CD138, CD171, CEA, CSPG4, EGFR, ErbB2(HER2)를 포함하는 EGFR 패밀리, EGFRvIII, EGP2, EGP40, EPCAM, EphA2, EpCAM, FAP, 태아 AchR, GD2, GD3, 글리피칸-3(GPC3), MAGE1, NY-ESO-1, IL-11R  $\alpha$ , IL-13R  $\alpha$  2, 람다, 루이스-Y, 카파, 메소텔린, Muc1, Muc16, NCAM, NKG2D 리간드, PRAME, PSCA, PSMA, ROR1, SSX, 서비빈, TAG72, TEMs, VEGFR2 및 WT-1로부터 가공된다.

특정 실시형태에서 제2 결합 도메인에 의해 인식되고 결합되는 면역 효과기 세포 상의 자극 또는 공자극 분자의 예시적 예는 CD3  $\gamma$ , CD3  $\delta$ , CD3  $\epsilon$ , CD3  $\zeta$ , CD28, CD134, CD137 및 CD278을 포함하지만, 이들로 제한되지 않는다.

특정 실시형태에서, DSB는 조작된 뉴클레아제에 의해 CBLB 유전자에서 유도되고, BiTE를 암호화하는 폴리뉴클레오타이드를 포함하는 공여자 수선 주형은 세포 내로 도입되며, 상동성 재조합에 의해 CBLB 유전자 내로 삽입된다.

[0357] b. 면역강화 인자

특정 실시형태에서, 본 명세서에서 상정된 게놈 편집된 면역 효과기 세포는 게놈 편집된 세포 또는 종양 미세환경 내 세포 중 하나에서 면역강화 인자를 증가시킴으로써 더 강하게 된다. 면역강화 인자는 면역 효과기 세포에서 면역 반응을 강화시키는 특정 사이토카인, 케모카인, 사이토톡신 및 사이토카인 수용체를 지칭한다. 일 실시

형태에서, T 세포는 사이토카인, 케모카인, 사이토톡신 또는 사이토카인 수용체를 암호화하는 폴리뉴클레오타이드를 포함하는 공여자 수선 주형의 존재 하에 CBLB 유전자에서 DSB를 도입함으로써 조작된다.

[0359] 특정 실시형태에서, 공여자 수선 주형은 IL-2, 인슐린, IFN- $\gamma$ , IL-7, IL-21, IL-10, IL-12, IL-15 및 TNF- $\alpha$ 로 이루어진 군으로부터 선택되는 사이토카인을 암호화하는 폴리뉴클레오타이드를 포함한다.

[0360] 특정 실시형태에서, 공여자 수선 주형은 MIP-1 $\alpha$ , MIP-1 $\beta$ , MCP-1, MCP-3 및 RANTES로 이루어진 군으로부터 선택되는 케모카인을 암호화하는 폴리뉴클레오타이드를 포함한다.

[0361] 특정 실시형태에서, 공여자 수선 주형은 퍼포린, 그랜자임 A 및 그랜자임 B로 이루어진 군으로부터 선택된 사이토톡신을 암호화하는 폴리뉴클레오타이드를 포함한다.

[0362] 특정 실시형태에서, 공여자 수선 주형은 IL-2 수용체, IL-7 수용체, IL-12 수용체, IL-15 수용체 및 IL-21 수용체로 이루어진 군으로부터 선택되는 사이토카인 수용체를 암호화하는 폴리뉴클레오타이드를 포함한다.

#### c. 플립 수용체

[0364] 특정 실시형태에서, 본 명세서에서 상정된 개념 편집된 면역 효과기 세포는 종양 미세환경에 의해 유발되는 면역억제 인자에 의한 면역억제 신호를 양성 면역자극 신호로 "플립핑(flapping)" 또는 "반전"시킴으로써 고갈에 대해 더 저항성으로 된다. 일 실시형태에서, 플립 수용체를 암호화하는 폴리뉴클레오타이드를 포함하는 공여자 수선 주형의 존재 하에 CBLB 유전자에 DSB를 도입함으로써 T 세포가 조작된다. 본 명세서에서 사용되는 용어 "플립 수용체"는 종양 미세환경으로부터의 면역억제 신호를 T 세포에서 면역자극 신호로 전환시키는 비천연 유래 폴리펩타이드를 지칭한다. 바람직한 실시형태에서, 플립 수용체는 면역억제 인자에 결합하는 엑소도메인, 막관통 도메인, 및 T 세포에 면역자극 신호를 전달하는 엔도도메인을 포함하는 폴리펩타이드를 지칭한다.

[0365] 일 실시형태에서, 공여자 수선 주형은 면역억제 사이토카인, 막관통 도메인 및 면역강화 사이토카인 수용체의 엔도도메인에 결합하는 엑소도메인 또는 세포외 결합 도메인을 포함하는 플립 수용체를 암호화한다.

[0366] 특정 실시형태에서, 플립 수용체는 IL-4 수용체, IL-6 수용체, IL-8 수용체, IL-10 수용체, IL-13 수용체, TGF $\beta$  수용체 1 또는 TGF $\beta$  수용체 2의 세포외 사이토카인 결합 도메인인 면역억제 사이토카인에 결합하는 엑소도메인; CD4, CD8 $\alpha$ , CD27, CD28, CD134, CD137, CD3, TGF $\beta$  수용체 1, TGF $\beta$  수용체 2, IL-2 수용체, IL-7 수용체, IL-12 수용체, IL-15 수용체 또는 IL-21 수용체로부터 단리된 막관통; 및 IL-2 수용체, IL-7 수용체, IL-12 수용체, IL-15 수용체 또는 IL-21 수용체로부터 단리된 엔도도메인을 포함한다.

[0367] 특정 실시형태에서, 플립 수용체는 IL-4, IL-6, IL-8, IL-10, IL-13, TGF $\beta$  수용체 1 또는 TGF $\beta$  수용체 2에 결합하는 항체 또는 이의 항원 결합 단편인 면역억제 사이토카인에 결합하는 엑소도메인; CD4, CD8 $\alpha$ , CD27, CD28, CD134, CD137, CD3, TGF $\beta$  수용체 1 또는 TGF $\beta$  수용체 2, IL-2 수용체, IL-7 수용체, IL-12 수용체, IL-15 수용체 또는 IL-21 수용체로부터 단리된 막관통; 및 IL-2 수용체, IL-7 수용체, IL-12 수용체, IL-15 수용체 또는 IL-21 수용체로부터 단리된 엔도도메인을 포함한다.

[0368] 일 실시형태에서, 공여자 수선 주형은 면역억제 인자에 결합하는 엑소도메인, 막관통 도메인 및 하나 이상의 세포내 공자극 신호전달 도메인 및/또는 1차 신호전달 도메인을 포함하는 플립 수용체를 포함한다.

[0369] 특정 실시형태에 상정된 플립 수용체의 특정 실시형태에서 사용하기에 적합한 엑소도메인의 예시적 예는 ITIM 및/또는 ITSM을 포함하는 수용체의 세포외 리간드 결합 도메인을 포함하지만, 이들로 제한되지 않는다.

[0370] 특정 실시형태에 상정된 플립 수용체의 특정 실시형태에서 사용하기에 적합한 엑소도메인의 추가적인 예시적 예는 PD-1, LAG-3, TIM-3, CTLA-4, BTLA, CEACAM1, TIGIT, TGF $\beta$  RI, TGF $\beta$  RII, IL4R, IL6R, CXCR1, CXCR2, IL10R, IL13R $\alpha$  2, TRAILR1, RCAS1R 및 FAS의 세포외 리간드 결합 도메인을 포함하지만, 이들로 제한되지 않는다.

[0371] 일 실시형태에서, 엑소도메인은 PD-1, LAG-3, TIM-3, CTLA-4, IL10R, TIGIT, TGF $\beta$  RI 및 TGF $\beta$  RII로 이루어진 군으로부터 선택된 수용체의 세포외 리간드 결합 도메인을 포함한다.

[0372] 일 실시형태에서, 공여자 수선 주형은 면역억제 사이토카인에 결합하는 엑소도메인, 막관통 도메인, 및 하나 이상의 세포내 공자극 신호전달 도메인 및/또는 1차 신호전달 도메인을 포함하는 플립 수용체를 포함한다.

[0373] 특정 실시형태에서 상정되는 플립 수용체의 특정 실시형태에서 사용하기에 적합한 막관통 도메인의 예시적 예는 다음의 단백질: PD-1, LAG-3, TIM-3, CTLA-4, IL10R, TIGIT, T-세포 수용체의 TGF $\beta$  RI 및 TGF $\beta$  RII 알파 또는

베타 쇄, CD $\delta$ , CD3 $\epsilon$ , CD3 $\zeta$ , CD4, CD5, CD8 $\alpha$ , CD9, CD 16, CD22, CD27, CD28, CD33, CD37, CD45, CD64, CD80, CD86, CD 134, CD137 또는 CD154를 포함하지만, 이들로 제한되지 않는다.

[0374] 다양한 실시형태에서, 플립 수용체는 면역자극 신호를 유발하는 엔도도메인을 포함한다. 본 명세서에서 사용되는 용어 "엔도도메인"은 면역수용체 타이로신 활성화 모티프(ITAM), 공자극 신호전달 도메인, 1차 신호전달 도메인 또는 T 세포에서 면역자극 신호의 유발과 관련된 다른 세포내 도메인을 포함하지만, 이들로 제한되지 않는 면역자극 모티프 또는 도메인을 지칭한다.

[0375] 특정 실시형태에서 상정된 플립 수용체의 특정 실시형태에서 사용하기에 적합한 엔도도메인의 예시적 예는 ITAM 모티프를 포함하는 도메인을 포함하지만, 이들로 제한되지 않는다.

[0376] 특정 실시형태에서 상정된 플립 수용체의 특정 실시형태에서 사용하기에 적합한 엔도도메인의 추가적인 예시적 예는 TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR8, TLR9, TLR10, CARD11, CD2, CD7, CD27, CD28, CD30, CD40, CD54 (ICAM), CD83, CD134 (OX40), CD137 (4-1BB), CD278 (ICOS), DAP10, LAT, NKG2C, SLP76, TRIM 또는 ZAP70로부터 단리된 공자극 신호전달 도메인을 포함하지만, 이들로 제한되지 않는다.

[0377] 특정 실시형태에서 상정된 플립 수용체의 특정 실시형태에서 사용하기에 적합한 엔도도메인의 추가적인 예시적 예는 IL-2 수용체, IL-7 수용체, IL-12 수용체 또는 IL-21 수용체로부터 단리된 엔도도메인을 포함하지만, 이들로 제한되지 않는다.

[0378] 특정 실시형태에서 상정된 플립 수용체의 특정 실시형태에서 사용하기에 적합한 엔도도메인의 추가적인 예시적 예는 FcR $\gamma$ , FcR $\beta$ , CD3 $\gamma$ , CD3 $\delta$ , CD3 $\epsilon$ , CD3 $\zeta$ , CD22, CD79a, CD79b 및 CD66d로부터 단리된 1차 신호전달 도메인을 포함하지만, 이들로 제한되지 않는다.

[0379] 특정 실시형태에서, 플립 수용체는 PD-1, LAG-3, TIM-3, CTLA-4, IL10R, TIGIT, TGF $\beta$ RI 또는 TGF $\beta$ RII의 세포 외 도메인을 포함하는 엑소도메인: IL-2 수용체, IL-7 수용체, IL-12 수용체, IL-15 수용체 또는 IL-21 수용체로부터 단리된 엔도도메인을 포함한다.

[0380] 특정 실시형태에서, 플립 수용체는 PD-1, LAG-3, TIM-3, CTLA-4, IL10R, TIGIT, TGF $\beta$ RI 또는 TGF $\beta$ RII의 세포 외 도메인을 포함하는 엑소도메인; CD3 폴리펩타이드, CD4, CD8 $\alpha$ , CD28, CD134, CD137, PD-1, LAG-3, TIM-3, CTLA-4, IL10R, TGF $\beta$ RI 및 TGF $\beta$ RII로부터의 막관통 도메인; 및 CD28, CD134, CD137, CD278 및/또는 CD3 $\zeta$ 로부터의 엔도도메인을 포함한다.

[0381] 특정 실시형태에서, 플립 수용체는 PD-1, LAG-3, TIM-3, CTLA-4, IL10R, TIGIT 또는 TGF $\beta$ RI로부터의 세포외 도메인을 포함하는 엑소도메인; CD3 폴리펩타이드, CD4, CD8 $\alpha$ , CD28, CD134 또는 CD137로부터의 막관통 도메인; 및 CD28, CD134, CD137, CD278 및/또는 CD3 $\zeta$ 로부터의 엔도도메인을 포함한다.

## 2. 면역억제 신호 램프

[0383] 존재하는 적응 세포 요법을 방해하는 한 가지 제한 또는 문제는 종양 미세환경에 의해 매개되는 고갈에 기인하는 면역 효과기 세포의 저장성이다. 고갈된 T 세포는 천연, 효과기 또는 기억 T 세포와 현저히 별개인 독특한 분자 서명을 가진다. 그들은 감소된 사이토카인 발현 및 효과기 기능을 갖는 T 세포로서 정의된다.

[0384] 특정 실시형태에서, 본 명세서에서 상정된 계놈 편집된 면역 효과기 세포는 면역억제 인자에 의해 신호전달을 감소시키거나 또는 램핑함으로써 고갈에 대해 더 내성으로 된다. 일 실시형태에서, T 세포는 면역억제 신호 램프를 암호화하는 폴리뉴클레오타이드를 포함하는 공여자 수선 주형의 존재 하에 CBLB 유전자에서 DSB를 도입함으로써 조작된다.

[0385] 본 명세서에서 사용되는 용어 "면역억제 신호 램프"는 종양 미세환경으로부터 T 세포까지 면역억제 신호의 형질 도입을 감소시키는 비천연 유래 폴리펩타이드를 지칭한다. 일 실시형태에서, 면역억제 신호 램프는 면역억제 인자에 결합하는 항체 또는 이의 항원 결합 단편이다. 바람직한 실시형태에서, 램프가 면역억제 인자에 결합하는 엑소도메인, 및 선택적으로, 막관통 도메인, 및 선택적으로, 변형된 엔도도메인(예를 들어, 세포내 신호전달 도메인)을 포함하기 때문에, 면역억제 신호 램프는 특정 면역억제 인자 또는 신호전달 경로에 대한 억제, 램핑 또는 우성 음성 효과를 유발하는 폴리펩타이드를 지칭한다.

[0386] 특정 실시형태에서, 엑소도메인은 면역억제 인자를 인식하고, 이에 결합하는 세포외 결합 도메인이다.

[0387] 특정 실시형태에서, 변형된 엔도도메인은 면역억제 신호를 감소시키거나 또는 저해하도록 돌연변이된다. 적합한 돌연변이 전략은 아미노산 치환, 첨가 또는 결실을 포함하지만, 이들로 제한되지 않는다. 적합한 돌연변이는 신

호전달 도메인을 제거하기 위한 엔도도메인 절단, 신호전달 모티프 활성에 중요한 잔기를 제거하기 위한 엔도도메인의 돌연변이 및 수용체 사이클링을 차단시키기 위한 엔도도메인 돌연변이를 추가로 포함하지만, 이들로 제한되지 않는다. 특정 실시형태에서, 엔도도메인이 존재할 때, 면역억제 신호를 전달하지 않거나, 또는 실질적으로 감소된 신호전달을 가진다.

[0388] 따라서, 일부 실시형태에서, 면역억제 신호 램퍼는 종양 미세환경으로부터의 하나 이상의 면역억제 인자에 대한 싱크(sink)로서 작용하며, T 세포에서 대응하는 면역억제 신호전달 경로를 저해한다.

[0389] 하나의 면역억제 신호는 트립토판 이화작용에 의해 매개된다. 암세포에서 인돌아민 2,3-다이옥시게나제(IDO)에 의한 트립토판 이화작용은 종양 미세환경에서 T 세포에 대한 면역억제 효과를 갖는 것으로 나타난 키뉴레닌의 생산을 야기한다. 예를 들어, 문헌[Platten et al. (2012) Cancer Res. 72(21):5435-40] 참조.

[0390] 일 실시형태에서, 공여자 수선 주형은 키뉴레니나제 활성을 갖는 효소를 포함한다.

[0391] 특정 실시형태에서 사용하기에 적합한 키뉴레니나제 활성을 갖는 효소의 예시적 예는 L-키뉴레닌 가수분해효소를 포함하지만, 이들로 제한되지 않는다.

[0392] 일 실시형태에서, 공여자 수선 주형은 면역억제 인자에 의해 매개되는 면역억제 신호전달을 감소시키거나 또는 차단시키는 면역억제 신호램퍼를 암호화하는 하나 이상의 폴리뉴클레오타이드를 포함한다.

[0393] 특정 실시형태에서 상정된 면역억제 신호 램퍼에 의해 표적화된 면역억제 인자의 예시적 예는 세포예정사 리간드 1(PD-L1), 세포예정사 리간드 2(PD-L2), 형질전환 성장인자  $\beta$  (TGF  $\beta$ ), 대식세포 집락 자극 인자 1(M-CSF1), 종양 괴사 인자 관련 세포자멸사 유도 리간드(TRAIL), SiSo 세포 리간드 상에서 발현된 수용체-결합 암 항원(RCAS1), Fas 리간드(FasL), CD47, 인터류킨-4(IL-4), 인터류킨-6(IL-6), 인터류킨-8(IL-8), 인터류킨-10(IL-10) 및 인터류킨-13(IL-13)을 포함하지만, 이들로 제한되지 않는다.

[0394] 다양한 실시형태에서, 면역억제 신호 램퍼는 면역억제 인자에 결합하는 항체 또는 이의 항원 결합 단편을 포함한다.

[0395] 다양한 실시형태에서, 면역억제 신호 램퍼는 면역억제 인자에 결합하는 엑소도메인을 포함한다.

[0396] 특정 실시형태에서, 면역억제 신호 램퍼는 면역억제 인자에 결합하는 엑소도메인 및 막관통 도메인을 포함한다.

[0397] 다른 실시형태에서, 면역억제 신호 램퍼는 형질도입하지 않거나 또는 면역억제 신호를 전달하는 실질적으로 감소된 능력을 갖는 면역억제 인자에 결합하는 엑소도메인, 막관통 도메인, 및 변형된 엔도도메인을 포함한다.

[0398] 본 명세서에서 사용되는 용어 "엑소도메인"은 항원 결합 도메인을 지칭한다. 일 실시형태에서, 엑소도메인은 종양 미세환경으로부터의 면역억제 신호를 T 세포에 형질도입하는 면역억제 수용체의 세포외 리간드 결합 도메인이다. 특정 실시형태에서, 엑소도메인은 면역수용체 타이로신 저해 모티프(ITIM) 및/또는 면역수용체 타이로신 스위치 모티프(ITSM)를 포함하는 수용체의 세포외 리간드 결합 도메인을 지칭한다.

[0399] 면역억제 신호 램퍼의 특정 실시형태에서 사용하기에 적합한 엑소도메인의 예시적 예는 다음의 폴리펩타이드로부터 단리된 항체 또는 이들의 항원 결합 단편, 또는 세포외 리간드 결합 도메인을 포함하지만, 이들로 제한되지 않는다: 세포예정사 단백질 1(PD-1), 림프구 활성화 유전자 3 단백질(LAG-3), T 세포 면역글로불린 도메인 및 뮤신 도메인 단백질 3(TIM3), 세포독성 T 림프구 항원-4(CTLA-4), 밴드 T 림프구 감쇠기(BTLA), T 세포 면역글로불린 및 면역수용체 타이로신-기반 저해 모티프 도메인(TIGIT), 형질전환 성장인자  $\beta$  수용체 1(TGF  $\beta$  RI), 형질전환 성장인자  $\beta$  수용체 2(TGF  $\beta$  RII), 대식세포 집락 자극 인자 1 수용체(CSF1R), 인터류킨 4 수용체 (IL4R), 인터류킨 6 수용체(IL6R), 케모카인(C-X-C 모티프) 수용체 1(CXCR1), 케모카인(C-X-C 모티프) 수용체 2(CXCR2), 인터류킨 10 수용체 서브유닛 알파(IL10R), 인터류킨 13 수용체 서브유닛 알파 2(IL13R  $\alpha$  2), 종양 괴사 인자 관련 세포자멸사 유도 리간드(TRAILR1), SiSo 세포 상에서 발현된 수용체-결합 암항원(RCAS1R) 및 Fas 세포 표면 사멸 수용체(FAS).

[0400] 일 실시형태에서, 엑소도메인은 PD-1, LAG-3, TIM3, CTLA-4, IL10R, TIGIT, CSF1R, TGF  $\beta$  RII 및 TGF  $\beta$  RII로 이루어진 군으로부터 선택되는 수용체의 세포외 리간드 결합 도메인을 포함한다.

[0401] 다수의 막관통 도메인이 특정 실시형태에서 사용될 수 있다. 특정 실시형태에서 상정된 면역억제 신호 램퍼의 특정 실시형태에서 사용하기에 적합한 막관통 도메인의 예시적 예는 다음의 단백질의 막관통 도메인을 포함하지만, 이들로 제한되지 않는다: T-세포 수용체의 알파 또는 베타쇄, CD8, CD3  $\epsilon$ , CD3  $\gamma$ , CD3  $\zeta$ , CD4, CD5, CD8  $\alpha$ , CD9, CD16, CD22, CD27, CD28, CD33, CD37, CD45, CD64, CD80, CD86, CD134, CD137, CD152, CD154 및

PD-1.

## [0402] 3. 조작된 항원 수용체

특정 실시형태에서, 본 명세서에서 상정된 개놈 편집된 면역 효과기 세포는 조작된 항원 수용체를 포함한다. 일 실시형태에서, 조작된 항원 수용체를 암호화하는 폴리뉴클레오타이드를 포함하는 T 세포는 CBLB 유전자에서 DSB를 도입함으로써 편집된다. 일 실시형태에서, T 세포는 조작된 항원 수용체를 암호화하는 폴리뉴클레오타이드를 포함하는 공여자 수선 주형의 존재 하에 CBLB 유전자 내에 DSB를 도입함으로써 조작된다.

특정 실시형태에서, 조작된 항원 수용체는 조작된 T 세포 수용체(TCR), 키메라 항원 수용체(CAR), Daric 수용체 또는 이들의 성분 또는 키메라 사이토카인 수용체이다.

## [0405] a. 조작된 TCR

특정 실시형태에서, 본 명세서에서 상정된 개놈 편집된 면역 효과기 세포는 조작된 TCR을 암호화하는 폴리뉴클레오타이드를 포함한다. 일 실시형태에서, 조작된 TCR을 암호화하는 폴리뉴클레오타이드를 포함하는 T 세포는 CBLB 유전자에서 DSB를 도입함으로써 조작된다. 하나의 실시형태에서, T 세포는 조작된 TCR을 암호화하는 공여자 수선 주형의 존재 하에 CBLB 유전자에서 DSB를 도입함으로써 조작된다. 특정 실시형태에서, 조작된 TCR은 단일 CBLB 대립유전자 내 DSB에서 삽입된다. 다른 실시형태에서, 조작된 TCR의 알파쇄는 하나의 CBLB 대립유전자 내 DSB 내로 삽입되고, 조작된 TCR의 베타쇄는 다른 CBLB 대립유전자 내 DSB 내로 삽입된다.

일 실시형태에서, 본 명세서에서 상정된 조작된 T 세포는 CBLB 유전자에서 삽입되지 않는 조작된 TCR을 암호화하는 폴리뉴클레오타이드를 포함하며, 면역억제 신호 램프, 플립 수용체, 조작된 T 세포 수용체(TCR)의 알파 및/또는 베타쇄, 키메라 항원 수용체(CAR), Daric 수용체 또는 이들의 성분, 또는 키메라 사이토카인 수용체 중 하나 이상은 하나 이상의 CBLB 대립유전자에서 DSB 내로 삽입된다.

천연 유래 T 세포 수용체는 2개의 서브유닛, 즉, 알파쇄 및 베타쇄 서브유닛을 포함하는데, 이들 각각은 각각의 T 세포 개놈에서의 재조합 사건에 의해 생산되는 독특한 단백질이다. TCR의 라이브러리는 특정 표적 항원에 대한 그들의 선택성에 대해 선별될 수 있다. 이런 방식으로, 표적 항원에 대해 높은 결합활성 및 반응성을 갖는 천연 TCR이 선택, 클로닝되고, 후속적으로 적응 면역요법을 위해 사용되는 T 세포 집단 내로 도입될 수 있다.

일 실시형태에서, T 세포는 하나 이상의 CBLB 대립유전자 내 DSB에서 TCR의 서브유닛을 암호화하는 폴리뉴클레오타이드를 포함하는 공여자 수선 주형을 도입함으로써 변형되며, TCR 서브유닛은 표적 항원을 발현시키는 종양 세포에 T 세포에 대한 특이성을 부여하는 TCR을 형성하는 능력을 가진다. 특정 실시형태에서, 서브유닛이 TCR을 형성하는 능력을 보유하고, 형질감염된 T 세포에 표적 세포로 귀소하는 능력을 부여하며, 면역학적으로 적절한 사이토카인 신호전달에 참여하는 한, 서브유닛은 천연 유래 서브유닛에 비해 하나 이상의 치환, 결실, 삽입 또는 변형을 가진다. 조작된 TCR은 바람직하게는 또한 높은 결합활성을 갖는 적절한 종양-관련 웨بت아이드를 나타내는 표적 세포에 결합하고, 선택적으로 생체내에서 적절한 웨بت아이드를 제시하는 표적 세포의 효율적인 사멸을 매개한다.

조작된 TCR을 암호화하는 핵산은 바람직하게는 T 세포의 (천연 유래) 염색체에서 그들의 천연 상황으로부터 단리되고, 본 명세서에 기재된 바와 같은 적합한 벡터 내로 흔입될 수 있다. 핵산과 그들을 포함하는 벡터는 둘다 세포, 바람직하게는 특정 실시형태에서 T 세포 내로 전달될 수 있다. 이어서, 변형된 T 세포는 형질도입된 핵산 또는 핵산들에 의해 암호화된 TCR의 하나 이상의 쇄를 발현시킬 수 있다. 바람직한 실시형태에서, 조작된 TCR은 특정 TCR을 정상적으로 발현시키지 않는 T 세포 내로 도입되기 때문에, 외인성 TCR이다. 조작된 TCR의 필수 양상은 주조직적 복합체(MHC) 또는 유사한 면역학적 성분에 의해 제시된 종양 항원에 대해 높은 결합활성을 가진다는 것이다. 조작된 TCR과 대조적으로, CAR은 MHC 독립적 방식으로 표적 항원에 결합하도록 조작된다.

TCR은 부착된 추가적인 폴리웨بت아이드가 작용성 T 세포 수용체 및 MHC 의존적 항언 인식을 형성하는 알파쇄 또는 베타쇄의 능력을 방해하지 않는 한, 본 발명의 TCR의 알파쇄 또는 베타쇄의 아미노-말단 또는 카복실-말단 부분에 부착된 추가적인 폴리웨بت아이드에 의해 발현될 수 있다.

특정 실시형태에서 상정된 조작된 TCR에 의해 인식되는 항원은 혈액학적 암 및 고형 종양 상의 항원을 포함하는, 암 항원을 포함하지만, 이들로 제한되지 않는다. 예시적인 항원은 알파 엽산 수용체, 5T4,  $\alpha\beta\gamma$  6 인 테그린, BCMA, B7-H3, B7-H6, CAIX, CD19, CD20, CD22, CD30, CD33, CD44, CD44v6, CD44v7/8, CD70, CD79a, CD79b, CD123, CD138, CD171, CEA, CSPG4, EGFR, ErbB2(HER2)를 포함하는 EGFR 패밀리, EGFRvIII, EGP2, EGP40, EPCAM, EphA2, EpCAM, FAP, 태아 AchR, GD2, GD3, 글리피칸-3(GPC3), HLA-A1+MAGE1, HLA-A2+MAGE1,

HLA-A3+MAGE1, HLA-A1+NY-ESO-1, HLA-A2+NY-ESO-1, HLA-A3+NY-ESO-1, IL-11R $\alpha$ , IL-13R $\alpha$  2, 람다, 루이스-Y, 카파, 메소텔린, Muc1, Muc16, NCAM, NKG2D 리간드, NY-ESO-1, PRAME, PSCA, PSMA, ROR1, SSX, 서비빈, TAG72, TEMs, VEGFR2 및 WT-1을 포함하지만, 이들로 제한되지 않는다.

[0413] 일 실시형태에서, 공여자 수선 주형은 RNA 중합효소 II 프로모터 또는 제1 자기-절단 바이러스 웨타이드를 암호화하는 폴리뉴클레오타이드 및 하나의 변형된 및/또는 비-작용성의 CBLB 유전자 내로 통합된 조작된 TCR의 알파쇄 및/또는 베타쇄를 암호화하는 폴리뉴클레오타이드를 포함한다.

[0414] 일 실시형태에서, 공여자 수선 주형은 RNA 중합효소 II 프로모터 또는 제1 자기-절단 바이러스 웨타이드를 암호화하는 폴리뉴클레오타이드 및 하나의 변형된 및 비-작용성의 CBLB 유전자 내로 통합된 조작된 TCR의 알파쇄 및/또는 베타쇄를 암호화하는 폴리뉴클레오타이드를 포함한다.

[0415] 특정 실시형태에서, 공여자 수선 주형은 5'에서 3'으로, 제1 자기-절단 바이러스 웨타이드를 암호화하는 폴리뉴클레오타이드, 조작된 TCR의 알파쇄를 암호화하는 폴리뉴클레오타이드, 제2 자기-절단 바이러스 웨타이드를 암호화하는 폴리뉴클레오타이드, 및 하나의 변형된 및/또는 비작용성의 CBLB 유전자 내로 통합된 조작된 TCR의 베타쇄를 암호화하는 폴리뉴클레오타이드를 포함한다. 이러한 경우에, 다른 CBLB 유전자는 작용성일 수 있거나 또는 감소된 기능을 가질 수 있거나 또는 DSB에 의한 비작용성 및 NHEJ에 의한 수선이 제공되었다. 일 실시형태에서, 다른 CBLB 유전자는 본 명세서에서 상정된 조작된 뉴클레아제에 의해 변형되었고, 감소된 기능을 갖거나 또는 비작용성으로 제공되었다.

[0416] 특정 실시형태에서, CBLB 유전자는 둘 다 변형되고, 감소된 기능을 갖거나 또는 비작용성이다: 제1 변형된 CBLB 유전자는 제1 자기-절단 바이러스 웨타이드를 암호화하는 폴리뉴클레오타이드 및 조작된 TCR의 알파쇄를 암호화하는 폴리뉴클레오타이드를 포함하는 핵산을 포함하고, 제2 변형된 CBLB 유전자는 제2 자기-절단 바이러스 웨타이드를 암호화하는 폴리뉴클레오타이드 및 조작된 TCR의 베타쇄를 암호화하는 폴리뉴클레오타이드를 포함한다.

b. 키메라 항원 수용체(CAR)

[0418] 특정 실시형태에서, 본 명세서에서 상정된 면역 효과기 세포는 하나 이상의 키메라 항원 수용체(CAR)를 포함한다. 일 실시형태에서, 하나 이상의 CAR을 암호화하는 폴리뉴클레오타이드를 포함하는 T 세포는 CBLB 유전자에서 DSB를 도입함으로써 편집된다. 일 실시형태에서, T 세포는 CAR을 암호화하는 공여자 수선 주형의 존재 하에서 하나 이상의 CBLB 유전자에서 DSB를 도입함으로써 조작된다. 특정 실시형태에서, CAR은 단일 CBLB 유전자 내 DSB에서 삽입된다.

[0419] 일 실시형태에서, 본 명세서에서 상정된 조작된 T 세포는 CBLB 유전자에서 삽입되지 않는 CAR을 포함하며, 면역 억제 신호 램프, 플립 수용체, 조작된 T 세포 수용체(TCR)의 알파 및/또는 베타쇄, 키메라 항원 수용체(CAR), Daric 수용체 또는 이들의 성분, 또는 키메라 사이토카인 수용체 중 하나 이상은 하나 이상의 CBLB 유전자에서 DSB 내로 삽입된다.

[0420] 다양한 실시형태에서, 계놈 편집된 T 세포는 종양 세포를 향해 세포독성을 다시 보내는 CAR을 발현시킨다. CAR은 특이적 항-종양 세포 면역 활성을 나타내는 키메라 단백질을 생성하기 위해 T 세포 수용체-활성화 세포내 도메인과 표적 항원(예를 들어, 종양 항원)에 대한 항원 기반 특이성을 결합하는 분자이다. 본 명세서에서 사용되는 용어, "키메라"는 상이한 단백질의 부분 또는 상이한 유래의 DNA로 구성 중이라는 것을 기재한다.

[0421] 다양한 실시형태에서, CAR은 특정 표적 항원에 결합하는 세포외 도메인(또한 결합 도메인 또는 항원-특이적 결합 도메인으로서 지칭됨), 막관통 도메인 및 세포내 신호전달 도메인을 포함한다. CAR의 주요 특징은 면역 효과기 세포 특이성을 다시 보냄으로써, 증식, 사이토카인 생산, 세포작용 또는 주조직적합성(MHC) 독립적 방식으로 표적 항원 발현 세포의 세포사를 매개할 수 있는 분자의 생산을 촉발하는 그들의 능력, 단클론성 항체의 세포 특이적 표적화 능력의 이용, 가용성 리간드 또는 세포 특이적 공수용체이다.

[0422] 특정 실시형태에서, CAR은 종양 세포 상에서 발현되는 표적 폴리웨타이드, 예를 들어, 표적 항원에 특이적으로 결합하는 세포외 결합 도메인을 포함한다. 본 명세서에서 사용되는 용어, "결합 도메인", "세포외 도메인", "세포외 결합 도메인", "항원 결합 도메인", "항원-특이적 결합 도메인" 및 "세포외 항원 특이적 결합 도메인"은 상호 호환적으로 사용되고, 관심 대상의 표적 항원에 특이적으로 결합하는 능력을 갖는 키메라 수용체, 예를 들어, CAR 또는 Daric을 제공한다. 결합 도메인은 임의의 단백질, 폴리웨타이드, 올리고웨타이드 또는 생물학적 분자를 특이적으로 인식하고 이에 결합하는 능력을 갖는 웨타이드(예를 들어, 세포 표면 수용체 또는 종양 단백질, 지질, 다당류, 또는 다른 세포 표면 표적 분자 또는 이의 성분)를 포함할 수 있다. 결합 도메인은 임의의

천연 유래, 합성, 반합성 또는 관심 대상의 생물학적 분자에 대한 재조합적으로 생산된 결합 항체를 포함한다.

[0423] 특정 실시형태에서, 세포외 결합 도메인은 항체 또는 이의 항원 결합 단편을 포함한다.

[0424] "항체"는 표적 항원, 예컨대 웹타이드, 지질, 다당류, 또는 항원 결정소, 예컨대 면역 세포에 의해 인식되는 것을 함유하는 핵산의 에피토프를 특이적으로 인식하고 이에 결합하는 적어도 경쇄 또는 중쇄 면역글로불린 가변 영역을 포함하는 폴리웹타이드인 결합 체제를 지칭한다. 항체는 항원 결합 단편, 예를 들어, Camel Ig(낙타 항체 또는 이의 VHH 단편), Ig NAR, Fab 단편, Fab' 단편, F(ab)'2 단편, F(ab)'3 단편, Fv, 단일쇄 Fv 항체 ("scFv"), 비스-scFv, (scFv)2, 미니바디, 다이어바디, 트라이어바디, 테트라바디, 이황화 안정화된 Fv 단백질 ("dsFv") 및 단일-도메인 항체(sdAb, 나노바디) 또는 이의 다른 항체 단편을 포함한다. 상기 용어는 또한 유전자 조작된 형태, 예컨대 키메라 항체(예를 들어, 인간화된 뮤린 항체), 이형접합체 항체(예컨대, 이중특이성 항체) 및 이의 항원 결합 단편을 포함한다. 또한, 문헌[Pierce Catalog and Handbook, 1994-1995 (Pierce Chemical Co., Rockford, IL); Kuby, J., Immunology, 3rd Ed., W. H. Freeman & Co., New York, 1997] 참조.

[0425] 하나의 바람직한 실시형태에서, 결합 도메인은 scFv이다.

[0426] 다른 바람직한 실시형태에서, 결합 도메인은 낙타과 항체이다.

[0427] 특정 실시형태에서, CAR은 알파 엽산 수용체, 5T4,  $\alpha\beta$ 6 인테그린, BCMA, B7-H3, B7-H6, CAIX, CD16, CD19, CD20, CD22, CD30, CD33, CD44, CD44v6, CD44v7/8, CD70, CD79a, CD79b, CD123, CD138, CD171, CEA, CSPG4, EGFR, ErbB2(HER2)를 포함하는 EGFR 패밀리, EGFRvIII, EGP2, EGP40, EPCAM, EphA2, EpCAM, FAP, 태아 AchR, GD2, GD3, 글리피칸-3(GPC3), HLA-A1+MAGE1, HLA-A2+MAGE1, HLA-A3+MAGE1, HLA-A1+NY-ESO-1, HLA-A2+NY-ESO-1, HLA-A3+NY-ESO-1, IL-11R $\alpha$ , IL-13R $\alpha$ 2, 람다, 루이스-Y, 카파, 메소텔린, Muc1, Muc16, NCAM, NKG2D 리간드, NY-ESO-1, PRAME, PSCA, PSMA, ROR1, SSX, 서비빈, TAG72, TEMs, VEGFR2 및 WT-1로 이루어진 군으로부터 선택된 항원에 결합하는 세포외 도메인을 포함한다.

[0428] 특정 실시형태에서, CAR은 세포외 결합 도메인, 예를 들어, 항원에 결합하는 항체 또는 이의 항원 결합 단편을 포함하되, 항원은 MHC-웹타이드 복합체, 예컨대 클래스 I MHC-웹타이드 복합체 또는 클래스 II MHC-웹타이드 복합체이다.

[0429] 특정 실시형태에서, CAR은 다양한 도메인 사이의 링커 잔기를 포함한다. "가변 영역 연결 서열"은 얻어진 폴리 웹타이드가 동일한 경쇄 및 중쇄 가변 영역을 포함하는 항체를 포함하는 항체와 동일한 표적 분자에 대해 특이적 결합 친화도를 보유하도록, 경쇄 가변 영역에 중쇄 가변 영역을 연결하고, 2개의 서브-결합 도메인의 상호 작용에 적합한 스페이서 기능을 제공하는 아미노산 서열이다. 특정 실시형태에서, CAR은 1, 2, 3, 4 또는 5개 이상의 링커를 포함한다. 특정 실시형태에서, 링커의 길이는 약 1 내지 약 25개의 아미노산, 약 5 내지 약 20개의 아미노산, 또는 약 10 내지 약 20개의 아미노산 또는 아미노산의 임의의 개체 길이이다. 일부 실시형태에서, 링커는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25개 이상의 아미노산 길이이다.

[0430] 특정 실시형태에서, CAR의 결합 도메인 다음에 하나 이상의 "스페이서 도메인"이 이어지는데, 이는 적절한 세포/세포 접촉, 항원 결합 및 활성화를 가능하게 하도록 효과기 세포 표면으로부터 떨어져서 항원 결합 도메인을 이동시키는 영역을 지칭한다(Patel et al., Gene Therapy, 1999; 6: 412-419). 스페이서 도메인은 천연, 합성, 반합성 또는 재조합 공급원으로부터 유래될 수 있다. 특정 실시형태에서, 스페이서 도메인은 하나 이상의 중쇄 불변 영역, 예를 들어, CH2 및 CH3을 포함하지만, 이들로 제한되지 않는 면역글로불린의 일부이다. 스페이서 도메인은 천연 유래 면역글로불린 힌지 영역 또는 변경된 면역글로불린 힌지 영역의 아미노산 서열을 포함할 수 있다.

[0431] 일 실시형태에서, 스페이서 도메인은 IgG1, IgG4 또는 IgD의 CH2 및 CH3을 포함한다.

[0432] 일 실시형태에서, CAR의 결합 도메인은 하나 이상의 "힌지 도메인"에 연결되는데, 이는 적절한 세포/세포 접촉, 항원 결합 및 활성화를 가능하게 하도록 효과기 세포 표면으로부터 떨어져서 항원 결합 도메인을 위치시키는 역할을 한다. CAR은 일반적으로 결합 도메인과 막관통 도메인(TM) 사이의 하나 이상의 힌지 도메인을 포함한다. 힌지 도메인은 천연, 합성, 반합성 또는 재조합 공급원으로부터 유래될 수 있다. 힌지 도메인은 천연 유래 면역글로불린 힌지 영역 또는 변경된 면역글로불린 힌지 영역의 아미노산 서열을 포함할 수 있다.

[0433] 본 명세서에 기재된 CAR에서 사용하기에 적합한 예시적인 힌지 도메인은 이들 분자로부터의 야생형 힌지 영역일 수 있거나 또는 변경될 수 있는 1형 막 단백질, 예컨대 CD8 $\alpha$ , 및 CD4의 세포외 영역으로부터 유래된 힌지 영역

을 포함한다. 다른 실시형태에서, 힌지 도메인은 CD8  $\alpha$  힌지 영역을 포함한다.

[0434] 일 실시형태에서, 힌지는 PD-1 힌지 또는 CD152 힌지이다.

[0435] "막관통 도메인"은 세포의 결합 부분에 융합하는 CAR의 부분 및 면역 효과기 세포의 혈장막에 CAR을 고정시키는 세포내 신호전달 도메인이다. TM 도메인은 천연, 합성, 반합성 또는 재조합 공급원으로부터 유래될 수 있다.

[0436] 예시적인 TM 도메인은 T-세포 수용체, CD3  $\delta$ , CD3  $\epsilon$ , CD3  $\gamma$ , CD3  $\zeta$ , CD4, CD5, CD8  $\alpha$ , CD9, CD16, CD22, CD27, CD28, CD33, CD37, CD45, CD64, CD80, CD86, CD134, CD137, CD152, CD154, AMN 및 PD-1의 알파 또는 베타쇄의 적어도 막관통 영역(들)으로부터 유래될 수 있다(즉, 포함한다).

[0437] 일 실시형태에서, CAR은 CD8  $\alpha$ 로부터 유래된 TM 도메인을 포함한다. 다른 실시형태에서, 본 명세서에서 상정된 CAR은 TM 도메인과 CAR의 세포내 신호전달 도메인을 연결하는 CD8  $\alpha$  및 짧은 올리고- 또는 폴리펩타이드 링커, 바람직하게는 길이가 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 아미노산으로부터 유래된 TM 도메인을 포함한다. 글리신-세린 링커는 특히 적합한 링커를 제공한다.

[0438] 특정 실시형태에서, CAR은 세포내 신호전달 도메인을 포함한다. "세포내 신호전달 도메인"은 CAR-결합 표적 세포에 대한 세포독성 인자의 방출, 또는 세포외 CAR 도메인에 대한 항원 결합에 의해 유발된 다른 세포 반응을 포함하는 효과기 세포 기능, 예를 들어, 활성화, 사이토카인 생산, 증식 및 세포독성 활성을 유발하기 위해 면역 효과기 세포의 내부로 표적 항원에 대한 효과적인 CAR 결합의 전령을 형질도입하는 데 참여하는 CAR의 부분을 지칭한다.

[0439] 용어 "효과기 기능"은 세포의 전문화된 기능을 지칭한다. T 세포의 효과기 기능은, 예를 들어, 세포용해 활성일 수 있거나 또는 사이토카인의 분비를 포함하는 활성을 드는다. 따라서, 용어 "세포내 신호전달 도메인"은 효과기 작용 신호를 전달하고 전문화된 기능을 수행하기 위해 세포를 보내는 단백질의 일부를 지칭한다. 보통 전체 세포내 신호전달 도메인이 사용될 수 있지만, 다수의 경우에, 전체 도메인을 사용하는 것을 필수적이지 않다. 세포내 신호전달 도메인의 절단된 일부가 사용되는 정도로, 이러한 절단된 부분은 효과기 기능 신호를 전달하는 한, 전체 도메인 대신에 사용될 수 있다. 용어 세포내 신호전달 도메인은 효과기 기능 신호를 전달하는 데 충분한 세포내 신호전달 도메인의 임의의 절단된 부분을 포함하는 것을 의미한다.

[0440] TCR 단독을 통해 생성된 신호는 T 세포의 완전한 활성화에 불충분하며, 2차 또는 공자극 신호가 또한 필요하다는 것이 알려져 있다. 따라서, T 세포 활성화는 세포내 신호전달 도메인의 2개의 별개의 부류(TCR(예를 들어, TCR/CD3 복합체)를 통해 항원-의존적 1차 활성화를 개시하는 1차 신호전달 도메인 및 2차 또는 공자극 신호를 제공하는 항원-독립적 방식으로 작용하는 공자극 신호전달 도메인)에 의해 매개되는 것으로 언급될 수 있다. 바람직한 실시형태에서, CAR은 하나 이상의 "공자극 신호전달 도메인" 및 "1차 신호전달 도메인"을 포함하는 세포내 신호전달 도메인을 포함한다.

[0441] 1차 신호전달 도메인은 자극 방법으로 또는 저해 방법으로 TCR 복합체의 1차 활성화를 조절한다. 자극 방식으로 작용하는 1차 신호전달 도메인은 면역수용체 타이로신계 활성화 모티프 또는 ITAM으로서 알려진 신호전달 모티프를 함유할 수 있다.

[0442] 특정 실시형태에서 상정된 CAR에서 사용하기에 적합한 1차 신호전달 도메인을 함유하는 ITAM의 예시적 예는 FcR  $\gamma$ , FcR  $\beta$ , CD3  $\gamma$ , CD3  $\delta$ , CD3  $\epsilon$ , CD3  $\zeta$ , CD22, CD79a, CD79b 및 CD66d로부터 유래되는 것을 포함한다. 특정 바람직한 실시형태에서, CAR은 CD3  $\zeta$  1차 신호전달 도메인 및 하나 이상의 공자극 신호전달 도메인을 포함한다. 세포내 1차 신호전달 및 공자극 신호전달 도메인은 막관통 도메인의 카복실 말단과 동시에 임의의 순서로 연결될 수 있다.

[0443] 특정 실시형태에서, CAR은 T 세포 발현 CAR 수용체의 효능 및 확장을 향상시키기 위해 하나 이상의 공자극 신호전달 도메인을 포함한다. 본 명세서에서 사용되는 용어, "공자극 신호전달 도메인" 또는 "공자극 도메인"은 공자극 분자의 세포내 신호전달 도메인을 지칭한다.

[0444] 특정 실시형태에서 상정된 CAR에서 사용하기에 적합한 이러한 공자극 분자의 예시적 예는 TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR8, TLR9, TLR10, CARD11, CD2, CD7, CD27, CD28, CD30, CD40, CD54(ICAM), CD83, CD134(OX40), CD137(4-1BB), CD278(ICOS), DAP10, LAT, NKG2C, SLP76, TRIM 및 ZAP70을 포함한다. 일 실시형태에서, CAR은 CD28, CD137, 및 CD134, 및 CD3  $\zeta$  1차 신호전달 도메인으로 이루어진 군으로부터 선택된 하나 이상의 공자극 신호전달 도메인을 포함한다.

[0445] 다양한 실시형태에서, CAR은 BCMA, CD19, CSPG4, PSCA, ROR1 및 TAG72로 이루어진 군으로부터 선택된 항원에

결합되는 세포의 도메인; CD4, CD8  $\alpha$ , CD154 및 PD-1로 이루어진 군으로부터 선택된 폴리펩타이드로부터 단리된 막관통 도메인; CD28, CD134 및 CD137로 이루어진 군으로부터 선택된 폴리펩타이드로부터 단리된 하나 이상의 세포내 공자극 신호전달 도메인; 및 FcR  $\gamma$ , FcR  $\beta$ , CD3  $\gamma$ , CD3  $\delta$ , CD3  $\epsilon$ , CD3  $\zeta$ , CD22, CD79a, CD79b 및 CD66d로 이루어진 군으로부터 선택된 폴리펩타이드로부터 단리된 신호전달 도메인을 포함한다.

[0446] c. Daric 수용체

[0447] 특정 실시형태에서, 조작된 면역 효과기 세포는 하나 이상의 Daric 수용체를 포함한다. 본 명세서에서 사용되는 용어 "Daric 수용체"는 다중쇄 조작된 항원 수용체를 지칭한다. 일 실시형태에서, 하나 이상의 Daric 수용체 성분을 암호화하는 폴리뉴클레오타이드를 포함하는 T 세포는 CBLB 유전자에서 DSB를 도입함으로써 편집된다. 일 실시형태에서, T 세포는 Daric의 하나 이상의 성분을 암호화하는 공여자 수선 주형의 존재 하에서 하나 이상의 CBLB 대립유전자에서 DSB를 도입함으로써 조작된다. 특정 실시형태에서, Daric 또는 이의 하나 이상의 성분은 단일 CBLB 대립유전자 내 DSB에서 삽입된다.

[0448] 일 실시형태에서, 조작된 T 세포는 CBLB 유전자에서 삽입되지 않는 Daric을 포함하며, 면역억제 신호 탬퍼, 플립 수용체, 조작된 T 세포 수용체(TCR)의 알파 및/또는 베타쇄, 키메라 항원 수용체(CAR), Daric 수용체 또는 이들의 성분 중 하나 이상은 하나 이상의 CBLB 대립유전자에서 DSB 내로 삽입된다.

[0449] Daric 구조 및 성분의 예시적 예는 국제 특허 출원 WO2015/017214 및 미국 특허 공개 제20150266973호에 개시되어 있으며, 이들 각각은 그의 전문이 본 명세서에 참고로 포함된다.

[0450] 일 실시형태에서, 공여자 수선 주형은 다음의 Daric 성분을 포함한다: 제1 다량체화 도메인, 제1 막관통 도메인, 및 하나 이상의 세포내 공자극 신호전달 도메인 및/또는 1차 신호전달 도메인을 포함하는 신호전달 폴리펩타이드; 및 결합 도메인, 제2 다량체화 도메인, 및 선택적으로 제2 막관통 도메인을 포함하는 결합 폴리펩타이드. 작용성 Daric은 신호전달 폴리펩타이드 및 결합 폴리펩타이드의 다량체화 도메인과 결합되고 이들 사이에 배치된 브리징 인자(bridging factor)와의 세포 표면 상의 Daric 수용체 복합체의 형성을 촉진하는 브리징 인자를 포함한다.

[0451] 특정 실시형태에서, 제1 및 제2 다량체화 도메인은 라파마이신 또는 이의 라팔로그, 쿠메르마이신 또는 이의 유도체, 지베렐린 또는 이의 유도체, 아비시스산(ABA) 또는 이의 유도체, 메토트렉세이트 또는 이의 유도체, 사이클로스포린 A 또는 이의 유도체, FKCsA 또는 이의 유도체, FKBP에 대한 트라이메토프림(Tmp)-합성 리간드(SLF) 또는 이의 유도체 및 이들의 임의의 조합물로 이루어진 군으로부터 선택된 브리징 인자와 결합된다.

[0452] 라파마이신 유도체(라팔로그)의 예시적 예는 미국 특허 제6,649,595호에 개시된 것을 포함하며, 라팔로그 구조는 본 명세서에 전문이 참고로 포함된다. 특정 실시형태에서, 브리징 인자는 라파마이신에 비해 실질적으로 감소된 면역억제 효과를 갖는 라팔로그이다. "실질적으로 감소된 면역억제 효과"는 임상적으로 또는 시험관내에서 적절하게(예를 들어, T 세포 증식의 저해) 또는 인간 면역억제 활성의 생체내 대용으로 측정하여, 등몰량의 라파마이신에 대해 관찰되거나 또는 예상된 면역억제 효과의 적어도 0.1 내지 0.005 미만을 갖는 라팔로그를 지칭한다. 일 실시형태에서, "실질적으로 감소된 면역억제 효과"는 동일한 분석에서 라파마이신에 대해 관찰된 EC50 값보다 적어도 10 내지 250배 더 큰 이러한 시험관내 분석에서 EC50 값을 갖는 라팔로그를 지칭한다.

[0453] 라팔로그의 다른 예시적 예는 에베롤리무스, 노볼리무스, 피메크롤리무스, 리다포롤리무스, 타크롤리무스, 템시롤리무스, 유미놀리무스 및 조타롤리무스를 포함하지만, 이들로 제한되지 않는다.

[0454] 특정 실시형태에서, 다량체화 도메인은 라파마이신 또는 이의 라팔로그인 브리징 인자와 결합될 것이다. 예를 들어, 제1 및 제2 다량체화 도메인은 FKBP 및 FRB로부터 선택된 쌍이다. FRB 도메인은 FKBP 단백질 및 라파마이신 또는 이의 라팔로그와 삼중 복합체를 형성할 수 있는 폴리펩타이드 영역(단백질 "도메인")이다. FRB 도메인은 인간 및 다른 종으로부터의 mTOR 단백질(또한 문헌에서 FRAP, RAPT1 또는 RAFT로서 지칭됨); Tor1 및 Tor2를 포함하는 효모 단백질; 및 칸디다 FRAP 상동체를 포함하는 다수의 천연 유래 단백질에 존재한다. 뉴클레오타이드 서열, 클로닝 및 이들 단백질의 다른 양상에 관한 정보는 이미 당업계에 공지되어 있다. 예를 들어, 인간 mTOR에 대한 단백질 서열 수탁 번호는 젠뱅크 수탁 번호 L34075.1이다(Brown et al., Nature 369:756, 1994).

[0455] 본 명세서에서 상정된 특정 실시형태에서 사용하기에 적합한 FRB 도메인은 일반적으로 적어도 약 85 내지 약 100개의 아미노산 잔기를 함유한다. 특정 실시형태에서, 본 개시내용의 융합 단백질에서 사용하기 위한 FRB 아미노산 서열은 젠뱅크 수탁 번호 L34075.1의 아미노산에 기반하여 93 아미노산 서열 Ile-2021 내지 Lys-2113 및 T2098L의 돌연변이를 포함할 것이다. 특정 실시형태에서 상정된 Daric에서 사용하기 위한 FRB 도메인은 라파마이신 또는 이의 라팔로그에 결합된 FKBP 단백질의 복합체에 결합할 것이다. 특정 실시형태에서, FRB 도메인의

펩타이드 서열은 (a) 인간 mTOR의 적어도 표시된 93개의 아미노산 영역에 걸쳐 있는 천연 유래 펩타이드 서열 또는 상동성 단백질의 대응하는 영역; (b) 천연 유래 펩타이드의 약 10개까지의 아미노산, 또는 약 1 내지 약 5 개의 아미노산 또는 약 1 내지 약 3개의 아미노산, 또는 일부 실시형태에서 단지 1개의 아미노산이 결실, 삽입 또는 치환된 천연 유래 FRB의 변이체; 또는 (c) 천연 유래 FRB 도메인을 암호화하는 DNA 분자에 선택적으로 혼성화할 수 있는 핵산 분자에 의해 또는 천연 유래 FRB 도메인을 암호화하는 DNA 분자에 선택적으로 혼성화할 수 있는(그러나 유전자 암호의 축퇴를 위한) DNA 서열에 의해 암호화된 펩타이드를 포함한다.

[0456] FKBP(FK506 결합 단백질)는 마크로라이드, 예컨대 FK506, FK520 및 라파마이신에 대한 사이토졸 수용체이고, 종 계통에 걸쳐 고도로 보존된다. FKBP는 라파마이신 또는 이의 라팔로그에 결합할 수 있고, 추가로 FRB-함유 단백 질 또는 융합 단백질과 삼중 복합체를 형성할 수 있는 단백질 또는 단백질 도메인이다. FKBP 도메인은 또한 "라 파마이신 결합 도메인"으로 지칭될 수 있다. 뉴클레오타이드 서열, 클로닝 및 다양한 FKBP 종의 다른 양상에 관한 정보는 당업계에 공지되어 있다(예를 들어, 문헌[Staendart et al., Nature 346:671, 1990(인간 FKBP12); Kay, Biochem. J. 314:361, 1996] 참조). 다른 포유류 종에서, 효모에서 그리고 다른 유기체에서 상동성 FKBP 단백질은 또한 당업계에 공지되어 있으며, 본 명세서에 개시된 융합 단백질에서 사용될 수 있다. 특정 실시형태에서 상정된 FKBP 도메인은 라파마이신 또는 이의 라팔로그에 결합할 수 있고, FRB-함유 단백질과 삼중 복합체로 참여할 수 있다(이러한 결합의 검출을 위해 임의의 수단에 의해, 직접 또는 간접적으로 결정될 수 있음).

[0457] 특정 실시형태에서 상정된 Daric에서 사용하기에 적합한 FKBP 도메인의 예시적 예는 천연 유래 FKBP 펩타이드 서열(바람직하게는 인간 FKBP12 단백질로부터 단리됨)(젠햄크 수탁 번호 AAA58476.1) 또는 이로부터, 다른 인간 FKBP로부터, 뮤린 또는 다른 포유류 FKBP로부터, 또는 일부 다른 동물, 효모 또는 진균 FKBP로부터 단리된 펩타이드 서열; 천연 유래 펩타이드의 약 10까지의 아미노산, 또는 약 1 내지 약 5개의 아미노산 또는 약 1 내지 약 3개의 아미노산, 또는 일부 실시형태에서 단지 1개의 아미노산이 결실, 삽입 또는 치환된 천연 유래 FKBP 서열의 변이체; 또는 천연 유래 FKBP를 암호화하는 DNA 분자에 선택적으로 혼성화할 수 있는 핵산 분자에 의해 또는 천연 유래 FKBP를 암호화하는 DNA 분자에 선택적으로 혼성화할 수 있는(그러나 유전자 암호의 축퇴를 위한) DNA 서열에 의해 암호화된 펩타이드 서열을 포함하지만, 이들로 제한되지 않는다.

[0458] 특정 실시형태에서 상정된 Daric에서 사용하기에 적합한 다량체화 도메인 쌍의 다른 예시적 예는 FKBP 및 FRB, FKBP 및 칼시뉴린, FKBP 및 사이클로필린, FKBP 및 박테리아 DHFR, 칼시뉴린 및 사이클로필린, PYL1 및 ABI1, 또는 GIB1 및 GAI, 또는 이의 변이체를 포함하지만, 이들로 제한되지 않는다.

[0459] 또 다른 실시형태에서, 항-브리징 인자는 신호전달 폴리펩타이드 및 결합 폴리펩타이드와 브리징 인자의 결합을 차단한다. 예를 들어, 사이클로스포린 또는 FK506은 라파마이신을 적정하기 위한 항-브리징 인자로서 사용될 수 있었고, 따라서 하나의 다량체화 도메인만이 결합되기 때문에 신호전달을 정지시킨다. 특정 실시형태에서, 항-브리징 인자(예를 들어, 사이클로스포린, FK506)는 면역억제제이다. 예를 들어, 면역억제 항-브리징 인자는 특정 실시형태에서 상정된 Daric 성분의 기능을 차단하거나 또는 최소화하기 위해 그리고 동시에 임상 상황에서 원치않는 또는 병리학적 염증 반응을 저해하거나 또는 차단하기 위해 사용될 수 있다.

[0460] 일 실시형태에서, 제1 다량체화 도메인은 FRB T2098L을 포함하고, 제2 다량체화 도메인은 FKBP12를 포함하며, 브리징 인자는 라팔로그 AP21967이다.

[0461] 다른 실시형태에서, 제1 다량체화 도메인은 FRB를 포함하며, 제2 다량체화 도메인은 FKBP12를 포함하고, 브리징 인자는 라파마이신, 템시롤리무스 또는 에베롤리무스이다.

[0462] 특정 실시형태에서, 신호전달 폴리펩타이드, 제1 막관통 도메인 및 결합 폴리펩타이드는 제2 막관통 도메인 또는 GPI 앵커를 포함한다. 제1 및 제2 막관통 도메인의 예시적 예는 CD3  $\delta$ , CD3  $\epsilon$ , CD3  $\gamma$ , CD3  $\zeta$ , CD4, CD5, CD8  $\alpha$ , CD9, CD16, CD22, CD27, CD28, CD33, CD37, CD45, CD64, CD80, CD86, CD134, CD137, CD152, CD154, AMN 및 PD-1로 이루어진 군으로부터 독립적으로 선택되는 폴리펩타이드로부터 단리된다.

[0463] 일 실시형태에서, 신호전달 폴리펩타이드는 하나 이상의 세포내 공자극 신호전달 도메인 및/또는 1차 신호전달 도메인을 포함한다.

[0464] 특정 실시형태에서 상정된 Daric 신호전달 성분에서 사용하기에 적합한 1차 신호전달 도메인의 예시적 예는 FcR  $\gamma$ , FcR  $\beta$ , CD3  $\gamma$ , CD3  $\delta$ , CD3  $\epsilon$ , CD3  $\zeta$ , CD22, CD79a, CD79b 및 CD66d로부터 유래된 것을 포함한다. 특정 바람직한 실시형태에서, Daric 신호전달 성분은 CD3  $\zeta$  1차 신호전달 도메인 및 하나 이상의 공자극 신호전달 도메인을 포함한다. 세포내 1차 신호전달 및 공자극 신호전달 도메인은 막관통 도메인의 카복실 말단과 동시에 임의의 순서로 연결될 수 있다.

- [0465] 특정 실시형태에서 상정된 Daric 신호전달 성분에서 사용하기에 적합한 이러한 공자극 분자의 예시적 예는 TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR8, TLR9, TLR10, CARD11, CD2, CD7, CD27, CD28, CD30, CD40, CD54(ICAM), CD83, CD134(OX40), CD137(4-1BB), CD278(ICOS), DAP10, LAT, NKG2C, SLP76, TRIM 및 ZAP70을 포함한다. 일 실시형태에서, Daric 신호전달 성분은 CD28, CD137 및 CD134, 및 CD3  $\zeta$  1차 신호전달 도메인으로 이루어진 군으로부터 선택되는 하나 이상의 공자극 신호전달 도메인을 포함한다.
- [0466] 특정 실시형태에서, Daric 결합 성분은 결합 도메인을 포함한다. 일 실시형태에서, 결합 도메인은 항체 또는 이의 항원 결합 단편이다.
- [0467] 항체 또는 이의 항원 결합 단편은 표적 항원, 예컨대 웹타이드, 지질, 다당류, 또는 항원 결정소, 예컨대 면역세포에 의해 인식되는 것을 함유하는 핵산의 에피토프를 특이적으로 인식하고 이에 결합하는 적어도 경쇄 또는 중쇄 면역글로불린 가변 영역을 포함한다. 항체는 항원 결합 단편, 예를 들어, Camel Ig(낙타 항체 또는 이의 VH 단편), Ig NAR, Fab 단편, Fab' 단편, F(ab)'2 단편, F(ab)'3 단편, Fv, 단일쇄 Fv 항체("scFv"), 비스-scFv, (scFv)2, 미니바디, 다이어바디, 트라이어바디, 테트라바디, 이황화 안정화된 Fv 단백질("dsFv") 및 단일-도메인 항체(sdB, 나노바디) 또는 이의 다른 항체 단편을 포함한다. 상기 용어는 또한 유전자 조작된 형태, 예컨대 키메라 항체(예를 들어, 인간화된 뮤린 항체), 이형접합체 항체(예컨대, 이중특이성 항체) 및 이의 항원 결합 단편을 포함한다. 또한, 문현[Pierce Catalog and Handbook, 1994-1995 (Pierce Chemical Co., Rockford, IL); Kuby, J., Immunology, 3rd Ed., W. H. Freeman & Co., New York, 1997] 참조.
- [0468] 하나의 바람직한 실시형태에서, 결합 도메인은 scFv이다.
- [0469] 다른 바람직한 실시형태에서, 결합 도메인은 낙타과 항체이다.
- [0470] 특정 실시형태에서, Daric 결합 성분은 알파 엽산 수용체, 5T4,  $\alpha\beta\gamma\delta$  인테그린, BCMA, B7-H3, B7-H6, CAIX, CD16, CD19, CD20, CD22, CD30, CD33, CD44, CD44v6, CD44v7/8, CD70, CD79a, CD79b, CD123, CD138, CD171, CEA, CSPG4, EGFR, Erbb2(HER2)를 포함하는 EGFR 패밀리, EGFRvIII, EGP2, EGP40, EPCAM, EphA2, EpCAM, FAP, 태아 AchR, GD2, GD3, 글리피칸-3(GPC3), HLA-A1+MAGE1, HLA-A2+MAGE1, HLA-A3+MAGE1, HLA-A1+NY-ESO-1, HLA-A2+NY-ESO-1, HLA-A3+NY-ESO-1, IL-11R  $\alpha$ , IL-13R  $\alpha$  2, 람다, 루이스-Y, 카파, 메소텔린, Muc1, Muc16, NCAM, NKG2D 리간드, NY-ESO-1, PRAME, PSCA, PSMA, ROR1, SSX, 서비빈, TAG72, TEMs, VEGFR2 및 WT-1로 이루어진 군으로부터 선택된 항원에 결합하는 세포외 도메인을 포함한다.
- [0471] 일 실시형태에서, Daric 결합 성분은 세포외 도메인, 예를 들어, MHC-웹타이드 복합체, 예컨대 클래스 I MHC-웹타이드 복합체 또는 클래스 II MHC-웹타이드 복합체에 결합하는 항체 또는 이의 항원 결합 단편을 포함한다.
- [0472] 특정 실시형태에서, 본 명세서에서 상정된 Daric 성분은 2개의 단백질, 폴리웹타이드, 웹타이드, 도메인, 영역 또는 모티프를 연결하는 링커 또는 스페이서를 포함한다. 특정 실시형태에서, 링커는 약 2 내지 약 35개의 아미노산, 또는 약 4 내지 약 20개의 아미노산 또는 약 8 내지 약 15개의 아미노산 또는 약 15 내지 약 25개의 아미노산을 포함한다. 다른 실시형태에서, 스페이서는 특정 구조, 예컨대, 항체 CH2CH3 도메인, 힌지 도메인 등을 가질 수 있다. 일 실시형태에서, 스페이서는 IgG1, IgG4 또는 IgD의 CH2 및 CH3 도메인을 포함한다.
- [0473] 특정 실시형태에서, 본 명세서에서 상정된 Daric 성분은 적절한 세포/세포 접촉, 항원 결합 및 활성화를 가능하게 하도록 도메인을 위치시키는 역할을 하는 하나 이상의 "힌지 도메인"을 포함한다. Daric은 결합 도메인과 다량체화 도메인 사이의 하나 이상의 힌지 도메인 및/또는 다량체화 도메인과 막관통 도메인 사이의 막관통 도메인(TM)을 포함할 수 있다. 힌지 도메인은 천연, 합성, 반합성 또는 재조합 공급원으로부터 유래될 수 있다. 힌지 도메인은 천연 유래 면역글로불린 힌지 영역 또는 변경된 면역글로불린 힌지 영역의 아미노산 서열을 포함할 수 있다. 특정 실시형태에서, 힌지는 CD8  $\alpha$  힌지 또는 CD4 힌지이다.
- [0474] 일 실시형태에서, Daric은 FRB T2098L의 제1 다량체화 도메인, CD8 막관통 도메인, 4-1BB 공자극 도메인, 및 CD3  $\zeta$  1차 신호전달 도메인을 포함하는 신호전달 폴리웹타이드를 포함하고; 결합 폴리웹타이드는 CD19에 결합하는 scFv, FKBP12의 제2 다량체화 도메인 및 CD4 막관통 도메인을 포함하고; 브리징 인자는 라팔로그 AP21967이다.
- [0475] 일 실시형태에서, Daric은 FRB, CD8 막관통 도메인, 4-1BB 공자극 도메인, 및 CD3  $\zeta$  1차 신호전달 도메인을 포함하는 제1 다량체화 도메인을 포함하는 신호전달 폴리웹타이드를 포함하고; 결합 폴리웹타이드는 CD19에 결합하는 scFv, FKBP12의 제2 다량체화 도메인 및 CD4 막관통 도메인을 포함하며; 브리징 인자는 라파마이신, 템시롤리무스 또는 에베롤리무스이다.

## [0476] d. 제타카인

[0477] 특정 실시형태에서, 본 명세서에서 상정된 면역 효과기 세포는 하나 이상의 키메라 사이토카인 수용체를 포함한다. 일 실시형태에서, 제타카인을 암호화하는 폴리뉴클레오타이드를 포함하는 T 세포는 CBLB 유전자에서 DSB를 도입함으로써 편집된다. 일 실시형태에서, T 세포는 CAR을 암호화하는 공여자 수선 주형의 존재 하에서 하나 이상의 CBLB 대립유전자에서 DSB를 도입함으로써 조작된다. 특정 실시형태에서, 키메라 사이토카인 수용체는 단일 CBLB 대립유전자 내 DSB에서 삽입된다.

[0478] 일 실시형태에서, 본 명세서에서 상정된 조작된 T 세포는 CBLB 유전자에서 삽입되지 않는 키메라 사이토카인 수용체를 포함하며, 면역억제 신호 댐퍼, 플립 수용체, 조작된 T 세포 수용체(TCR)의 알파 및/또는 베타쇄, 키메라 항원 수용체(CAR), Daric 수용체 또는 이들의 성분, 또는 키메라 사이토카인 수용체 중 하나 이상은 하나 이상의 CBLB 대립유전자에서 DSB 내로 삽입된다.

[0479] 다양한 실시형태에서, 계놈 편집된 T 세포는 종양 세포를 향해 세포독성을 다시 보내는 키메라 사이토카인 수용체를 발현시킨다. 제타카인은 세포 표면, 막관통 영역 및 세포내 신호전달 도메인에 세포외 도메인을 테더링할 수 있는 지지 영역에 연결된 가용성 수용체 리간드를 포함하는 세포외 도메인을 포함하는 키메라 막관통 면역수용체이다. T 립프구의 표면 상에서 발현된 제타카인은 T 세포 활성을 가용성 수용체 리간드에 특이적인 수용체를 발현시키는 해당 세포로 보낸다. 제타카인 키메라 면역수용체는, 특히 인간 악성종양에 의해 이용되는 자가분비/근거리분비 사이토카인 시스템을 통한 다양한 암 치료에 대한 적용에 의해 T 세포의 항원 특이성을 다시 보낸다.

[0480] 특정 실시형태에서, 키메라 사이토카인 수용체는 면역억제 사이토카인 또는 이의 사이토카인 수용체 결합 변이체, 링커, 막관통 도메인, 및 세포내 신호전달 도메인을 포함한다.

[0481] 특정 실시형태에서, 사이토카인 또는 이의 사이토카인 수용체 결합 변이체는 인터류킨-4(IL-4), 인터류킨-6(IL-6), 인터류킨-8(IL-8), 인터류킨-10(IL-10) 및 인터류킨-13(IL-13)으로 이루어진 군으로부터 선택된다.

[0482] 특정 실시형태에서, 링커는 CH2CH3 도메인, 헌지 도메인 등을 포함한다. 일 실시형태에서, 링커는 IgG1, IgG4 또는 IgD의 CH2 및 CH3 도메인을 포함한다. 일 실시형태에서, 링커는 CD8  $\alpha$  또는 CD4 헌지 도메인을 포함한다.

[0483] 특정 실시형태에서, 막관통 도메인은 T-세포 수용체, CD3  $\delta$ , CD3  $\epsilon$ , CD3  $\gamma$ , CD3  $\zeta$ , CD4, CD5, CD8  $\alpha$ , CD9, CD16, CD22, CD27, CD28, CD33, CD37, CD45, CD64, CD80, CD86, CD134, CD137, CD152, CD154, AMN 및 PD-1의 알파 또는 베타쇄로 이루어진 군으로부터 선택된다.

[0484] 특정 실시형태에서, 세포내 신호전달 도메인은 1차 신호전달 도메인 및/또는 공자극 도메인을 함유하는 ITAM으로 이루어진 군으로부터 선택된다.

[0485] 특정 실시형태에서, 세포내 신호전달 도메인은 FcR  $\gamma$ , FcR  $\beta$ , CD3  $\gamma$ , CD3  $\delta$ , CD3  $\epsilon$ , CD3  $\zeta$ , CD22, CD79a, CD79b 및 CD66d로 이루어진 군으로부터 선택된다.

[0486] 특정 실시형태에서, 세포내 신호전달 도메인은 TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR8, TLR9, TLR10, CARD11, CD2, CD7, CD27, CD28, CD30, CD40, CD54(ICAM), CD83, CD134(OX40), CD137(4-1BB), CD278(ICOS), DAP10, LAT, NKD2C, SLP76, TRIM 및 ZAP70로 이루어진 군으로부터 선택된다.

[0487] 일 실시형태에서, 키메라 사이토카인 수용체는 CD28, CD137, 및 CD134, 및 CD3  $\zeta$  1차 신호전달 도메인으로 이루어진 군으로부터 선택된 하나 이상의 공자극 신호전달 도메인을 포함한다.

## [0488] F. 폴리펩타이드

[0489] 호밍 엔도뉴클레아제 변이체, 메가TAL 및 융합 폴리펩타이드를 포함하지만, 이들로 제한되지 않는 다양한 폴리펩타이드가 본 명세서에서 상정된다. 바람직한 실시형태에서, 폴리펩타이드는 서열번호 1 내지 19 및 38에 제시된 아미노산 서열을 포함한다. "폴리펩타이드", "폴리펩타이드 단편", "펩타이드" 및 "단백질"은, 달리 대조적으로 구체화되지 않는 한, 상호 호환적으로 그리고 통상적인 의미에 따라, 즉, 아미노산 서열로서 사용된다. 일 실시형태에서, "폴리펩타이드"는 융합 폴리펩타이드 및 다른 변이체를 포함한다. 폴리펩타이드는 특정 길이로 제한되지 않으며, 예를 들어, 그들은 전장 단백질 서열, 전장 단백질의 단편 또는 융합 단백질을 포함할 수 있고, 폴리펩타이드의 번역후 변형, 예를 들어, 글리코실화, 아세틸화, 포스포릴화 등뿐만 아니라 당업계에 공지된 다른 변형, 즉, 천연 유래와 비천연 유래 둘 다를 포함할 수 있다.

- [0490] 본 명세서에서 사용된 바와 같은 "단리된 단백질", "단리된 웨타이드" 또는 "단리된 폴리웨타이드" 등은 세포 환경으로부터 그리고 세포의 다른 성분과의 결합으로부터의 웨타이드 또는 폴리웨타이드 분자의 시험관내 합성, 단리 및/또는 정제를 지칭하며, 즉, 이는 생체내 물질과 상당하게 관련되지 않는다.
- [0491] 특정 실시형태에서 상정된 폴리웨타이드의 예시적 예는 호밍 엔도뉴클레아제 변이체, 메가TAL, BiTE, 사이토카인, 케모카인, 사이토톡신 및 사이토카인 수용체, 플립 수용체, 면역억제 신호 램프, CAR, DARIC, TCR 및 제타 카인을 포함하지만, 이들로 제한되지 않는다.
- [0492] 폴리웨타이드는 "폴리웨타이드 변이체"를 포함한다. 폴리웨타이드 변이체는 하나 이상의 아미노산 치환, 결실, 첨가 및/또는 삽입에서 천연 유래 폴리웨타이드와 상이할 수 있다. 이러한 변이체는 천연 유래일 수 있거나 또는, 예를 들어, 상기 폴리웨타이드 서열의 하나 이상의 아미노산을 변형시킴으로써 합성에 의해 생성될 수 있다. 예를 들어, 특정 실시형태에서, 폴리웨타이드 내로 하나 이상의 치환, 결실, 첨가 및/또는 삽입을 도입함으로써 인간 CBLB 유전자 내 표적 부위에 결합하고, 이를 절단하는 호밍 엔도뉴클레아제, 메가TAL 등의 생물학적 특성을 개선시키는 것이 바람직할 수 있다. 특정 실시형태에서, 폴리웨타이드는 본 명세서에서 상정된 임의의 기준 서열에 대해 적어도 약 65%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99%의 아미노산 서열을 갖는 폴리웨타이드를 포함하며, 전형적으로 여기서 변이체는 기준 서열의 적어도 하나의 생물학적 활성을 유지한다.
- [0493] 폴리웨타이드 변이체는 생물학적 활성 "폴리웨타이드 단편"을 포함한다. 생물학적 활성 폴리웨타이드 단편의 예시적 예는 DNA 결합 도메인, 뉴클레아제 도메인 등을 포함한다. 본 명세서에서 사용되는 용어 "생물학적 활성 단편" 또는 "최소 생물학적 활성 단편"은 천연 유래 폴리웨타이드 활성의 적어도 100%, 적어도 90%, 적어도 80%, 적어도 70%, 적어도 60%, 적어도 50%, 적어도 40%, 적어도 30%, 적어도 20%, 적어도 10% 또는 적어도 5%를 보유하는 폴리웨타이드 단편을 지칭한다. 바람직한 실시형태에서, 생물학적 활성은 표적 서열에 대한 결합 친화도 및/또는 절단 활성이다. 특정 실시형태에서, 단편은 적어도 5 내지 약 1700개의 아미노산 길이의 아미노산 쇄를 포함할 수 있다. 특정 실시형태에서, 단편은 적어도 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700개 이상의 아미노산 길이라는 것이 인식될 것이다. 특정 실시형태에서, 폴리웨타이드는 호밍 엔도뉴클레아제 변이체의 생물학적 활성 단편을 포함한다. 특정 실시형태에서, 본 명세서에 제시된 폴리웨타이드는 "X"로서 나타내는 하나 이상의 아미노산을 포함할 수 있다. "X"가 아미노산 서열 번호에 존재한다면, 임의의 아미노산을 지칭한다. 하나 이상의 "X" 잔기는 본 명세서에서 상정된 특정 서열번호에 제시된 아미노산 서열의 N- 및 C-말단에 존재할 수 있다. "X" 아미노산이 존재하지 않는다면, 서열번호에 제시된 남아있는 아미노산 서열은 생물학적 활성 단편이 고려될 수 있다.
- [0494] 특정 실시형태에서, 폴리웨타이드는 호밍 엔도뉴클레아제 변이체의 생물학적 활성 단편, 예를 들어, 서열번호 3 내지 12, 또는 메가TAL(서열번호 13 내지 19)을 포함한다. 생물학적 활성 단편은 N-말단의 절단 및/또는 C-말단의 절단을 포함할 수 있다. 특정 실시형태에서, 생물학적 활성 단편은 대응하는 야생형 호밍 엔도뉴클레아제 서열에 비해 호밍 엔도뉴클레아제 변이체의 1, 2, 3, 4, 5, 6, 7 또는 8개의 N-말단 아미노산의 결실, 더 바람직하게는 대응하는 야생형 호밍 엔도뉴클레아제 서열에 비해 호밍 엔도뉴클레아제 변이체의 4개의 N-말단 아미노산의 결실을 결여하거나 또는 포함한다. 특정 실시형태에서, 생물학적 활성 단편은 대응하는 야생형 호밍 엔도뉴클레아제 서열에 비해 호밍 엔도뉴클레아제 변이체의 1, 2, 3, 4 또는 5개의 C-말단 아미노산의 결실, 더 바람직하게는 대응하는 야생형 호밍 엔도뉴클레아제 서열에 비해 호밍 엔도뉴클레아제 변이체의 2개의 C-말단 아미노산의 결실을 결여하거나 또는 포함한다. 특정 바람직한 실시형태에서, 생물학적 활성 단편은 대응하는 야생형 호밍 엔도뉴클레아제 서열에 비해 호밍 엔도뉴클레아제 변이체의 4개의 N-말단 아미노산 및 2개의 C-말단 아미노산의 결실을 결여하거나 또는 포함한다.
- [0495] 특정 실시형태에서, I-OnuI 변이체는 다음의 N-말단 아미노산: M, A, Y, M, S, R, R, E의 1, 2, 3, 4, 5, 6, 7 또는 8개의 결실; 및/또는 다음의 1, 2, 3, 4 또는 5개의 C-말단 아미노산: R, G, S, F, V의 결실을 포함한다.
- [0496] 특정 실시형태에서, I-OnuI 변이체는 다음의 N-말단 아미노산: M, A, Y, M, S, R, R, E의 1, 2, 3, 4, 5, 6, 7 또는 8개의 결실 또는 치환; 및/또는 다음의 1, 2, 3, 4 또는 5개의 C-말단 아미노산: R, G, S, F, V의 결실 또는 치환을 포함한다.

[0497]

상기 언급한 바와 같이, 폴리펩타이드는 아미노산 치환, 결실, 절단 및 삽입을 포함하는 다양한 방법으로 변경될 수 있다. 이러한 조작을 위한 방법은 일반적으로 당업계에 공지되어 있다. 예를 들어, 기준 폴리펩타이드의 아미노산 서열 변이체는 DNA에서의 돌연변이에 의해 제조될 수 있다. 돌연변이유발 및 뉴클레오타이드 서열 변경을 위한 방법은 당업계에 잘 공지되어 있다. 예를 들어, 문헌[Kunkel (1985, Proc. Natl. Acad. Sci. USA. 82: 488-492), Kunkel et al., (1987, Methods in Enzymol., 154: 367-382)], 미국 특허 제4,873,192호, 문헌 [Watson, J. D. et al., (Molecular Biology of the Gene, Fourth Edition, Benjamin/Cummings, Menlo Park, Calif., 1987)] 및 그에 인용된 참고문헌을 참조. 관심 대상의 단백질의 생물학적 활성에 영향을 미치지 않는 적절한 아미노산 치환에 대한 가이드는 문헌[Dayhoff et al., (1978) Atlas of Protein Sequence and Structure (Natl. Biomed. Res. Found., Washington, D.C.)]의 모델에서 찾을 수 있다.

[0498]

특정 실시형태에서, 변이체는 하나 이상의 보존적 치환을 함유할 것이다. "보존적 치환"은 펩타이드 화학 분야의 당업자가 폴리펩타이드의 2차 구조 및 소수성 특성이 실질적으로 변하지 않는 것으로 예상하도록 아미노산이 유사한 특성을 갖는 다른 아미노산으로 치환되는 것이다. 변형은 특정 실시형태에서 상정된 폴리뉴클레오타이드 및 폴리펩타이드의 구조에서 만들어질 수 있으며, 폴리펩타이드는 적어도 폴리펩타이드를 포함하고, 바람직한 특징을 갖는 변이체 또는 유도체 폴리펩타이드를 암호화하는 작용성 분자를 여전히 얻는다. 동등한, 또는 심지어 개선된 변이체 폴리펩타이드를 생성하기 위해 폴리펩타이드의 아미노산 서열을 변경시킬 때, 당업자는, 예를 들어, 표 1에 따라, 예를 들어 암호화 DNA 서열의 코돈 중 하나 이상을 변화시킬 수 있다.

표 1

## 아미노산 코돈

아미노산	1글자 암호	3글자 암호	코돈					
알라닌	A	Ala	GCA	GCC	GCG	GCU		
시스테인	C	Cys	UGC	UGU				
아스파르트산	D	Asp	GAC	GAU				
글루탐산	E	Glu	GAA	GAG				
페닐알라닌	F	Phe	UUC	UUU				
글리신	G	Gly	GGA	GGC	GGG	GGU		
히스티딘	H	His	CAC	CAU				
아이소류신	I	Iso	AUA	AUC	AUU			
라이신	K	Lys	AAA	AAG				
류신	L	Leu	UUA	UUG	CUA	CUC	CUG	CUU
메티오닌	M	Met	AUG					
아스파라긴	N	Asn	AAC	AAU				
프롤린	P	Pro	CCA	CCC	CCG	CCU		
글루타민	Q	Gln	CAA	CAG				
알기닌	R	Arg	AGA	AGG	CGA	CGC	CGG	CGU
세린	S	Ser	AGC	AGU	UCA	UCC	UCG	UCU
트레오닌	T	Thr	ACA	ACC	ACG	ACU		
발린	V	Val	GUU	GUC	GUG	GUU		
트립토판	W	Trp	UGG					
타이로신	Y	Tyr	UAC	UAU				

[0499]

[0500] 생물학적 활성을 없애는 일 없이 아미노산 잔기가 치환, 삽입 또는 결실되는 것을 결정하는 가이드는 당업계에 잘 공지된 컴퓨터 프로그램, 예컨대 DNASTAR, GCG, DNA Strider, Geneious, Mac Vector, 또는 Vector NTI 소프트웨어를 이용하여 찾을 수 있다. 바람직하게는, 본 명세서에 개시된 단백질 변이체에서 아미노산 변화는 보존적 아미노산 변화, 즉, 유사하게 하전된 또는 비하전된 아미노산의 치환이다. 보존적 아미노산 변화는 그들의 측쇄와 관련된 아미노산 패밀리 중 하나의 치환을 수반한다. 천연 유래 아미노산은 일반적으로 4가지 패밀리로 분류된다: 산성(아스파르테이트, 글루타메이트), 염기성(라이신, 알기닌, 히스티딘), 비극성(알라닌, 발린, 류신, 아이소류신, 프롤린, 페닐알라닌, 메티오닌, 트립토판) 및 비하전 극성(글리신, 아스파라긴, 글루타민, 시스테인, 세린, 트레오닌, 타이로신) 아미노산. 페닐알라닌, 트립토판 및 타이로신은 때때로 공통적으로 방향족 아미노산으로서 분류된다. 웨타이드 또는 단백질에서, 아미노산의 적합한 보존적 치환은 당업자에게 공지되어 있으며, 일반적으로 얻어진 분자의 생물학적 활성을 변경시키는 일 없이 이루어질 수 있다. 당업자라면 일반적으로, 폴리웨타이드의 비필수 영역 내 단일 아미노산 치환이 생물학적 활성을 실질적으로 변경시키지 않는다는 것을 인식한다(예를 들어, 문헌[Watson et al. Molecular Biology of the Gene, 4th Edition, 1987, The

Benjamin/Cummings Pub. Co., p.224] 참조).

- [0501] 일 실시형태에서, 2 이상의 폴리펩타이드의 발현이 요망되는 경우, 그들을 암호화하는 폴리뉴클레오타이드 서열은 본 명세서의 다른 곳에서와 같이 IRES에 의해 분리될 수 있다.
- [0502] 특정 실시형태에서 상정된 폴리펩타이드는 융합 폴리펩타이드를 포함한다. 특정 실시형태에서, 융합 폴리펩타이드 및 융합 폴리펩타이드를 암호화하는 폴리뉴클레오타이드가 제공된다. 융합 폴리펩타이드 및 융합 단백질은 적어도 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 폴리펩타이드 세그먼트를 갖는 폴리펩타이드를 지칭한다.
- [0503] 다른 실시형태에서, 2 이상의 폴리펩타이드는 본 명세서의 다른 곳에 개시된 바와 같은 하나 이상의 자기-절단성 폴리펩타이드 서열을 포함하는 융합 단백질로서 발현될 수 있다.
- [0504] 일 실시형태에서, 본 명세서에서 상정된 융합 단백질은 하나 이상의 DNA 결합 도메인 및 하나 이상의 뉴클레아제, 및 하나 이상의 링커 및/또는 자기-절단성 폴리펩타이드를 포함한다.
- [0505] 일 실시형태에서, 본 명세서에서 상정된 융합 단백질은 뉴클레아제 변이체; 링커 또는 자기-절단성 웨პ타이드; 및 5'-3' 엑소뉴클레아제, 5'-3' 알칼리성 엑소뉴클레아제, 및 3'-5' 엑소뉴클레아제(예를 들어, Trex2)를 포함하지만, 이들로 제한되지 않는, 말단-가공 효소를 포함한다.
- [0506] 융합 폴리펩타이드는 신호 웨პ타이드, 세포 침투성 웨პ타이드 도메인(CPP), DNA 결합 도메인, 뉴클레아제 도메인 등, 에피토프 태그(예를 들어, 말토스 결합 단백질("MBP")), 글루타티온 S 트랜스퍼라제(GST), HIS6, MYC, FLAG, V5, VSV-G 및 HA), 폴리펩타이드 링커 및 폴리펩타이드 절단 신호를 포함하지만, 이들로 제한되지 않는 하나 이상의 폴리펩타이드 도메인 또는 세그먼트를 포함할 수 있다. 융합 폴리펩타이드는, 그들이 또한 C-말단에서 C-말단으로, N-말단에서 N-말단으로, 또는 N-말단에서 C-말단으로 연결될 수 있지만, 전형적으로 C-말단에서 N-말단으로 연결된다. 특정 실시형태에서, 융합 단백질의 폴리펩타이드는 임의의 순서일 수 있다. 융합 폴리펩타이드 또는 융합 단백질은 또한, 융합 폴리펩타이드의 목적으로 하는 활성이 보존되는 한, 보존적으로 변형된 변이체, 다형성 변이체, 대립유전자, 돌연변이체, 하위서열 및 중간 상동체를 포함할 수 있다. 융합 폴리펩타이드는 화학적 합성 방법에 의해 또는 두 모이어티 사이의 화학적 결합에 의해 생산될 수 있거나 또는 일반적으로 다른 표준 기법을 이용하여 제조될 수 있다. 융합 폴리펩타이드를 포함하는 결찰된 DNA 서열은 본 명세서의 다른 곳에 개시된 바와 같은 적합한 전사 또는 번역 제어 요소에 작동 가능하게 연결된다.
- [0507] 융합 폴리펩타이드는 하나 이상의 폴리펩타이드 또는 폴리펩타이드 내의 도메인을 연결하기 위해 사용될 수 있는 링커를 선택적으로 포함할 수 있다. 웨პ타이드 링커 서열은 폴리펩타이드 도메인이 그들의 목적으로 하는 기능을 발휘하도록 각각의 폴리펩타이드가 그의 적절한 2차 및 3차 구조로 폴딩되는 것을 보장하는 데 충분한 거리만큼 임의의 2개 이상의 폴리펩타이드 성분을 분리시키는 데 사용될 수 있다. 이러한 웨პ타이드 링커 서열은 당업계의 표준 기법을 이용하여 융합 폴리펩타이드 내로 혼입된다. 적합한 웨პ타이드 링커 서열은 다음의 인자에 기반하여 선택될 수 있다: (1) 가요성의 연장된 입체배좌를 채택하는 그들의 능력; (2) 제1 및 제2 폴리펩타이드 상의 작용성 에피토프와 상호작용할 수 있는 2차 구조를 채택하는 그들의 불능; 및 (3) 폴리펩타이드 작용성 에피토프와 반응하는 소수성 또는 하전된 잔기의 결여. 바람직한 웨პ타이드 링커 서열은 Gly, Asn 및 Ser 잔기를 함유한다. 다른 근처의 중성 아미노산, 예컨대 Thr 및 Ala은 또한 링커 서열에서 사용될 수 있다. 링커로서 유용하게 사용될 수 있는 아미노산 서열은 문헌[Maratea et al., Gene 40:39-46, 1985; Murphy et al., Proc. Natl. Acad. Sci. USA 83:8258-8262, 1986]; 미국 특허 제4,935,233호 및 미국 특허 제4,751,180호에 개시된 것을 포함한다. 특정 융합 폴리펩타이드 세그먼트가 작용성 도메인을 분리시키고 입체 방해를 방지하기 위해 사용될 수 있는 비필수 N-말단 아미노산 영역을 함유할 때 링커 서열은 필요하지 않다. 바람직한 링커는 재조합 융합 단백질의 부분으로서 합성되는 전형적으로 가요성의 아미노산 하위서열이다. 링커 폴리펩타이드는 길이가 1 내지 200개의 아미노산, 길이가 1 내지 100개의 아미노산, 또는 길이가 1 내지 50개의 아미노산(그 사이의 모든 정수를 포함)일 수 있다.
- [0508] 예시적인 링커는 다음의 아미노산 서열을 포함하지만, 이들로 제한되지 않는다: 글리신 중합체(G)n; 글리신-세린 중합체(G1-5S1-5)n(여기서, n은 적어도 1, 2, 3, 4 또는 5의 정수임); 글리신-알라닌 중합체; 알라닌-세린 중합체; GGG(서열번호 39); DGGGS(서열번호 40); TGEKP(서열번호 41)(예를 들어, 문헌[Liu et al., PNAS 5525-5530(1997)] 참조); GGRR(서열번호 42)(Pomerantz et al. 1995, 상기 참조); (GGGGS)n(여기서, n = 1, 2, 3, 4 또는 5임)(서열번호 43)(Kim et al., PNAS 93, 1156-1160(1996.); EGKSSGGSESKVD(서열번호 44)(Chaudhary et al., 1990, Proc. Natl. Acad. Sci. U.S.A. 87:1066-1070); KESGSVSSEQLAQFRSLD(서열번호 45)(Bird et al., 1988, Science 242:423-426), GGRRGGGS(서열번호 46); LRQRDGERP(서열번호 47); LRQKDGGGSER(서열번호 48);

LRQKD(GGGS)2ERP(서열번호 49). 대안적으로, 가요성 링커는 DNA-결합 부위와 웹타이드 그 자체를 둘 다 모델링 할 수 있는 컴퓨터 프로그램을 이용하여(문헌[Desjarlais & Berg, PNAS 90:2256-2260(1993), PNAS 91:11099-11103 (1994)]) 또는 과시 디스플레이 방법에 의해 합리적으로 설계될 수 있다.

[0509] 융합 폴리웹타이드는 본 명세서에 기재된 각각의 폴리웹타이드 사이에 또는 내인성 오픈 리딩 프레임과 공여자 수선 주형에 의해 암호화된 폴리웹타이드 사이에 폴리웹타이드 절단 신호를 추가로 포함할 수 있다. 추가로, 폴리웹타이드 절단 부위는 임의의 링커 웹타이드 서열 내로 들어갈 수 있다. 예시적인 폴리웹타이드 절단 신호는 폴리웹타이드 절단 인식 부위, 예컨대 프로테아제 절단 부위, 뉴클레아제 절단 부위(예를 들어, 희귀 제한 효소 인식 부위, 자기-절단성 리보자임 인식 부위) 및 자기-절단성 바이러스 올리고웹타이드를 포함한다(문헌[deFelipe and Ryan, 2004. *Traffic*, 5(8); 616-26] 참조).

[0510] 적합한 프로테아제 절단 부위 및 자기-절단성 웹타이드는 당업자에게 공지되어 있다(예를 들어, 문헌[Ryan et al., 1997. *J. Gener. Virol.* 78, 699-722; Scymczak et al. (2004) *Nature Biotech.* 5, 589-594] 참조). 예시적인 프로테아제 절단 부위는 포티바이러스 NIa 프로테아제(예를 들어, 담배 식각 바이러스), 포티바이러스 HC 프로테아제, 포티바이러스 P1(P35) 프로테아제, 비오바이러스 NIa 프로테아제, 비오바이러스 RNA-2-암호화된 프로테아제, 아프타바이러스 L 프로테아제, 엔테로바이러스 2A 프로테아제, 리노바이러스 2A 프로테아제, 피코르나 3C 프로테아제, 코모바이러스 24K 프로테아제, 네포바이러스 24K 프로테아제, RTSV (벼 통그로 구형 바이러스) 3C-유사 프로테아제, PYVF(파스닙 황색 얼룩 바이러스(parsnip yellow fleck virus)) 3C-유사 프로테아제, 헤파린, 트롬빈, 인자 Xa 및 엔테로키나제의 절단 부위를 포함하지만, 이들로 제한되지 않는다. 그의 고절단 염격성에 기인하여, TEV(담배 식각 바이러스) 프로테아제 절단 부위는 일 실시형태, 예를 들어, EXXYXQ(G/S)(서열번호 50), 예를 들어, ENLYFQG(서열번호 51) 및 ENLYFQS(서열번호 52)에서 바람직하되, X는 임의의 아미노산을 나타낸다(TEV에 의한 절단은 Q와 G 또는 Q와 S 사이에서 생김).

[0511] 특정 실시형태에서, 자기-절단성 폴리웹타이드 부위는 2A 또는 2A-유사 부위, 서열 또는 도메인을 포함한다 (Donnelly et al., 2001. *J. Gen. Virol.* 82:1027-1041). 특정 실시형태에서, 바이러스 2A 웹타이드는 아프타 바이러스 2A 웹타이드, 포티바이러스 2A 웹타이드, 또는 카디오바이러스 2A 웹타이드이다.

[0512] 일 실시형태에서, 바이러스 2A 웹타이드는 구제역바이러스(FMDV)(F2A) 웹타이드, 마 비염 A 바이러스 (ERAV)(E2A) 웹타이드, 토세아 아시그나 바이러스(Thosea asigna virus: TaV)(T2A) 웹타이드, 돼지 테스코바이러스-1(PTV-1)(P2A) 웹타이드, 테일로바이러스 2A 웹타이드 및 뇌심근염바이러스 2A 웹타이드로 이루어진 군으로부터 선택된다.

[0513] 2A 부위의 예시적인 예를 표 2에 제공한다.

## 표 2

예시적인 2A 부위는 다음의 서열을 포함한다:

서열번호 53	GSGATNFSLLKQAGDVEENPGP
서열번호 54	ATNFSLLKQAGDVEENPGP
서열번호 55	LLKQAGDVEENPGP
서열번호 56	GSGEGRGSLLTCGDVEENPGP
서열번호 57	EGRGSLLTCGDVEENPGP
서열번호 58	LLTCGDVEENPGP
서열번호 59	GSGQCTNYALLKLAGDVESNPGP
서열번호 60	QCTNYALLKLAGDVESNPGP
서열번호 61	LLKLAGDVESNPGP
서열번호 62	GSGVKQTLNFDLLKLAGDVESNPGP
서열번호 63	VKQTLNFDLLKLAGDVESNPGP
서열번호 64	LLKLAGDVESNPGP
서열번호 65	LLNFDLLKLAGDVESNPGP
서열번호 66	TLNFDLLKLAGDVESNPGP
서열번호 67	LLKLAGDVESNPGP
서열번호 68	NFDLLKLAGDVESNPGP
서열번호 69	QLNFDLLKLAGDVESNPGP
서열번호 70	APVKQTLNFDLLKLAGDVESNPGP
서열번호 71	VTELYRMKRAETYCPRPLLAIHPTEARHKQKIVAPVKQT
서열번호 72	LNF DLLKLAGDVESNPGP
서열번호 73	LLAIHPTEARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP
서열번호 74	EARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP

[0514]

## G. 폴리뉴클레오타이드

특정 실시형태에서, 본 명세서에서 상정된 하나 이상의 호밍 엔도뉴클레아제 변이체, 메가TAL, 말단-가공 효소 및 융합 폴리펩타이드 또는 하나 이상의 생물학적 활성 단편 또는 변이체를 암호화하는 폴리뉴클레오타이드가 제공된다. 본 명세서에서 사용되는 용어 "폴리뉴클레오타이드" 또는 "핵산"은 데옥시리보핵산(DNA), 리보핵산(RNA) 및 DNA/RNA 혼성체를 지칭한다. 폴리뉴클레오타이드는 단일 가닥 또는 이중 가닥이고 재조합, 합성이거나, 또는 단리될 수 있다. 폴리뉴클레오타이드는 전-전령 RNA(전-mRNA), 전령 RNA(mRNA), RNA, 짧은 간섭 RNA(siRNA), 짧은 헤어핀 RNA(shRNA), 마이크로RNA(miRNA), 리보자임, 게놈 RNA(gRNA), 플러스 가닥 RNA(RNA(+)), 마이너스 가닥 RNA(RNA(-)), tracrRNA, crRNA, 단일 가이드 RNA(sgRNA), 합성 RNA, 합성 mRNA, 게놈 DNA(gDNA), PCR 증폭된 DNA, 상보성 DNA(cDNA), 합성 DNA 또는 재조합 DNA를 포함하지만, 이들로 제한되

[0515]

[0516]

지 않는다. 특정 실시형태에서, 폴리뉴클레오타이드는 하나 이상의 생물학적 활성 폴리펩타이드 단편 또는 변이체를 암호화하는 폴리뉴클레오타이드 단편이다. 폴리뉴클레오타이드는 리보뉴클레오타이드 또는 데옥시뉴클레오타이드 또는 뉴클레오타이드 유형의 변형된 형태의 길이가 적어도 5, 적어도 10, 적어도 15, 적어도 20, 적어도 25, 적어도 30, 적어도 40, 적어도 50, 적어도 100, 적어도 200, 적어도 300, 적어도 400, 적어도 500, 적어도 1000, 적어도 5000, 적어도 10000, 또는 적어도 15000개 이상뿐만 아니라 모든 중간 길이인 뉴클레오타이드의 중합체 형태를 지칭한다. 이와 관련하여 "중간 길이"는 인용된 값 사이의 임의의 길이, 예컨대 6, 7, 8, 9 등, 101, 102, 103 등; 151, 152, 153 등; 201, 202, 203 등을 의미한다. 특정 실시형태에서, 폴리뉴클레오타이드 또는 변이체는 기준 서열에 대해 적어도 또는 약 50%, 55%, 60%, 65%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 가진다.

[0517] 특정 실시형태에서, 폴리뉴클레오타이드는 코돈-최적화될 수 있다. 본 명세서에서 사용되는 용어 "코돈-최적화된"은 폴리펩타이드의 빌현, 안정성 및/또는 활성을 증가시키기 위해 폴리펩타이드를 암호화하는 폴리뉴클레오타이드 내 코돈을 치환하는 것을 지칭한다. 코돈 최적화에 영향을 미치는 인자는 (i) 둘 이상의 유기체 사이의 코돈 편향의 변형 또는 유전자 또는 합성으로 작제된 편향 표, (ii) 유기체, 유전자 또는 유전자 세트 내에서 코돈 편향 정도의 변형, (iii) 맥락을 포함하는 코돈의 체계적 변형, (iv) 그들의 탈암호 tRNA에 따른 코돈의 변형, (v) 전반적으로 또는 삼중 중 하나의 위치에서 GC%에 따른 코돈의 변형, (vi) 기준 서열, 예를 들어 천연 유래 서열과의 유사성 정도의 변형, (vii) 코돈 빈도 컷오프에서의 변형, (viii) DNA 서열로부터 전사된 mRNA의 구조적 특성, (ix) 코돈 치환 세트의 설계에 기반한 DNA 서열의 기능에 관한 사전 지식, (x) 각각의 아미노산에 대한 코돈 세트의 체계적 변형, 및/또는 (xi) 비논리적 번역 개시 부위의 단리된 제거를 포함하지만, 이들로 제한되지 않는다.

[0518] 본 명세서에서 사용되는 용어 "뉴클레오타이드"는 포스포릴화된 당과의 N-글리코사이드 결합에서의 복소환식 질소 염기를 지칭한다. 뉴클레오타이드는 천연 염기 및 매우 다양한 기술-인식된 변형 염기를 포함하는 것으로 이해된다. 이러한 염기는 일반적으로 뉴클레오타이드 당 모이어티의 1' 위치에 위치된다. 뉴클레오타이드는 일반적으로 염기, 당 및 인산염기를 포함한다. 리보핵산(RNA)에서, 당은 리보스이며, 데옥시리보핵산(DNA)에서, 당은 데옥시리보스, 즉, 리보스에 존재하는 하이드록실기가 없는 당이다. 예시적인 천연 질소 염기는 퓨린, 아데노신(A) 및 구아니딘(G), 및 피리미딘, 사이티딘(C) 및 티미딘(T)(또는 RNA와 관련하여, 유라실(U))을 포함한다. 데옥시리보스의 C-1 원자는 피리미딘의 N-1 또는 퓨린의 N-9에 결합된다. 뉴클레오타이드는 보통 일, 이- 또는 삼인산염이다. 뉴클레오타이드는 당, 인산염 및/또는 염기 모이어티에서 비변형 또는 변형될 수 있다(또한 뉴클레오타이드 유사체, 뉴클레오타이드 유도체, 변형된 뉴클레오타이드, 비천연 뉴클레오타이드 및 비표준 뉴클레오타이드와 상호 호환적으로 지칭됨; 예를 들어, WO 92/07065 및 WO 93/15187). 변형된 핵산 염기의 예는 문헌[Limbach et al., (1994, Nucleic Acids Res. 22, 2183-2196)]에 의해 요약된다.

[0519] 뉴클레오타이드는 또한 당의 C-5에 부착된 하이드록실기 상에서 생기는 에스터화에 의한 뉴클레오사이드의 인산 염 에스터로서 간주될 수 있다. 본 명세서에서 사용되는 용어 "뉴클레오사이드"는 당과의 N-글리코사이드 결합에서의 복소환식 질소 염기를 지칭한다. 뉴클레오사이드는 천연 염기를 포함하고 잘 공지된 변형된 염기를 포함하는 것으로 당업계에서 인식된다. 이러한 염기는 일반적으로 뉴클레오사이드 당 모이어티의 1' 위치에 위치된다. 뉴클레오사이드는 일반적으로 염기 및 당기를 포함한다. 뉴클레오사이드는 비변형이거나 또는 당 및/또는 염기 모이어티에서 변형될 수 있다(또한 뉴클레오사이드 유사체, 뉴클레오사이드 유도체, 변형된 뉴클레오사이드, 비천연 뉴클레오사이드 또는 비표준 뉴클레오사이드로서 상호 호환적으로 지칭됨). 또한 상기 언급한 바와 같이, 변형된 핵산 염기의 예는 문헌[Limbach et al., (1994, Nucleic Acids Res. 22, 2183-2196)]에 의해 요약된다.

[0520] 폴리뉴클레오타이드의 예시적 예는 서열번호 1 내지 19 및 38을 암호화하는 폴리뉴클레오타이드, 및 서열번호 20 내지 37에 제시된 폴리뉴클레오타이드 서열을 포함하지만, 이들로 제한되지 않는다.

[0521] 다양한 예시적 실시형태에서, 본 명세서에서 상정된 폴리뉴클레오타이드는 호밍 엔도뉴클레아제 변이체, 메가TAL, 말단-가공 효소, 융합 폴리펩타이드, 및 빌현 벡터, 바이러스 벡터, 및 본 명세서에서 상정된 폴리뉴클레오타이드를 포함하는 전달 폴리펩타이드를 암호화하는 폴리뉴클레오타이드를 포함하지만, 이들로 제한되지 않는다.

[0522] 본 명세서에서 사용되는 용어 "폴리뉴클레오타이드 변이체" 및 "변이체" 등은 기준 폴리뉴클레오타이드 서열과 실질적인 서열 동일성을 나타내는 폴리뉴클레오타이드 또는 본 명세서에서 이후에 나타내는 엄격한 조건 하에

기준 서열과 혼성화되는 폴리뉴클레오타이드를 지칭한다. 이들 용어는 또한 적어도 하나의 뉴클레오타이드의 첨가, 결실, 치환 또는 변형에 의해 기준 폴리뉴클레오타이드와 구별되는 폴리뉴클레오타이드를 포함한다. 따라서, 용어 "폴리뉴클레오타이드 변이체" 및 "변이체"는 하나 이상의 뉴클레오타이드가 첨가되거나 또는 결실되거나, 또는 변형되거나 또는 상이한 뉴클레오타이드로 대체된 폴리뉴클레오타이드를 포함한다. 이와 관련하여, 돌연변이, 첨가, 결실 및 치환을 포함하는 특정 변경은 기준 폴리뉴클레오타이드에 대해 이루어질 수 있고, 이에 의해 변경된 폴리뉴클레오타이드는 기준 폴리뉴클레오타이드의 생물학적 기능 또는 활성을 보유한다는 것이 잘 이해된다.

[0523] 폴리뉴클레오타이드 변이체는 생물학적 활성 폴리펩타이드 단편 또는 변이체를 암호화하는 폴리뉴클레오타이드 단편을 포함한다. 본 명세서에서 사용되는 용어 "폴리뉴클레오타이드 단편"은 천연 유래 폴리펩타이드 활성의 적어도 100%, 적어도 90%, 적어도 80%, 적어도 70%, 적어도 60%, 적어도 50%, 적어도 40%, 적어도 30%, 적어도 20%, 적어도 10% 또는 적어도 5%를 보유하는 폴리펩타이드 변이체를 암호화하는 길이가 적어도 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700개 이상의 뉴클레오타이드의 폴리뉴클레오타이드 단편을 지칭한다. 폴리뉴클레오타이드 단편은 천연 유래 또는 재조합 생성된 폴리펩타이드의 하나 이상의 아미노-말단 결실, 카복실-말단 결실 및/또는 내부 결실 또는 치환을 갖는 폴리펩타이드를 암호화하는 폴리뉴클레오타이드를 지칭한다.

[0524] 일 실시형태에서, 폴리뉴클레오타이드는 염격한 조건 하에 표적 핵산 서열에 혼성화하는 뉴클레오타이드 서열을 포함한다. "엄격한 조건" 하에 혼성화하는 것은 서로에 대해 적어도 60% 동일한 뉴클레오타이드 서열이 혼성화된 채로 남아있는 혼성화 프로토콜을 기재한다. 일반적으로, 염격한 조건은 정해진 이온 강도 및 pH에서 특정 서열에 대해 열 융점( $T_m$ )보다 약 5°C 더 낮아지도록 선택된다.  $T_m$ 은 (정해진 이온 강도, pH 및 핵산 농도 하에) 표적 서열에 대해 상보성인 프로브의 50%가 평형 상태에서 표적 서열에 혼성화하는 온도이다. 표적 서열은 일반적으로 과량으로 존재하기 때문에,  $T_m$ 에서, 프로브의 50%는 평형상태에서 점유된다.

[0525] 설명 "서열 동일성" 또는, 예를 들어, 본 명세서에서 사용되는 바와 같이 "에 대해 50% 동일한 서열"을 포함하는 것은 서열이 비교창에 걸쳐 뉴클레오타이드 대 뉴클레오타이드 염기 기준 또는 아미노산 대 아미노산 기준과 동일한 정도를 지칭한다. 따라서, "서열 동일성의 백분율"은 비교창에 걸쳐 2개의 최적으로 정렬된 서열을 비교하는 단계, 매칭된 위치의 수를 수득하기 위해 동일한 핵산 염기(예를 들어, A, T, C, G, I) 또는 동일한 아미노산 잔기(예를 들어, Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys 및 Met)가 서열 둘 다에서 생기는 위치 수를 결정하는 단계, 매칭된 위치의 수를 비교창에서의 위치의 총수(즉, 창 크기)로 나누는 단계, 및 결과에 100을 곱하여서 서열 동일성의 백분율을 수득하는 단계에 의해 계산될 수 있다. 본 명세서에 기재된 임의의 기준 서열에 대해 적어도 약 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 뉴클레오타이드 및 폴리펩타이드가 포함되며, 전형적으로 폴리펩타이드 변이체는 기준 폴리펩타이드의 적어도 하나의 생물학적 활성을 유지한다.

[0526] 2 이상의 폴리뉴클레오타이드 또는 폴리펩타이드 사이의 서열 관계를 기재하기 위해 사용되는 용어는 "기준 서열", "비교창", "서열 동일성", "서열 동일성의 백분율" 및 "실질적인 동일성"을 포함한다. "기준 서열"은 길이가 적어도 12개이지만, 빈번하게는 15 내지 18개 및 종종 적어도 25개의 단량체 단위(뉴클레오타이드 및 아미노산 잔기를 포함)이다. 두 폴리뉴클레오타이드는 (1) 두 폴리뉴클레오타이드 사이의 유사한 서열(즉, 완전한 폴리뉴클레오타이드 서열의 일부만), 및 (2) 두 폴리뉴클레오타이드 간에 다른 서열을 각각 포함할 수 있기 때문에, 2(이상의) 폴리뉴클레오타이드 사이의 서열 비교는 전형적으로 서열 유사성의 국소 영역을 확인하고 비교하기 위해 "비교창"에 걸쳐 두 폴리뉴클레오타이드의 서열을 비교함으로써 수행된다. "비교창"은 적어도 6개의 인접한 위치, 보통 약 50 내지 약 100, 더 보통으로는 약 100 내지 약 150개의 개념적 세그먼트를 지칭하며, 이때 서열은 두 서열이 최적으로 정렬된 후에 동일한 수의 인접한 위치의 기준 서열과 비교된다. 비교창은 두 서열의 최적의 정렬을 위해 기준 서열(첨가 또는 결실을 포함하지 않음)에 의해 약 20% 이하의 첨가 또는 결실(즉, 캡)을 포함할 수 있다. 비교창을 정렬하기 위한 서열의 최적의 정렬은 알고리즘의 컴퓨터화된 실행(GAP, BESTFIT, FASTA 및 TFASTA, 예를 들어, DNASTAR, GCG, DNA Strider, Geneious, Mac Vector 또는 Vector NTI 소프트웨어)에 의해 또는 검사 및 선택된 임의의 다양한 방법에 의해 생성된 최적의 정렬(즉, 비교창에 걸쳐 가장 큰 백분율의 상동성을 초래)에 의해 수행될 수 있다. 또한, 예를 들어 문헌[Altschul et al., 1997, Nucl. Acids Res. 25:3389]에 의해 개시되는 바와 같이 프로그램의 BLAST 패밀리가 언급될 수 있다. 서열 분석의 상세

한 논의는 문헌[Unit 19.3 of Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons Inc., 1994-1998, Chapter 15]에서 찾을 수 있다.

[0527] 본 명세서에서 사용되는 "단리된 폴리뉴클레오타이드"는 천연-유래 상태, 예를 들어, 단편에 정상적으로 인접한 서열로부터 제거된 DNA 단편에서 그것에 측접하는 서열로부터 정제된 폴리뉴클레오타이드를 지칭한다. 특정 실시형태에서, "단리된 폴리뉴클레오타이드"는 상보성 DNA(cDNA), 재조합 폴리뉴클레오타이드, 합성 폴리뉴클레오타이드 또는 천연에 존재하지 않고 사람의 손으로 만들어지는 다른 폴리뉴클레오타이드를 지칭한다.

[0528] 다양한 실시형태에서, 폴리뉴클레오타이드는 호밍 엔도뉴클레아제 변이체, 메가TAL 및 말단-가공 효소를 포함하지만, 이들로 제한되지 않는 본 명세서에서 상정된 폴리펩타이드를 암호화하는 mRNA를 포함한다. 특정 실시형태에서, mRNA는 캡, 하나 이상의 뉴클레오타이드 및 폴리(A) 꼬리를 포함한다.

[0529] 본 명세서에서 사용되는 용어 "5' 캡" 또는 "5' 캡 구조" 또는 "5' 캡 모이어티"는 mRNA의 5' 말단에 혼입된 화학적 변형을 지칭한다. 5' 캡은 핵 유출, mRNA 안정성 및 변역에 연루된다.

[0530] 특정 실시형태에서, 본 명세서에서 상정된 mRNA는 말단의 구아노신 캡 잔기와 mRNA 분자의 5'-말단의 전사된 센스 뉴클레오타이드 사이의 5'-PPP-5'-삼인산염 결합을 포함하는 5' 캡을 포함한다. 이어서, 이 5'-구아닐레이트 캡은 메틸화되어 N7-메틸-구아닐레이트 잔기를 생성할 수 있다.

[0531] 본 명세서에서 상정된 mRNA 폴리뉴클레오타이드의 특정 실시형태에서 사용하기에 적합한 5' 캡의 예시적 예는 비메틸화된 5' 캡 유사체, 예를 들어, G(5')PPP(5')G, G(5')PPP(5')C, G(5')PPP(5')A; 메틸화된 5' 캡 유사체, 예를 들어,  $\text{m}^7\text{G}(5')\text{PPP}(5')\text{G}$ ,  $\text{m}^7\text{G}(5')\text{PPP}(5')\text{C}$ , 및  $\text{m}^7\text{G}(5')\text{PPP}(5')\text{A}$ ; 다이메틸화된 5' 캡 유사체, 예를 들어,  $\text{m}^{2,7}\text{G}(5')\text{PPP}(5')\text{G}$ ,  $\text{m}^{2,7}\text{G}(5')\text{PPP}(5')\text{C}$  및  $\text{m}^{2,7}\text{G}(5')\text{PPP}(5')\text{A}$ ; 트라이메틸화된 5' 캡 유사체, 예를 들어,  $\text{m}^{2,2,7}\text{G}(5')\text{PPP}(5')\text{G}$ ,  $\text{m}^{2,2,7}\text{G}(5')\text{PPP}(5')\text{C}$ , 및  $\text{m}^{2,2,7}\text{G}(5')\text{PPP}(5')\text{A}$ ; 다이메틸화된 대칭적 5' 캡 유사체, 예를 들어,  $\text{m}^7\text{G}(5')\text{PPPM7}(5')\text{G}$ ,  $\text{m}^7\text{G}(5')\text{PPPM7}(5')\text{C}$ , 및  $\text{m}^7\text{G}(5')\text{PPPM7}(5')\text{A}$ ; 및 안티-리버스 5' 캡 유사체, 예를 들어, 안티-리버스 캡 유사체(ARCA) 캡, 표기된 3'0-Me- $\text{m}^7\text{G}(5')\text{PPP}(5')\text{G}$ , 2'0-Me- $\text{m}^7\text{G}(5')\text{PPP}(5')\text{G}$ , 2'0-Me- $\text{m}^7\text{G}(5')\text{PPP}(5')\text{C}$ , 2'0-Me- $\text{m}^7\text{G}(5')\text{PPP}(5')\text{A}$ ,  $\text{m}^7\text{2}'\text{d}(5')\text{PPP}(5')\text{G}$ ,  $\text{m}^7\text{2}'\text{d}(5')\text{PPP}(5')\text{C}$ ,  $\text{m}^7\text{2}'\text{d}(5')\text{PPP}(5')\text{A}$ , 3'0-Me- $\text{m}^7\text{G}(5')\text{PPP}(5')\text{C}$ , 3'0-Me- $\text{m}^7\text{G}(5')\text{PPP}(5')\text{A}$ ,  $\text{m}^7\text{3}'\text{d}(5')\text{PPP}(5')\text{G}$ ,  $\text{m}^7\text{3}'\text{d}(5')\text{PPP}(5')\text{C}$ ,  $\text{m}^7\text{3}'\text{d}(5')\text{PPP}(5')\text{A}$  및 이들의 사인산염 유도체(예를 들어, 문헌[Jemielity et al., RNA, 9: 1108-1122 (2003)] 참조)를 포함하지만, 이들로 제한되지 않는다.

[0532] 특정 실시형태에서, mRNA는 제1 전사된 뉴클레오타이드의 5'-말단에 삼인산염 브리지를 통해 연결된 7-메틸 구아닐레이트("m<sup>7</sup>G")인 5' 캡을 포함하여,  $\text{m}^7\text{G}(5')\text{PPP}(5')\text{N}$ 을 초래하며, 여기서 N은 임의의 뉴클레오사이드이다.

[0533] 일부 실시형태에서, mRNA는 5' 캡을 포함하여, 캡은 Cap0 구조(Cap0 구조는 염기 1 및 2에 부착된 리보스의 2'-0-메틸 잔기를 결여함), Cap1 구조(Cap1 구조는 염기 2에서 2'-0-메틸 잔기를 가짐), 또는 Cap2 구조(Cap2 구조는 염기 2와 3 둘 다에 부착된 2'-0-메틸 잔기를 가짐)이다.

[0534] 일 실시형태에서, mRNA는  $\text{m}^7\text{G}(5')\text{PPP}(5')\text{G}$  캡을 포함한다.

[0535] 일 실시형태에서, mRNA는 ARCA 캡을 포함한다.

[0536] 특정 실시형태에서, 본 명세서에서 상정된 mRNA는 하나 이상의 변형된 뉴클레오사이드를 포함한다.

[0537] 일 실시형태에서, mRNA는 슈도유리딘, 피리딘-4-온 리보뉴클레오사이드, 5-아자-유리딘, 2-티오-5-아자-유리딘, 2-티오유리딘, 4-티오-슈도유리딘, 2-티오-슈도유리딘, 5-하이드록시유리딘, 3-메틸유리딘, 5-카복시메틸-유리딘, 1-카복시메틸-슈도유리딘, 5-프로핀일-유리딘, 1-프로핀일-슈도유리딘, 5-타우리노메틸유리딘, 1-타우리노메틸-슈도유리딘, 5-타우리노메틸-2-티오-유리딘, 1-타우리노메틸-4-티오-유리딘, 5-메틸-유리딘, 1-메틸-슈도유리딘, 4-티오-1-메틸-슈도유리딘, 2-티오-1-메틸-슈도유리딘, 1-메틸-1-데아자-슈도유리딘, 2-티오-1-메틸-1-데아자-슈도유리딘, 다이하이드로유리딘, 다이하이드로슈도유리딘, 2-티오-다이하이드로유리딘, 2-티오-다이하이드로슈도유리딘, 2-메톡시유리딘, 2-메톡시-4-티오-유리딘, 4-메톡시-슈도유리딘, 4-메톡시-2-티오-슈도유리딘, 5-아자-사이티딘, 슈도아이소사이티딘, 3-메틸-사이티딘, N4-아세틸사이티딘, 5-폼일사이티딘, N4-메틸사이티딘, 5-하이드록시메틸사이티딘, 1-메틸-슈도아이소사이티딘, 피롤로-사이티딘, 피롤로-슈도아이소사이티딘, 2-티오-사이티딘, 2-티오-5-메틸-사이티딘, 4-티오-슈도아이소사이티딘, 4-티오-1-메틸-슈도아이소사이티딘, 4-

티오-1-메틸-1-데아자-슈도아이소사이티딘, 1-메틸-1-데아자-슈도아이소사이티딘, 제불라린, 5-아자-제불라린, 5-메틸-제불라린, 5-아자-2-티오-제불라린, 2-티오-제불라린, 2-메톡시-사이티딘, 2-메톡시-5-메틸-사이티딘, 4-메톡시-슈도아이소사이티딘, 4-메톡시-1-메틸-슈도아이소사이티딘, 2-아미노퓨린, 2,6-다이아미노퓨린, 7-데아자-아데닌, 7-데아자-8-아자-아데닌, 7-데아자-2-아미노퓨린, 7-데아자-8-아자-2-아미노퓨린, 7-데아자-2,6-다이아미노퓨린, 7-데아자-8-아자-2,6-다이아미노퓨린, 1-메틸아데노신, N6-메틸아데노신, N6-아이소펜텐일아데노신, N6-(시스-하이드록시아이소펜텐일)아데노신, 2-메틸티오-N6-(시스-하이드록시아이소펜텐일)아데노신, N6-글리신일카바모일아데노신, N6-트레오닐카바모일아데노신, 2-메틸티오-N6-트레오닐 카바모일아데노신, N6,N6-다이메틸아데노신, 7-메틸아데닌, 2-메틸티오-아데닌, 2-메톡시-아데닌, 이노신, 1-메틸-이노신, 위오신, 위부토신, 7-데아자-구아노신, 7-데아자-8-아자-구아노신, 6-티오-구아노신, 6-티오-7-데아자-구아노신, 6-티오-7-데아자-8-아자-구아노신, 7-메틸-구아노신, 6-티오-7-메틸-구아노신, 7-메틸이노신, 6-메톡시-구아노신, 1-메틸구아노신, N2-메틸구아노신, N2,N2-다이메틸구아노신, 8-옥소-구아노신, 7-메틸-8-옥소-구아노신, 1-메틸-6-티오-구아노신, N2-메틸-6-티오-구아노신 및 N2,N2-다이메틸-6-티오-구아노신으로 이루어진 군으로부터 선택된 하나 이상의 변형된 뉴클레오사이드를 포함한다.

[0538]

일 실시형태에서, mRNA는 슈도유리딘, 피리딘-4-온 리보뉴클레오사이드, 5-아자-유리딘, 2-티오-5-아자-유리딘, 2-티오유리딘, 4-티오-슈도유리딘, 2-티오-슈도유리딘, 5-하이드록시유리딘, 3-메틸유리딘, 5-카복시메틸-유리딘, 1-카복시메틸-슈도유리딘, 5-프로핀일-유리딘, 1-프로핀일-슈도유리딘, 5-타우리노메틸유리딘, 1-타우리노메틸-슈도유리딘, 5-타우리노메틸-2-티오-유리딘, 1-타우리노메틸-4-티오-유리딘, 5-메틸-유리딘, 1-메틸-슈도유리딘, 4-티오-1-메틸-슈도유리딘, 2-티오-1-메틸-슈도유리딘, 1-메틸-1-데아자-슈도유리딘, 2-티오-1-메틸-1-데아자-슈도유리딘, 다이하이드로유리딘, 다이하이드로슈도유리딘, 2-티오-다이하이드로유리딘, 2-메톡시유리딘, 2-메톡시-4-티오-유리딘, 4-메톡시-슈도유리딘 및 4-메톡시-2-티오-슈도유리딘으로 이루어진 군으로부터 선택된 하나 이상의 변형된 뉴클레오사이드를 포함한다.

[0539]

일 실시형태에서, mRNA는 5-아자-사이티딘, 슈도아이소사이티딘, N4-아세틸사이티딘, 5-폼일사이티딘, N4-메틸사이티딘, 5-하이드록시메틸사이티딘, 1-메틸-슈도아이소사이티딘, 피롤로-사이티딘, 피롤로-슈도아이소사이티딘, 2-티오-사이티딘, 2-티오-5-메틸-사이티딘, 4-티오-슈도아이소사이티딘, 4-티오-1-메틸-슈도아이소사이티딘, 4-티오-1-메틸-1-데아자-슈도아이소사이티딘, 1-메틸-1-데아자-슈도아이소사이티딘, 제불라린, 5-아자-제불라린, 5-메틸-제불라린, 5-아자-2-티오-제불라린, 2-티오-제불라린, 2-메톡시-사이티딘, 2-메톡시-5-메틸-사이티딘, 4-메톡시-슈도아이소사이티딘 및 4-메톡시-1-메틸-슈도아이소사이티딘으로 이루어진 군으로부터 선택된 하나 이상의 변형된 뉴클레오사이드를 포함한다.

[0540]

일 실시형태에서, mRNA는 2-아미노퓨린, 2,6-다이아미노퓨린, 7-데아자-아데닌, 7-데아자-8-아자-아데닌, 7-데아자-2-아미노퓨린, 7-데아자-8-아자-2-아미노퓨린, 7-데아자-2,6-다이아미노퓨린, 1-메틸아데노신, N6-메틸아데노신, N6-(시스-하이드록시아이소펜텐일)아데노신, 2-메틸티오-N6-(시스-하이드록시아이소펜텐일)아데노신, N6-글리신일카바모일아데노신, N6-트레오닐카바모일아데노신, 2-메틸티오-N6-트레오닐 카바모일아데노신, N6,N6-다이메틸아데노신, 7-메틸아데닌, 2-메틸티오-아데닌 및 2-메톡시-아데닌으로 이루어진 군으로부터 선택된 하나 이상의 변형된 뉴클레오사이드를 포함한다.

[0541]

일 실시형태에서, mRNA는 이노신, 1-메틸-이노신, 위오신, 위부토신, 7-데아자-구아노신, 7-데아자-8-아자-구아노신, 6-티오-구아노신, 6-티오-7-데아자-구아노신, 6-티오-7-데아자-8-아자-구아노신, 7-메틸-구아노신, 6-티오-7-메틸-구아노신, 7-메틸이노신, 6-메톡시-구아노신, 1-메틸구아노신, N2-메틸구아노신, N2,N2-다이메틸구아노신, 8-옥소-구아노신, 7-메틸-8-옥소-구아노신, 1-메틸-6-티오-구아노신, N2-메틸-6-티오-구아노신 및 N2,N2-다이메틸-6-티오-구아노신으로 이루어진 군으로부터 선택된 하나 이상의 변형된 뉴클레오사이드를 포함한다.

[0542]

일 실시형태에서, mRNA는 하나 이상의 슈도유리딘, 하나 이상의 5-메틸-사이토신, 및/또는 하나 이상의 5-메틸-사이티딘을 포함한다.

[0543]

일 실시형태에서, mRNA는 하나 이상의 슈도유리딘을 포함한다.

[0544]

일 실시형태에서, mRNA는 하나 이상의 5-메틸-사이티딘을 포함한다.

[0545]

일 실시형태에서, mRNA는 하나 이상의 5-메틸-사이토신을 포함한다.

[0546]

특정 실시형태에서, 본 명세서에서 상정된 mRNA는 엑소뉴클레아제 분해로부터 mRNA를 보호하고, mRNA를 안정화시키며, 번역을 용이하게 하도록 돋는 폴리(A) 꼬리를 포함한다. 특정 실시형태에서, mRNA는 3' 폴리(A) 꼬리 구조를 포함한다.

[0547]

특정 실시형태에서, 폴리(A) 꼬리의 길이는 적어도 약 10, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450 또는 적어도 약 500개 이상의 아데닌 뉴클레오타이드 또는 임의의 개재된 수의 아데닌 뉴클레오타이드이다. 특정 실시형태에서, 폴리(A) 꼬리의 길이는 적어도 약 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 202, 203, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 또는 275개 이상의 아데닌 뉴클레오타이드이다.

[0548]

특정 실시형태에서, 폴리(A) 꼬리의 길이는 약 10 내지 약 500개의 아데닌 뉴클레오타이드, 약 50 내지 약 500개의 아데닌 뉴클레오타이드, 약 100 내지 약 500개의 아데닌 뉴클레오타이드, 약 150 내지 약 500개의 아데닌 뉴클레오타이드, 약 200 내지 약 500개의 아데닌 뉴클레오타이드, 약 250 내지 약 500개의 아데닌 뉴클레오타이드, 약 300 내지 약 500개의 아데닌 뉴클레오타이드, 약 50 내지 약 450개의 아데닌 뉴클레오타이드, 약 50 내지 약 400개의 아데닌 뉴클레오타이드, 약 50 내지 약 350개의 아데닌 뉴클레오타이드, 약 100 내지 약 500개의 아데닌 뉴클레오타이드, 약 100 내지 약 450개의 아데닌 뉴클레오타이드, 약 100 내지 약 400개의 아데닌 뉴클레오타이드, 약 100 내지 약 350개의 아데닌 뉴클레오타이드, 약 100 내지 약 300개의 아데닌 뉴클레오타이드, 약 150 내지 약 500개의 아데닌 뉴클레오타이드, 약 150 내지 약 450개의 아데닌 뉴클레오타이드, 약 150 내지 약 400개의 아데닌 뉴클레오타이드, 약 150 내지 약 350개의 아데닌 뉴클레오타이드, 약 150 내지 약 300개의 아데닌 뉴클레오타이드, 약 200 내지 약 500개의 아데닌 뉴클레오타이드, 약 200 내지 약 450개의 아데닌 뉴클레오타이드, 약 200 내지 약 400개의 아데닌 뉴클레오타이드, 약 200 내지 약 350개의 아데닌 뉴클레오타이드, 약 200 내지 약 300개의 아데닌 뉴클레오타이드, 약 250 내지 약 500개의 아데닌 뉴클레오타이드, 약 250 내지 약 450개의 아데닌 뉴클레오타이드, 약 250 내지 약 400개의 아데닌 뉴클레오타이드, 또는 약 250 내지 약 300개의 아데닌 뉴클레오타이드 또는 아데닌 뉴클레오타이드의 임의의 개재 범위이다.

[0549]

폴리뉴클레오타이드의 배향을 기재하는 용어는 하기를 포함한다: 5'(정상적으로는 유리 인산염기를 갖는 폴리뉴클레오타이드의 말단) 및 3'(정상적으로는 유리 하이드록실(OH) 기를 갖는 폴리뉴클레오타이드의 말단). 폴리뉴클레오타이드 서열은 5'에서 3' 배향 또는 3'에서 5' 배향으로 주석이 달릴 수 있다. DNA 및 mRNA에 대해, 5'에서 3' 가닥은 "센스", "플러스" 또는 "암호화" 가닥으로 표기되는데, 그의 서열이 전-전령(전-mRNA)의 서열과 동일하기 때문이다[DNA에서 티민(T)을 대신하는 RNA에서의 유라실(U)을 제외함]. DNA 및 mRNA에 대해, RNA 중합효소에 의해 전사되는 가닥인 상보성 3'에서 5' 가닥은 "주형", "안티센스", "마이너스" 또는 "비암호화" 가닥으로 표기된다. 본 명세서에서 사용되는 용어 "역 배향"은 3'에서 5' 배향으로 기재된 5'에서 3' 서열 또는 5'에서 3' 배향으로 기재된 3'에서 5' 서열을 지칭한다.

[0550]

용어 "상보적" 및 "상보성"은 염기쌍 규칙과 관련된 폴리뉴클레오타이드(즉, 뉴클레오타이드의 서열)를 지칭한다. 예를 들어, DNA 서열 5' A G T C A T G 3'의 상보성 가닥은 3' T C A G T A C 5'이다. 후자의 서열은 종종 좌측 상에서 5' 말단 및 우측 상에서 3' 말단과 역 상보체로서 기재된다, 5' C A T G A C T 3'. 역 상보체와 동일한 서열은 회문구조 서열인 것으로 언급된다. 상보성은 "부분적"일 수 있는데, 이때 핵산의 염기 중 일부만이 염기쌍 규칙에 따라 매칭된다. 또는, 핵산 사이의 "완전한" 또는 "총" 상보성이 있을 수 있다.

[0551]

본 명세서에서 사용되는 용어 "핵산 카세트" 또는 "발현 카세트"는 RNA, 및 후속적으로 폴리펩타이드를 발현시킬 수 있는 벡터 내의 유전자 서열을 지칭한다. 일 실시형태에서, 핵산 카세트는 관심 대상의 유전자(들), 예를 들어, 관심 대상의 폴리뉴클레오타이드(들)를 함유한다. 다른 실시형태에서, 핵산 카세트는 하나 이상의 발현 제어 서열, 예를 들어, 프로모터, 인핸서, 폴리(A) 서열 및 관심대상의 유전자(들), 예를 들어, 관심대상의 폴리뉴클레오타이드(들)를 함유한다. 벡터는 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개 이상의 핵산 카세트를 포함할 수 있다. 카세트 내 핵산이 RNA로 전사될 수 있고, 필요할 때, 단백질 또는 폴리펩타이드로 번역되며, 형질전환 세포에서 활성에 필요한 적절한 번역후 변형을 겪으며, 적절한 세포내 구획에 대한 표적화 또는 세포외 구획으로의 분비에 의해 생물학적 활성을 위한 적절한 구획으로 전좌되도록, 핵산 카세트는 벡터 내에서 위치적으로 그리고 순차적으로 배향된다. 바람직하게는, 일부 실시형태에서, 카세트는 벡터 및/또는 게놈 내로 용이한 삽입

에 적합한 그의 3' 및 5' 말단을 가지며, 예를 들어, 그는 각각의 말단에서 제한 엔도뉴클레아제 부위를 가진다. 바람직한 실시형태에서, 핵산 카세트는 유전 장애를 치료하거나, 예방하거나 또는 개선시키기 위해 사용되는 치료 유전자의 서열을 함유한다. 카세트는 제거될 수 있고, 단일 단위로서 플라스미드 또는 바이러스 벡터 내로 삽입된다.

[0552] 폴리뉴클레오타이드는 관심 대상의 폴리뉴클레오타이드(들)를 포함한다. 본 명세서에서 사용되는 용어 "관심 대상의 폴리뉴클레오타이드"는 폴리펩타이드 또는 융합 폴리펩타이드를 암호화하는 폴리뉴클레오타이드 또는 본 명세서에서 상정된 저해 폴리뉴클레오타이드의 전사를 위한 주형으로서 작용하는 폴리뉴클레오타이드를 지칭한다.

[0553] 게다가, 유전자 암호 축퇴의 결과로서, 본 명세서에서 상정되는 폴리펩타이드, 또는 이의 단편 또는 변이체를 암호화할 수 있는 다수의 뉴클레오타이드 서열이 있다는 것은 당업자에 의해 인식될 것이다. 이를 폴리뉴클레오타이드 중 일부는 임의의 천연 유전자의 뉴클레오타이드 서열에 대해 최소 상동성을 보유한다. 그럼에도 불구하고, 코돈출현빈도의 차이에 기인하여 변하는 폴리뉴클레오타이드, 예를 들어 인간 및/또는 영장류 코돈 선택을 위해 최적화된 폴리뉴클레오타이드가 특정 실시형태에서 특이적으로 상정된다. 일 실시형태에서, 특정 대립유전자 서열을 포함하는 폴리뉴클레오타이드가 제공된다. 대립유전자는 하나 이상의 돌연변이, 예컨대 뉴클레오타이드의 결실, 첨가 및/또는 치환의 결과로서 변경된 내인성 폴리뉴클레오타이드 서열이다.

[0554] 특정 실시형태에서, 관심 대상의 폴리뉴클레오타이드는 공여자 수선 주형을 포함한다.

[0555] 특정 실시형태에서, 관심 대상의 폴리뉴클레오타이드는 siRNA, miRNA, shRNA, 리보자임 또는 다른 저해 RNA를 포함하지만, 이들로 제한되지 않는 저해 폴리뉴클레오타이드를 포함한다.

[0556] 일 실시형태에서, 저해 RNA를 포함하는 공여자 수선 주형은 본 명세서의 다른 곳에 기재된 바와 같은 하나 이상의 조절 서열, 예를 들어, 강한 구성적 pol III, 예를 들어, 인간 또는 마우스 U6 snRNA 프로모터, 인간 및 마우스 H1 RNA 프로모터, 또는 인간 tRNA-val 프로모터, 또는 강한 구성적 pol II 프로모터를 포함한다.

[0557] 암호 서열 그 자체의 길이와 상관없이, 특정 실시형태에서 상정된 폴리뉴클레오타이드는 그들의 전체 길이가 상당히 다를 수 있도록, 본 명세서의 다른 곳에 개시되거나 또는 당업계에 공지된 바와 같이 다른 DNA 서열, 예컨대 프로모터 및/또는 인핸서, 비번역 영역(UTR), 코작 서열, 폴리아데닐화 신호, 추가적인 제한 효소 부위, 다중 클로닝 부위, 내부 리보솜 유입 부위(IRES), 재조합효소 인식 부위(예를 들어, LoxP, FRT 및 Att 부위), 종결 코돈, 전사 종결 신호, 전사후 반응 요소, 및 자기-절단성 폴리펩타이드를 암호화하는 폴리뉴클레오타이드, 에피토프 태그와 조합될 수 있다. 따라서 특정 실시형태에서 거의 임의의 길이의 폴리뉴클레오타이드 단편이 사용될 수 있으며, 총 길이가 제조의 용이함 및 의도된 재조합 DNA 프로토콜에서의 사용에 의해 바람직하게 제한된다는 것이 상정된다.

[0558] 폴리뉴클레오타이드는 당업계에 공지되고 이용 가능한 임의의 다양한 잘 확립된 기법을 이용하여 제조되고/되거나 조작되고/되거나 발현되고/되거나 전달될 수 있다. 목적으로 하는 폴리펩타이드를 발현시키기 위해, 폴리펩타이드를 암호화하는 뉴클레오타이드 서열은 적절한 벡터 내로 삽입될 수 있다. 목적으로 하는 폴리펩타이드는 또한 폴리펩타이드를 암호화하는 mRNA를 세포 내로 전달함으로써 발현될 수 있다.

[0559] 벡터의 예시적 예는 플라스미드, 자율적 복제 서열 및 전위요소, 예를 들어, 잠자는 미녀(Sleeping Beauty), 피기백(PiggyBac)을 포함하지만, 이들로 제한되지 않는다.

[0560] 벡터의 추가적인 예시적 예는 플라스미드, 파지미드, 코스미드, 인공 염색체, 예컨대 효모 인공 염색체(yeast artificial chromosome: YAC), 박테리아 인공 염색체(bacterial artificial chromosome: BAC) 또는 P1-유래 인공 염색체(PAC), 박테리오파지, 예컨대 람다 파지 또는 M13 파지 및 동물 바이러스를 포함하지만, 이들로 제한되지 않는다.

[0561] 벡터로서 유용한 바이러스의 예시적 예는 레트로바이러스(렌티바이러스를 포함), 아데노바이러스, 아데노-연관 바이러스, 헤르페스바이러스(예를 들어, 단순 포진 바이러스), 폭스바이러스, 바클로바이러스, 유두종바이러스 및 파보바이러스(예를 들어, SV40)을 포함하지만, 이들로 제한되지 않는다.

[0562] 발현 벡터의 예시적 예는 포유류 세포에서 발현을 위한 pClneo 벡터(프로메가(Promega)); 포유류 세포에서 렌티바이러스-매개 유전자 전달 및 발현을 위한 pLenti4/V5-DEST(상표명), pLenti6/V5-DEST(상표명) 및 pLenti6.2/V5-GW/lacZ(인비트로젠(Invitrogen))를 포함하지만, 이들로 제한되지 않는다. 특정 실시형태에서, 본 명세서에 개시된 폴리펩타이드의 암호 서열은 포유류 세포에서 폴리펩타이드의 발현을 위해 이러한 발현 벡터

내로 결찰될 수 있다.

[0563] 특정 실시형태에서, 백터는 에피솜 백터 또는 염색체외로 유지되는 백터이다. 본 명세서에서 사용되는 용어 "에피솜"은 숙주의 염색체 DNA 내로의 통합 없이 그리고 분할 숙주 세포로부터의 점진적 상실 없이 복제할 수 있는 백터를 지칭하는데, 또한 이는 상기 백터가 염색체외로 또는 에피솜으로 복제한다는 것을 의미한다.

[0564] 발현 백터에 존재하는 "발현 제어 서열", "제어 요소" 또는 "조절 서열"은 전사 및 번역을 수행하기 위해 숙주 세포 단백질화 상호작용하는 백터의 해당 비번역 영역(복제기점, 선택 카세트, 프로모터, 인핸서, 번역 개시 신호(샤인 달가노 서열 또는 코작 서열) 인트론, 전사후 조절 요소, 폴리아데닐화 서열, 5' 및 3' 비번역 영역)이다. 이러한 요소는 그들의 강도 및 특이성이 다를 수 있다. 이용되는 백터 시스템 및 숙주에 따라서, 편재성 프로모터 및 유도성 프로모터를 포함하는 다수의 적합한 전사 및 번역 요소가 사용될 수 있다.

[0565] 특정 실시형태에서, 폴리뉴클레오타이드는 발현 백터 및 바이러스 백터를 포함하지만, 이들로 제한되지 않는 백터를 포함한다. 백터는 하나 이상의 외인성, 내인성, 또는 이종성 제어 서열, 예컨대 프로모터 및/또는 인핸서를 포함할 수 있다. "내인성 제어 서열"은 게놈 내 주어진 유전자와 자연적으로 연결된 것이다. "외인성 제어 서열"은 해당 유전자의 전사가 연결된 인핸서/프로모터에 의해 지시되도록 유전자 조작 수단(즉, 분자 생물학 기법)에 의해 유전자에 대한 별별 배치로 위치되는 것이다. "이종성 제어 서열"은 유전자 조작 중인 세포와 상이한 종으로부터 유래된 외인성 서열이다. "합성" 제어 서열은 하나 이상의 내인성 및/또는 외인성 서열의 요소, 및/또는 특정 요법에 대해 최적의 프로모터 및/또는 인핸서 활성을 제공하는 시험관내 또는 인실리코로 결정된 서열을 포함할 수 있다.

[0566] 본 명세서에서 사용되는 용어 "프로모터"는 RNA 중합효소에 결합하는 폴리뉴클레오타이드(DNA 또는 RNA)의 인식 부위를 지칭한다. RNA 중합효소는 프로모터에 작동 가능하게 연결된 폴리뉴클레오타이드를 개시하고 전사시킨다. 특정 실시형태에서, 포유류 세포에서 작동성인 프로모터는 전사가 개시되는 부위로부터 상류의 대략 25 내지 30개 염기에 위치된 AT-풍부 영역 및/또는 전사 시작으로부터 상류의 70 내지 80개 염기로 발현되는 다른 서열, N이 임의의 뉴클레오타이드일 수 있는 CNCAAT 영역을 포함한다.

[0567] 용어 "인핸서"는 향상된 전사를 제공할 수 있는 서열을 함유하는 DNA의 세그먼트를 지칭하며, 일부 예에서 다른 제어 서열에 비해 그들의 배향과 독립적으로 작용할 수 있다. 인핸서는 프로모터 및/또는 다른 인핸서 요소와 협응하여 또는 상가적으로 작용할 수 있다. 용어 "프로모터/인핸서"는 프로모터와 인핸서 기능을 둘 다 제공할 수 있는 서열을 함유하는 DNA의 세그먼트를 지칭한다.

[0568] 용어 "작동 가능하게 연결된"은 별별배치를 지칭하되, 기재된 성분인 그들의 의도된 방식으로 작용하도록 하는 관계이다. 일 실시형태에서, 상기 용어는 핵산 발현 제어 서열(예컨대 프로모터, 및/또는 인핸서)과 제2 폴리뉴클레오타이드 서열, 예를 들어, 관심 대상의 폴리뉴클레오타이드 사이의 작용성 결합을 지칭하되, 발현 제어 서열은 제2 서열에 대응하는 핵산의 전사를 지시한다.

[0569] 본 명세서에서 사용되는 용어 "구성적 발현 제어 서열"은 작동 가능하게 연결된 서열의 전사를 계속해서 또는 지속적으로 허용하는 프로모터, 인핸서, 또는 프로모터/인핸서를 지칭한다. 구성적 발현 제어 서열은 매우 다양한 세포 및 조직 유형에서의 발현을 허용하는 "편재성" 프로모터, 인핸서, 또는 프로모터/인핸서 또는 제한된 각각 다양한 세포 및 조직 유형에서의 발현을 허용하는 "세포 특이적", "세포 유형 특이적", "세포 계통 특이적" 또는 "조직 특이적" 프로모터, 인핸서 또는 프로모터/인핸서일 수 있다.

[0570] 특정 실시형태에서 사용하는 데 적합한 예시적 편재성 발현 제어 서열은 거대세포바이러스(CMV) 급초기 프로모터, 바이러스 유인원 바이러스 40(SV40)(예를 들어, 초기 또는 후기), 몰로니 뮤린 백혈병 바이러스(MoMLV) LTR 프로모터, 라우스 육종 바이러스(RSV) LTR, 단순 포진 바이러스(HSV)(티미딘 키나제) 프로모터, 백시니아 바이러스로부터의 H5, P7.5 및 P11 프로모터, 짧은 신장 인자 1-알파(EF1a-짧은) 프로모터, 긴 신장 인자 1-알파(EF1a-긴) 프로모터, 초기 성장 반응 1(EGR1), 페리틴 H(FerH), 페리틴 L(FerL), 글리세르알데하이드 3-인산염 탈수소효소(GAPDH), 진핵 번역 개시 인자 4A1(EIF4A1), 열 충격 70kDa 단백질 5(HSP5), 열 충격 단백질 90kDa 베타, 구성원 1(HSP90B1), 열 충격 단백질 70kDa(HSP70),  $\beta$ -카네신( $\beta$ -KIN), 인간 ROSA 26 좌위(Irions et al., Nature Biotechnology 25, 1477 - 1482 (2007)), 유비퀴틴 C 프로모터(UBC), 포스포글리세레이트 키나제-1(PGK) 프로모터, 거대세포바이러스 인핸서/짧은  $\beta$ -액틴(CAG) 프로모터,  $\beta$ -액틴 프로모터 및 골수증식성 육종 바이러스 인핸서, 결실된 음성 대조 영역, d1587rev 프라이머-결합 부위 치환된(MND) 프로모터를 포함하지만, 이들로 제한되지 않는다(Challita et al., J Virol. 69(2):748-55 (1995)).

[0571] 특정 실시형태에서, 목적으로 하는 폴리뉴클레오타이드 서열의 세포 유형 특이적, 계통 특이적 또는 조직 특이

적 발현을 달성하기 위해(예를 들어, 세포 유형, 세포 계통 또는 조직의 서브세트만의 또는 발생의 특정 단계 동안 폴리펩타이드를 암호화하는 특정 핵산을 발현시키기 위해) 세포, 세포 유형, 세포 계통 또는 조직 특이적 발현 제어 서열을 사용하는 것이 바람직할 수 있다.

[0572] 본 명세서에서 사용되는, "조건적 발현"은 유도성 발현; 억제성 발현; 특정 생리적, 생물학적 또는 질환 상태 등을 갖는 세포 또는 조직에서의 발현을 포함하지만, 이들로 제한되지 않는 임의의 유형의 조건적 발현을 지칭할 수 있다. 이 정의는 세포 유형 또는 조직 특이적 발현을 제외하는 것으로 의도되지 않는다. 특정 실시형태는 관심 대상의 폴리뉴클레오타이드의 조건적 발현을 제공하며, 예를 들어, 발현은 세포, 조직, 유기체 등에 폴리뉴클레오타이드가 발현되도록 야기하거나 또는 관심 대상의 폴리뉴클레오타이드에 의해 암호화된 폴리뉴클레오타이드의 증가 또는 감소를 야기하는 처리 또는 조건을 실시함으로써 제어된다.

[0573] 유도성 프로모터/시스템의 예시적 예는 스테로이드-유도성 프로모터, 예컨대 글루코코르티코이드 또는 에스트로겐 수용체를 암호화하는 유전자에 대한 프로모터(대응하는 호르몬에 의한 처리에 의한 유도성), 메탈로티오닌 프로모터(다양한 중금속에 의한 처리에 의한 유도성), MX-1 프로모터(인터페론에 의한 유도성), "진스위치(GeneSwitch)" 미페프리스톤-조절 가능 시스템(Sirin et al., 2003, Gene, 323:67), 큐메이트 유도성 유전자 스위치(WO 2002/088346), 테트라사이클린-의존적 조절 시스템 등을 포함하지만, 이들로 제한되지 않는다.

[0574] 조건적 발현은 또한 부위 특이적 DNA 재조합효소를 이용함으로써 달성될 수 있다. 특정 실시형태에 따르면, 폴리뉴클레오타이드는 부위 특이적 재조합효소에 의해 매개된 재조합을 위한 적어도 하나의(전형적으로 2개의) 부위(들)를 포함한다. 본 명세서에서 사용되는 용어 "재조합효소" 또는 "부위 특이적 재조합효소"는 하나 이상의 재조합 부위(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상)를 수반하는 재조합 반응에 연루된 절단 또는 통합적 단백질, 효소, 보조 요소 또는 관련 단백질을 포함하는데, 이는 야생형 단백질(문헌[Landy, Current Opinion in Biotechnology 3:699-707(1993)] 참조) 또는 돌연변이체, 유도체(예를 들어, 재조합 단백질 서열 또는 이의 단편을 함유하는 융합 단백질), 이의 단편 및 변이체일 수 있다. 특정 실시형태에서 사용하기에 적합한 재조합효소의 예시적 예는 Cre, Int, IHF, Xis, Flp, Fis, Hin, Gin,  $\Phi$ C31, Cin, Tn3 위치특이성 재조합촉진효소, TndX, XerC, XerD, TnpX, Hjc, Gin, SpCCE1 및 ParA를 포함하지만, 이들로 제한되지 않는다.

[0575] 폴리뉴클레오타이드는 임의의 매우 다양한 부위 특이적 재조합효소에 대한 하나 이상의 재조합 부위를 포함할 수 있다. 부위 특이적 재조합효소에 대한 표적 부위는 백터, 예를 들어, 레트로바이러스 백터 또는 렌티바이러스 백터의 통합에 필요한 임의의 부위(들)에 추가된다는 것이 이해되어야 한다. 본 명세서에서 사용되는 용어 "재조합 서열", "재조합 부위" 또는 "부위-특이적 재조합 부위"는 재조합효소가 인식하고 결합하는 특정 핵산 서열을 지칭한다.

[0576] 예를 들어, Cre 재조합효소에 대한 하나의 재조합 부위는 8개의 염기쌍 코어 서열에 측접하는 2개의 13개 염기쌍 역위 반복부(재조합효소 결합 부위로서 작용)를 포함하는 34개의 염기쌍 서열인 loxP이다(문헌[Sauer, B., Current Opinion in Biotechnology 5:521-527(1994)]). 다른 예시적인 loxP 부위는 lox511(Hoess et al., 1996; Bethke and Sauer, 1997), lox5171(Lee and Saito, 1998), lox2272 (Lee and Saito, 1998), m2(Langer et al., 2002), lox71(Albert et al., 1995) 및 lox66(Albert et al., 1995)을 포함하지만, 이들로 제한되지 않는다.

[0577] FLP 재조합효소에 대한 적합한 인식 부위는 적합한 인식 부위는 FRT (McLeod, et al., 1996), F1, F2, F3(Schlake and Bode, 1994), F4, F5(Schlake and Bode, 1994), FRT(LE)(Senecoff et al., 1988), FRT(RE)(Senecoff et al., 1988)를 포함하지만, 이들로 제한되지 않는다.

[0578] 인식 서열의 다른 예는 재조합효소  $\lambda$  인테그라제, 예를 들어, phi-c31에 의해 인식되는 attB, attP, attL, 및 attR 서열이다.  $\Phi$ C31 SSR은 이종성 부위 attB(길이가 34 bp)와 attP(길이가 39 bp)사이에서만 재조합을 매개한다(Groth et al., 2000). 각각 박테리아 및 과지 계놈 상의 과지 인테그라제에 대한 부착 부위에 대해 명명되는 attB 및 attP는 둘 다  $\Phi$ C31 동종이량체에 의해 결합될 가능성이 있는 불완전한 역위 반복부를 함유한다(Groth et al., 2000). 산물 부위인 attL 및 attR은 추가적인  $\Phi$ C31-매개 재조합으로 효과적으로 비활성이 되어(Belteki et al., 2003), 반응을 비가역적으로 만든다. 삽입을 측매하기 위해, attB-보유 DNA는 attP가 계놈 attB 부위 내로 삽입되는 것보다 더 용이하게 계놈 attP 부위 내로 삽입한다는 것을 발견하였다(Thyagarajan et al., 2001; Belteki et al., 2003). 따라서, 전형적인 전략은 정해진 좌위 내로 attP-보유 "도킹 부위"의 상동성 재조합에 의해 위치가 정해지며, 이는 이어서, 삽입을 위한 attB-보유 유입 서열과 짹지어진다.

[0579] 일 실시형태에서, 본 명세서에서 상정된 폴리뉴클레오타이드는 재조합효소 인식 부위의 쌍에 의해 측접되는 공

여자 수선 주형 폴리뉴클레오타이드를 포함한다. 특정 실시형태에서, 수선 주형 폴리뉴클레오타이드는 LoxP 부위, FRT 부위 또는 att 부위에 의해 측접된다.

[0580] 특정 실시형태에서, 본 명세서에서 상정된 폴리뉴클레오타이드는 하나 이상의 폴리펩타이드를 암호화하는 관심 대상의 폴리뉴클레오타이드를 포함한다. 특정 실시형태에서, 각각의 복수의 폴리펩타이드의 효율적인 번역을 달성하기 위해, 폴리뉴클레오타이드 서열은 자기-절단성 폴리펩타이드를 암호화하는 하나 이상의 IRES 서열 또는 폴리뉴클레오타이드 서열에 의해 분리될 수 있다.

[0581] 본 명세서에서 사용되는, "내부 리보솜 유입 부위" 또는 "IRES"는 시스트론(단백질 암호화 영역)의 개시 코돈, 예컨대 ATG에 대한 직접 내부 리보솜 유입을 촉진시킴으로써 유전자의 캡-독립적 번역을 야기하는 요소를 지칭한다. 예를 들어, 문헌[Jackson et al., 1990. Trends Biochem Sci 15(12):477-83] 및 Jackson and Kaminski. 1995. RNA 1(10):985-1000] 참조. 당업자에 의해 일반적으로 사용되는 IRES의 예는 미국 특허 제6,692,736호에 기재된 것을 포함한다. 당업계에 공지된 "IRES"의 추가적인 예는 피코르나 바이러스로부터 얻을 수 있는 IRES(Jackson et al., 1990) 및 바이러스 또는 세포 mRNA 공급원, 예를 들어, 면역글로불린 중쇄 결합 단백질 (BiP), 혈관 내피 성장인자(VEGF)(Huez et al. 1998. Mol. Cell. Biol. 18(11):6178-6190), 섬유아세포 성장인자 2(FGF-2), 및 인슐린-유사 성장 인자(IGFII), 번역 개시 인자 eIF4G 및 효모 전사 인자 TFIID 및 HAP4, 노바겐(Novagen)으로부터 상업적으로 입수 가능한 뇌심근염 바이러스(EMCV)(Duke et al., 1992. J. Virol 66(3):1602-9)로부터 얻을 수 있는 IRES 및 VEGF IRES(Huez et al., 1998. Mol Cell Biol 18(11):6178-90)를 포함하지만, 이들로 제한되지 않는다. IRES는 또한 피코르나비리대(Picornaviridae), 디시스트로비리대(Dicistroviridae) 및 플라리비리대(Flaviviridae) 종의 바이러스 계놈에서 그리고 HCV, 프랜드 뮤린 백혈병 바이러스(FrMLV) 및 몰로니 뮤린 백혈병 바이러스(MoMLV)에서 보고되었다.

[0582] 일 실시형태에서, 본 명세서에서 상정된 폴리뉴클레오타이드에서 사용되는 IRES는 EMCV IRES이다.

[0583] 특정 실시형태에서, 폴리뉴클레오타이드는 공통 코작 서열을 갖고 목적으로 하는 폴리펩타이드를 암호화하는 폴리뉴클레오타이드를 포함한다. 본 명세서에서 사용되는 용어 "코작 서열"은 리보솜에 대한 mRNA의 초기 결합을 크게 용이하게 하고 번역을 증가시키는 짧은 뉴클레오타이드 서열을 지칭한다. 공통 코작 서열은 (GCC)RCCATGG(서열번호 75)이며, 여기서 R은 퓨린(A 또는 G)이다(Kozak, 1986. Cell. 44(2):283-92, 및 Kozak, 1987. Nucleic Acids Res. 15(20):8125-48).

[0584] 이종성 핵산 전사체의 효율적인 종결 및 폴리아데닐화를 지시하는 요소는 이종성 유전자 발현을 증가시킨다. 전사 종결 신호는 일반적으로 폴리아데닐화 신호의 하류에서 발견된다. 특정 실시형태에서, 백터는 발현될 폴리펩타이드를 암호화하는 폴리뉴클레오타이드의 폴리아데닐화 서열 3'을 포함한다. 본 명세서에서 사용되는 용어 "폴리A 부위", "폴리A 서열"은 RNA 중합효소 II에 의한 초기 RNA 전사체의 종결과 폴리아데닐화를 둘 다 지시하는 DNA 서열을 의미한다. 폴리아데닐화 서열은 암호 서열의 3' 말단에 대한 폴리A 꼬리의 첨가에 의해 mRNA 안정성을 촉진시킬 수 있고, 따라서, 증가된 번역 효율에 기여한다. 절단 및 폴리아데닐화는 RNA에서 폴리(A) 서열에 의해 지시된다. 포유류 프레-mRNA에 대한 코어 폴리(A) 서열은 절단-폴리아데닐화 부위에 측접하는 2개의 인식 요소를 가진다. 전형적으로, 대부분의 비변이체 AAUAAA 6량체는 U 또는 GU 잔기가 풍부한 더 가변적인 요소의 상류에 20 내지 50개의 뉴클레오타이드가 놓인다. 초기 전사체의 절단은 이들 두 요소 사이에서 생기며, 5' 절단 산물에 대한 250개까지의 아데노산의 첨가에 결합된다. 특정 실시형태에서, 코어 폴리(A) 서열은 이상적인 폴리A 서열(예를 들어, AATAAA, ATTAAA, AGTAAA)이다. 특정 실시형태에서, 폴리(A) 서열은 SV40 폴리A 서열, 소 성장 호르몬 폴리A 서열(BGHpA), 토끼  $\beta$ -글로빈 폴리A 서열(r  $\beta$  gpa), 이의 변이체 또는 당업계에 공지된 다른 적합한 이종성 또는 내인성 폴리A 서열이다.

[0585] 일부 실시형태에서, 폴리뉴클레오타이드 또는 폴리뉴클레오타이드를 보유하는 세포는 직접적인 독성 및/또는 비제어 증식의 위험을 감소시키기 위해 유도성 자살 유전자를 포함하는 자살 유전자를 이용한다. 구체적 실시형태에서, 자살 유전자는 폴리뉴클레오타이드 또는 세포를 보유하는 숙주에 대해 면역원성이 아니다. 사용될 수 있는 자살 유전자의 특정 예는 카스파제-9 또는 카스파제-8 또는 사이토신 데아미나제이다. 카스파제-9는 이량체화(CID)의 특정 화학적 유도자를 이용하여 활성화될 수 있다.

[0586] 특정 실시형태에서, 폴리뉴클레오타이드는 본 명세서에서 상정된 유전자 변형된 세포가 생체내 음성 선택을 허용하게 하는 유전자 세그먼트를 포함한다. "음성 선택"은 개체의 생체내 조건의 변화와 결과로서 제거될 수 있는 주입 세포를 지칭한다. 음성 선택 가능한 표현형은 투여되는 제제, 예를 들어, 화합물에 대한 민감성을 부여하는 유전자의 삽입으로부터 초래될 수 있다. 음성 선택 유전자는 당업계에 공지되어 있으며, 간사이클로비어 민감성을 부여하는 단순 포진 바이러스 I형 티미딘 키나제(HSV-I TK) 유전자; 세포의 하이포크산틴 포스포리보실

트랜스페라제(HPRT) 유전자, 세포의 아데닌 포스포리보실트랜스페라제(APRT) 유전자 및 박테리아 사이토신 탈아미나제를 포함하지만, 이들로 제한되지 않는다.

[0587] 일부 실시형태에서, 유전자 변형된 세포는 시험관내에서 음성 선택 가능한 표현형의 세포 선택을 가능하게 하는 양성 마커를 더 포함하는 폴리뉴클레오타이드를 포함한다. 양성 선택 가능 마커는 숙주 세포 내로 도입 시, 유전자를 운반하는 세포의 양성 선택을 허용하는 우성 표현형을 발현시키는 유전자일 수 있다. 이 유형의 유전자는 공지되어 있으며, 하이그로마이신 B에 대한 내성을 부여하는 하이그로마이신-B 포스포트랜스페라제 유전자(hph), 항생제 G418에 대한 내성을 암호화하는 Tn5로부터의 아미노 글리코사이드 포스포트랜스페라제 유전자(neo 또는 aph), 다이하이드로엽산 환원효소(DHFR) 유전자, 아데노신 탈아미나제 유전자(ADA) 및 다제내성(MDR) 유전자를 포함하지만, 이들로 제한되지 않는다.

[0588] 일 실시형태에서, 양성 선택 가능 마커 및 음성 선택 가능 요소는 음성 선택 가능 요소의 상실이 또한 양성 선택 가능 마커의 상실을 수반하도록 연결된다. 특정 실시형태에서, 양성 및 음성 선택 가능 마커는 하나의 상실이 다른 것의 상실을 필수적으로 야기하도록 융합된다. 발현 산물로서 상기 기재한 목적으로 하는 양성 및 음성 선택 특징을 부여하는 폴리펩타이드를 수득하는 융합된 폴리뉴클레오타이드 예는 하이그로마이신 포스포트랜스페라제 티미딘 키나제 융합 유전자(HyTK)이다. 이 유전자의 발현은 시험관내 양성 선택을 위해 하이그로마이신 B 내성, 및 생체내 음성 선택을 위해 간사이클로비어 민감성을 부여하는 폴리펩타이드를 수득한다. 또한 우성 양성 선택 가능 마커를 음성 선택 가능 마커와 융합시키는 것으로부터 유래된 2작용성 선택 가능 유전자의 용도를 기재하는 S. D. Lupton에 의한 PCT US91/08442 및 PCT/US94/05601의 공보를 참조한다.

[0589] 바람직한 양성 선택 가능 마커는 hph, nco 및 gpt로 이루어진 군으로부터 선택되는 유전자로부터 유래되고, 바람직한 음성 선택 가능 마커는 사이토신 탈아미나제, HSV-I TK, VZV TK, HPRT, APRT 및 gpt로 이루어진 군으로부터 선택되는 유전자로부터 유래된다. 특정 실시형태에서 상정된 예시적인 2작용성 선택 가능한 융합 유전자는 유전자를 포함하지만, 이것으로 제한되지 않되, 양성 선택 가능 마커는 hph 또는 neo로부터 유래되고, 음성 선택 가능 마커는 사이토신 탈아미나제 또는 TK 유전자 또는 선택 가능한 마커로부터 유래된다.

[0590] 특정 실시형태에서, 하나 이상의 뉴클레아제 변이체, 메가TAL, 말단-가공 효소 또는 융합 폴리펩타이드를 암호화하는 폴리뉴클레오타이드는 조혈세포, 예를 들어, T 세포 내로, 비바이러스 및 바이러스적 방법에 의해 도입될 수 있다. 특정 실시형태에서, 뉴클레아제 및/또는 공여자 수선 주형을 암호화하는 하나 이상의 폴리뉴클레오타이드의 전달은 동일한 방법에 의해 또는 상이한 방법에 의해, 그리고/또는 동일한 벡터에 의해 또는 상이한 벡터에 의해 제공될 수 있다.

[0591] 용어 "벡터"는 다른 핵산 분자를 전달하거나 또는 수송할 수 있는 핵산 분자를 지칭하기 위해 본 명세서에서 사용된다. 전달된 핵산은 일반적으로 벡터 핵산 분자에 연결되고, 예를 들어, 이에 삽입된다. 벡터는 세포 내 자율 복제를 자시하는 서열을 포함할 수 있거나, 또는 숙주 세포 DNA 내로의 통합을 허용하기에 충분한 서열을 포함할 수 있다. 특정 실시형태에서, 비바이러스 벡터는 본 명세서에서 상정된 하나 이상의 폴리뉴클레오타이드를 T 세포에 전달하는 데 사용된다.

[0592] 비바이러스 벡터의 예시적 예는 플라스미드(예를 들어, DNA 플라스미드 또는 RNA 플라스미드), 트랜스포존, 코스미드 및 박테리아 인공 염색체를 포함하지만, 이들로 제한되지 않는다.

[0593] 특정 실시형태에서 상정된 폴리뉴클레오타이드의 비바이러스 전달의 예시적 방법은 전기천공법, 초음파천공법, 리포펙션, 미량주사법, 유전자총법, 바이로좀, 리포좀, 면역리포좀, 나노입자, 다양이온 또는 지질:핵산 접합체, 네이키드 DNA, 인공 비리온, DEAE-덱스트란-매개 전달, 유전자 총 및 열-충격을 포함하지만, 이들로 제한되지 않는다.

[0594] 특정 실시형태에 상정된 특정 실시형태에서 사용하기에 적합한 폴리뉴클레오타이드 전달 시스템의 예시적 예는 아막사 바이오시스템즈(Amaxa Biosystems), 맥스사이트 인코포레이티드(Maxcyte, Inc.), BTX 몰래큘러 딜리버리 시스템즈(BTX Molecular Delivery Systems) 및 코페르니쿠스 쎄라퓨틱스 인코포레이티드(Copernicus Therapeutics Inc.)에 의해 제공된 것을 포함하지만, 이들로 제한되지 않는다. 리포펙션 시약은 상업적으로 시판된다(예를 들어, 트랜스펙탐(Transfectam)(상표명) 및 리포펙틴(Lipofectin)(상표명)). 폴리뉴클레오타이드의 효율적인 수용체-인식 리포펙션에 적합한 양이온성 및 중성 지질은 문헌에 기재되어 있다. 예를 들어, 문헌[Liu et al. (2003) Gene Therapy. 10:180-187; 및 Balazs et al. (2011) Journal of Drug Delivery. 2011:1-12] 참조. 항체-표적화된, 박테리아 유래된, 살아있지 않은 나노세포-기반 전달이 또한 특정 실시형태에서 상정된다.

- [0595] 특정 실시형태에서 상정된 폴리뉴클레오타이드를 포함하는 바이러스 벡터는 개개 환자에 대한 투여에 의해, 전형적으로 이하에 기재하는 바와 같이 전신 투여(예를 들어, 정맥내, 복강내, 근육내, 피하 또는 두개내 주입) 또는 국소 적용에 의해 생체내로 전달될 수 있다. 대안적으로, 벡터는 생체외 세포, 예컨대 개개 환자로부터 외식된 세포(예를 들어, 동원된 말초 혈액, 텁프구, 골수 흡입물, 조직 생검 등) 또는 보편적 공여자 조혈 세포 세포에 전달될 수 있고, 다음에 환자 내로 세포의 재이식이 이어진다.
- [0596] 일 실시형태에서, 뉴클레아제 변이체 및/또는 공여자 수선 주형을 포함하는 바이러스 벡터는 생체내 세포의 형질도입을 위해 유기체에 직접적으로 투여된다. 대안적으로, 네이키드 DNA가 투여될 수 있다. 투여는 주사, 주입, 국소 적용 및 전기천공법을 포함하지만, 이들로 제한되지 않는 혈액 또는 조직 세포와의 궁극적 접촉으로 분자를 도입하기 위해 정상적으로 사용되는 임의의 경로에 의한다. 특정 조성물을 투여하기 위해 하나 초과의 경로가 사용될 수 있지만, 이러한 혁신을 투여하는 적합한 방법은 당업계에서 이용되고 잘 공지되어 있으며, 특정 경로는 종종 다른 경로보다 더 즉각적이고 더 효과적인 반응을 제공할 수 있다.
- [0597] 본 명세서에서 상정된 특정 실시형태에서 사용하기에 적합한 바이러스 벡터 시스템의 예시적 예는 아데노-연관 바이러스(AAV), 레트로바이러스, 단순 포진 바이러스, 아데노바이러스 및 백시니아 바이러스 벡터를 포함하지만, 이들로 제한되지 않는다.
- [0598] 다양한 실시형태에서, 뉴클레아제 변이체 및/또는 공여자 수선 주형을 암호화하는 하나 이상의 폴리뉴클레오타이드는 하나 이상의 폴리뉴클레오타이드를 포함하는 재조합 아데노-연관 바이러스(rAAV)를 이용하여 세포를 형질도입함으로써 조혈세포, 예를 들어, T 세포 내로 도입된다.
- [0599] AAV는 작은(대략 26nm) 복제-결합, 주로 에피솜, 비외피 바이러스이다. AAV는 분할세포와 비분할 세포를 둘 다 감염시킬 수 있고, 숙주 세포 내로 그의 계놈을 혼입시킬 수 있다. 재조합 AAV(rAAV)는 전형적으로 최소, 이식 유전자 및 그의 조절 서열, 및 5' 및 3' AAV 역위 말단 반복부(ITR)로 구성된다. ITR 서열은 길이가 약 145bp이다. 특정 실시형태에서, rAAV는 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9 및 AAV10으로부터 단리된 ITR 및 캡시드 서열을 포함한다.
- [0600] 일부 실시형태에서, 키메라 rAAV가 사용되며, ITR 서열은 하나의 AAV 혈청형으로부터 단리되고, 캡시드 서열은 상이한 AAV 혈청형으로부터 단리된다. 예를 들어, AAV2로부터 유래된 ITR 서열 서열 및 AAV6으로부터 유래된 캡시드 서열을 갖는 rAAV는 AAV2/AAV6로서 지칭된다. 특정 실시형태에서, rAAV 벡터는 AAV2로부터의 ITR, 및 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9 또는 AAV10 중 임의의 하나로부터의 캡시드 단백질을 포함할 수 있다. 바람직한 실시형태에서, rAAV는 AAV2로부터 유래된 ITR 서열 및 AAV6으로부터 유래된 캡시드 서열을 포함한다. 바람직한 실시형태에서, rAAV는 AAV2로부터 유래된 ITR 서열 및 AAV2으로부터 유래된 캡시드 서열을 포함한다.
- [0601] 일부 실시형태에서, 조작 및 선택 방법은 관심 대상의 세포를 형질도입할 가능성을 더 만들기 위해 AAV 캡시드에 적용될 수 있다.
- [0602] rAAV 벡터의 구성, 이의 생산 및 정제는, 예를 들어, 미국 특허 제9,169,494호; 제9,169,492호; 제9,012,224호; 제8,889,641호; 제8,809,058호; 및 제8,784,799호에 개시되어 있으며, 이들 각각은 본 명세서에 그의 전문이 참고로 포함된다.
- [0603] 뉴클레아제 변이체 및/또는 공여자 수선 주형을 암호화하는 다양한 실시형태에서, 하나 이상의 폴리뉴클레오타이드는 레트로바이러스, 예를 들어, 하나 이상의 폴리뉴클레오타이드를 포함하는 렌티바이러스를 이용하여 세포를 형질도입함으로써 조혈세포 내로 도입된다.
- [0604] 본 명세서에서 사용되는 용어 "레트로바이러스"는 그의 계놈 RNA를 선형 이중 가닥 DNA 복제물로 역전사시키고, 후속적으로 그의 계놈 DNA를 숙주 계놈 내로 공유적으로 통합시키는 RNA 바이러스를 지칭한다. 특정 실시형태에서 사용하기에 적합한 예시적인 레트로바이러스는 몰로니 뮤린 백혈병 바이러스(M-MuLV), 몰로니 뮤린 육종 바이러스(MoMSV), 하비 뮤린 육종 바이러스(HaMuSV), 뮤린 유방 종양 바이러스(MuMTV), 갑본 유인원 백혈병 바이러스(GaLV), 고양이 백혈병 바이러스(FLV), 스푸마바이러스, 프렌드 뮤린 백혈병 바이러스, 뮤린 줄기 세포 바이러스(MSCV) 및 라우스 육종 바이러스(RSV)) 및 렌티바이러스를 포함하지만, 이들로 제한되지 않는다.
- [0605] 본 명세서에서 사용되는 용어 "렌티바이러스"는 복잡한 레트로바이러스의 군(또는 속)을 지칭한다. 예시적인 렌티바이러스는 HIV(인간 면역결핍 바이러스; HIV 1형, 및 HIV 2를 포함); 비스나-메디 바이러스(VMV) 바이러스; 염소 관절염-뇌염바이러스(CAEV); 마 전염성 빈혈 바이러스(EIAV); 고양이 면역결핍 바이러스(FIV); 소 면역 결

펩 바이러스(BIV); 및 시미안 면역결핍 바이러스(SIV)를 포함하지만, 이들로 제한되지 않는다. 일 실시형태에서, HIV 기반 벡터 골격(즉, HIV 시스-작용성 서열 요소)이 바람직하다.

[0606] 다양한 실시형태에서, 본 명세서에서 상정된 렌티바이러스 벡터는 하나 이상의 LTR, 및 다음의 부속 요소: cPPT/FLAP, Psi(Ψ) 패키징 신호, 유출 요소, 폴리(A) 서열 중 하나 이상 또는 모두를 포함하고, 선택적으로 본 명세서의 다른 곳에 논의된 바와 같이 WPRE 또는 HPRE, 절연체 요소, 선택 가능한 마커, 및 세포 자살 유전자를 포함할 수 있다.

[0607] 특정 실시형태에서, 본 명세서에서 상정된 렌티바이러스 벡터는 통합적 또는 비통합적 또는 통합 결합 렌티바이러스일 수 있다. 본 명세서에서 사용되는 용어 "통합 결합 렌티바이러스" 또는 "IDLV"는 숙주 세포의 게놈 내로 바이러스 게놈을 통합하는 능력을 결여하는 인테그라제를 갖는 렌티바이러스를 지칭한다. 통합-부적격 바이러스 벡터는 본 명세서에 전문이 참고로 포함된 특허 출원 WO 2006/010834에 기재되어 있다.

[0608] 인테그라제 활성을 감소시키는 데 적합한 HIV-1 pol 유전자에서의 예시적인 돌연변이는 H12N, H12C, H16C, H16V, S81 R, D41A, K42A, H51A, Q53C, D55V, D64E, D64V, E69A, K71A, E85A, E87A, D116N, D116I, D116A, N120G, N120I, N120E, E152G, E152A, D35E, K156E, K156A, E157A, K159E, K159A, K160A, R166A, D167A, E170A, H171A, K173A, K186Q, K186T, K188T, E198A, R199c, R199T, R199A, D202A, K211A, Q214L, Q216L, Q221L, W235F, W235E, K236S, K236A, K246A, G247W, D253A, R262A, R263A 및 K264H를 포함하지만, 이들로 제한되지 않는다.

[0609] 일 실시형태에서, HIV-1 인테그라제 결여 pol 유자는 D64V, D116I, D116A, E152G 또는 E152A 돌연변이; D64V, D116I 및 E152G 돌연변이; 또는 D64V, D116A 및 E152A 돌연변이를 포함한다.

[0610] 일 실시형태에서, HIV-1 인테그라제 결여 pol 유자는 D64V 돌연변이를 포함한다.

[0611] 용어 "긴 말단 반복부(LTR)"는 그들의 천연 서열과 관련하여, 반복을 지시하고 U3, R 및 U5 영역을 함유하는 레트로바이러스 DNA의 말단에 위치된 염기쌍의 도메인을 지칭한다.

[0612] 본 명세서에서 사용되는 용어 "FLAP 요소" 또는 "cPPT/FLAP"는 서열이 레트로바이러스, 예를 들어, HIV-1 또는 HIV-2의 중심 폴리펩타이드 및 중심 종결 서열(cPPT 및 CTS)을 포함하는 핵산을 지칭한다. 적합한 FLAP 요소는 미국 특허 제6,682,907호에 그리고 문헌[Zennou, et al., 2000, Cell, 101:173]에 기재되어 있다. 다른 실시형태에서, 렌티바이러스 벡터는 cPPT 및/또는 CTS 요소에서 하나 이상의 돌연변이를 갖는 FLAP 요소를 함유한다. 또 다른 실시형태에서, 렌티바이러스 벡터는 cPPT 또는 CTS 요소 중 하나를 포함한다. 또 다른 실시형태에서, 렌티바이러스 벡터는 cPPT 또는 CTS 요소를 포함하지 않는다.

[0613] 본 명세서에서 사용되는 용어 "패키징 신호" 또는 "패키징 서열"은 바이러스 캡시드 또는 입자 내로의 바이러스 RNA의 삽입에 필요한 레트로바이러스 게놈 내에 위치된 싸이[Ψ] 서열을 지칭한다, 예를 들어, 문헌[Clever et al., 1995. J. of Virology, Vol. 69, No. 4; pp. 2101-2109] 참조.

[0614] 용어 "유출 요소"는 세포의 핵으로부터 세포질까지 RNA 전사체의 수송을 조절하는 시스-작용성 전사후 조절 요소를 지칭한다. RNA 유출 요소의 예는 인간 면역결핍 바이러스(HIV) rev 반응 요소(RRE)(예를 들어, 문헌[Cullen et al., 1991. J. Virol. 65: 1053; 및 Cullen et al., 1991. Cell 58: 423] 참조), 및 B형 간염 바이러스 전사후 조절 요소(HPRE)를 포함하지만, 이들로 제한되지 않는다.

[0615] 특정 실시형태에서, 바이러스 벡터 내 이종성 서열의 발현은 전사후 조절 요소, 효율적인 폴리아데닐화 부위, 및 선택적으로, 전사 종결 신호를 벡터 내로 혼입함으로써 증가된다. 다양한 전사후 조절 요소, 예를 들어, 마우스 간염 바이러스 전사후 조절 요소(WPRE; Zufferey et al., 1999, J. Virol., 73:2886); B형 간염 바이러스에 존재하는 전사후 조절 요소(HPRE)(Huang et al., Mol. Cell. Biol., 5:3864); 등(Liu et al., 1995, Genes Dev., 9:1766)은 단백질에서 이종성 핵산의 발현을 증가시킬 수 있다.

[0616] 렌티바이러스 벡터는 바람직하게는 LTR을 변형시키는 결과로서 몇몇 안전한 향상을 함유한다. "자기-비활성화"(SIN) 벡터는 복제-결합 벡터를 지칭하며, 예를 들어, 이때 U3 영역으로서 알려진 우측(3') LTR 인핸서-프로모터 영역은 바이러스 복제의 제1 라운드 이후의 바이러스 전사를 방지하기 위해 (예를 들어, 결절 또는 치환에 의해) 변형되었다. 추가적인 안전성 향상은 바이러스 입자의 생산 동안 바이러스 게놈의 전사를 유도하기 위해 5' LTR의 U3 영역을 이종성 프로모터로 대체함으로써 제공된다. 사용될 수 있는 이종성 프로모터의 예는, 예를 들어, 바이러스 유인원 바이러스 40(SV40)(예를 들어, 초기 또는 후기), 거대세포바이러스(CMV)(예를 들어, 급 초기), 몰로니 뮤런 백혈병 바이러스(MoMLV), 라우스 육종 바이러스(RSV) 및 단순 포진 바이러스(HSV)(티미딘

키나제) 프로모터를 포함한다.

- [0617] 본 명세서에서 사용되는 용어 "위형" 또는 "위형별(pseudotyping)"은 바이러스 외피 단백질이 바람직한 특징을 갖는 다른 바이러스의 외피 단백질로 치환된 바이러스를 지칭한다. 예를 들어, HIV 외피 단백질(env 유전자에 의해 암호화)이 정상적으로는 바이러스를 CD4+ 세포로 표적화하기 때문에, HIV가 매우 다양한 세포를 감염시키도록 허용하는 수포성 구내염 바이러스 G-단백질(VSV-G) 외피 단백질에 의해 HIV는 위형화될 수 있다.
- [0618] 특정 실시형태에서, 렌티바이러스 벡터는 공지된 방법에 따라 생산된다. 예를 들어, 문헌[Kutner et al., BMC Biotechnol. 2009;9:10. doi: 10.1186/1472-6750-9-10; Kutner et al. Nat. Protoc. 2009;4(4):495-505. doi: 10.1038/nprot.2009.22] 참조.
- [0619] 본 명세서에서 상정된 특정 구체적 실시형태에 따르면, 대부분의 또는 모든 바이러스 벡터 골격 서열은 렌티바이러스, 예를 들어, HIV-1로부터 유래된다. 그러나, 레트로바이러스 및/또는 렌티바이러스 서열의 다수의 상이한 공급원이 사용되거나 또는 조합될 수 있고, 특정 렌티바이러스 서열의 수많은 치환 및 변경은 본 명세서에 기재된 기능을 수행하는 전달 벡터의 능력을 손상시키는 일 없이 수용될 수 있다는 것이 이해되어야 한다. 게다가, 다양한 렌티바이러스 벡터는 당업계에 공지되어 있으며(문헌[Naldini et al., (1996a, 1996b 및 1998); Zufferey et al., (1997); Dull et al., 1998, 미국 특허 제6,013,516호; 및 제5,994,136호 참조], 이들 중 다수는 본 명세서에서 상정된 바이러스 벡터 또는 전달 플라스미드를 생산하기에 적합할 수 있다.
- [0620] 뉴클레아제 변이체 및/또는 공여자 수선 주형을 암호화하는 다양한 실시형태에서, 하나 이상의 폴리뉴클레오타이드는 하나 이상의 폴리뉴클레오타이드를 포함하는 아데노바이러스를 이용하여 세포를 형질도입함으로써 조혈세포 내로 도입된다.
- [0621] 아데노바이러스 기반 벡터는 다수의 세포 유형에서 매우 높은 형질도입 효율 능력이 있으며, 세포 분할을 필요로 하지 않는다. 이러한 벡터에 의해, 고역가 및 고수준의 발현이 얻어졌다. 이 벡터는 상대적으로 단순한 시스템에서 다양으로 생산될 수 있다. 이식유전자가 Ad E1a, E1b 및/또는 E3 유전자를 대체하고; 후속적으로 복제결합 벡터가 도중에 결실된 유전자 기능을 공급하는 인간 293 세포에서 증식되도록 대부분의 아데노바이러스 벡터는 조작된다. Ad 벡터는 비분할, 분화 세포, 예컨대 간, 신장 및 근육에서 발견되는 것을 포함하는 생체내 다중 유형의 조직을 형질도입할 수 있다. 통상적인 Ad 벡터는 큰 운반 능력을 가진다.
- [0622] 복제 결여인 현재의 아데노바이러스 벡터의 생성 및 증식은 293으로 표기되는 독특한 헬퍼 세포주를 이용할 수 있는데, 이는 Ad5 DNA 단편에 의해 인간 배아 신장 세포로부터 형질전환되고, E1 단백질을 구성적으로 발현시킨다(Graham et al., 1977). E3 영역은 아데노바이러스 계보으로부터 필요 없기 때문에(Jones & Shenk, 1978), 293 세포의 도움으로 본 아데노바이러스 벡터는 E1, D3 중 하나 또는 영역 둘 다에서 외래 DNA를 운반한다(Graham & Prevec, 1991). 아데노바이러스 벡터는 진핵 유전자 발현(Levrero et al., 1991; Gomez-Foix et al., 1992) 및 백신 개발에서 사용되었다(Grunhaus & Horwitz, 1992; Graham & Prevec, 1992). 상이한 조직에 재조합 아데노바이러스를 투여하는 연구는 기관 점적(Rosenfeld et al., 1991; Rosenfeld et al., 1992), 근육주사(Ragot et al., 1993), 말초 정맥내 주사(Herz & Gerard, 1993) 및 뇌 내로의 정위 접종(Le Gal La Salle et al., 1993)을 포함한다. 임상 시험에서 Ad 벡터사용의 예는 근육내 주사에 의한 항종양 면역화를 위한 폴리뉴클레오타이드 요법을 수반하였다(Sterman et al., Hum. Gene Ther. 7:1083-9 (1998)).
- [0623] 뉴클레아제 변이체 및/또는 공여자 수선 주형을 암호화하는 다양한 실시형태에서, 하나 이상의 폴리뉴클레오타이드는 하나 이상의 폴리뉴클레오타이드를 포함하는 단순 포진 바이러스, 예를 들어, HSV-1, HSV-2를 이용하여 세포를 형질도입함으로써 조혈세포 내로 도입된다.
- [0624] 성숙 HSV 비리온은 152kb인 선형 이중 가닥 DNA 분자로 이루어진 바이러스 계보와 함께 외피 이십면체 캡시드로 이루어진다. 일 실시형태에서, HSV 기반 바이러스 벡터는 하나 이상의 필수 또는 비필수 HSV 유전자에서 결여된다. 일 실시형태에서, HSV 기반 바이러스 벡터는 복제 결여이다. 대부분의 복제 결여 HSV 벡터는 복제를 방지하기 위해 하나 이상의 급초기, 초기 또는 후기 HSV 유전자를 제거하는 결실을 함유한다. 예를 들어, HSV 벡터는 ICP4, ICP22, ICP27, ICP47 및 이들의 조합물로 이루어진 군으로부터 선택된 급초기 유전자가 결여될 수 있다. HSV 벡터의 이점은 장기간 DNA 발현을 초래할 수 있는 잠복기로 유입되는 그의 능력 및 25kB까지의 외인성 DNA 삽입물을 수용할 수 있는 그의 거대 바이러스 DNA 계보이다. HSV-기반 벡터는, 예를 들어, 미국 특허 제5,837,532호, 제5,846,782호 및 제5,804,413호 및 국제 특허 출원 WO 91/02788, WO 96/04394, WO 98/15637 및 WO 99/06583에 기재되어 있고, 이들 각각은 그의 전문이 본 명세서에 참고로 포함된다.
- [0625] H. 계보 편집 세포

- [0626] 특정 실시형태에서 상정된 방법에 의해 제조된 계놈 편집된 세포는 CBLB 유전자에서의 하나 이상의 유전자 편집을 포함하고, 암, GVHD, 감염성 질환, 자가면역 질환, 면역결핍 또는 이와 관련된 병태의 적어도 하나의 증상의 예방, 치료 또는 개선을 위한 개선된 세포 기반 치료를 제공한다. 임의의 특정 이론에 의해 구속되는 일 없이, 본 명세서에서 상정된 조성물 및 방법은 부분적으로, 치료제를 면역억제 신호 및 고갈에 대해 더 지속적이고 더 내성으로 만듦으로써 적응 세포 요법의 효능을 증가시킨다.
- [0627] 특정 실시형태에서 상정된 계놈 편집된 세포는 자가/자율("자기") 또는 비-자가("비-자기", 예를 들어, 동종이계, 동계 또는 이종성)일 수 있다. 본 명세서에서 사용되는 "자가"는 동일한 대상체로부터의 세포를 지칭한다. 본 명세서에 사용되는 "동종이계"는 비교하는 세포와 유전적으로 상이한 동일한 종의 세포를 지칭한다. 본 명세서에서 사용되는 "동계"는 비교하는 세포와 유전적으로 동일한 상이한 대상체의 세포를 지칭한다. 본 명세서에서 사용되는 "이종성"은 비교하는 세포와 상이한 종의 세포를 지칭한다. 바람직한 실시형태에서, 세포는 포유류 대상체로부터 얻는다. 더 바람직한 실시형태에서, 세포는 영장류 대상체, 선택적으로는 비인간 영장류로부터 얻는다. 가장 바람직한 실시형태에서, 세포는 인간 대상체로부터 얻는다.
- [0628] "단리된 세포"는 비-천연 유래 세포, 예를 들어, 생체내 조직 또는 인간으로부터 얻거나 또는 세포의 기질이 실질적으로 없는 천연에서 존재하지 않는 세포, 변형된 세포, 조작된 세포 등을 지칭한다.
- [0629] 본 명세서에서 사용되는 용어 "세포의 집단"은 본 명세서의 다른 곳에 기재된 바와 같이 다수의 그리고/또는 상동성 또는 이종성 세포 유형의 조합을 구성할 수 있는 복수의 세포를 지칭한다. 예를 들어, T 세포의 형질도입을 위해, 세포의 집단은 말초 혈액으로부터 단리되거나 또는 얻어질 수 있다. 세포의 집단은 편집될 표적 세포 유형의 약 10%, 약 20%, 약 30%, 약 40%, 약 50%, 약 60%, 약 70%, 약 80%, 약 90%, 또는 약 100%를 포함할 수 있다. 특정 실시형태에서, T 세포는 당업계에 공지된 방법을 이용하여 이종성 세포 집단으로부터 단리되거나 또는 정제될 수 있다.
- [0630] 계놈이 본 명세서에서 상정된 조성물 및 방법을 이용하여 편집될 수 있는 세포 유형의 예시적 예는 세포주, 1차 세포, 줄기 세포, 조상 세포, 및 분화된 세포, 및 이들의 혼합물을 포함하지만, 이들로 제한되지 않는다.
- [0631] 바람직한 실시형태에서, 계놈 편집 조성물 및 방법은 조혈세포, 더 바람직하게는 면역 세포, 및 더욱더 바람직하게는 T 세포를 편집하기 위해 사용된다.
- [0632] 용어 "T 세포" 또는 "T 림프구"는 당업계에 인식되어 있고, 흥선세포, 면역 효과기 세포, 조절 T 세포, 미경험 T 림프구, 미성숙 T 림프구, 성숙 T 림프구, 휴지 T 림프구 또는 활성화된 T 림프구를 포함하는 것으로 의도된다. T 세포는 T 헬퍼(Th) 세포, 예를 들어 T 헬퍼 1(Th1) 또는 T 헬퍼 2(Th2) 세포일 수 있다. T 세포는 헬퍼 T 세포(HTL; CD4+ T 세포), 세포독성 T 세포(CTL; CD8+ T 세포), 종양 침윤성 세포독성 T 세포(TIL; CD8+ T 세포), CD4+CD8+ T 세포, CD4-CD8- T 세포, 또는 T 세포의 임의의 다른 서브세트일 수 있다. 일 실시형태에서, T 세포는 면역 효과기 세포이다. 일 실시형태에서, T 세포는 NKT 세포이다. 특정 실시형태에서 사용하기에 적합한 T 세포의 다른 예시적 집단은 미경험 T 세포 및 기억 T 세포를 포함한다.
- [0633] 다양한 실시형태에서, 계놈 편집된 세포는 본 명세서에서 상정된 조성물 및 방법에 의해 편집된 CBLB 유전자를 포함하는 면역 효과기 세포를 포함한다. "면역 효과기 세포"는 하나 이상의 효과기 기능(예를 들어, 세포독성 세포 사멸 활성, 사이토카인의 분비, ADCC 및/또는 CDC의 유도)을 갖는 면역계의 임의의 세포이다. 특정 실시형태에서 상정된 예시적인 면역 효과기 세포는 T 림프구, 특히 세포독성 T 세포(CTL; CD8+ T 세포), TIL 및 헬퍼 T 세포(HTL; CD4+ T 세포)이다. 일 실시형태에서, 면역 효과기 세포는 T 세포를 포함한다. 일 실시형태에서, 면역 효과기 세포는 자연 살해(NK) 세포를 포함한다. 일 실시형태에서, 면역 효과기 세포는 자연 살해 T(NKT) 세포를 포함한다.
- [0634] "강한 T 세포" 및 "어린 T 세포"는 특정 실시형태에서 상호호환적으로 사용되며, T 세포 표현형을 지칭하되, T 세포는 증식 및 분화의 동시 감소를 가능하게 한다. 특정 실시형태에서, 어린 T 세포는 "미경험 T 세포"의 표현형을 가진다. 특정 실시형태에서, 어린 T 세포는 다음의 생물학적 마커: CD62L, CCR7, CD28, CD27, CD122, CD127, CD197 및 CD38 중 하나 이상, 또는 모두를 포함한다. 일 실시형태에서, 어린 T 세포는 다음의 생물학적 마커: CD62L, CD127, CD197 및 CD38 중 하나 이상 또는 모두를 포함한다. 일 실시형태에서, 어린 T 세포는 CD57, CD244, CD160, PD-1, CTLA4, PD-1 및 LAG3의 발현을 결여한다.
- [0635] T 세포는 말초혈액 단핵세포, 골수, 림프절 조직, 제대혈, 흥선 조직, 감염 부위로부터의 조직, 복수, 흉수, 비장 조직 및 종양을 포함하지만, 이들로 제한되지 않는 다수의 공급원으로부터 얻을 수 있다.

- [0636] 특정 실시형태에서, 면역 효과기 세포 또는 T 세포를 포함하는 세포 집단은 편집된 CBLB 유전자를 포함하되, 편집은 NHEJ에 의해 수선된 DSB이다. 특정 실시형태에서, 면역 효과기 세포 또는 T 세포는 편집된 CBLB 유전자를 포함하되, 편집은 NHEJ에 의해 수선된 DSB이다. 특정 실시형태에서, 편집은 CBLB 유전자의 암호 서열에서, 바람직하게는 CBLB 유전자의 엑손 6에서, 더 바람직하게는 CBLB 유전자의 엑손 6에서의 서열번호 20(또는 서열번호 22)에서 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25개 이상의 뉴클레오타이드의 삽입 또는 결실(INDEL)이다.
- [0637] 특정 실시형태에서, 면역 효과기 세포 또는 T 세포를 포함하는 세포의 집단은 HDR에 의해 수선되는 DSB에서 혼입된 공여자 수선 주형을 포함하는 편집된 CBLB유전자를 포함한다.
- [0638] 특정 실시형태에서, 면역 효과기 세포 또는 T 세포를 포함하는 세포의 집단은 CBLB 유전자 또는 이의 일부를 포함하는 공여자 수선 주형을 포함하는 편집된 CBLB 유전자를 포함하고, CBLB 발현 활성 또는 신호전달을 변형시키기 위해, 바람직하게는, CBLB 발현 활성 및/또는 신호전달을 감소 또는 제거하기 위해 게놈 CBLB 서열에서 하나 이상의 돌연변이를 도입하도록 설계된다.
- [0639] 다양한 실시형태에서, 게놈 편집 세포는 CBLB 유전자에서의 편집을 포함하고, 플립 수용체, 이중특이성 T 세포 관여자(BiTE) 분자; 사이토카인(예를 들어, IL-2, 인슐린, IFN- $\gamma$ , IL-7, IL-21, IL-10, IL-12, IL-15 및 TNF- $\alpha$ ), 케모카인(예를 들어, MIP-1 $\alpha$ , MIP-1 $\beta$ , MCP-1, MCP-3 및 RANTES), 사이토톡신(예를 들어, 퍼포린, 그랜자임 A 및 그랜자임 B), 사이토카인 수용체(예를 들어, IL-2 수용체, IL-7 수용체, IL-12 수용체, IL-15 수용체, 및 IL-21 수용체), 또는 조작된 항원 수용체(예를 들어, 조작된 T 세포 수용체(TCR), 키메라 항원 수용체(CAR), Daric 수용체 또는 이의 성분, 또는 키메라 사이토카인 수용체)를 암호화하는 폴리뉴클레오타이드를 추가로 포함한다. 일 실시형태에서, 폴리뉴클레오타이드 및 뉴클레아제 변이체를 포함하는 공여자 수선 주형은 세포 내로 도입되고, 폴리뉴클레오타이드는 HDR 수선에 의해 CBLB 유전자 내 DSB 부위에서 세포의 게놈 내로 혼입된다. 폴리뉴클레오타이드는 또한, 예를 들어, 폴리뉴클레오타이드를 포함하는 벡터를 이용하여 세포를 형질 도입함으로써 CBLB 유전자 이외의 부위에서 세포 내로 도입될 수 있다.
- [0640] I. 조성물 및 제형
- [0641] 특정 실시형태에서 상정된 조성물은 본 명세서에서 상정된 바와 같이 하나 이상의 폴리펩타이드, 폴리뉴클레오타이드, 이를 포함하는 벡터, 및 게놈 편집 조성물 및 게놈 편집 세포 조성물을 포함할 수 있다. 특정 실시형태에서 상정된 게놈 편집 조성물 및 방법은 세포 또는 세포 집단에서 인간 CBLB 유전자 내 표적 부위를 편집하는 데 유용하다. 바람직한 실시형태에서, 게놈 편집 조성물은 조혈세포, 예를 들어, T 세포 또는 면역 효과기 세포에서 CBLB 유전자를 편집하는 데 사용될 수 있다.
- [0642] 다양한 실시형태에서, 본 명세서에서 상정된 조성물은 뉴클레아제 변이체, 및 선택적으로 말단-가공 효소, 예를 들어, 3'-5' 엑소뉴클레아제(Trex2)를 포함한다. 뉴클레아제 변이체는 상기 개시된 폴리뉴클레오타이드 전달 방법, 예를 들어, 전기천공법, 지질 나노입자 등을 통해 세포 내로 도입되는 mRNA 형태일 수 있다. 일 실시형태에서, 호밍 엔도뉴클레아제 변이체 또는 메가TAL을 암호화하는 mRNA, 및 선택적으로 3'-5' 엑소뉴클레아제를 포함하는 조성물은 상기 개시한 폴리뉴클레오타이드 전달 방법을 통해 세포에 도입된다. 조성물은 오류유발 NHEJ에 의해 게놈 편집된 세포 또는 게놈 편집된 세포의 집단을 생성하는 데 사용될 수 있다.
- [0643] 다양한 실시형태에서, 본 명세서에서 상정된 조성물은 공여자 수선 주형을 포함한다. 조성물은 뉴클레아제 변이체, 및 선택적으로 말단-가공 효소를 발현시키거나 또는 발현시킬 세포에 전달될 수 있다. 일 실시형태에서, 조성물은 호밍 엔도뉴클레아제 변이체 또는 메가TAL, 및 선택적으로 3'-5' 엑소뉴클레아제를 발현시키거나 또는 발현시킬 세포에 전달될 수 있다. 공여자 수선 주형의 존재 하에 유전자 편집 효소의 발현은 HDR에 의해 게놈 편집된 세포 또는 게놈 편집된 세포의 집단을 생성하는 데 사용될 수 있다.
- [0644] 특정 실시형태에서, 본 명세서에서 상정된 조성물은 세포 집단, 뉴클레아제 변이체, 및 선택적으로, 공여자 수선 주형을 포함한다. 특정 실시형태에서, 본 명세서에서 상정된 조성물은 세포 집단, 뉴클레아제 변이체, 말단-가공 효소, 및 선택적으로, 공여자 수선 주형을 포함한다. 뉴클레아제 변이체 및/또는 말단-가공 효소는 상기 개시된 폴리뉴클레오타이드 전달 방법을 통해 세포 내로 도입되는 mRNA의 형태일 수 있다.
- [0645] 특정 실시형태에서, 본 명세서에서 상정된 조성물은 세포 집단, 호밍 엔도뉴클레아제 변이체 또는 메가TAL, 및 선택적으로, 공여자 수선 주형을 포함한다. 특정 실시형태에서, 본 명세서에서 상정된 조성물은 세포 집단, 호밍 엔도뉴클레아제 변이체 또는 메가TAL, 3'-5' 엑소뉴클레아제, 및 선택적으로, 공여자 수선 주형을 포함한다. 호밍 엔도뉴클레아제 변이체, 메가TAL 및/또는 3'-5' 엑소뉴클레아제는 상기 개시한 폴리뉴클레오타이드 전달

방법, 예를 들어, 전기천공법을 통해 세포 내로 도입되는 mRNA의 형태일 수 있다.

[0646] 특정 실시형태에서, 세포 접단은 유전자 변형된 면역 효과기 세포를 포함한다.

[0647] 조성물은 약제학적 조성물을 포함하지만, 이들로 제한되지 않는다. "약제학적 조성물"은 단독으로 또는 한 가지 이상의 치료 약상과 조합하여 세포 또는 동물에게 투여를 위한 약제학적으로 허용 가능한 또는 생리적으로 허용 가능한 용액 중에서 제형화된 조성물을 지칭한다. 또한 원한다면, 조성물은 마찬가지로 다른 제제, 예컨대 사이토카인, 성장 인자, 호르몬, 소분자, 화학치료제, 프로드러그, 약물, 항체 또는 다른 다양한 약제학적 활성제와 조합하여 투여될 수 있다는 것이 이해될 것이다. 또한 조성물 중에 포함될 수 있는 다른 성분에 대한 제한은 사실상 없으며, 단, 추가적인 제제는 조성물에 유해하게 영향을 미치지 않는다.

[0648] 어구 "약제학적으로 허용 가능한"은 타당한 의학적 판단의 범주 내에서, 합리적인 유해/유익비에 비례하여 과도한 독성, 자극, 알레르기 반응 또는 다른 문제 또는 합병증 없이 인간 및 동물의 조직과 접촉하여 사용하기에 적합한 본 명세서에서 해당 화합물, 물질, 조성물 및/또는 투약 형태를 지칭하기 위해 사용된다.

[0649] 용어 "약제학적으로 허용 가능한 담체"는 치료 세포와 함께 투여되는 희석제, 보조제, 부형제 또는 비히클을 지칭한다. 약제학적 담체의 예시적 예는 멸균 액체, 예컨대 세포 배양 배지, 물 및 석유, 동물성, 식물성 또는 합성 유래의 오일, 예컨대 땅콩유, 대두유, 광유, 참깨유 등일 수 있다. 식염수 용액 및 수성 텍스트로스 및 글리세롤 용액은 또한 액체 담체로서, 특히 주사용 용액으로 사용될 수 있다. 특정 실시형태에서, 적합한 약제학적 부형제는 전분, 글루코스, 락토스, 수크로스, 젤라틴, 맥아, 쌀, 밀가루, 백악, 실리카겔, 스테아르산나트륨, 글리세롤 모노스테아레이트, 탈크, 염화나트륨, 건조탈지유, 글리세롤, 프로필렌, 글리콜, 물, 에탄올 등을 포함한다. 임의의 통상적인 배지 또는 제제가 활성 성분에 적합하지 않은 경우를 제외하고, 치료 조성물에서의 그의 용도가 상정된다. 보충적 활성 성분이 또한 조성물 내로 혼입될 수 있다.

[0650] 일 실시형태에서, 약제학적으로 허용 가능한 담체를 포함하는 조성물은 대상체에 대한 투여에 적합하다. 특정 실시형태에서, 담체를 포함하는 조성물은 비경구 투여, 예를 들어, 혈관내(정맥내 또는 동맥내), 복강내 또는 근육내 투여에 적합하다. 특정 실시형태에서, 약제학적으로 허용 가능한 담체를 포함하는 조성물은 심실내, 척수내 또는 척추강내에 적합하다. 약제학적으로 허용 가능한 담체 멸균 수용액, 세포 배양 배지 또는 분산물을 포함한다. 약제학적으로 활성인 물질에 대한 이러한 배지 및 제제의 사용은 당업계에 잘 공지되어 있다. 임의의 통상적인 배지 또는 제제가 형질도입 세포에 적합하지 않은 경우를 제외하고, 약제학적 조성물 중의 이의 용도가 상정된다.

[0651] 특정 실시형태에서, 본 명세서에서 상정되는 조성물은 유전자 변형된 T 세포 및 약제학적으로 허용 가능한 담체를 포함한다. 본 명세서에서 상정된 세포 기반 조성물을 포함하는 조성물은 장용 또는 비경구 투여 방법에 의해 별개로 또는 목적으로 하는 치료 목적을 달성하기 위한 다른 적합한 화합물과 조합하여 투여될 수 있다.

[0652] 약제학적으로 허용 가능한 담체는 치료 중인 인간 대상체에 대한 투여에 적합하도록 충분히 고순도 그리고 충분히 낮은 독성을 가져야 한다. 추가로 조성물의 안정성을 유지하거나 또는 증가시켜야 한다. 약제학적으로 허용 가능한 담체는 액체 또는 고체일 수 있고, 조성물의 다른 성분과 조합될 때 목적으로 하는 벌크, 점조도 등을 제공하기 위해, 유념해둔 계획한 투여 방식으로 선택된다. 예를 들어, 약제학적으로 허용 가능한 담체는 결합제(예를 들어, 전호화된 메이즈 전분, 폴리비닐필로리돈 또는 하이드록시프로필 메틸셀룰로스 등), 충전제(예를 들어, 락토스 및 다른 당, 미정질 셀룰로스, 펙틴, 젤라틴, 황산칼슘, 에틸 셀룰로스, 폴리아크릴레이트, 인산 수소칼슘 등), 윤활제(예를 들어, 스테아르산마그네슘, 탈크, 실리카, 콜로이드 이산화규소, 스테아르산, 금속 스테아레이트, 수소화된 식물성유, 옥수수 전분, 폴리에틸렌 글리콜, 벤조산나트륨, 아세트산나트륨 등), 봉해제(예를 들어, 전분, 글리콜산나트륨전분 등) 또는 습윤제(예를 들어, 라우릴황산나트륨 등)일 수 있지만, 이들로 제한되지 않는다. 본 명세서에서 상정된 조성물에 대한 다른 적합한 약제학적으로 허용 가능한 담체는 물, 염 용액, 알코올, 폴리에틸렌 글리콜, 젤라틴, 아밀로스, 스테아르산마그네슘, 탈크, 규산, 점성 파라핀, 하이드록시메틸셀룰로스, 폴리비닐피롤리돈 등을 포함하지만, 이들로 제한되지 않는다.

[0653] 이러한 담체 용액은 또한 완충제, 희석제 및 다른 적합한 첨가제를 함유할 수 있다. 본 명세서에서 사용되는 용어 "완충제"는 화학적 구성이 pH의 상당한 변화없이 산 또는 염기를 중화시키는 용액 또는 액체를 지칭한다. 본 명세서에서 상정된 완충제의 예는 둘베코 인산염 완충 식염수(PBS), 렁거 용액, 수 중 5% 텍스트로스(D5W), 정상/생리식염수(0.9% NaCl)를 포함하지만, 이들로 제한되지 않는다.

[0654] 약제학적으로 허용 가능한 담체는 조성물의 pH를 약 7로 유지하기에 충분한 양으로 존재할 수 있다. 대안적으로, 조성물은 약 6.8 내지 약 7.4, 예를 들어, 6.8, 6.9, 7.0, 7.1, 7.2, 7.3 및 7.4 범위의 pH를 가

진다. 또 다른 실시형태에서, 조성물은 약 7.4의 pH를 가진다.

[0655] 본 명세서에서 상정된 조성물은 비독성의 약제학적으로 허용 가능한 배지를 포함할 수 있다. 조성물은 혼탁액일 수 있다. 본 명세서에서 사용되는 용어 "혼탁액"은 세포가 고체 지지체에 부착되지 않는 비부착 상태를 지칭한다. 예를 들어, 혼탁액으로서 유지된 세포는 교반되거나 또는 휘저어질 수 있고, 지지체, 예컨대 배양 접시에 부착되지 않는다.

[0656] 특정 실시형태에서, 본 명세서에서 상정된 조성물은 계놈 편집된 T 세포가 정맥내 (IV) 등에서 허용 가능한 액체 배지 또는 용액, 예를 들어, 식염수 또는 무혈청 배지 내에서 분산되는 혼탁액 중에서 제형화된다. 허용 가능한 희석제는 물, 립거용액, 등장성 염화나트륨(식염수) 용액, 무혈청 세포 배양 배지 및 저온 저장에 적합한 배지, 예를 들어, Cryostor(등록상표) 배지를 포함하지만, 이들로 제한되지 않는다.

[0657] 특정 실시형태에서, 약제학적으로 허용 가능한 담체는 인간 또는 동물 유래의 천연 단백질이 실질적으로 없고, 계놈 편집된 T 세포 접단을 포함하는 조성물을 저장하는 데 적합하다. 치료 조성물은 인간 환자에게 투여되도록 의도되고, 따라서 세포 배양 성분, 예컨대 소 혈청 알부민, 말 혈청 및 소태아 혈청이 실질적으로 없다.

[0658] 일부 실시형태에서, 조성물은 약제학적으로 허용 가능한 세포 배양 배지에서 제형화된다. 이러한 조성물은 인간 대상체에 대한 투여에 적합하다. 특정 실시형태에서, 약제학적으로 허용 가능한 세포 배양 배지는 무혈청 배지이다.

[0659] 무혈청 배지는 단순화된 그리고 더 잘 한정된 조성물, 감소된 정도의 오염물질, 감염제의 잠재적 공급원의 제거 및 보다 저비용을 포함하는, 혈청 함유 배지 이상의 몇몇 이점을 가진다. 다양한 실시형태에서, 무혈청 배지는 무동물이며, 선택적으로 무단백질일 수 있다. 선택적으로, 배지는 생약제학적으로 허용 가능한 재조합 단백질을 함유할 수 있다. "무 동물" 배지는 성분이 비동물 공급원으로부터 유래된 배지를 지칭한다. 재조합 단백질은 무동물 배지에서 천연 동물 단백질을 대체하며, 영양분은 합성, 식물 또는 미생물 공급원으로부터 얻어진다. "무 단백질" 배지는 대조적으로 단백질이 실질적으로 없는 것으로 정의된다.

[0660] 특정 조성물에서 사용되는 무혈청 배지의 예시적 예는 QBSF-60(웰러티 바이올로지컬 인코포레이티드(Quality Biological, Inc.)), 스템프로-34(StemPro-34)(라이프 테크놀로지즈(Life Technologies)) 및 X-VIVO 10을 포함하지만, 이들로 제한되지 않는다.

[0661] 바람직한 실시형태에서, 계놈 편집된 T 세포를 포함하는 조성물은 PlasmaLyte에서 제형화된다.

[0662] 다양한 실시형태에서, 계놈 편집된 T 세포를 포함하는 조성물은 동결보존 배지에서 제형화된다. 예를 들어, 동결보존제가 있는 동결보존 배지는 해동 후 높은 세포 생존도 결과를 유지하기 위해 사용될 수 있다. 특정 조성물에서 사용되는 동결보존 배지의 예시적 예는 CryoStor CS10, CryoStor CS5 및 CryoStor CS2를 포함하지만, 이들로 제한되지 않는다.

[0663] 일 실시형태에서, 조성물은 50:50 PlasmaLyte A 대 CryoStor CS10을 포함하는 용액에서 제형화된다.

[0664] 특정 실시형태에서, 조성물은 마이코플라스마, 내독소 및 미생물 오염물질이 실질적으로 없다. 내독소와 관련하여 "실질적으로 없는"은 생물제제에 대해 FDA에 의해 허용되는 것(5 EU/kg 체중/일의 총 내독소, 평균 70kg의 사람은 세포의 총 용량 당 350 EU임)보다 세포의 용량 당 내독소가 더 적다는 것을 의미한다. 특정 실시형태에서, 본 명세서에서 상정된 형질도입된 조혈 줄기 또는 조상 세포를 포함하는 조성물은 약 0.5 EU/ml, 내지 약 5.0EU/ml, 또는 약 0.5EU/ml, 1.0EU/ml, 1.5EU/ml, 2.0EU/ml, 2.5EU/ml, 3.0EU/ml, 3.5EU/ml, 4.0EU/ml, 4.5 EU/ml, 또는 5.0 EU/ml, 를 함유한다.

[0665] 특정 실시형태에서, 하나 이상의 채프로그래밍된 뉴클레아제를 암호화하는 하나 이상의 mRNA, 및 선택적으로 말단-가공 효소를 포함하지만, 이들로 제한되지 않는 폴리뉴클레오타이드의 전달에 적합한 조성물 및 제형이 상정된다.

[0666] 생체외 전달을 위한 예시적인 제형은 또한 당업계에 공지된 다양한 형질감염제, 예컨대 인산칼슘, 전기천공법, 열 충격 및 다양한 리포좀 제형(즉, 지질-매개 형질감염)의 사용을 포함할 수 있다. 이하에 더 상세하게 기재하는 리포좀은 수성 유치의 분획을 불잡는 지질 이중층이다. DNA는 양이온성 리포좀의 외면에 (그의 하전에 의해) 결합하고, 이들 리포좀은 세포막과 상호작용할 것이다.

[0667] 특정 실시형태에서, 약제학적으로 허용 가능한 담체 용액은, 예를 들어, 장용 및 비경구, 예를 들어, 혈관내, 정맥내, 동맥내, 골내, 심실내, 대뇌내, 두개내, 척수내, 척추강내 및 골수내 투여 및 제형을 포함하는 다양한

치료 요법에서 본 명세서에 기재된 특정 조성물을 이용하기 위한 적합한 투약 및 치료 요법의 개발과 같이, 당업자에게 잘 공지되어 있다. 본 명세서에서 상정된 특정 실시형태는 다른 제형, 예컨대 약제학적 분야에 잘 공지되어 있고, 예를 들어, 본 명세서에 전문이 참고로 포함된 문헌[Remington: The Science and Practice of Pharmacy, volume I and volume II. 22nd Edition. Edited by Loyd V. Allen Jr. Philadelphia, PA: Pharmaceutical Press; 2012]에 기재되어 있는 것을 포함할 수 있다는 것은 당업자에 의해 이해될 것이다.

[0668] J. 계놈 편집 세포 요법

본 명세서에서 상정된 조성물 및 방법에 의해 제조되는 편집된 CBLB 유전자를 포함하는 세포는 암, GVHD, 감염성 질환, 자가면역 질환, 염증성 질환 또는 면역결핍 중 적어도 하나의 증상의 예방, 치료 또는 개선에서 사용하기 위한 개선된 약물 제품을 제공한다. 본 명세서에서 사용되는 용어 "약물 제품"은 본 명세서에서 상정된 조성물 및 방법을 이용하여 생산된 유전자 변형된 세포를 지칭한다. 특정 실시형태에서, 약물 제품은 유전자 편집된 면역 효과기 세포 또는 T 세포를 포함한다. 바람직한 실시형태에서, 약물 제품은 조작된 TCR 또는 CAR 또는 다른 조작된 항원 수용체를 발현시키는 유전자 편집된 면역 효과기 세포 또는 T 세포를 포함한다. 게다가, 특정 실시형태에서 상정된 계놈 편집된 T 세포는 더 안전하고 더 효능있는 적응 세포 요법을 제공하는데, 그들이 T 세포 고갈에 대해 내성이 있고, 지속된 요법을 야기할 수 있는 증가된 내구성 및 지속성을 나타내기 때문이다.

[0669] [0670] 특정 실시형태에서, 편집된 CBLB 유전자를 포함하는 유효량의 계놈 편집된 면역 효과기 세포 또는 T 세포는 암, GVHD, 감염성 질환, 자가면역 질환, 염증성 질환 또는 면역결핍 중 적어도 하나의 증상을 예방하거나, 치료하거나 또는 개선시키기 위해 대상체에게 투여된다.

[0671] 특정 실시형태에서, CBLB 편집 세포는 CBLB을 실질적으로 발현시키지 않거나 또는 발현을 결여하고, 따라서 작용성 CBLB 발현 및/또는 활성을 결여하고, 예를 들어, T 세포 고갈을 증가시키고 전염증 사이토카인의 발현을 저해하는 능력을 결여한다. 특정 실시형태에서, CBLB을 결여하는 계놈 편집된 면역 효과기 세포는 종양 미세환경으로부터의 면역억제 신호에 대해 더 내성이 있고, T 세포 고갈에 대해 증가된 지속성 및 내성을 나타낸다.

[0672] 특정 실시형태에서, 암의 적어도 하나의 증상을 예방하거나, 치료하거나 또는 개선시키는 방법은 종양 또는 암에 세포를 다시 보내기 위해 대상체에게 유효량의 계놈 편집된 면역 효과기 세포 또는 편집된 CBLB 유전자 및 조작된 TCR, CAR 또는 Daric 또는 다른 치료 이식유전자를 포함하는 T 세포를 투여하는 단계를 포함한다. 유전자 변형된 세포는 내구성 있고 지속적인 약물 제품인데, 세포가 CBLB 발현을 감소 또는 제거하는 CBLB 유전자의 편집 때문에 종양 미세환경으로부터의 면역억제 신호에 대해 더 내성이 있기 때문이다.

[0673] 특정 실시형태에서, 본 명세서에서 상정된 계놈 편집된 세포는 고형 종양 또는 암의 치료에서 사용된다.

[0674] 특정 실시형태에서, 본 명세서에서 상정된 계놈 편집된 세포는 부신암, 부신피질암종, 항문암, 맹장암, 성상세포종, 비정형 유기형/간상 종양, 기저세포암, 담도암, 방광암, 끌암, 뇌/CNS 암, 유방암, 기관지 종양, 심장 종양, 자궁경부암, 담관암종, 연골육종, 척색종, 결장암, 결장직장암, 두개인두종, 관상피내암(DCIS) 자궁내막암, 상의세포종, 식도암, 각각신경모세포종, 유잉 육종, 두개외 생식세포종양, 성선외생식세포종양, 안암, 나팔관암, 섬유성 조직육종, 섬유육종, 담낭암, 위암, 위장유암종, 위장관기질 종양(GIST), 생식 세포 종양, 신경교종, 교모세포종, 두경부암, 혈관모세포종, 간세포 암, 식도암, 안구내 흑색종, 카포시 육종, 신장암, 후두암, 평활근육종, 구순암, 지방육종, 간암, 폐암, 비소세포 폐암, 폐유암종, 악성 중피종, 수질암종, 수도세포종, 뇌수막종, 흑색종, 머켈 세포 암종, 정중선관 암종, 구강암, 점액육종, 골수이형성 증후군, 골수증식성 신생물, 비강 및 부비동암, 비인두암, 신경아세포종, 펩지교종, 구강암, 구강암, 인두암, 골육종, 난소암, 췌장암, 췌장섬 세포 종양, 유두 암종, 부신결정종, 부갑상선암, 음경암, 인두암, 크롬친화세포종, 송과체종, 하수체종양, 흉막폐아세포종, 원발성 복막암, 전립선암, 직장암, 망막모세포종, 신세포 암종, 신우 및 요관암, 횡문근육종, 침샘암, 피지선암종, 피부암, 연조직 육종, 편평세포암종, 소세포폐암, 소장암, 위암, 땀샘 암종, 혈막종, 고환암, 인후암, 흉선암, 갑상선암, 요도암, 자궁암, 자궁 육종, 질암, 혈관암, 외음부암 및 월름 종양을 포함하지만, 이들로 제한되지 않는 고형암 또는 암의 치료에서 사용된다.

[0675] 특정 실시형태에서, 본 명세서에서 상정된 계놈 편집된 세포는 간암, 췌장암, 폐암, 유방암, 방광암, 뇌암, 골암, 갑상선암, 신장암 또는 피부암을 포함하지만, 이들로 제한되지 않는 고형 종양 또는 암의 치료에서 사용된다.

[0676] 특정 실시형태에서, 본 명세서에서 상정된 계놈 편집된 세포는 췌장, 방광 및 폐를 포함하지만, 이들로 제한되지 않는 다양한 암의 치료에서 사용된다.



kg 또는  $6 \times 10^6$  개의 세포/kg 내지 약  $8 \times 10^6$  개의 세포/kg(세포의 모든 개체 용량을 포함)이다.

[0686] 당업자는 특정 실시형태에서 상정된 조성물의 다중 투여가 목적으로 하는 요법을 달성하는 데 필요할 수 있다는 것을 인식할 것이다. 예를 들어, 조성물은 1주, 2주, 3주, 1개월, 2개월, 3개월, 4개월, 5개월, 6개월, 1년, 2년, 5년, 10년 이상의 기간에 걸쳐 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10회 이상 투여될 수 있다.

[0687] 특정 실시형태에서, 대상체에게 활성화된 T 세포를 투여하는 것이 바람직할 수 있으며, 이어서, 후속적으로 혈액을 재채혈하고(또는 성분채집술이 수행됨), 이로부터 T 세포를 활성화시키며, 이들 활성화되고 확장된 T 세포를 환자에 재주입한다. 이 과정은 몇 주마다 다회 수행될 수 있다. 특정 실시형태에서, T 세포는 10cc 내지 400cc의 채혈로부터 활성화될 수 있다. 특정 실시형태에서, T 세포는 20cc, 30cc, 40cc, 50cc, 60cc, 70cc, 80cc, 90cc, 100cc, 150cc, 200cc, 250cc, 300cc, 350cc 또는 400cc 이상의 채혈로부터 활성화된다. 이론에 의해 구속되지 않고, 이런 다회 채혈/다중 재주입 프로토콜을 이용하는 것은 T 세포의 특정 집단을 선택하는 작용을 할 수 있다.

[0688] 특정 실시형태에서 상정된 조성물의 투여는 에어로졸 흡입, 주사, 섭취, 수혈, 주입 또는 이식에 의하는 것을 포함하는 임의의 편리한 방법으로 수행될 수 있다. 바람직한 실시형태에서, 조성물은 비경구로 투여된다. 본 명세서에서 사용되는 어구 "비경구 투여" 및 "비경구로 투여되는"은 장용 및 국소 투여 이외의, 보통 주사에 의하는 투여 방식을 지칭하며, 혈관내, 정맥내, 근육내, 동맥내, 척추강내, 낭내, 안와내, 종양내, 심장내, 피내, 복강내, 기관지경, 피하, 피부밑, 관절강내, 피막하, 지주막하, 척주내 및 흉골내 주사 및 주입을 포함하지만, 이들로 제한되지 않는다. 일 실시형태에서, 본 명세서에서 상정된 조성물은 종양, 림프절 또는 감염 부위에 직접 주사에 의해 대상체에게 투여된다.

[0689] 일 실시형태에서, 암으로 진단된 대상체를 치료하는 방법은 대상체로부터 면역 효과기 세포를 제거하는 단계, 상기 면역 효과기 세포 계놈을 편집하는 단계 및 계놈 편집된 면역 효과기 세포 집단을 생산하는 단계, 및 동일한 대상체에게 계놈 편집된 면역 효과기 세포 집단을 투여하는 단계를 포함한다. 바람직한 실시형태에서, 면역 효과기 세포는 T 세포를 포함한다.

[0690] 특정 실시형태에서 상정된 세포 조성물을 투여하기 위한 방법은 생체외 계놈 편집된 면역 효과기 세포의 재도입에서 또는 대상체에 도입 시 성숙 면역 효과기 세포로 분화하는 면역 효과기 세포의 계놈 편집된 전구체의 재도입에 대한 결과에 효과적인 임의의 방법을 포함한다. 한 가지 방법은 생체외 말초 혈액 T 세포를 계놈 편집하는 단계 및 형질도입된 세포를 대상체에게 복귀시키는 단계를 포함한다.

[0691] 본 명세서의 모든 간행물, 특히 출원 및 공개된 특허는 각각의 개개 간행물, 특히 출원 또는 공개된 특허가 구체적이고 개별적으로 참고로 포함되도록 포함되는 것과 같이 본 명세서에 참고로 포함된다.

[0692] 앞서 언급한 실시형태는 이해의 분명함의 목적을 위해 예시 및 실시예로 일부 상세하게 기재하였지만, 특정 변화 및 변형이 첨부된 청구범위의 정신 또는 범주로부터 벗어나는 일 없이 이루어질 수 있다는 것은 본 명세서에서 상정된 교시에 비추어 당업자에게 용이하게 분명할 것이다. 다음의 실시예는 단지 예로서 제한의 방법에 의하지 않고 제공한다. 당업자는 본질적으로 유사한 결과를 수득하기 위해 변화 또는 변형될 수 있는 다양한 중요하지 않은 매개변수를 용이하게 인식할 것이다.

#### 실시예

##### 실시예 1

[0695] 인간 카시타스 B-계통(CBL) 림프종 원발암유전자 B(CBLB) 유전자를 봉괴시키기 위한 재프로그래밍 I-OnuI

[0696] DNA 인식 계면에서 가변 아미노산 잔기를 함유하는 모듈 라이브러리를 삭제함으로써 CBLB 유전자의 엑손 6을 표적화하도록 I-OnuI을 재프로그래밍하였다(도 1). 변이체를 삭제하기 위해, 올리고뉴클레오타이드를 이용하여 축퇴 코돈을 I-OnuI DNA 결합 도메인 내로 혼입시켰다. 축퇴 코돈을 암호화하는 올리고뉴클레오타이드를 PCR 주형으로서 사용하여 효모 균주 사카로마이세스 세레비시애(S. cerevisiae)에서 캡 재조합에 의해 변이체 라이브러리를 생성하였다. 각각의 변이체 라이브러리는 N- 또는 C-말단 I-OnuI DNA 인식 도메인 중 하나에 걸쳐있고, 대략  $10^7$  내지  $10^8$ 개의 독특한 형질전환체를 함유하였다. 얻어진 표면 디스플레이 라이브러리는, 도 2에 나타내는 바와 같이, 대응하는 도메인의 "절반-부위(서열번호 23 내지 29)를 포함하는 표적 부위에 대한 절단 활성에 대해 유세포분석에 의해 선별하였다.

[0697] N- 및 C-말단의 도메인 재프로그래밍된 I-OnuI HE를 나타내는 효모를 정제하고 나서, 플라스미드 DNA를 추출하

였다. PCR 반응을 수행하여 재프로그래밍된 도메인을 증폭시키고, 이를 후속적으로 융합시키고 사카로마이세스 세레비시애로 형질전환시켜 재프로그래밍된 도메인 조합의 라이브러리를 생성하였다. CBLB 유전자의 엑손 6에 존재하는 완전한 표적 부위(서열번호 20)를 인식하는 완전히 재프로그래밍된 I-OnuI 변이체를 이 라이브러리로부터 동정하고, 정제하였다.

[0698]

## 실시예 2

[0699] CBLB 유전자의 엑손 6을 표적화하는 재프로그래밍된 I-OnuI 호밍 엔도뉴클레아제

[0700]

염색체로 통합된 형광 수용체 시스템을 이용하여 CBLB 유전자의 엑손 6을 표적화하는 재프로그래밍된 I-OnuI HE의 활성을 측정하였다(Certo et. al., 2011). CBLB ITSM 표적 서열(서열번호 20)에 결합하고, 이를 절단하는 완전히 재프로그래밍된 I-OnuI HE를 포유류 발현 플라스미드 새로 클로닝시켰고, 이어서, 형광 iRFP 단백질을 암호화하는 프레임 밖 유전자 상류의 CBLB 표적 서열이 함유된 HEK 293T 섬유아세포 세포주 새로 개별적으로 형질감염시켰다. 포매된 표적 부위의 HE에 의한 절단 및 비상동성 말단 결합(NHEJ) 경로를 통한 DNA 수선 후 삽입 결실의 축적은 3개의 수선된 좌위 중 대략 하나가 형광 수용체 유전자 뒤 "프레임 내"에 위치하도록 초래한다. 따라서 iRFP 형광 HEK 293T 세포의 백분율은 염색체로 포매된 표적 서열에서 엔도뉴클레아제 활성의 판독을 위해 사용한다. CBLB 표적 서열에 결합하고 이를 절단하는 완전히 재프로그래밍된 I-OnuI HE는 세포의 염색체 내용에서 고빈도의 iRFP 발현을 나타내었는데, 이는 HE 변이체가 세포의 염색체 내용에서 높은 편집 효율의 요망 되는 특성을 가진다는 것을 나타낸다. 도 3.

[0701]

CBLB.E3 HE 변이체는 엑손 6 표적 부위에 대해 서브-나노몰 친화도를 가졌다. 도 4. 도 5는 대표적인 I-OnuI 변이체의 상대적 정렬뿐만 아니라 DNA 인식 계면을 포함하는 잔기의 위치 정보를 나타낸다.

[0702]

## 실시예 3

[0703] CBLB 엑손 6을 표적화하는 메가TAL의 특성규명

[0704]

CBLB.E3, CBLB.A8, CBLB.B5, CBLB.D2 및 CBLB.F6 HE 변이체를 메가뉴클레아제 도메인의 N-말단에 대해 12개 염기쌍 TAL 어레이 표적 부위(서열번호 21)에 대응하는 11.5 단위 TAL 어레이를 현수함으로써 메가TAL(서열번호 13 내지 19)로서 형식화시켰다(문헌[Boissel et al., 2013]에 기재한 바와 같음). 도 6. 메가TAL 표적 부위 서열을 서열번호 22에 제시한다.

[0705]

48 내지 72시간 동안 사이토카인-보충 배지에서 항-CD3 및 항-CD28 항체를 이용하여 1차 인간 T 세포를 사전 자극함으로써 메가TAL 편집 효율을 평가하였고, 이어서, 메가TAL를 암호화하는 시험관내 전사되고(IVT), 캡핑되고, 폴리아데닐화된 mRNA를 이용하여 세포를 전기천공시켰다. 추가적으로, 3'-5' 엑소뉴클레아제 Trex2를 암호화하는 IVT-mRNA를 첨가하여 비상동성 말단 결합(NHEJ) 경로에 의해 파손 가공을 향상시켰다(문헌[Certo et al., 2012] 참조).

[0706]

한 세트의 실험에서, 세포를 전기천공법 후 7일 동안(총 10일 배양) 배양시키고 나서, 배양의 제5일, 제7일 및 제10일에 게놈 DNA를 단리시킴으로써 그리고 삽입결실 빈도를 측정하기 위해 CBLB 표적 부위에 걸쳐 서열분석함으로써 편집 효율을 측정하였다. 도 7. CD3+ T 세포의 세포내 FACS 분석에 의해 CBLB 단백질 수준을 측정하였다. 도 8.

[0707]

CBLB 메가TAL로 편집한 T 세포 또는 CAR T 세포는 활성 메가TAL(비활성 메가TAL로 전기천공)을 결여하거나 또는 메가TAL을 함께 결여한(EP 없음) CAR T 세포 및 비편집 T 세포에 비해 항-CD3 작용성 항체 또는 CAR-특이적 항원으로 각각 자극될 때 증가된 사이토카인 생성 수준을 나타내었다. 도 9.

[0708]

다른 세트의 실험에서, 세포를 사이토카인-보충 배지에서 7 내지 10일동안 배양시켰다. 전기천공법 후 제7일에, 게놈 DNA를 단리시킨 후에, CBLB 엑손 6 표적 부위에 대해 PCR 증폭시켰다. 분해에 의한 삽입결실의 추적(분해에 의한 삽입결실의 추적)(TIDE, 문헌[Brinkman et al., 2014] 참조)을 이용하여 삽입결실 빈도를 측정하였다. 도 10은 대표적인 TIDE 분석을 나타낸다.

[0709]

## 실시예 4

[0710] CBLB 메가TAL를 이용하는 예시적인 상동 직접 수선 전략

[0711]

골수증식성 육종 바이러스 인핸서, 결실된 음성 제어 영역, 형광 단백질을 암호화하는 이식유전자(GFP)에 작동 가능하게 연결된 치환된 d1587rev 프라이머-결합 부위(MND) 프로모터, 우드체크 간염 바이러스(WPRE)의 전사후 조절 요소 및 SV40 후기 폴리아데닐화 신호를 포함하는 이식유전자 카세트를 함유하는 아데노-연관 바이러스

(AAV) 플라스미드를 설계하고 작제하였다(도 11, AAV 공여자). XmaI 분해 및 서열분석으로 AAV 플라스미드의 완전성을 확인하였다. CBLB 메가TAL 절단 부위(서열번호 20, 22)에 측정하는 두 0.3kb 상동성 영역 사이에 이식유전자 카세트를 두었다. 상동성 영역 중 어떤 것도 완전한 메가TAL 표적 부위를 함유하지 않았다.

[0712] 하나 이상의 플라스미드와 함께 HEK 293T 세포를 일시적으로 공동형질감염시켜 복제, 캡시드 및 아데노바이러스 헬퍼 요소를 필수로 제공함으로써 재조합 AAV(rAAV)를 제조하였다. 이오디산을 기반 구배에서 초원심분리를 이용하여 공동형질감염된 HEK 293T 세포 배양물로부터 재조합 AAV를 정제하였다.

[0713] 메가TAL-유도 상동성 재조합을 CD3 및 CD28로 활성화시킨 1차 인간 T 세포에서 평가하였고, IL-2로 보충한 완전 배지에서 배양시켰다. 3일 후에, T 세포를 세척하고 나서, CBLB.A8 메가TAL를 암호화하는 시험관내 전사된 mRNA(서열번호 32)로 전기천공시키고, 후속적으로 상기 기재한 MND-GFP 이식유전자 카세트를 포함하는 DNA 공여자 수선 주형을 암호화하는 정제된 재조합 AAV로 형질도입하였다. 형광 단백질을 발현시키는 T 세포의 빈도를 측정하기 위해 그리고 비통합 rAAV 표적화 벡터로부터 형광 단백질의 일시적 발현을 분화시키기 위해 다중 시점에 유세포분석을 사용하였다.

[0714] 메가TAL과 rAAV 표적화 벡터 둘 다로 처리한 T 세포의 21 내지 24%에서 장기간 이식유전자 발현을 관찰하였다. 대조적으로, 비처리 대조군 샘플은 검출 가능한 수준의 형광 단백질을 발현시키지 않았다. 독립적 공여자로부터 단리시킨 T 세포 상에서 수행한 실험에서 결과를 확인하였다.

#### 실시예 5

[0716] CBL-B 편집된 항-EGFR CAR T 세포는 시험관내 종양 모델에서 증가된 항-종양 효능을 나타낸다

[0717] 항-EGFR CAR T 세포를 사용하여 마우스에서 생체내 A549 종양 모델에서의 CAR T 세포 약학에 대한 CBLB 유전자 편집의 효과를 시험하였다.

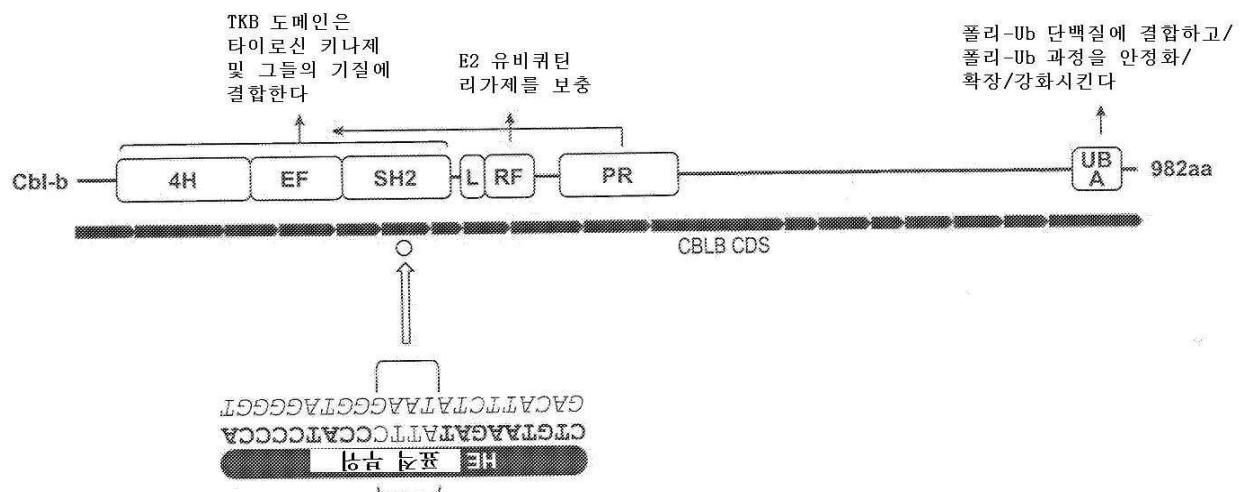
[0718] 인간 PBMC( $1 \times 10^6$  개의 세포/ml)를 제0일에 가용성 항-CD3 및 항-CD28 항체로 활성화시켰다. 24시간의 인큐베이션 후에, 항-EGFR CAR 렌티바이러스로  $1 \times 10^6$  개의 세포를 형질도입하였다. 항체 자극 후 72시간에, 형질도입된 PBMC 또는 비형질도입 대조군 PBMC를 CBLB.E3 메가TAL(서열번호 31)을 암호화하는 시험관내 전사된 mRNA와 함께 또는 이것 없이 전기천공시켰다(론자 뉴클레오펙터(Lonza Nucleofector)). 전기천공된 세포를 IL-2를 함유하는 배지에서 10일 동안 배양시켰다. 배양 후에, 차세대 서열분석(next generation sequencing: NGS) 삽입결실 분석에 의해 평가하여 T 세포의 71%는 CBLB-음성이었다.

[0719] 캘리퍼 측정을 이용하여 상피 암종 세포주(A549) 및 종양 성장의 피하 투여를 받은 마우스를 모니터링하였다. 평균 종양 용적이 120mm<sup>3</sup>에 도달되었을 때, 마우스에 동등한 CAR+ T 세포 용량( $3 \times 10^7$  개의 CAR+ T 세포)로 주입하였다. 비히를 처리된 T 세포 및 비형질도입 대조군 T 세포는 종양 성장에 대해 최소의 영향을 가졌다. CBLB 편집 없이 항-EGFR CAR T 세포로 처리한 동물은 대조군에 비해 종양 증식을 지연시킬 수 있었고; 7마리 마우스 중 2마리는 완전한 종양 제어를 나타내었다. 대조적으로, CBLB 편집된 항-EGFR CAR T 세포를 받은 7마리 마우스는 모두 A549 종양이 완전히 클리어런스되었고, 연구의 지속기간 동안(CAR T 세포 주사 후 35일) 무 종양 상태로 남아있었다.

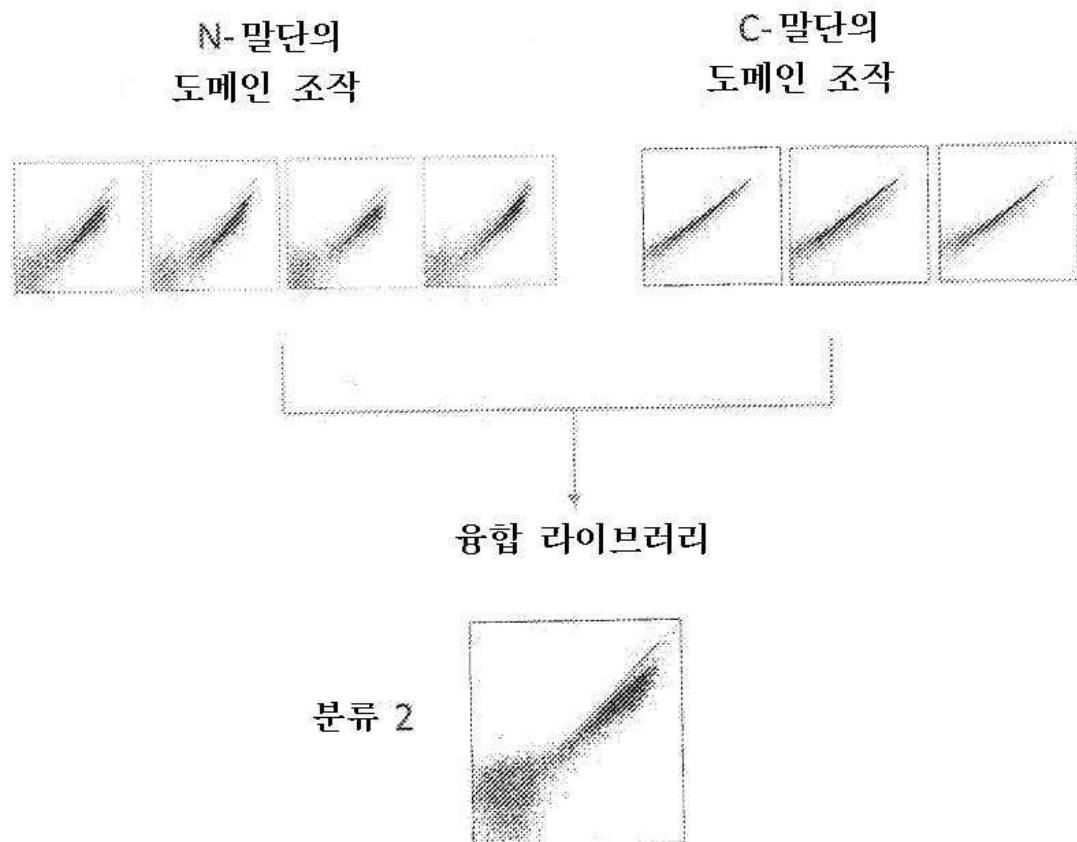
[0720] 일반적으로, 다음의 청구범위에서, 사용하는 용어는 본 명세서 및 청구범위에 개시된 구체적 실시형태로 청구범위를 제한하는 것으로 해석되어서는 안 되며, 이러한 청구범위가 자격을 부여하는 균등론의 완전한 범주와 함께 모든 가능한 실시형태를 포함하는 것으로 해석되어야 한다. 따라서, 청구범위는 본 개시내용에 의해 제한되지 않는다.

## 도면

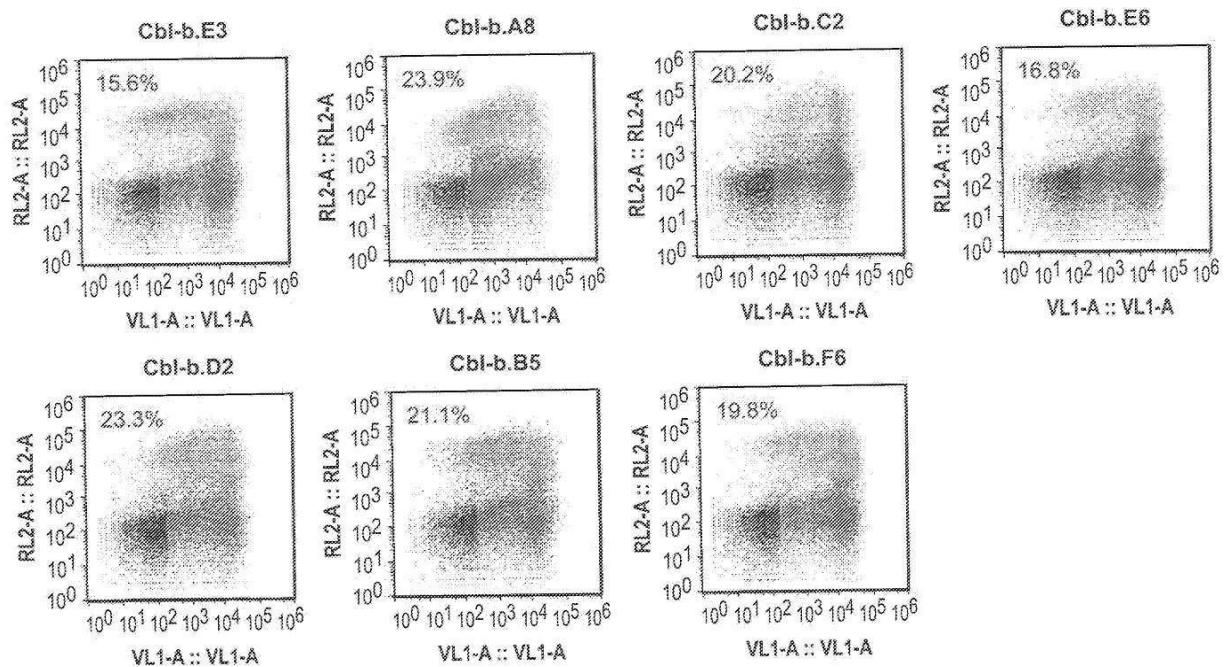
## 도면1



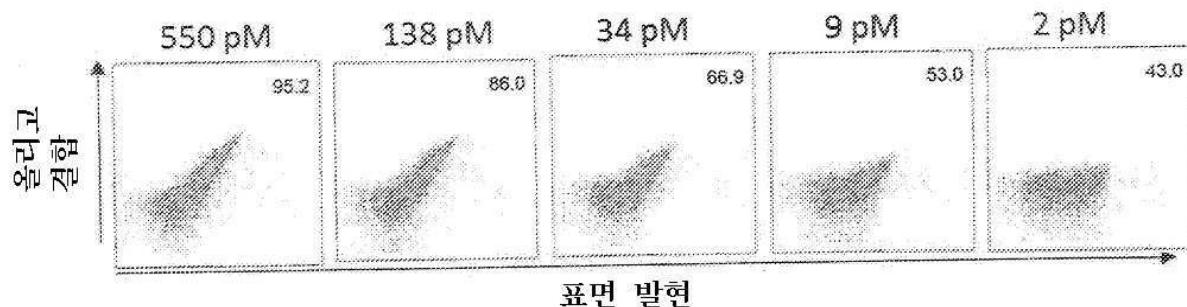
## 도면2



## 도면3



## 도면4



## 도면5a

	1	10	20	30	40	50	60	70
	SINPWILTGFADAEGCFRLDIRNANDLRAGYRTRLSFEIVLHNKDKSILENIQSTWKVGTIYNAGDNAVRL							
1. I-Onul	SINPWILTGFADAEGSFLRIRNNNNKSSVGYSTELGFQITLHNKDKSILENIQSTWKVGVIANSGDNAVSL	1	10	20	30	40	50	60
2. Cblb.E3	SINPWILTGFADAEGCFRLDIRNANDLRAGYRTRLAFEIVLHNKDKSILENIQSTWKVGTIYNAGDNAVRL							70
3. Cblb.F6	SINPWILTGFADAEGCFRLDIRNANDLRAGYRTRLAFEIVLHNKDKSILENIQSTWKVGTIYNAGDNAVRL							
4. Cblb.E6	SINPWILTGFADAEGCFRLDIRNANDLRAGYRTRLSFEISLHNKDKSILENIQSTWKVGTIYNAGDNAVRL							
5. Cblb.D2	SINPWILTGFADAEGCFRLDIHNNANVLRSGYRTRLSFEIVLHNKDKSILENIQSTWKVGTIYNAGDNAVRL							
6. Cblb.C2	SINPWILTGFADAEGCFGLYIHNSNVLRSGYRTRLSFEISLHNKDKSILENIQSTWKVGTIYNAGDNAVRL							
7. Cblb.B5	SINPWILTGFADAEGCFRLDIHNNANVLRSGYRTRLSFEIVLHNKDKSILENIQSTWKVGTIYNAGDNAVRL							
8. Cblb.A8	SINPWILTGFADAEGCFRLDIRNANDLRAGYRTRLAFEIVLHNKDKSILENIQSTWKVGTIYNAGDNAVRL							
	80	90	100	110	120	130	140	
	QVTRFEDLKVIIDHFEKYPLITQKLGDYKLFQAFSVMENKEHLKENGIKELVRIAKAKMNWGLNDELKKAF							
	80	90	100	110	120	130	140	
1. I-Onul	KVTRFEDLKVIIDHFEKYPLITQKLGDYKLFQAFSVMENKEHLKENGIKELVRIAKAKMNWGLNDELKKAF							
2. Cblb.E3	QVTRFEDLKVIIDHFEKYPLITQKLGDYKLFQAFSVMENKEHLKENGIKELVRIAKAKMNWGLNDELKKAF							
3. Cblb.F6	QVTRFEDLKVIIDHFEKYPLITQKLGDYKLFQAFSVMENKEHLKENGIKELVRIAKAKMNWGLNDELKKAF							
4. Cblb.E6	QVTRFEDLKVIIDHFEKYPLITQKLGDYKLFQAFSVMENKEHLKENGIKELVRIAKAKMNWGLNDELKKAF							
5. Cblb.D2	QVTRFEDLKVIIDHFEKYPLITQKLGDYKLFQAFSVMENKEHLKENGIKELVRIAKAKMNWGLNDELKKAF							
6. Cblb.C2	QVTRFEDLKVIIDHFEKYPLITQKLGDYKLFQAFSVMENKEHLKENGIKELVRIAKAKMNWGLNDELKKAF							
7. Cblb.B5	QVTRFEDLKVIIDHFEKYPLITQKLGDYKLFQAFSVMENKEHLKENGIKELVRIAKAKMNWGLNDELKKAF							
8. Cblb.A8	QVTRFEDLKVIIGHFEKYPLITQKLGDYKLFQAFSVMENKEHLKENGIKELVRIAKAKMNWGLNDELKKAF							

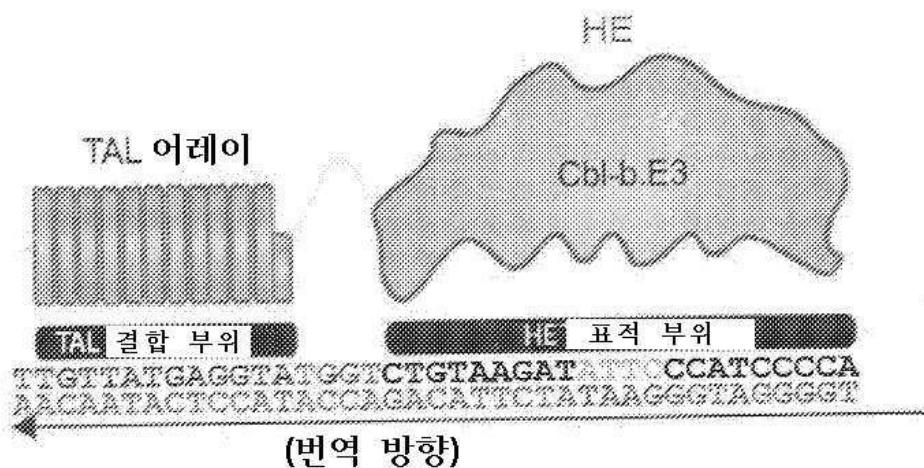
## 도면5b

	150	160	170	180	190	200	210
	PENISKERPLINKIIPNLKWLAFSGDSFMVELMKNKNNIVVRVRLRFSISQHIRDKNLMSLITYLGC						
1. I-Onul	PENISKERSLINKNIPNFKWLAFSGDSFMVELKKRSPVKVVRRLRFSITQHIRDKNLMSLITYLGC	150	160	170	180	190	200
2. Cblb.E3	PENISKERSLINKNIPNFKWLAFSGDSFMVELKKRSPVKVVRRLRFSITQHIRDKNLMSLITYLGC						
3. Cblb.F6	PENISKERPLINKNIPNFKWLAFSGDSFMVELMKNKNNIVVRVRLRFSISQHIRDKNLMSLITYLGC						
4. Cblb.E6	PENISKERPLINKNIPNFKWLAFSGDSFMVELMKNKNNIVVRVRLRFSISQHIRDKNLMSLITYLGC						
5. Cblb.D2	PENISKERPLINKNIPNFKWLAFSGDSFMVELMKNKNNIVVRVRLRFSISQHIRDKNLMSLITYLGC						
6. Cblb.C2	PENISKERPLINKNIPNFKWLAFSGDSFMVELMKNKNNIVVRVRLRFSISQHIRDKNLMSLITYLGC						
7. Cblb.B5	PENISKERPLINKNIPNFKWLAFSGDSFMVELMKNKNNIVVRVRLRFSISQHIRDKNLMSLITYLGC						
8. Cblb.A8	PENISKERPLINKNIPNFKWLAFSGDSFMVELMKNKNNIVVRVRLRFSISQHIRDKNLMSLITYLGC						
	220	230	240	250	260	270	280
	GRIENNPKSEHSLWLEFIVTKFSDINDKIIIPVFOENTLIGVKLEDfedWCKVAKLIEEKKHLLTESGLDEIKK						
	220	230	240	250	260	270	280
1. I-Onul	GYIKEKNKSEFSWLDVVTKFSINDKIIIPVFOENTLIGVKLEDfedWCKVAKLIEEKKHLLTESGLDEIKK						
2. Cblb.E3	GRIENNPKSEHSLWLEFIVTKFSDINDKIIIPVFOENTLIGVKLEDfedWCKVAKLIEEKKHLLTESGLDEIKK						
3. Cblb.F6	GRIENNPKSEHSLWLEFIVTKFSDINDKIIIPVFOENTLIGVKLEDfedWCKVAKLIEEKKHLLTESGLDEIKK						
4. Cblb.E6	GRIENNPKSEHSLWLEFIVTKFSDINDKIIIPVFOENTLIGVKLEDfedWCKVAKLIEEKKHLLTESGLDEIKK						
5. Cblb.D2	GRIENNPKSEHSLWLEFIVTKFSDINDKIIIPVFOENTLIGVKLEDfedWCKVAKLIEEKKHLLTESGLDEIKK						
6. Cblb.C2	GRIENNPKSEHSLWLEFIVTKFSDINDKIIIPVFOENTLIGVKLEDfedWCKVAKLIEEKKHLLTESGLDEIKK						
7. Cblb.B5	GRIENNPKSEHSLWLEFIVTKFSDINDKIIIPVFOENTLIGVKLEDfedWCKVAKLIEEKKHLLTESGLDEIKK						
8. Cblb.A8	GRIENNPKSEHSLWLEFIVTKFSDINDKIIIPVFOENTLIGVKLEDfedWCKVAKLIEEKKHLLTESGLDEIKK						

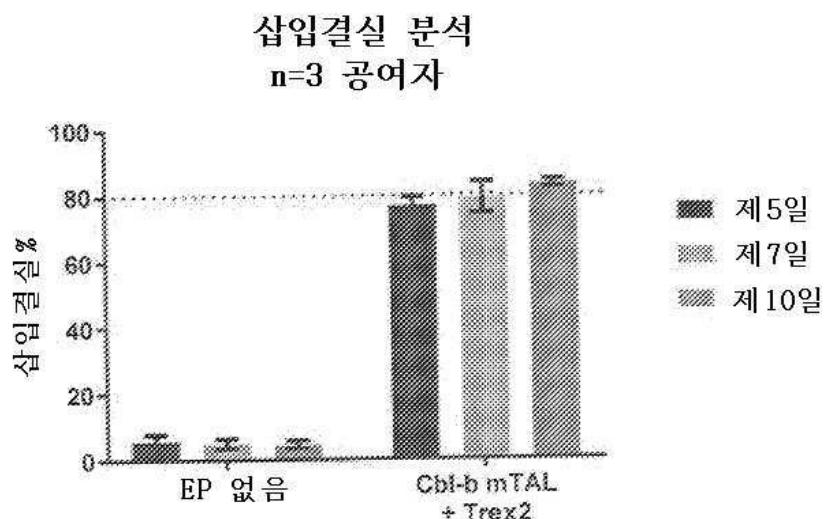
## 도면5c

	290	300
	IKNMNKGRVFSGR**	
	290 293	
1. I-Onul	IKNMNKGR	
2. Cblb.E3	IKNMNKGRVFSGR**	
3. Cblb.F6	IKNMNKGRVFSGR**	
4. Cblb.E6	IKNMNKGRVFSGR**	
5. Cblb.D2	IKNMNKGRVFSGR**	
6. Cblb.C2	IKNMNKGRVFSGR**	
7. Cblb.B5	IKNMNKGRVFSGR**	
8. Cblb.A8	IKNMNKGRVFSGR**	

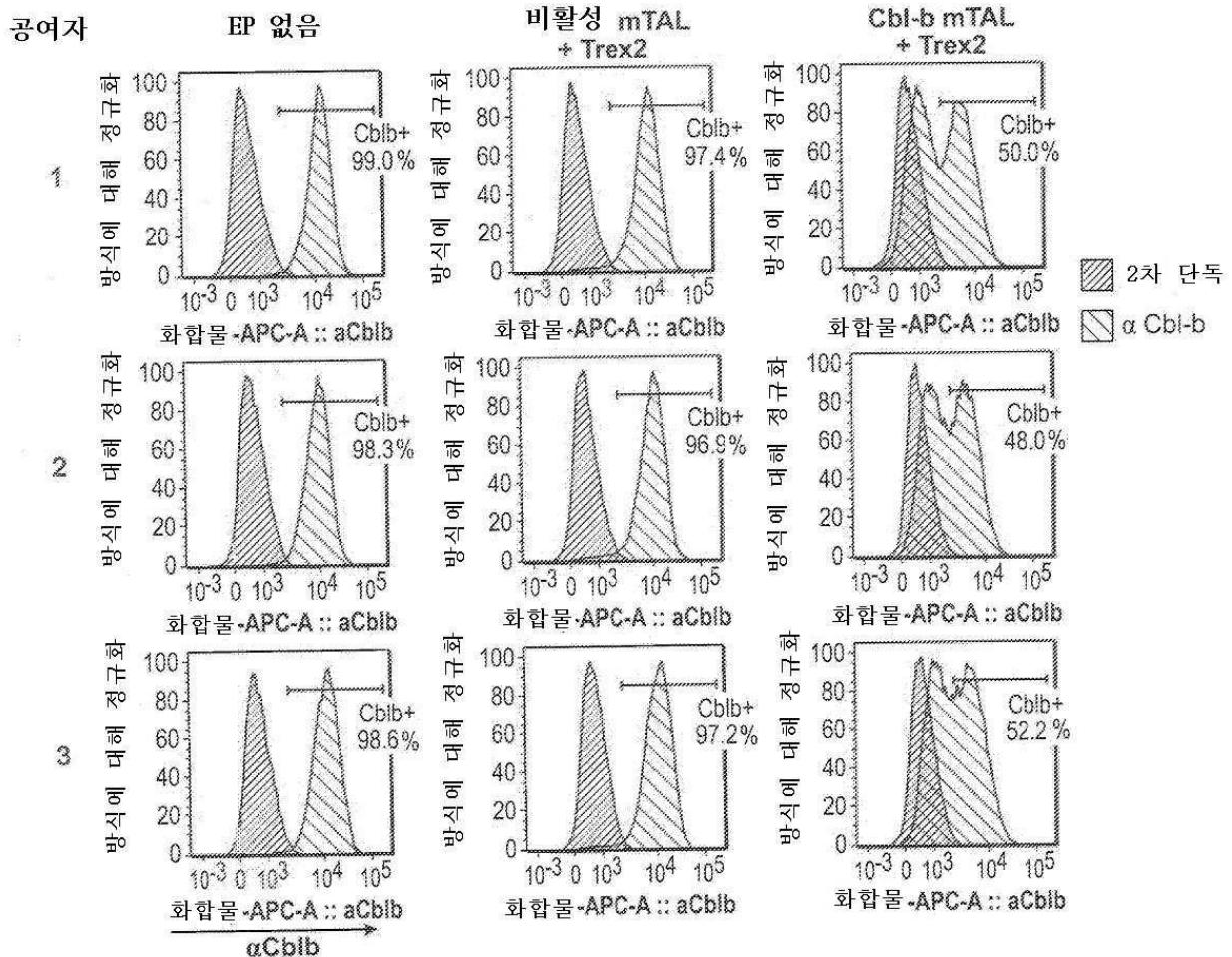
도면6



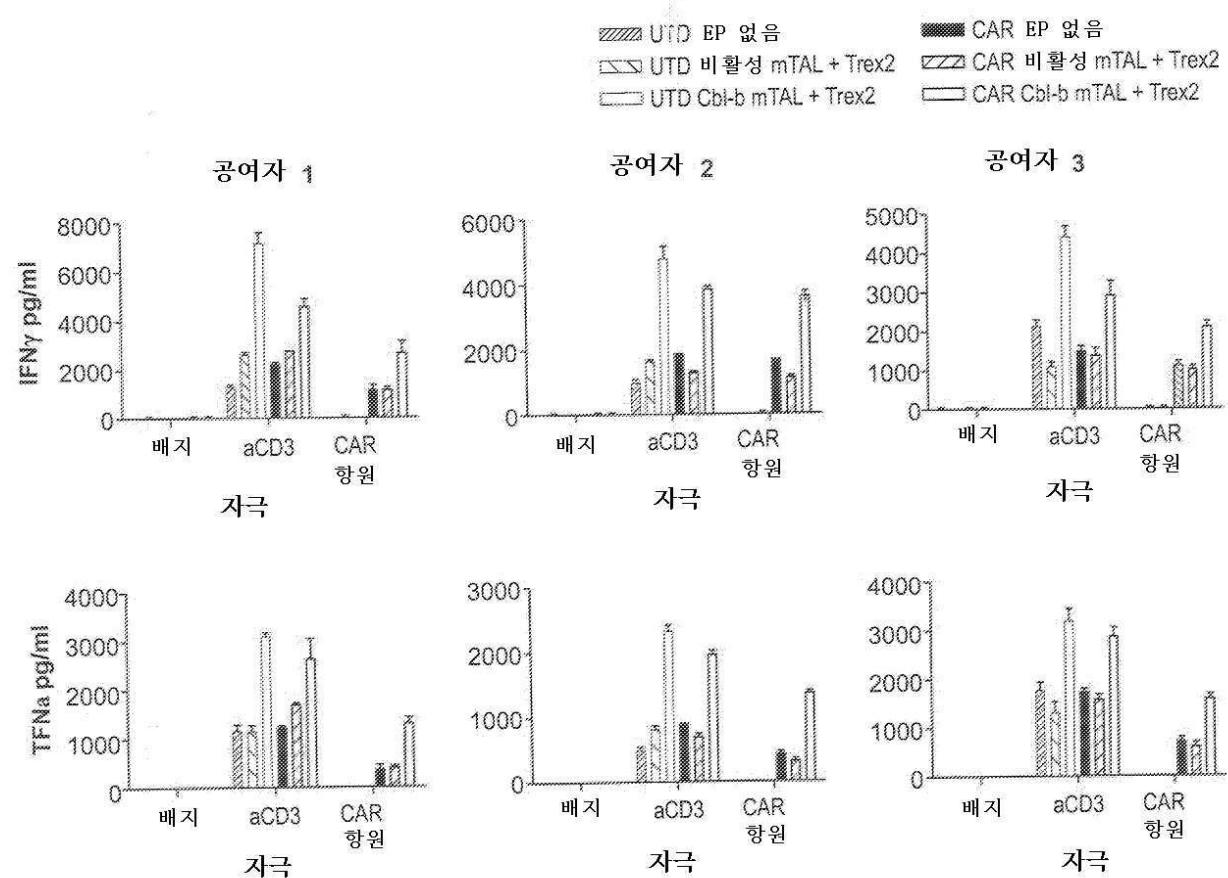
도면7



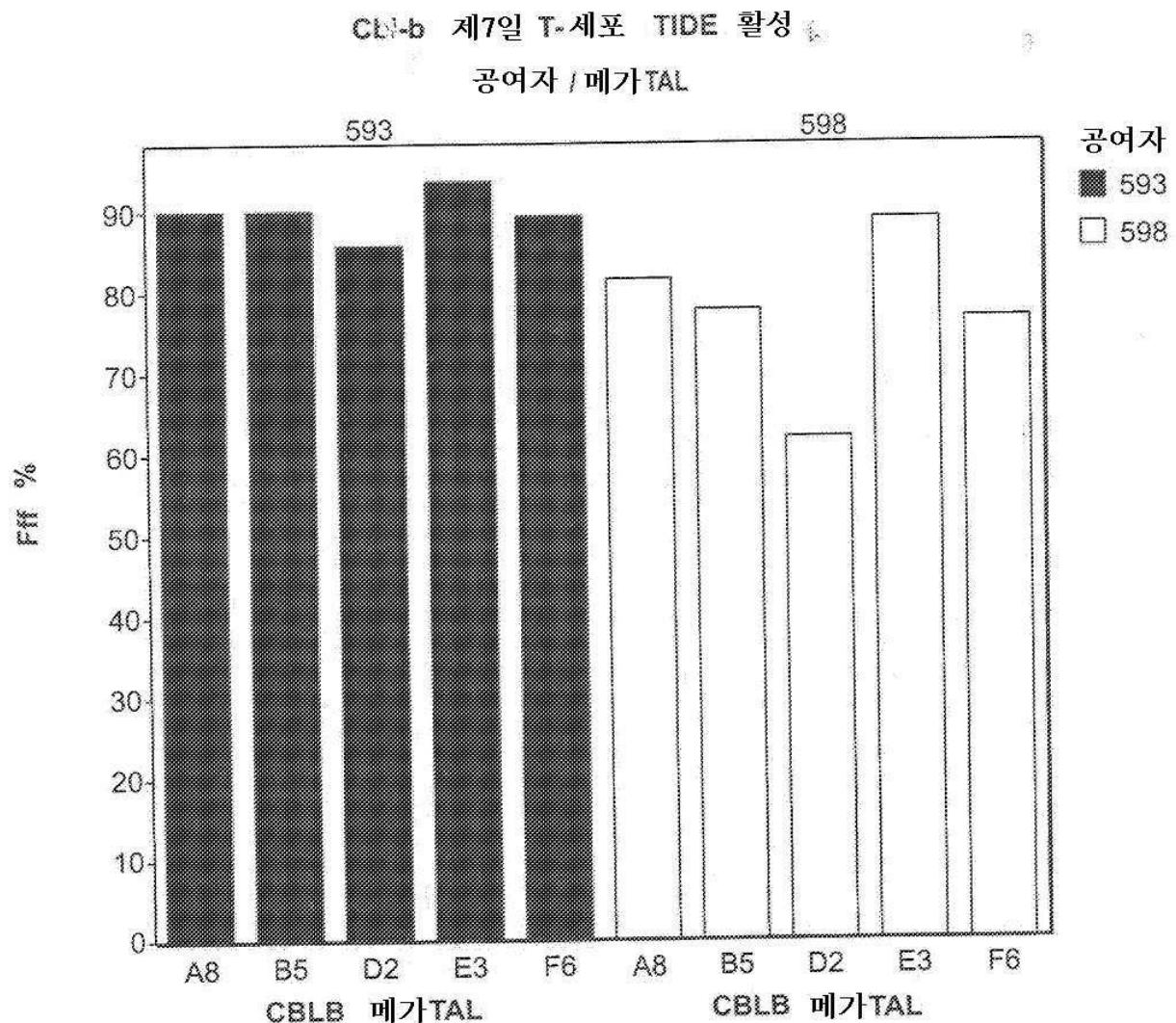
## 도면8



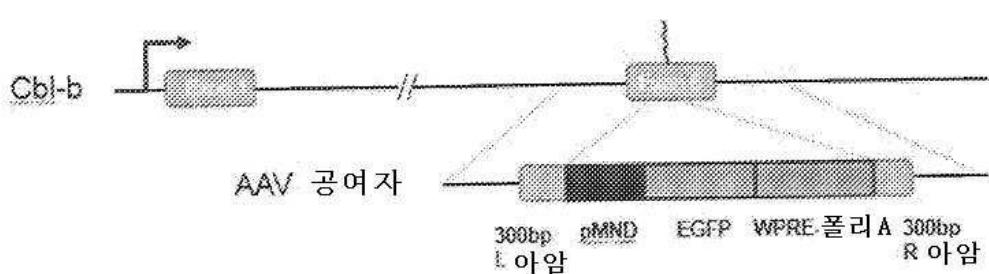
## 도면9



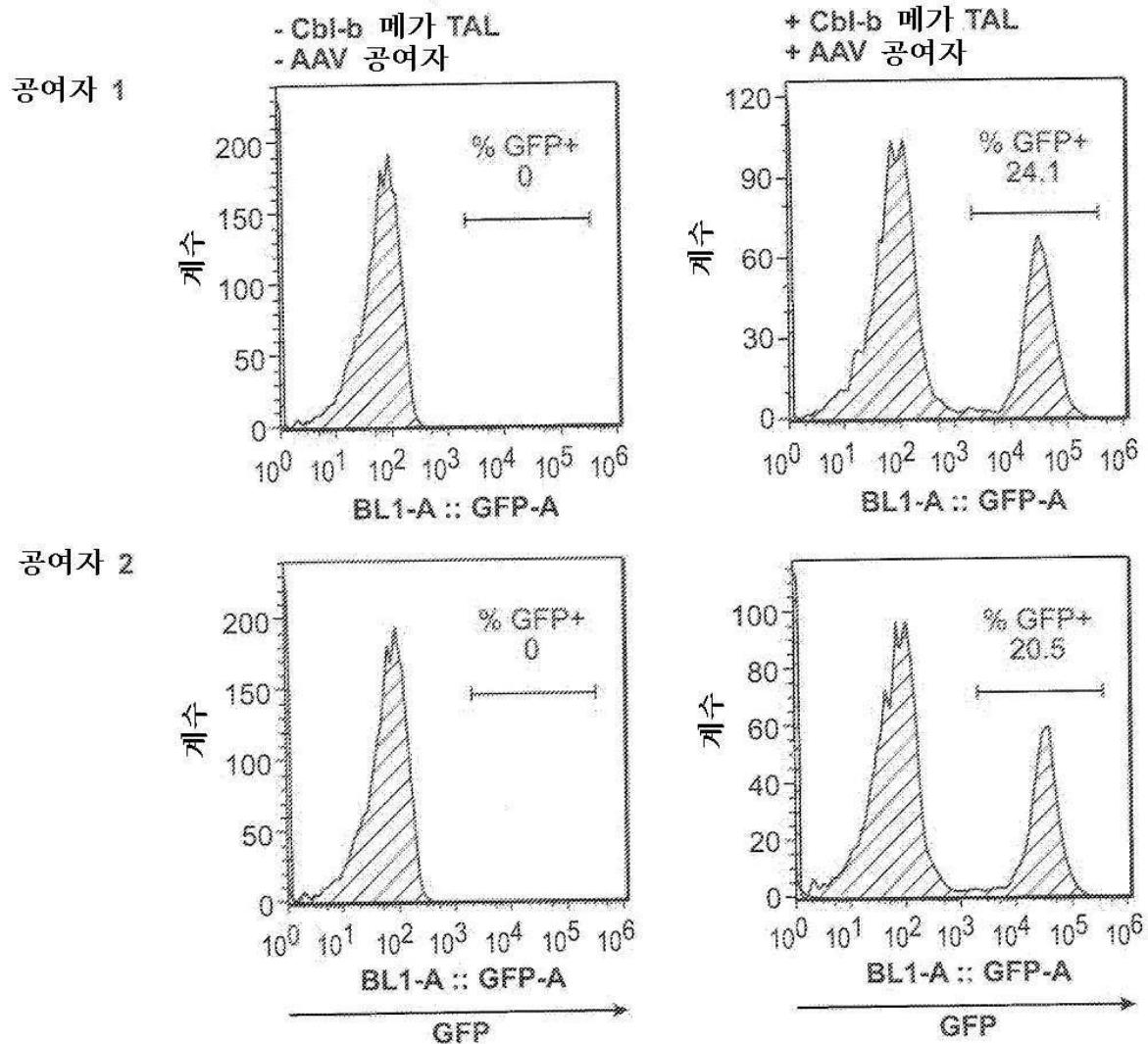
## 도면10



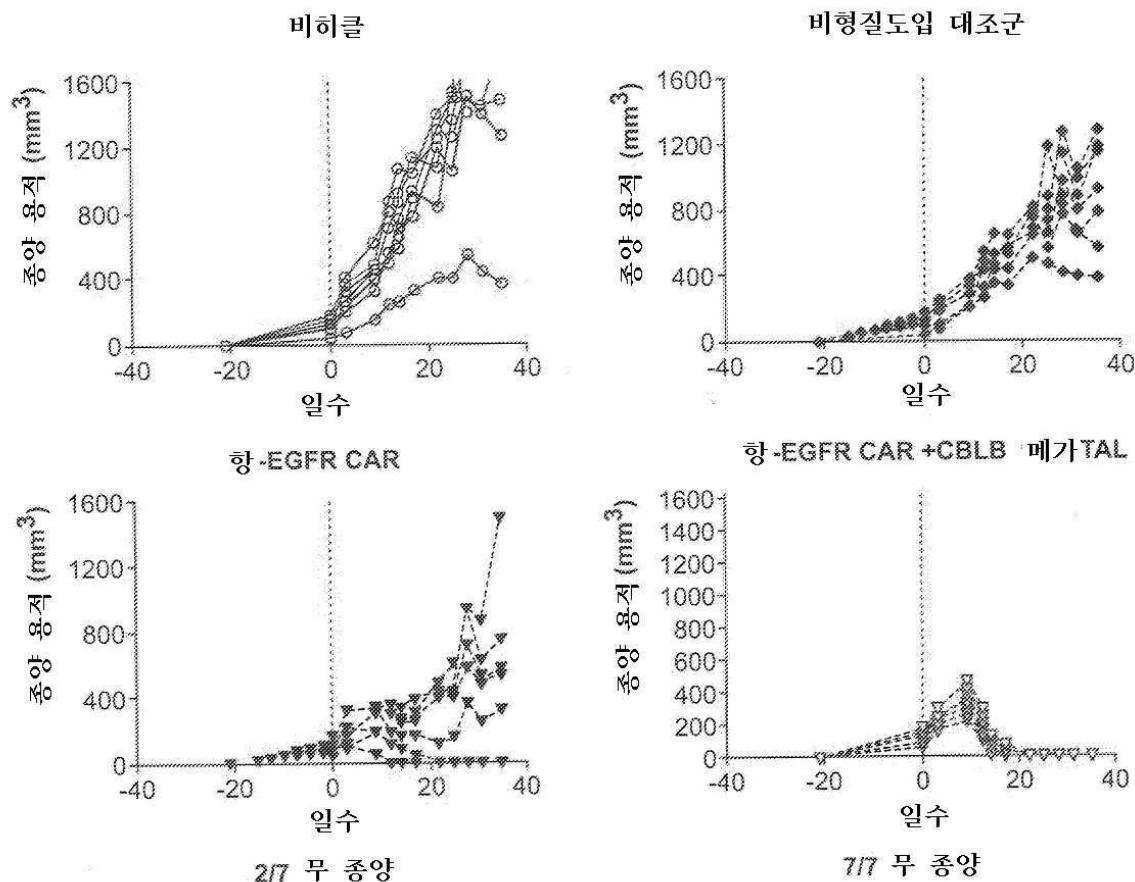
## 도면11



## 도면12



## 도면13



## 서 열 목록

- <110> bluebird bio, Inc.  
 Jarjour, Jordan  
 Havens, Kyle  
 Krostag, Anne-Rachel
- <120> CBLB ENDONUCLEASE VARIANTS, COMPOSITIONS, AND METHODS OF USE
- <130> IPA191517-US-D1
- <150> US 62/567,417
- <151> 2017-10-30
- <150> US 62/511,194
- <151> 2017-05-25
- <160> 84
- <170> Patent In version 3.5
- <210> 1
- <211> 303
- <212> PRT

&lt;213&gt; Ophiostoma novo-ulmi

&lt;400&gt; 1

Met Ala Tyr Met Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr

1 5 10 15

Gly Phe Ala Asp Ala Glu Gly Ser Phe Leu Leu Arg Ile Arg Asn Asn

20 25 30

Asn Lys Ser Ser Val Gly Tyr Ser Thr Glu Leu Gly Phe Gln Ile Thr

35 40 45

Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp

50 55 60

Lys Val Gly Val Ile Ala Asn Ser Gly Asp Asn Ala Val Ser Leu Lys

65 70 75 80

Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys

85 90 95

Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Met Leu Phe Lys Gln

100 105 110

Ala Phe Cys Val Met Glu Asn Lys Glu His Leu Lys Ile Asn Gly Ile

115 120 125

Lys Glu Leu Val Arg Ile Lys Ala Lys Leu Asn Trp Gly Leu Thr Asp

130 135 140

Glu Leu Lys Lys Ala Phe Pro Glu Ile Ile Ser Lys Glu Arg Ser Leu

145 150 155 160

Ile Asn Lys Asn Ile Pro Asn Phe Lys Trp Leu Ala Gly Phe Thr Ser

165 170 175

Gly Glu Gly Cys Phe Phe Val Asn Leu Ile Lys Ser Lys Ser Lys Leu

180 185 190

Gly Val Gln Val Gln Leu Val Phe Ser Ile Thr Gln His Ile Lys Asp

195 200 205

Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Tyr Ile

210 215 220

Lys Glu Lys Asn Lys Ser Glu Phe Ser Trp Leu Asp Phe Val Val Thr

225	230	235	240
Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn			
245	250	255	
Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val			
260	265	270	
Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp			
275	280	285	
Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Val Phe			
290	295	300	

<210>	2
<211>	303
<212>	PRT
<213>	Ophiostoma novo-ulmi
<400>	2

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

50	55	60	
Lys Val Gly Val Ile Ala Asn Ser Gly Asp Asn Ala Val Ser Leu Lys			
65	70	75	80
Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys			
85	90	95	
Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln			
100	105	110	
Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile			
115	120	125	

Lys Glu Leu Val Arg Ile Lys Ala Lys Leu Asn Trp Gly Leu Thr Asp			
130	135	140	

Glu Leu Lys Lys Ala Phe Pro Glu Asn Ile Ser Lys Glu Arg Ser Leu  
 145 150 155 160  
 Ile Asn Lys Asn Ile Pro Asn Phe Lys Trp Leu Ala Gly Phe Thr Ser  
 165 170 175  
 Gly Glu Gly Cys Phe Phe Val Asn Leu Ile Lys Ser Lys Ser Lys Leu  
 180 185 190  
 Gly Val Gln Val Gln Leu Val Phe Ser Ile Thr Gln His Ile Lys Asp  
  
 195 200 205  
 Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Tyr Ile  
 210 215 220  
 Lys Glu Lys Asn Lys Ser Glu Phe Ser Trp Leu Asp Phe Val Val Thr  
 225 230 235 240  
 Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn  
 245 250 255  
 Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val  
 260 265 270  
  
 Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp  
 275 280 285  
 Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Val Phe  
 290 295 300  
 <210> 3  
 <211> 303  
 <212> PRT  
 <213> Ophiostoma novo-ulmi  
 <220><221> MOD\_RES  
 <222> (1)..(3)  
 <223> Any amino acid or absent  
 <400> 3  
  
 Xaa Xaa Xaa Met Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr  
 1 5 10 15

Gly Phe Ala Asp Ala Glu Gly Ser Phe Leu Leu Arg Ile Arg Asn Asn  
 20 25 30

Asn Lys Ser Ser Val Gly Tyr Ser Thr Glu Leu Gly Phe Gln Ile Thr  
 35 40 45  
 Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp  
 50 55 60  
 Lys Val Gly Val Ile Ala Asn Ser Gly Asp Asn Ala Val Ser Leu Lys  
 65 70 75 80  
 Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys  
  
 85 90 95  
 Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln  
 100 105 110  
 Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile  
 115 120 125  
 Lys Glu Leu Val Arg Ile Lys Ala Lys Leu Asn Trp Gly Leu Thr Asp  
 130 135 140  
 Glu Leu Lys Ala Phe Pro Glu Asn Ile Ser Lys Glu Arg Ser Leu  
 145 150 155 160  
  
 Ile Asn Lys Asn Ile Pro Asn Phe Lys Trp Leu Ala Gly Phe Thr Ser  
 165 170 175  
 Gly Glu Gly Cys Phe Phe Val Asn Leu Ile Lys Ser Lys Ser Lys Leu  
 180 185 190  
 Gly Val Gln Val Gln Leu Val Phe Ser Ile Thr Gln His Ile Lys Asp  
 195 200 205  
 Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Tyr Ile  
 210 215 220  
 Lys Glu Lys Asn Lys Ser Glu Phe Ser Trp Leu Asp Phe Val Val Thr  
  
 225 230 235 240  
 Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn  
 245 250 255  
 Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val  
 260 265 270  
 Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp  
 275 280 285

Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Val Phe

290 295 300

<210> 4

<211> 303

<212> PRT

<213> Ophiostoma novo-ulmi

<220><221> MOD\_RES

<222> (1)..(4)

<223> Any amino acid or absent

<220><221> MOD\_RES

<222> (302)..(303)

<223> Any amino acid or absent

<400> 4

Xaa Xaa Xaa Xaa Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr

1 5 10 15

Gly Phe Ala Asp Ala Glu Gly Ser Phe Leu Leu Arg Ile Arg Asn Asn

20 25 30

Asn Lys Ser Ser Val Gly Tyr Ser Thr Glu Leu Gly Phe Gln Ile Thr

35 40 45

Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp

50 55 60

Lys Val Gly Val Ile Ala Asn Ser Gly Asp Asn Ala Val Ser Leu Lys

65 70 75 80

Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys

85 90 95

Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln

100 105 110

Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile

115 120 125

Lys Glu Leu Val Arg Ile Lys Ala Lys Leu Asn Trp Gly Leu Thr Asp

130 135 140

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

145 150 155 160  
 Ile Asn Lys Asn Ile Pro Asn Phe Lys Trp Leu Ala Gly Phe Thr Ser  
 165 170 175  
 Gly Glu Gly Cys Phe Phe Val Asn Leu Ile Lys Ser Lys Ser Lys Leu  
  
 180 185 190  
 Gly Val Gln Val Gln Leu Val Phe Ser Ile Thr Gln His Ile Lys Asp  
 195 200 205  
 Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Tyr Ile  
 210 215 220  
 Lys Glu Lys Asn Lys Ser Glu Phe Ser Trp Leu Asp Phe Val Val Thr  
 225 230 235 240  
 Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn  
 245 250 255  
  
 Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val  
 260 265 270  
 Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp  
 275 280 285  
 Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Xaa Xaa  
 290 295 300  
 <210> 5  
 <211> 303  
 <212> PRT  
 <213> Ophiostoma novo-ulmi  
 <220><221> MOD\_RES  
 <222> (1)..(8)  
 <223> Any amino acid or absent  
 <220>  
 <221> MOD\_RES  
 <222> (302)..(303)  
 <223> Any amino acid or absent  
 <400> 5  
  
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Ile Asn Pro Trp Ile Leu Thr  
 1 5 10 15

Gly Phe Ala Asp Ala Glu Gly Ser Phe Leu Leu Arg Ile Arg Asn Asn

20 25 30

Asn Lys Ser Ser Val Gly Tyr Ser Thr Glu Leu Gly Phe Gln Ile Thr

35 40 45

Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp

50 55 60

Lys Val Gly Val Ile Ala Asn Ser Gly Asp Asn Ala Val Ser Leu Lys

65 70 75 80

Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys

85 90 95

Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln

100 105 110

Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile

115 120 125

Lys Glu Leu Val Arg Ile Lys Ala Lys Leu Asn Trp Gly Leu Thr Asp

130 135 140

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

145 150 155 160

Ile Asn Lys Asn Ile Pro Asn Phe Lys Trp Leu Ala Gly Phe Thr Ser

165 170 175

Gly Glu Gly Cys Phe Phe Val Asn Leu Ile Lys Ser Lys Ser Lys Leu

180 185 190

Gly Val Gln Val Gln Leu Val Phe Ser Ile Thr Gln His Ile Lys Asp

195 200 205

Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Tyr Ile

210 215 220

Lys Glu Lys Asn Lys Ser Glu Phe Ser Trp Leu Asp Phe Val Val Thr

225 230 235 240

Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn

245 250 255

Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val

260 265 270

Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp

275 280 285

Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Xaa Xaa

290 295 300

&lt;210&gt; 6

&lt;211&gt; 306

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Synthesized I-OnuI variant CBLB.E3

&lt;220&gt;&lt;221&gt; MOD\_RES

&lt;222&gt; (1)..(8)

&lt;223&gt; Any amino acid or absent

&lt;220&gt;&lt;221&gt; MOD\_RES

&lt;222&gt; (302)..(306)

&lt;223&gt; Any amino acid or absent

&lt;400&gt; 6

Xaa Xaa Xaa Xaa Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr

1 5 10 15

Gly Phe Ala Asp Ala Glu Gly Cys Phe Arg Leu Asp Ile Arg Asn Ala

20 25 30

Asn Asp Leu Arg Ala Gly Tyr Arg Thr Arg Leu Ala Phe Glu Ile Val

35 40 45

Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp

50 55 60

Lys Val Gly Thr Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln

65 70 75 80

Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys

85 90 95

Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln

100 105 110

Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile

115	120	125
Lys Glu Leu Val Arg Ile Lys Ala Lys Met Asn Trp Gly Leu Asn Asp		
130	135	140
Glu Leu Lys Lys Ala Phe Pro Glu Asn Ile Ser Lys Glu Arg Ser Leu		
145	150	155
Ile Asn Lys Asn Ile Pro Asn Leu Lys Trp Leu Ala Gly Phe Thr Ser		
165	170	175
Gly Asp Gly Ser Phe Val Val Glu Leu Lys Lys Arg Arg Ser Pro Val		
180	185	190
Lys Val Gly Val Arg Leu Arg Phe Ser Ile Thr Gln His Ile Arg Asp		
195	200	205
Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile		
210	215	220
Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr		
225	230	235
Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn		
245	250	255
Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val		
260	265	270
Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp		
275	280	285
Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Xaa Xaa Xaa		
290	295	300
Xaa Xaa		
305		
<210>	7	
<211>	306	
<212>	PRT	
<213>	Artificial Sequence	
<220><223>	Synthesized I-OnuI variant CBLB.F6	
<220><221>	MOD_RES	
<222>	(1)..(8)	

<223> Any amino acid or absent

<220><221> MOD\_RES

<222> (302)..(306)

<223> Any amino acid or absent

<400> 7

Xaa Xaa Xaa Xaa Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr

1	5	10	15
---	---	----	----

Gly Phe Ala Asp Ala Glu Gly Cys Phe Arg Leu Asp Ile Arg Asn Ala

20	25	30
----	----	----

Asn Asp Leu Arg Ala Gly Tyr Arg Thr Arg Leu Ala Phe Glu Ile Val

35	40	45
----	----	----

Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp

50	55	60
----	----	----

Lys Val Gly Thr Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln

65	70	75	80
----	----	----	----

Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys

85	90	95
----	----	----

Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln

100	105	110
-----	-----	-----

Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile

115	120	125
-----	-----	-----

Lys Glu Leu Val Arg Ile Lys Ala Lys Met Asn Trp Gly Leu Asn Asp

130	135	140
-----	-----	-----

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

145	150	155	160
-----	-----	-----	-----

Ile Asn Lys Asn Ile Pro Asn Leu Lys Trp Leu Ala Gly Phe Thr Ser

165	170	175
-----	-----	-----

Gly Asp Gly Ser Phe Met Val Glu Leu Met Lys Asn Lys Asn Asn Val

180	185	190
-----	-----	-----

Ile Val Arg Val Arg Leu Arg Phe Ser Ile Ser Gln His Ile Arg Asp

195	200	205
-----	-----	-----

Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile

210	215	220	
Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr			
225	230	235	240
Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn			
245	250	255	
Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val			
260	265	270	
Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp			
275	280	285	
Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Xaa Xaa Xaa			
290	295	300	

Xaa Xaa

305

<210> 8

<211> 306

<212> PRT

<213> Artificial Sequence

<220><223> Synthesized I-OnuI variant CBLB.E6

<220><221> MOD\_RES

<222> (1)..(8)

<223> Any amino acid or absent

<220><221> MOD\_RES

<222> (302)..(306)

<223> Any amino acid or absent

<400> 8

Xaa Xaa Xaa Xaa Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr

1	5	10	15
---	---	----	----

Gly Phe Ala Asp Ala Glu Gly Cys Phe Arg Leu Asp Ile Arg Asn Ala

20	25	30	
Asn Asp Leu Arg Ala Gly Tyr Arg Thr Arg Leu Ser Phe Glu Ile Ser			
35	40	45	
Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp			

50	55	60
Lys Val Gly Thr Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln		
65	70	75
Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys		
85	90	95
Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln		
100	105	110
Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile		
115	120	125
Lys Glu Leu Val Arg Ile Lys Ala Lys Met Asn Trp Gly Leu Asn Asp		
130	135	140
Glu Leu Lys Ala Phe Pro Glu Asn Ile Ser Lys Glu Arg Pro Leu		
145	150	155
Ile Asn Lys Asn Ile Pro Asn Leu Lys Trp Leu Ala Gly Phe Thr Ser		
165	170	175
Gly Asp Gly Ser Phe Met Val Glu Leu Met Lys Asn Lys Asn Val		
180	185	190
Ile Val Arg Val Arg Leu Arg Phe Ser Ile Ser Gln His Ile Arg Asp		
195	200	205
Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile		
210	215	220
Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr		
225	230	235
Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn		
245	250	255
Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val		
260	265	270
Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp		
275	280	285
Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Xaa Xaa Xaa		
290	295	300
Xaa Xaa		

305

&lt;210&gt;

9

&lt;211&gt; 306

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Synthesized I-OnuI variant CBLB.D2

&lt;220&gt;&lt;221&gt; MOD\_RES

&lt;222&gt; (1)..(8)

&lt;223&gt; Any amino acid or absent

&lt;220&gt;&lt;221&gt; MOD\_RES

&lt;222&gt; (302)..(306)

&lt;223&gt; Any amino acid or absent

&lt;400&gt; 9

Xaa Xaa Xaa Xaa Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr

1 5 10 15

Gly Phe Ala Asp Ala Glu Gly Cys Phe Arg Leu Asp Ile His Asn Ala

20 25 30

Asn Val Leu Arg Ser Gly Tyr Arg Thr Arg Leu Ser Phe Glu Ile Val

35 40 45

Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp

50 55 60

Lys Val Gly Lys Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln

65 70 75 80

Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys

85 90 95

Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln

100 105 110

Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile

115 120 125

Lys Glu Leu Val Arg Ile Lys Ala Lys Met Asn Trp Gly Leu Asn Asp

130 135 140

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

145	150	155	160
Ile Asn Lys Asn Ile Pro Asn Leu Lys Trp Leu Ala Gly Phe Thr Ser			
165	170	175	
Gly Asp Gly Ser Phe Met Val Glu Leu Met Lys Asn Lys Asn Asn Val			
180	185	190	
Ile Val Arg Val Arg Leu Arg Phe Ser Ile Ser Gln His Ile Arg Asp			
195	200	205	
Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile			
210	215	220	
Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr			
225	230	235	240
Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn			
245	250	255	
Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val			
260	265	270	
Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp			
275	280	285	
Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Xaa Xaa Xaa			
290	295	300	
Xaa Xaa			
305			
<210>	10		
<211>	306		
<212>	PRT		
<213>	Artificial Sequence		

<220><223> Synthesized I-OnuI variant CBLB.C2

<220><221> MOD\_RES

<222> (1)..(8)

<223> Any amino acid or absent

<220><221> MOD\_RES

<222> (302)..(306)

<223> Any amino acid or absent

&lt;400&gt; 10

Xaa Xaa Xaa Xaa Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr

1 5 10 15

Gly Phe Ala Asp Ala Glu Gly Cys Phe Gly Leu Tyr Ile His Asn Ser

20 25 30

Asn Val Leu Arg Ser Gly Tyr Arg Thr Arg Leu Ser Phe Glu Ile Ser

35 40 45

Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp

50 55 60

Lys Val Gly Thr Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln

65 70 75 80

Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys

85 90 95

Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln

100 105 110

Ala Phe Ser Leu Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile

115 120 125

Lys Glu Leu Val Arg Ile Lys Ala Lys Met Asn Trp Gly Leu Asn Asp

130 135 140

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

145 150 155 160

Ile Asn Lys Asn Ile Pro Asn Leu Lys Trp Leu Ala Gly Phe Thr Ser

165 170 175

Gly Asp Gly Ser Phe Met Val Glu Leu Met Lys Asn Lys Asn Val

180 185 190

Ile Val Arg Val Arg Leu Arg Phe Ser Ile Ser Gln His Ile Arg Asp

195 200 205

Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile

210 215 220

Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr

225 230 235 240

Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn

245 250 255

Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val

260 265 270

Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp

275 280 285

Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Xaa Xaa Xaa

290 295 300

Xaa Xaa

305

&lt;210&gt; 11

&lt;211&gt; 306

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Synthesized I-OnuI variant CBLB.B5

&lt;220&gt;&lt;221&gt;

&gt; MOD\_RES

&lt;222&gt; (1)..(8)

&lt;223&gt; Any amino acid or absent

&lt;220&gt;&lt;221&gt; MOD\_RES

&lt;222&gt; (302)..(306)

&lt;223&gt; Any amino acid or absent

&lt;400&gt; 11

Xaa Xaa Xaa Xaa Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr

1 5 10 15

Gly Phe Ala Asp Ala Glu Gly Cys Phe Arg Leu Asp Ile His Asn Ala

20 25 30

Asn Val Leu Arg Ser Gly Tyr Arg Thr Arg Leu Ser Phe Glu Ile Val

35 40 45

Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp

50 55 60

Lys Val Gly Thr Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln

65 70 75 80

Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys

85	90	95	
Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln			
100	105	110	
Ala Phe Ser Leu Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile			
115	120	125	
Lys Glu Leu Val Arg Ile Lys Ala Lys Met Asn Trp Gly Leu Asn Asp			
130	135	140	
Glu Leu Lys Ala Phe Pro Glu Asn Ile Ser Lys Glu Arg Pro Leu			
145	150	155	160
Ile Asn Lys Asn Ile Pro Asn Leu Lys Trp Leu Ala Gly Phe Thr Ser			
165	170	175	
Gly Asp Gly Ser Phe Val Val Glu Leu Lys Lys Arg Arg Ser Pro Val			
180	185	190	
Lys Val Gly Val Arg Leu Arg Phe Gly Ile Thr Gln His Ile Arg Asp			
195	200	205	
Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile			
210	215	220	
Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr			
225	230	235	240
Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn			
245	250	255	
Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val			
260	265	270	
Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp			
275	280	285	
Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Xaa Xaa Xaa			
290	295	300	
Xaa Xaa			
305			
<210>	12		
<211>	306		
<212>	PRT		

<213> Artificial Sequence

<220><223> Synthesized I-OnuI variant CBLB.A8

<220><221> MOD\_RES

<222> (1)..(8)

<223> Any amino acid or absent

<220><221> MOD\_RES

<222> (302)..(306)

<223> Any amino acid or absent

<400> 12

Xaa Xaa Xaa Xaa Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr

1 5 10 15

Gly Phe Ala Asp Ala Glu Gly Cys Phe Arg Leu Asp Ile Arg Asn Ala

20 25 30

Asn Asp Leu Arg Ala Gly Tyr Arg Thr Arg Leu Ala Phe Glu Ile Val

35 40 45

Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp

50 55 60

Lys Val Gly Thr Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln

65 70 75 80

Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Gly His Phe Glu Lys

85 90 95

Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln

100 105 110

Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile

115 120 125

Lys Glu Leu Val Arg Ile Lys Ala Lys Met Asn Trp Gly Leu Asn Asp

130 135 140

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

145 150 155 160

Ile Asn Lys Asn Ile Pro Asn Leu Lys Trp Leu Ala Gly Phe Thr Ser

165 170 175

Gly Asp Gly Ser Phe Met Val Glu Leu Met Lys Asn Lys Asn Val

180	185	190
Ile Val Arg Val Arg Leu Arg Phe Ser Ile Ser Gln His Ile Arg Asp		
195	200	205
Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile		
210	215	220
Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr		
225	230	235
Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn		
245	250	255
Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val		
260	265	270
Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp		
275	280	285
Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Xaa Xaa Xaa		
290	295	300
Xaa Xaa		
305		
<210> 13		
<211> 914		
<212> PRT		
<213> Artificial Sequence		
<220><223> Made in Lab - synthesized megaTAL CBLB.E3 construct		
<400> 13		
Met Gly Ser Ala Pro Pro Lys Lys Arg Lys Val Val Asp Leu Arg		
1	5	10
Thr Leu Gly Tyr Ser Gln Gln Gln Glu Lys Ile Lys Pro Lys Val		
20	25	30
Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe		
35	40	45
Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly		
50	55	60
Thr Val Ala Val Thr Tyr Gln His Ile Ile Thr Ala Leu Pro Glu Ala		

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

Pro Asp Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala

325 330 335

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

340 345 350

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys

355 360 365

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp

370 375 380

His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Ile Gly

385 390 395 400

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys

405 410 415

Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn

420 425 430

Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val

435 440 445

Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala

450 455 460

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

465 470 475 480

Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala

485 490 495

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

500 505 510

Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val

515 520 525

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Ser Ile

530 535 540

Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn

545 550 555 560

Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Met Asp

565 570 575

Ala Val Lys Lys Gly Leu Pro His Ala Pro Glu Leu Ile Arg Arg Val

580 585 590

Asn Arg Arg Ile Gly Glu Arg Thr Ser His Arg Val Ala Ile Ser Arg

595 600 605

Val Gly Gly Ser Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr

610 615 620

Gly Phe Ala Asp Ala Glu Gly Cys Phe Arg Leu Asp Ile Arg Asn Ala

625 630 635 640

Asn Asp Leu Arg Ala Gly Tyr Arg Thr Arg Leu Ala Phe Glu Ile Val

645 650 655

Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp

660 665 670

Lys Val Gly Thr Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln

675 680 685

Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys

690 695 700

Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln

705 710 715 720

Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile

725 730 735

Lys Glu Leu Val Arg Ile Lys Ala Lys Met Asn Trp Gly Leu Asn Asp

740 745 750

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

755 760 765

Ile Asn Lys Asn Ile Pro Asn Leu Lys Trp Leu Ala Gly Phe Thr Ser

770 775 780

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

785 790 795 800

Lys Val Gly Val Arg Leu Arg Phe Ser Ile Thr Gln His Ile Arg Asp

805 810 815

Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile

820	825	830
Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr		
835	840	845
Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn		
850	855	860
Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val		
865	870	875
880		
Ala Lys Leu Ile Glu Glu Lys His Leu Thr Glu Ser Gly Leu Asp		
885	890	895
Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Val Phe Ser		
900	905	910
Gly Arg		

<210>	14		
<211>	914		
<212>	PRT		
<213>	Artificial Sequence		
<220><223>	Made in Lab - synthesized megaTAL CBLB.F6 construct		
<400>	14		
Met Gly Ser Ala Pro Pro Lys Lys Arg Lys Val Val Asp Leu Arg			
1	5	10	15

Thr Leu Gly Tyr Ser Gln Gln Gln Glu Lys Ile Lys Pro Lys Val			
20	25	30	
Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe			
35	40	45	
Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly			
50	55	60	
Thr Val Ala Val Thr Tyr Gln His Ile Ile Thr Ala Leu Pro Glu Ala			
65	70	75	80
Thr His Glu Asp Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg			
85	90	95	
Ala Leu Glu Ala Leu Leu Thr Asp Ala Gly Glu Leu Arg Gly Pro Pro			

100	105	110	
Leu Gln Leu Asp Thr Gly Gln Leu Val Lys Ile Ala Lys Arg Gly Gly			
115	120	125	
Val Thr Ala Met Glu Ala Val His Ala Ser Arg Asn Ala Leu Thr Gly			
130	135	140	
Ala Pro Leu Asn Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn			
145	150	155	160
Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val			
165	170	175	
Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala			
180	185	190	
Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu			
195	200	205	
Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala			
210	215	220	
Ile Ala Ser Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg			
225	230	235	240
Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val			
245	250	255	
Val Ala Ile Ala Ser Asn Gly Gly Lys Gln Ala Leu Glu Thr Val			
260	265	270	
Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp			
275	280	285	
Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu			
290	295	300	
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr			
305	310	315	320
Pro Asp Gln Val Val Ala Ile Ala Ser Asn Gly Gly Lys Gln Ala			
325	330	335	
Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly			
340	345	350	
Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys			

355 360 365

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp

370 375 380

His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Ile Gly

385 390 395 400

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys

405 410 415

Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn

420 425 430

Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val

435 440 445

Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala

450 455 460

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

465 470 475 480

Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala

485 490 495

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

500 505 510

Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val

515 520 525

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Ser Ile

530 535 540

Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Leu Thr Asn

545 550 555 560

Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Met Asp

565 570 575

Ala Val Lys Lys Gly Leu Pro His Ala Pro Glu Leu Ile Arg Arg Val

580 585 590

Asn Arg Arg Ile Gly Glu Arg Thr Ser His Arg Val Ala Ile Ser Arg

595 600 605

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

660 665 670  
 Lys Val Gly Thr Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln  
 675 680 685  
 Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys  
 690 695 700  
 Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln  
 705 710 715 720  
 Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile  
 725 730 735

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

805 810 815  
 Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile  
 820 825 830  
 Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr  
 835 840 845  
 Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn  
 850 855 860

Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val  
 865 870 875 880

Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp  
 885 890 895  
 Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Val Phe Ser  
 900 905 910  
 Gly Arg

<210> 15  
 <211> 914  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Made in Lab - synthesized megaTAL CBLB.E6 construct  
 <400> 15  
 Met Gly Ser Ala Pro Pro Lys Lys Arg Lys Val Val Asp Leu Arg  
 1 5 10 15

Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val  
 20 25 30  
 Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe  
 35 40 45  
 Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly  
 50 55 60  
 Thr Val Ala Val Thr Tyr Gln His Ile Ile Thr Ala Leu Pro Glu Ala  
 65 70 75 80  
 Thr His Glu Asp Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg

85 90 95  
 Ala Leu Glu Ala Leu Leu Thr Asp Ala Gly Glu Leu Arg Gly Pro Pro  
 100 105 110  
 Leu Gln Leu Asp Thr Gly Gln Leu Val Lys Ile Ala Lys Arg Gly Gly  
 115 120 125  
 Val Thr Ala Met Glu Ala Val His Ala Ser Arg Asn Ala Leu Thr Gly  
 130 135 140

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

385 390 395 400  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 405 410 415  
 Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn  
 420 425 430  
 Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 435 440 445  
  
 Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala  
 450 455 460  
 Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 465 470 475 480  
 Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala  
 485 490 495  
 Ile Ala Ser Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 500 505 510  
 Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val  
  
 515 520 525  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Ser Ile  
 530 535 540  
 Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn  
 545 550 555 560  
 Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Met Asp  
 565 570 575  
 Ala Val Lys Lys Gly Leu Pro His Ala Pro Glu Leu Ile Arg Arg Val  
 580 585 590  
  
 Asn Arg Arg Ile Gly Glu Arg Thr Ser His Arg Val Ala Ile Ser Arg  
 595 600 605  
 Val Gly Gly Ser Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr  
 610 615 620  
 Gly Phe Ala Asp Ala Glu Gly Cys Phe Arg Leu Asp Ile Arg Asn Ala  
 625 630 635 640  
 Asn Asp Leu Arg Ala Gly Tyr Arg Thr Arg Leu Ser Phe Glu Ile Ser

645	650	655
Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp		
660	665	670
Lys Val Gly Thr Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln		
675	680	685
Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys		
690	695	700
Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln		
705	710	715
Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile		
725	730	735
Lys Glu Leu Val Arg Ile Lys Ala Lys Met Asn Trp Gly Leu Asn Asp		
740	745	750
Glu Leu Lys Lys Ala Phe Pro Glu Asn Ile Ser Lys Glu Arg Pro Leu		
755	760	765
Ile Asn Lys Asn Ile Pro Asn Leu Lys Trp Leu Ala Gly Phe Thr Ser		
770	775	780
Gly Asp Gly Ser Phe Met Val Glu Leu Met Lys Asn Lys Asn Asn Val		
785	790	795
Ile Val Arg Val Arg Leu Arg Phe Ser Ile Ser Gln His Ile Arg Asp		
805	810	815
Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile		
820	825	830
Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr		
835	840	845
Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn		
850	855	860
Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val		
865	870	875
Ala Lys Leu Ile Glu Glu Lys His Leu Thr Glu Ser Gly Leu Asp		
885	890	895

Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Val Phe Ser

900 905 910

Gly Arg

<210> 16

<211> 914

<212> PRT

<213> Artificial Sequence

<220><223> Made in Lab - synthesized megaTAL CBLB.D2 construct

<400> 16

Met Gly Ser Ala Pro Pro Lys Lys Lys Arg Lys Val Val Asp Leu Arg

1 5 10 15

Thr Leu Gly Tyr Ser Gln Gln Gln Glu Lys Ile Lys Pro Lys Val

20 25 30

Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe

35 40 45

Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly

50 55 60

Thr Val Ala Val Thr Tyr Gln His Ile Ile Thr Ala Leu Pro Glu Ala

65 70 75 80

Thr His Glu Asp Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg

85 90 95

Ala Leu Glu Ala Leu Leu Thr Asp Ala Gly Glu Leu Arg Gly Pro Pro

100 105 110

Leu Gln Leu Asp Thr Gly Gln Leu Val Lys Ile Ala Lys Arg Gly Gly

115 120 125

Val Thr Ala Met Glu Ala Val His Ala Ser Arg Asn Ala Leu Thr Gly

130 135 140

Ala Pro Leu Asn Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn

145 150 155 160

Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val

165 170 175

Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala

180 185 190

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

195 200 205

Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala

210 215 220

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

225 230 235 240

Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val

245 250 255

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

260 265 270

Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp

275 280 285

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu

290 295 300

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr

305 310 315 320

Pro Asp Gln Val Val Ala Ile Ala Ser Asn Gly Gly Lys Gln Ala

325 330 335

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

340 345 350

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys

355 360 365

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp

370 375 380

His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Ile Gly

385 390 395 400

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys

405 410 415

Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn

420 425 430

Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val

435 440 445

Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala

450 455 460

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

465 470 475 480

Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala

485 490 495

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

500 505 510

Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val

515 520 525

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Ser Ile

530 535 540

Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn

545 550 555 560

Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Met Asp

565 570 575

Ala Val Lys Lys Gly Leu Pro His Ala Pro Glu Leu Ile Arg Arg Val

580 585 590

Asn Arg Arg Ile Gly Glu Arg Thr Ser His Arg Val Ala Ile Ser Arg

595 600 605

Val Gly Gly Ser Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr

610 615 620

Gly Phe Ala Asp Ala Glu Gly Cys Phe Arg Leu Asp Ile His Asn Ala

625 630 635 640

Asn Val Leu Arg Ser Gly Tyr Arg Thr Arg Leu Ser Phe Glu Ile Val

645 650 655

Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp

660 665 670

Lys Val Gly Lys Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln

675	680	685
Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys		
690	695	700
Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln		
705	710	715
Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile		
725	730	735
Lys Glu Leu Val Arg Ile Lys Ala Lys Met Asn Trp Gly Leu Asn Asp		
740	745	750
Glu Leu Lys Lys Ala Phe Pro Glu Asn Ile Ser Lys Glu Arg Pro Leu		
755	760	765
Ile Asn Lys Asn Ile Pro Asn Leu Lys Trp Leu Ala Gly Phe Thr Ser		
770	775	780
Gly Asp Gly Ser Phe Met Val Glu Leu Met Lys Asn Lys Asn Asn Val		
785	790	795
Ile Val Arg Val Arg Leu Arg Phe Ser Ile Ser Gln His Ile Arg Asp		
805 810 815		
Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile		
820	825	830
Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr		
835	840	845
Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn		
850	855	860
Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val		
865	870	875
Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp		
885	890	895
Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Val Phe Ser		
900	905	910
Gly Arg		

&lt;211&gt; 914

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Made in Lab - synthesized megaTAL CBLB.C2 construct

&lt;400&gt; 17

Met Gly Ser Ala Pro Pro Lys Lys Lys Arg Lys Val Val Asp Leu Arg

1 5 10 15

Thr Leu Gly Tyr Ser Gln Gln Gln Glu Lys Ile Lys Pro Lys Val

20 25 30

Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe

35 40 45

Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly

50 55 60

Thr Val Ala Val Thr Tyr Gln His Ile Ile Thr Ala Leu Pro Glu Ala

65 70 75 80

Thr His Glu Asp Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg

85 90 95

Ala Leu Glu Ala Leu Leu Thr Asp Ala Gly Glu Leu Arg Gly Pro Pro

100 105 110

Leu Gln Leu Asp Thr Gly Gln Leu Val Lys Ile Ala Lys Arg Gly Gly

115 120 125

Val Thr Ala Met Glu Ala Val His Ala Ser Arg Asn Ala Leu Thr Gly

130 135 140

Ala Pro Leu Asn Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn

145 150 155 160

Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val

165 170 175

Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala

180 185 190

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

195 200 205

Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala

210	215	220	
Ile Ala Ser Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg			
225	230	235	240
Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val			
245	250	255	
Val Ala Ile Ala Ser Asn Gly Gly Lys Gln Ala Leu Glu Thr Val			
260	265	270	
Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp			
275	280	285	
Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu			
290	295	300	
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr			
305	310	315	320
Pro Asp Gln Val Val Ala Ile Ala Ser Asn Gly Gly Lys Gln Ala			
325	330	335	
Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly			
340	345	350	
Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys			
355	360	365	
Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp			
370	375	380	
His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Ile Gly			
385	390	395	400
Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys			
405	410	415	
Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn			
420	425	430	
Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val			
435	440	445	
Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala			
450	455	460	

Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 465 470 475 480

Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala  
 485 490 495

Ile Ala Ser Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 500 505 510

Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val  
 515 520 525

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Ser Ile  
 530 535 540

Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn  
 545 550 555 560

Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Met Asp  
 565 570 575

Ala Val Lys Lys Gly Leu Pro His Ala Pro Glu Leu Ile Arg Arg Val  
 580 585 590

Asn Arg Arg Ile Gly Glu Arg Thr Ser His Arg Val Ala Ile Ser Arg  
 595 600 605

Val Gly Gly Ser Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr  
 610 615 620

Gly Phe Ala Asp Ala Glu Gly Cys Phe Gly Leu Tyr Ile His Asn Ser  
 625 630 635 640

Asn Val Leu Arg Ser Gly Tyr Arg Thr Arg Leu Ser Phe Glu Ile Ser  
 645 650 655

Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp  
 660 665 670

Lys Val Gly Thr Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln  
 675 680 685

Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys  
 690 695 700

Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln  
 705 710 715 720

Ala Phe Ser Leu Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile

725 730 735

Lys Glu Leu Val Arg Ile Lys Ala Lys Met Asn Trp Gly Leu Asn Asp

740 745 750

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

755 760 765

Ile Asn Lys Asn Ile Pro Asn Leu Lys Trp Leu Ala Gly Phe Thr Ser

770 775 780

Gly Asp Gly Ser Phe Met Val Glu Leu Met Lys Asn Lys Asn Asn Val

785 790 795 800

Ile Val Arg Val Arg Leu Arg Phe Ser Ile Ser Gln His Ile Arg Asp

805 810 815

Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile

820 825 830

Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr

835 840 845

Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn

850 855 860

Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val

865 870 875 880

Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp

885 890 895

Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Val Phe Ser

900 905 910

Gly Arg

<210> 18

<211> 914

<212> PRT

<213> Artificial Sequence

<220><223> Made in Lab - synthesized megaTAL CBLB.B5 construct

<400> 18

Met Gly Ser Ala Pro Pro Lys Lys Arg Lys Val Val Asp Leu Arg

1 5 10 15

Thr Leu Gly Tyr Ser Gln Gln Gln Glu Lys Ile Lys Pro Lys Val

20 25 30

Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe

35 40 45

Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly

50 55 60

Thr Val Ala Val Thr Tyr Gln His Ile Ile Thr Ala Leu Pro Glu Ala

65 70 75 80

Thr His Glu Asp Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg

85 90 95

Ala Leu Glu Ala Leu Leu Thr Asp Ala Gly Glu Leu Arg Gly Pro Pro

100 105 110

Leu Gln Leu Asp Thr Gly Gln Leu Val Lys Ile Ala Lys Arg Gly Gly

115 120 125

Val Thr Ala Met Glu Ala Val His Ala Ser Arg Asn Ala Leu Thr Gly

130 135 140

Ala Pro Leu Asn Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn

145 150 155 160

Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val

165 170 175

Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala

180 185 190

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

195 200 205

Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala

210 215 220

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

225 230 235 240

Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val

245	250	255
Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val		
260	265	270
Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp		
275	280	285
Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu		
290	295	300
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr		
305	310	315
Pro Asp Gln Val Val Ala Ile Ala Ser Asn Gly Gly Lys Gln Ala		
325	330	335
Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly		
340	345	350
Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys		
355	360	365
Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp		
370	375	380
His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Ile Gly		
385	390	395
Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys		
405	410	415
Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn		
420	425	430
Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val		
435	440	445
Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala		
450	455	460
Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu		
465	470	475
Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala		
485	490	495
Ile Ala Ser Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg		

500	505	510
Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val		
515	520	525
Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Ser Ile		
530	535	540
Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn		
545	550	555
Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Met Asp		
565	570	575
Ala Val Lys Lys Gly Leu Pro His Ala Pro Glu Leu Ile Arg Arg Val		
580	585	590
Asn Arg Arg Ile Gly Glu Arg Thr Ser His Arg Val Ala Ile Ser Arg		
595	600	605
Val Gly Gly Ser Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr		
610	615	620
Gly Phe Ala Asp Ala Glu Gly Cys Phe Arg Leu Asp Ile His Asn Ala		
625	630	635
Asn Val Leu Arg Ser Gly Tyr Arg Thr Arg Leu Ser Phe Glu Ile Val		
645	650	655
Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp		
660	665	670
Lys Val Gly Thr Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln		
675	680	685
Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys		
690	695	700
Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln		
705	710	715
Ala Phe Ser Leu Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile		
725	730	735
Lys Glu Leu Val Arg Ile Lys Ala Lys Met Asn Trp Gly Leu Asn Asp		
740	745	750

Glu Leu Lys Lys Ala Phe Pro Glu Asn Ile Ser Lys Glu Arg Pro Leu  
 755 760 765

Ile Asn Lys Asn Ile Pro Asn Leu Lys Trp Leu Ala Gly Phe Thr Ser  
 770 775 780

Gly Asp Gly Ser Phe Val Val Glu Leu Lys Lys Arg Arg Ser Pro Val  
 785 790 795 800  
 Lys Val Gly Val Arg Leu Arg Phe Gly Ile Thr Gln His Ile Arg Asp

805 810 815  
 Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile  
 820 825 830

Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr  
 835 840 845  
 Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn

850 855 860  
 Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val  
 865 870 875 880

Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp  
 885 890 895  
 Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Val Phe Ser  
 900 905 910  
 Gly Arg

<210> 19  
 <211> 914  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Made in Lab - synthesized megaTAL CBLB.A8 construct  
 <400> 19

Met Gly Ser Ala Pro Pro Lys Lys Lys Arg Lys Val Val Asp Leu Arg  
 1 5 10 15

Thr Leu Gly Tyr Ser Gln Gln Gln Glu Lys Ile Lys Pro Lys Val  
 20 25 30

Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe  
 35 40 45  
 Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly  
 50 55 60  
 Thr Val Ala Val Thr Tyr Gln His Ile Ile Thr Ala Leu Pro Glu Ala  
 65 70 75 80  
 Thr His Glu Asp Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg  
 85 90 95  
 Ala Leu Glu Ala Leu Leu Thr Asp Ala Gly Glu Leu Arg Gly Pro Pro  
 100 105 110  
 Leu Gln Leu Asp Thr Gly Gln Leu Val Lys Ile Ala Lys Arg Gly Gly  
 115 120 125  
 Val Thr Ala Met Glu Ala Val His Ala Ser Arg Asn Ala Leu Thr Gly  
 130 135 140  
 Ala Pro Leu Asn Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn  
 145 150 155 160  
  
 Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 165 170 175  
 Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala  
 180 185 190  
 Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 195 200 205  
 Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala  
 210 215 220  
 Ile Ala Ser Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 225 230 235 240  
  
 Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val  
 245 250 255  
 Val Ala Ile Ala Ser Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
 260 265 270  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp  
 275 280 285

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
 290 295 300

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr  
 305 310 315 320

Pro Asp Gln Val Val Ala Ile Ala Ser Asn Gly Gly Lys Gln Ala  
 325 330 335

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
 340 345 350

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Asn Gly Lys  
 355 360 365

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp

370 375 380  
 His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Ile Gly  
 385 390 395 400

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 405 410 415

Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn  
 420 425 430

Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 435 440 445

Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala  
 450 455 460

Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 465 470 475 480

Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val Val Ala  
 485 490 495

Ile Ala Ser Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 500 505 510

Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val

515 520 525

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Ser Ile

530	535	540
Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn		
545	550	555
Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Met Asp		
565	570	575
Ala Val Lys Lys Gly Leu Pro His Ala Pro Glu Leu Ile Arg Arg Val		
580	585	590
Asn Arg Arg Ile Gly Glu Arg Thr Ser His Arg Val Ala Ile Ser Arg		
595	600	605
Val Gly Gly Ser Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr		
610	615	620
Gly Phe Ala Asp Ala Glu Gly Cys Phe Arg Leu Asp Ile Arg Asn Ala		
625	630	635
Asn Asp Leu Arg Ala Gly Tyr Arg Thr Arg Leu Ala Phe Glu Ile Val		
645	650	655
Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp		
660 665 670		
Lys Val Gly Thr Ile Tyr Asn Ala Gly Asp Asn Ala Val Arg Leu Gln		
675	680	685
Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Gly His Phe Glu Lys		
690	695	700
Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Lys Leu Phe Lys Gln		
705	710	715
Ala Phe Ser Val Met Glu Asn Lys Glu His Leu Lys Glu Asn Gly Ile		
725	730	735
Lys Glu Leu Val Arg Ile Lys Ala Lys Met Asn Trp Gly Leu Asn Asp		
740	745	750
Glu Leu Lys Lys Ala Phe Pro Glu Asn Ile Ser Lys Glu Arg Pro Leu		
755	760	765
Ile Asn Lys Asn Ile Pro Asn Leu Lys Trp Leu Ala Gly Phe Thr Ser		
770	775	780
Gly Asp Gly Ser Phe Met Val Glu Leu Met Lys Asn Lys Asn Val		

785 790 795 800

Ile Val Arg Val Arg Leu Arg Phe Ser Ile Ser Gln His Ile Arg Asp

805 810 815

Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Arg Ile

820 825 830

Val Glu Asn Asn Lys Ser Glu His Ser Trp Leu Glu Phe Ile Val Thr

835 840 845

Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn

850 855 860

Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val

865 870 875 880

Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp

885 890 895

Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Val Phe Ser

900 905 910

Gly Arg

<210> 20

<211> 22

<212> DNA

<213> Homo sapiens

<400> 20

ctgtaagata ttcccatccc ca 22

<210> 21

<211> 13

<212> DNA

<213> Homo sapiens

<400> 21

ttgttatgag gta 13

<210> 22

<211> 39

<212> DNA

<213>	Homo sapiens	
<400>	22	
ttgttatgag gtaggtctg taagatattc ccatccca		39
<210>	23	
<211>	22	
<212>	DNA	
<213>	Homo sapiens	
<400>	23	
ctgtaagata ttcaacctt ta		22
<210>	24	
<211>	22	
<212>	DNA	
<213>	Homo sapiens	
<400>	24	
tttccactta ttccatccc ca		22
<210>	25	
<211>	22	
<212>	DNA	
<213>	Homo sapiens	
<400>	25	
ctgtaagata ttctacgtct gc		22
<210>	26	
<211>	22	
<212>	DNA	
<213>	Homo sapiens	
<400>	26	
cttccagggaa ttccatccc ca		22
<210>	27	
<211>	22	
<212>	DNA	
<213>	Homo sapiens	
<400>	27	
ctgtaagata ttccacaggc tc		22
<210>	28	

&lt;211&gt; 22

&lt;212&gt;

PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 28

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

1 5 10 15

Thr Cys Cys Cys Cys Ala

20

&lt;210&gt; 29

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 29

ctgtaagata ttacccgatt tt 22

&lt;210&gt; 30

&lt;211&gt; 7244

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Made in Lab - synthesized I-OnuI variant CBLB.E3 surface display

plasmid

&lt;400&gt; 30

gacgaaaggg cctcgtata cgcttatttt tataggtaa tgtcatgata ataatggttt 60

cttaggacgg atcgcttgc ttaacttac acgcgcctcg tatctttaa ttagtggata 120

atttggaaat ttactctgtt tttatttttttt gtatggat ttttagaaagt 180

aaataaaagaa ggtagaagag ttacgaaatg aaaaaaaaaa aataaaca aaaa 240

atttcaacaa aaagcgtaact ttacatataat atttattaga caagaaaagc agattaaata 300

gatatacatt cgattaacga taagtaaaat gtaaaatcac aggatttcg tgtgtggct 360

tctcacacaga caagatgaaa caattcggca ttaatacctg agagcaggaa gagcaagata 420

aaaggttagta ttgttggcg atccccctag agtctttac atcttcggaa aacaaaaact 480

atttttctt taatttctttttt ttttacttac tatttttaat ttatataattt atattaaaa 540

atttaaatata taattttttt tatagcacgt gatgaaaagg acccaggtgg cactttcgg 600

ggaaatgtgc gcgaaacccc tattgttta ttttctaaa tacattcaaa tatgtatccg 660

ctcatgagac aataaccctg ataaatgctt caataatatt gaaaaaggaa gagtatgagt	720
attcaacatt tccgtgcgc ccttattccc tttttgcgg cattttgcct tcctgtttt	780
gctcacccag aaacgctggt gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg	840
ggttacatcg aactggatct caacagcgg aagatcctt agagtttcg cccgaagaa	900
cgtttccaa ttagtggcac ttttaagtt ctgctatgtg gcgcggattt atccgtatt	960
gacgccggc aagagcaact cggtcgccc atacactatt ctcaaatga cttgggttag	1020
tactcaccag tcacagaaaa gcatcttacg gatggcatga cagtaagaga attatgcgt	1080
gctgccataa ccatgagtga taacactgctg gccaacttac ttctgacaac gatcgagga	1140
ccgaaggagc taaccgctt tttcacaac atggggatc atgtaactcg ctttgatcgt	1200
tggaaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgt	1260
gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactt agctccgg	1320
caacaattaa tagactggat ggaggcggat aaagttgcag gaccacttct gcgcgtggcc	1380
cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg gtctcgccgt	1440
atcattgcag cactggggcc agatggtaag ccctccgtt tcgtagttat ctacacgacg	1500
ggcagtccagg caactatgga tgaacgaaat agacagatcg ctgagatagg tgcctactg	1560
attaaggattt ggttaactgtc agaccaagt tactcatata tacttttagat tgatttaaaa	1620
cttcattttt aatttaaaag gatcttaggtg aagatcctt ttgataatct catgaccaaa	1680
atcccttaac gtgagtttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagga	1740
tcttcttgag atcctttttt tctgcgtta atctgtct tgcaaaacaaa aaaaccaccg	1800
ctaccagcgg tggttgttt gccggatcaa gagctaccaa ctctttcc gaaggtaact	1860
ggcttcagca gagcgcagat accaaatact gtccttctag ttagccgtt gttaggccac	1920
cacttcaaga actctgttagc accgcctaca tacctcgctc tgctaatctt gttaccagt	1980
gctgtcgccaa gtggcgatata gtcgtgtttt accgggttgg actcaagacg atagttaccg	2040
gataaggcgc acgggtcggtt ctgaacgggg gttcggtca cacagcccgat cttggagcga	2100
acgacacctaca ccgaactgag atacctacag cgtgagcatt gagaaagcgc cacgcttccc	2160
gaagggagaa aggccggacag gtatccgtta agccggcagg tcggaaacagg agagcgcacg	2220
agggagcttc cagggggaa cgcctggat ctttatagtc ctgtcggtt tcgcaccc	2280
tgacttgcgc gtcgtttttt gtgtatgtcg tcagggggc cggccatgt gaaaaacgcc	2340
agcaacgcgg ccttttacg gttccctggcc ttttgcgttca cttttgttca catgttctt	2400
cctgcgttat cccctgattt tttggataac cgtattacgg ctttgagtg agctgatacc	2460

gctcgccgca gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc	2520
ccaatacgca aaccgcctct ccccgcgctg tggccgattc attaatgcag ctggcacgac	2580
aggttcccg actggaaagc gggcagttag cgcaacgcaa ttaatgttag ttacccact	2640
cattaggcac cccaggctt acactttatg cttccggctc ctatgttgt tggaatttgt	2700
agcgataac aatttcacac agggaaacgcgat tatgaccatg attacccaa gctcggatt	2760
aaccctact aaagggaaaca aaagctgggt acccgacagg ttatcagcaa caacacagtc	2820
atatccattc tcaatttagct ctaccacagt gtgtgaacca atgtatccag caccacctgt	2880
aaccaaaaaca atttagaag tactttact ttgttaactga gctgtcattt atattgaatt	2940
ttcaaaaatt ctactttt tttggatgg acgcaaaagaa gttataataat catattacat	3000
gcattacca ccatatacat atccatatac atatccatat ctaatcttac ttatgttg	3060
tggaaatgtt aagagccca ttatcttagc ctaaaaaaaac cttcttttgcgaaatccat	3120
taatacgctt aactgctcat tgctatattg aagtacggat tagaagccgc cgagcgggtg	3180
acagccctcc gaaggaagac tctccctcggt gcgtccctgt ctccaccggc cgcgttcctg	3240
aaacgcagat gtgcctcgcg ccgcactgtc cggacaata aagattctac aatactagct	3300
tttatggtta tgaagaggaa aaattggcag taacctggcc ccacaaacct tcaaataaac	3360
gaatcaaatt aacaaccata ggatgataat gcgatttagtt ttttagcctt atttctgggg	3420
taattaatca gcaagcgat gatttttgtt ctttaacag atatataaaat gcaaaaaactg	3480
cataaccact ttaactaata cttcaacat tttcggtttg tattacttct tattcaaatg	3540
taataaaaga tcgaatccta cttcatacat tttcaattaa gatgcagttt cttcgcttt	3600
tttcaatatt ttctgttatt gcttcagttt tagcacagga actgacaact atatgcgagc	3660
aaatccctc accaacttta gaatcgacgc cgtactctt gtcaacgact actatttgg	3720
ccaacggaa ggcataatgcaaa ggagtttttgcgaaatattacaa atcgttaacg tttgtcgtt	3780
attgcggttc tcacccctca acaacttagca aaggcagccc cataaacaaca cagtatgttt	3840
ttaaggacaa tagtcgacg attgaaggta gataccata cgacgttcca gactacgtc	3900
tgcaggctag tggtggagga ggctctgggt gaggcggtag cggaggcggaa gggtcggcta	3960
gctccatcaa cccatggatt ctgactggtt tcgctgtatgc cgaaggatgc ttccgactag	4020
acatccgcaa cggcaacgat ttaagagccg gatacagaac tagactggcc ttcaatcg	4080
tactgcacaa caaggacaaa tcgattctgg agaatatcca gtcgacttgg aaggtcgca	4140
caatctacaa cggggcgac aacgcgtca gactgcaagt cacacgttgc gaaatgttgc	4200

aagtgattat cgaccacttc gagaaatatac cgctgataac acagaaattt ggcgattaca	4260
agttgttaa acaggcatc agcgtcatgg agaacaaga acatctaag gagaatggga	4320
ttaaggagct cgtacgaatc aaagctaaga tgaattgggg tctcaatgac gaattgaaaa	4380
aagcattcc agagaacatt agcaaagagc gctccctat caataagaac attccgaatc	4440
tcaaattggct ggctggattc acatctggtg acggctcggt cgtggggaa ctaaagaaga	4500
gaagaagccc cgtcaaggta ggagtgcggc tgcgatttag catcacccag cacatcagag	4560
acaagaacct gatgaattca ttgataacat acctaggctg tggcgatc gttgagaata	4620
acaatctga gcacagttgg ctcgaattca ttgtaacaaa attcagcgat atcaacgaca	4680
agatcattcc ggtattccag gaaaatactc tgatggcgt caaactcgag gacttigaag	4740
attggtgcaa ggttgccaaa ttgatcgaag agaagaaaca cctgaccgaa tccggtttg	4800
atgagattaa gaaaatcaag ctgaacatga acaaaggctg tgtcttctt agaggcggtt	4860
ccagaagcgg atctggtaact ggcaacaga aactcataag cgaagaagac cttagcggga	4920
ctggagagca aaagttgatt tctgaggagg atttgcggg aaccggggag cagaagttaa	4980
tcagtgaaga ggtctcagt ggaacggcg aacaaaagtt gatctcgag gaagacttat	5040
aatgaagatc tgataacaac agttagatg taacaaaatc gactttgttc ccactgtact	5100
tttagcttgt acaaaataca atatacttt catttctcg taaacaacat gtttccat	5160
gtaatatcct tttctattt tcgttccgtt accaacttta cacatactt atatagctat	5220
tcacttctat acactaaaaa actaagacaa ttttaattt gctgcctgcc atattcaat	5280
ttgttataaa ttccctataat ttatcctatt agtagctaa aaaagatgaa tgtgaatcga	5340
atcctaagag aattgagctc caattcgccc tatagtgagt cgtattacaa ttcaactggcc	5400
gtcggtttac aacgtcgtga ctggaaaac cctggcgta cccaaacttaa tcgccttgca	5460
gcacatcccc cttcgccag ctggcgtaat agcgaagagg cccgcaccga tcgccttcc	5520
caacagtgc gcagcctgaa tggcgatgg acgcgcctg tagcgccgca ttaagcgccg	5580
cgggtgtggt gtttacgcgc agcgtgacgg ctacacttgc cagcgcctt cgcggcgctc	5640
cttcgcctt ctcccttcc ttctcgcca cgttcggccg cttccctgt caagctctaa	5700
atcggggct cccttaggg ttccgattta gtgcattacg gcacctcgac cccaaaaaac	5760
ttgattaggg ttaggttca cgtatgggc catgcgcctg atagacgggtt ttccctt	5820
tgacgttggaa gtccacgttc ttaatagtg gactcttgc ccaaactgga acaacactca	5880
accctatctc ggtctattct ttgatttt aaggatttt gccgatttcg gcctattgg	5940
taaaaaatga gctgattaa caaaaatata acgcgaattt taacaaaata ttaacgttta	6000
caatttcctg atgcgttatt ttctccttac gcatctgtgc ggtatttcac accgcaggca	6060

agtgcacaaa caatacttaa ataaatacta ctcagtaata acctatttct tagcatttt 6120  
gacgaaattt gctattttgt tagagtcctt tacaccattt gtctccacac ctccgcttac 6180  
atcaacacca ataacccat ttaatctaag cgcatcacca acatttctg gcgtcagttcc 6240  
accagctaac ataaaatgtta agcttcggg gctctttgc cttccaaccc agtcagaaat 6300  
  
cgagttccaa tccaaaagtt cacctgtccc acctgcttct gaatcaaaca agggaaataaa 6360  
cgaatgaggt ttctgtgaag ctgcacttag tagtatgtt cagtttttgc gaaatacgag 6420  
tcttttaata actggcaaac cgaggaactc ttggatttct tgccacgact catctccatg 6480  
cagttggacg acatcaatgc cgtaatcatt gaccagagcc aaaacatctt ccttaggtt 6540  
attacgaaac acgccaacca agtatttcgg agtgcctgaa ctattttat atgctttac 6600  
aagacttggaa atttcccttgc caataaccgg gtcatttttttgc ctctttctat tgggcacaca 6660  
tataataccca agcaagtcag catcgaaatc aagagcacat tctgcggcct ctgtgtctg 6720  
  
  
caagccgcaa actttcacca atggaccaga actacctgtt aaattaataa cagacataact 6780  
ccaagctgcc tttgtgtgtt taatcaatgt tactcacgtt ctcaatagtc accaatgccc 6840  
tccctcttgg ccctctccctt ttctttttc gaccgaatttatttcaatc ggcaaaaaaa 6900  
gaaaagctcc ggatcaagat tgtacgttgc gtcataactg caaagtacac atatattacg atgctgtcta 6960  
ttccactact gccatctggc gtcataactg caaagtacac atatattacg atgctgtcta 7020  
ttaaatgctt cctatattat atatataatgtt atgtcggttta tggcactc tcagttacat 7080  
ctgctctgtat gcccataatgtt taagccagcc ccgacaccccg ccaacaccccg ctgacgcgcc 7140  
  
  
ctgacgggct tgcgtgtcc cggcatccgc ttacagacaa gctgtgaccg tctccggag 7200  
ctgcatgtgtt cagagggtttt caccgtcatc accgaaacgc gcga 7244  
<210> 31  
<211> 2748  
<212> RNA  
<213> Artificial Sequence  
<220><223> Made in Lab - synthesized megaTAL CBLB.E3 mRNA  
<400> 31  
  
augggaucggc cgccaccuaa gaagaaacgc aaagucgugg aucuacgcac gcucggcuac 60  
agucagcagc agcaagagaaa gaucaaaccg aaggugcguu cgacaguggc gcagcaccac 120  
gagggcacugg uggcccaugg guuuacacac ggcacacuacg uugcgcucag ccaacacccg 180  
  
  
gcagcguuag ggaccgucgc uguacguau cagcacauaa ucacggcguu gccagaggcg 240  
acacacgaag acaucguugg cgucggcaaa cagugguccg ggcacacgcgc ccuggaggcc 300

uugcucacgg augcggggga guugagaggu ccgcccguuac aguuggacac aggccaacuu	360
gugaagauug caaaaacgugg cggcgugacc gcaauuggagg cagugcaugc aucgcgcaau	420
gcacugacgg gugcccccu gaacuuaaca cccgaucaag ugguagcgau agcgucaaac	480
ggcgaaaaa aacaggcuuu ggagacggua cagcgguuau ugccguacu cugccaggac	540
cacggauuga caccggacca agugguggcg auugcgucca acaacggagg caagcaggca	600
cuagagacug uccaacggcu ucuuuccguu cuuugucagg aucaugggcu aaccccugau	660
cagguagucg cuauagcuuc aaauggaggg ggcaagcaag cacuggagac uguuacaacga	720
cuccugccag ugcucugcca agaccacgga cuuacaccag aucaaguggu ugcuaauugcc	780
uccaauggug gcgaaaaaca agcguuggaa acugugcaga gacuguuacc ugcuuugugu	840
caagaccacg gccucacgccc agaucaggug guagccauag cgucgaacau ugugguaag	900
caagccccuug aaacggucca gcgucuucug ccggugugu gccaggacca cggacuaacg	960
ccggaucagg ucguaagccau ugcuucaaa gggggcggca aacaggcgcu agagacaguc	1020
cagcgccucu ugccuguguu augccaggau cacggcuuaa ccccagacca aguuguggcu	1080
auugcaucua acaauuggugg caaacaaggcc uuggagacag ugcaacgauu acugccuguc	1140
uuauugucagg aucauggccu gacgcccgau cagguagugg caaucgcauc uaacauugga	1200
gguaaagcaag cacuggagac uguccagaga uuguuacccg uacuauguca agaucauggu	1260
uugacgcccug aucagguugu ugcgauagcc agcaauaaacg gagggaaaca ggcucuugaa	1320
accguacagc gaciuucuccc agucuugugc caagauacg ggcuuacucc ugaucuaguc	1380
guagcuaucg ccagcaauaa cgguggaaaa caggcccugg aaaccguaca acgucuccuc	1440
ccaguacuuu gucaagacca cggguugacu ccggaucaag ucgucgcgau cgcgagcaau	1500
gguggggggga agcaggcgcu ggaaacuguu cagagacugc ugccuguacu uugucaggac	1560
cauggucuga caccugacca aguuguggcg auagccagua auaucggggg aaaacaggca	1620
cuagagagca uuguggccca gcugagccgg ccugauccgg cguuggccgc guugaccaac	1680
gaccaccucg ucgccuuggc cugccucggc ggacgucugc ccauggaugc agugaaaaag	1740
ggauugccgc acgcgcccgg aaugaucaga agagucaauc gccguauugg cgaacgcacg	1800
ucccaucgca uugcgauauc uagagugggaa ggaagcucgc gcagagaguc caucaaccca	1860
uggauucuga cugguuucgc ugaugccgaa ggaugcucc gacuagacau ccgcaacgca	1920
aacgauuuua gagccggaua cagaacuaga cuggccuuugc aaaucguacu gcacaacaag	1980
gacaaauucga uucuggagaa uaucaggugc acuuggaagg ucggcacaau cuacaacgcg	2040
ggcgacaacg cagucagacu gcaagucaca cguuucgaaag auuugaaaagu gauuaucgac	2100
cacuucgaga aauauccgcu gauaacacag aaauuggcg auuacaagu guuuuaacag	2160

gcauucagcg ucauggagaa caaagaacau cuuaaggaga augggauuaa ggagcucgua	2220
cgaaucaaag cuaagaugaa uuggggucuc aaugacgaau ugaaaaaaagc auuuccagag	2280
aacauuagca aagagcgcuc ccuuaucaau aagaacauuc cgaaucucaa auggcuggcu	2340
ggauucacau cuggugacgg cucguucgug guggaacuaa agaagagaag aagccccguc	2400
aaggauaggag ugcggcugcg auucagcauc acccagcaca ucagagacaa gaaccugaug	2460
aauucauuga uaacauuaccu aggcuguggu cguaucguug agaauaacaa aucugagcac	2520
aguuggcucg aauucauugu aacaaaauuc agcgauauca acgacaagau cauuccggua	2580
uuccaggaaa auacucugau uggcgucaaa cucgaggacu uugaagauug gugcaagguu	2640
gccaauuuga ucgaagagaa gaaacaccug accgaauccg guuuggauga gauuaagaaa	2700
aucaagcuga acaugaacaa aggucguguc uucagcggcc gcugauaa	2748
<210> 32	
<211> 2748	
<212> RNA	
<213> Artificial Sequence	
<220><223> Made in Lab - synthesized megaTAL CBLB.A8 mRNA	
<400> 32	
augggaauccg cgccaccuaa gaagaaacgc aaagucgugg aucuacgcac gcucggcuac	60
agucagcagc agcaagagaa gauaaacccg aaggugcguu cgacaguggc gcagcaccac	120
gaggcacugg ugcccuaagg guuuuacacac ggcacacaucg uugcgcucag ccaacacccg	180
gcagcguuag ggaccgcugc ugcacguau cagcacauua ucacggcguu gccagaggcg	240
acacacgaag acaucguugg cgucggcaaa cagugguccg ggcacgcgc ccuggaggcc	300
uugcucacgg augcggggga guugagaggu ccgcccguuac aguuggacac aggccaacuu	360
gugaagauug caaaacgugg cggcgugacc gcaauggagg cagugcaugc aucgcgcaau	420
gcacugacgg gugcccccua acuuuaaca cccgaucaag uggugcguau agcgucaaac	480
ggcgggggua aacaggcuuu ggagacggua cagcgguuau ugcccguacu cugccaggac	540
cacggauuga caccggatca agugguggcg auugcgucca acaacggagg caagcaggca	600
cuagagacug uccaacggcu ucuuuccguu cuuugucagg aucaugggcu aacccugau	660
cagguagucg cuauagcuuc aaauggaggg ggcaagcaag cacuggagac uguuacaacga	720
cuccugccag ugcucugcca agaccacggc cuuacaccag aucaaguggu ugcuaauugcc	780
uccaauggug gcgaaaaca agcguuggaa acugugcaga gacuguuacc ugcuuugugu	840
caagaccacg gccucacgccc agaucaggug guagccauag cgucgaacau ugugguaag	900

caagccuug aaacggucca gcgucuucug ccgguguugu gccaggacca cggacuaacg	960
ccggauccagg ucguagccau ugcuucaaau gggggcggca aacaggcgcu agagacaguc	1020
cagcgccucu ugccuguguu augccaggau cacggcuuaa ccccagacca aguuguggcu	1080
auugcaucua acaauggugg caaacaagcc uuggagacag ugcaacgauu acugccuguc	1140
uuauugucagg aucauggccu gacgcccgau cagguagugg caaucgcauc uaacauugga	1200
gguaagcaag cacuggagac uguccagaga uuguuacccg uacuauguca agaucauggu	1260
uugacgcccug aucagguugu ugcgauagcc agcaauaacg gagggaaaca ggcucuugaa	1320
accguacagc gaciuucuccc agucuugugc caagauacg ggcuuacucc ugaucaaguc	1380
guagcuaucg ccagcaauaa cgguggaaaa caggcccugg aaaccguaca acgucuccuc	1440
ccaguacuuu gucaagacca cggguugacu ccggaucaag ucgucgcgau cgcgagcaau	1500
gguggggggga agcaggcgcu ggaaacuguu cagagacugc ugccuguacu uugucaggac	1560
cauggucuga caccugacca aguuguggcg auagccagua auaucggggg aaaacaggca	1620
cuagagagca uuguggccca gcugagccgg ccugauccgg cguuggccgc guugaccaac	1680
gaccaccucg ucgcuuggc cugccucggc ggacguccug ccauggaugc agugaaaaag	1740
ggauugccgc acgcgcccga auugaucaga agagucaauc gccguauugg cgaacgcacg	1800
ucccaucgcg uugcgauauc uagaguggga ggaagcucgc gcagagaguc caucaaccca	1860
uggauucuga cugguuucgc ugaugccgaa ggaugcuucc gacuagacau ccgcaacgca	1920
aacgauuuua gagccggaua cagaacuaga cuggccuucg aaaucguacu gcacaacaag	1980
gacaaaucga uucuggagaa uauccagucg aciuuggaagg ucggcacaau cuacaacgcg	2040
ggcgacaacg cagucagacu gcaagucaca cguuucgaag auuugaaagu gauuaucggc	2100
cacuucgaga aauauccgcu gauaacacag aaauuggcg auuacaaguu guuuuaacag	2160
gcauucagcg ucauggagaa caaagaacau cuuaaggaga augggauua ggagcucgua	2220
cgaaucaaag cuaagaugaa uuggggucuc aaugacgaaau ugaaaaaagc auuuuccagag	2280
aacauuagca aagagcgccc ccuuauccaa aagaacaucc caaaucucaa auggcuggcu	2340
ggauucacau cuggugacgg cucguucaug guggaacuaa ugaagaauaa gaauaacguu	2400
auuuguacgug ugcgucugag auucucaauc ucccagcaca ucagagacaa gaaccugaug	2460
aauucauuga uaacauaccu aggcuguggu cguaucguug agaauaacaa aucugagcac	2520
aguuggcucg aauucauugu aacaaaauuc agcgauauca acgacaagau cauuccggua	2580
uuccaggaaa auacucugau uggcguaaa cucgaggacu uugaagauug gugcaagguu	2640
gccaaauuga ucgaagagaa gaaacaccug accgaauccg guuuggauga gauuaagaaa	2700

aucaagcuga acaugaacaa	aggucguguc	uucagcggcc	gcugauaa	2748		
<210> 33						
<211> 2748						
<212> RNA						
<213> Artificial Sequence						
<220><223> Made in Lab - codon optimized megaTAL CBLB.E3 mRNA						
<400> 33						
augggaucg	ccccccgaa	aaagaagcgg	aaaguggucg	accuuagaac	ccugggauac	60
ucccagcagc	agcaggaaaa	gauuaagccg	aagguccgg	caacaguggc	ccagcaccac	120
gaggcacuug	ugggacacgg	auuuacccau	gcccauaucg	ucgcccuiuc	ccagcauccg	180
gcggcucugg	gaaccgucgc	cgugaccuac	cagcauuaa	ucacugcgcu	gccugaagcu	240
acccacgagg	auuuuguggg	agugggaaag	caguggagcg	gagcgagagc	gcuggaagcc	300
cuccugacccg	acgcaggaga	gcuccgggga	ccuccacucc	aacucgacac	uggacagcuu	360
gugaagauug	ccaagagagg	cgccgucacc	gcuauggagg	cggugcacgc	uucccggAAC	420
gcucugacccg	gagccccucu	gaaucugacu	ccagaccaag	ucguggcaau	cgcauccaac	480
ggcgggggca	agcaagcucu	ggaaacagug	caacgccucc	ugcccgugcu	gugccaggac	540
cacggacuca	cccccgauca	aguaguggcc	aucgcuucca	acaacggugg	aaaacaggcu	600
cucgaaacag	uccagcggcu	guugccaguc	cugugccaag	accacgguuu	gaccccugac	660
caaguggugg	ccauagcgcuc	gaauggugga	ggcaagcaag	ccuuggaaac	cguucagagg	720
uugcuuccag	ugcuuugcca	ggaucacgga	cugacaccug	accagguggu	cgcaauugcc	780
uccaacggag	gggggaagca	ggccucgaa	acggugcaac	ggcugcugcc	ugugcucugu	840
caggaccacg	gcuugacccc	cgaccaagug	guggccaucg	cgagcaauau	ugguggaaag	900
cagggcuugg	agacugugca	gagacuucug	ccggugcucu	gccaagauca	uggccuuacc	960
ccagaucaag	uggucgcgau	cgcuucgaaac	ggaggcggca	aacaggccu	ggagacugug	1020
caaagacugc	ugccgguguu	gugucaggau	cauggauuga	cucccgauca	ggugguggcg	1080
aucgccucaa	auaacggagg	aaaacaagcc	cucgagacug	uucagcggcu	acugccgcuc	1140
cugugucagg	accauggcuu	gacacccgac	cagguagucg	ccaucgcgc	caacaucggc	1200
ggcaagcagg	cacucgaaac	cgugcagaga	cugcuccccc	ugcugugcca	ggaccacggu	1260
cugacuccgg	aucaggucgu	ugcuaucgccc	ucgaacaacg	ggggaaagca	agcgcuugag	1320
acuguccaaac	gccuucugcc	cgucuugugu	caagauacg	ggcuaucucc	ggaccagguc	1380
guggcuauug	ccagcaacaa	cgccgggaag	caagcccugg	aaacugugca	gcccugug	1440

ccggucuugu gccaagacca cgggcucacc ccggaucaag ucgucgcuau cgcaucgaac 1500  
gguggaggua aacaagcacu ggaaaccguc cagaggcugc uuccgguccu cugucaagac 1560

cauggccuca cuccugauca aguggucgcc auagccagca auaucggugg gaaacaggca 1620  
cuggagagca ucgucgccc gcuuucccg ccagauccgg cccuggccgc uuugacgaac 1680  
gaccaucugg ugcccuggc uugccugggg ggaaggccag cgauggacgc cgugaagaag 1740  
ggacugcccc augcccuga gcugauucgg cgagugaaca ggccggaucgg cgaacggacc 1800  
ucccaccgcg uggcgaucuc cagagucggg ggaucgucca gacgcgaguc caucaaccca 1860  
uggauccuca ccggauucgc cgaugcagaa ggguguucc ggcuggauau cagaaacgcg 1920  
aacgaucucc gugcuggua ccggaccaga cuggccuucg agaucgugc gcacaacaag 1980

gacaagagca uucucgaaaa cauccagucg accuggaaag uggaacuau cuacaacg 2040  
ggagauaacg cggugcgccu gcaagucacc agguucgagg auciacaagg gaucaucgau 2100  
cacuucgaga aguaccucu gaucacccaa aagcuggug auuacaagcu guucaagcag 2160  
gcguucuccg ucauggagaa caaggaacac cucaaggaaa acggcauua agagcucgug 2220  
agaaucaagg ccaagaugaa cuggggccug aacgacgaac ucaagaaagc cuuccccag 2280  
aacaucucca aggagcgguc ccugaucaac aagaacaucc ccaaccugaa guggcucg 2340  
ggcuucacca gcgaggauugg cucguucgug gucgaacuga agaagaggcg gucccccagug 2400

aaggugggag ugcgacugcg guucuccauu acucagcaca uccgcgacaa gaaccug 2460  
aacucgcuga ucaccuaccu uggaugcggc agaaucgugg agaaauaaca guccgagcac 2520  
uccuggcugg agiuucaucgu gaccaaguu ucggauauua augacaagau uauuccggug 2580  
uuccaggaaa acacccugau ugugucaaa cuggaggacu uugaggacug gugcaaggug 2640  
gccaaagcuua ucgaaagagaa gaagcaccug accgaguc 60  
gaccaaggacaa ggucgcgug uucuccgguc gcugauag 2700  
aucaagcuga acaugaacaa 2748

&lt;210&gt; 34

&lt;211&gt; 2748

&lt;212&gt; RNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Made in Lab - synthesized megaTAL CBLB.D2 mRNA

&lt;400&gt; 34

augggauccg cgccaccuaa gaagaaacgc aaagucgugg auciucgcac gcucggcuac 60  
agucagcgc agcaagagaa gaucaaaaccg aaggugcguu cgacaguggc gcagcaccac 120  
gaggcacugg ugcccuaagg guuuacacac ggcacacucg uugcgcucag ccaacacccg 180

gcagcguuag ggaccgucgc ugucacguau cagcacauaa ucacggcguu gccagaggcg	240
acacacgaag acaucguugg cgucggcaaa cagugguccg ggcacgcgc ccuggaggcc	300
uugcucacgg augcggggga guugagaggu ccgcccguuac aguuggacac aggccaacuu	360
gugaagauug caaaacgugg cggcgugacc gcaaugaggagg cagugcaugc aucgcgcaau	420
gcacugacgg gugcccccua gaacuuuaaca cccgaucaag ugguagcguu agcguuacac	480
ggcggggua aacaggcuuu ggagacggua cagcgguuau ugcccguuac cugccaggac	540
cacggauuga caccggacca agugguggcg auugcguuca acaacggagg caagcaggca	600
cuagagacug uccaacggcu uciuuccguu cuuugucagg aucaugggcu aaccccugau	660
cagguagucg cuauagcuc aaaauggaggg ggcaagcaag cacuggagac uguucaacga	720
cuccugccag ugcucugcca agaccacgga cuuacaccag aucaaguggu ugcuaauugcc	780
uccaauggug gcgggaaaca agcguuggaa acugugcaga gacuguuacc uguuuugugu	840
caagaccacg gccucacgccc agaucaggug guagccauag cgucgaacau uggugguaag	900
caagccuug aaacggucca ggcguuucug ccgguguugu gccaggacca cggacuaacg	960
ccggauacagg ucguaagccau ugcuuaauu gggggcgca aacaggcguu agagacaguc	1020
cagcgccucu ugcccuguu augccaggau cacggcuuaa cccagacca aguuguggcu	1080
auugcaucua acaauggugg caaacaagcc uuggagacag ugcaacgauu acugccuguc	1140
uuaugucagg aucaugggcuu gacgcccguu cagguagugg caaucgcauc uaacauugga	1200
gguaagcaag cacuggagac uguccagaga uuguuacccg uacuauguca agaucauggu	1260
uugacgcccug aucagguugu ugcgauagcc agcaauaacf gagggaaaca ggcucuugaa	1320
accguacagc gacuucuccc agucuugugc caagauacag ggcguuacucc ugaucaaguc	1380
guagcuaucg ccagcaauaa cgguggaaaa caggcccugg aaaccguaca acgucuccuc	1440
ccaguacuuu gucaagacca cggguugacu ccggaucaag ugcgucgcguu cgcgagcaau	1500
ggugggggga agcaggcguu ggaaacuguu cagagacugc ugcccuguu uugucaggac	1560
cauggucuga caccugacca aguuguggcg auagccagua auaucggggg aaaacaggca	1620
cuagagagca uuguggccca gcugagccgg ccugauccgg cguuggccgc guugaccaac	1680
gaccaccucg ugcguuuggc cugccucggc ggacguuccug ccauggaugc agugaaaaag	1740
ggauugccgc acgcgcccggaa auugaucaga agagucaauu ggcguuauugg cgaacgcacg	1800
ucccaucgcg uugcguauac uagaguggga ggaagcucgc gcagagaguc caucaaccca	1860
uggauucuga cugguuucgc ugaugccgaa ggaugcguuac gacuagacau ccacaacgc	1920
aacguauuga ggucugguua cagaacuaga cuguccuuacg aaaucguuac gcacaacaag	1980

gacaaaucga uucuggagaa uauccagucg acuuggaagg ucggcaaaau cuacaacg 2040

ggcgacaacg cagucagacu gcaagucaca cguuucgaag auuugaaagu gauuaucgac 2100  
 cacuucgaga aauauccgcu gauaacacag aaauuggcg auuacaaguu guuuaacag 2160  
 gcauucagcg ucauggagaa caaagaacau cuuaaggaga augggauua ggagcucgua 2220  
 cgaaucaaaag cuaagaugaa uuggggucuc aaugacgaaau ugaaaaaaagc auuuccagag 2280  
 aacauuagca aagagcgccc ccuuaucaau aagaacauuc cgaaucucaa auggcuggcu 2340  
 ggauucacau cuggugacgg cucguucaug guggaacuaa ugaagaaaua gaauaacuu 2400  
 auuguacgug ugcgucugag auucucaauc ucccagcaca ucagagacaa gaaccugau 2460

aauucauuga uaacauaccu aggcuguggu cguaucguug agaauaacaa aucugagcac 2520  
 aguuggcucg aauucauugu aacaaaauuc agcgauauca acgacaagau cauuccggua 2580  
 uuuccaggaaa auacucugau uggcguaaa cucgaggacu uugaagauug gugcaagguu 2640  
 gccaaauuga ucgaagagaa gaaacaccug accgaaucg guuuggauga gauuaagaaa 2700  
 aucaagcuga acaugaacaa aggucguguc uucagcggcc gcugauaa 2748

<210> 35

<211> 2748

<212> RNA

<213> Artificial Sequence

<220><223> Made in Lab - synthesized megaTAL CBLB.B5 mRNA

<400> 35

augggauccg cgccaccuaa gaagaaacgc aaagucgugg aucuacgcac gcucggcuac 60  
 agucagcagc agcaagagaa gauaaacccg aaggugcgguu cgacaguggc gcagcaccac 120  
 gaggcacugg uggcccaugg guuuacacac ggcacacaucg uugcgcucag ccaacacccg 180  
 gcagcguuag ggaccgucgc uguacguau cagcacauaa ucacggcguu gccagaggcg 240  
 acacacgaag acaucguugg cgucggcaaa cagugguccg ggcacgcgc ccuggaggcc 300  
 uugcucacgg augcggggga guugagaggu ccgcccguuac aguuggacac aggccaacuu 360  
 gugaagauug caaaacgugg cggcgugacc gcaauuggagg cagugcaugc aucgcgcaau 420

gcacugacgg gugcccccua gaacuuaaca cccgaucaag ugguagcgau agcgucaaac 480  
 ggcggggua aacaggcguu ggagacggua cagcgguuau ugccgguacu cugccaggac 540  
 cacgauuga cacggacca agugguggcg auugcgucca acaacggagg caagcaggca 600  
 cuagagacug uccaacggcu ucuuuccguu cuuugucagg aucaugggcu aacccugau 660  
 cagguagucg cuauagcguuc aaauggaggg ggcaagcaag cacuggagac uguuacaacga 720

cuccugccag ugcucugcca agaccacgga cuuacaccag aucaaguggu ugcuaauugcc	780
uccaauggug gcgggaaaca agcguuggaa acugugcaga gacuguuacc ugucuugugu	840
caagaccacg gccucacgccc agaucaggug guagccauag cgucgaacau ugugguaag	900
caagccuug aaacggucca gcgucuucug ccgguguugu gccaggacca cggacuaacg	960
ccggauccagg ucguagccau ugcuuaauu gggggcggca aacaggcgcu agagacaguc	1020
cagcgccuuc uggcuguguu augccaggau cacggcuuaa cccagacca aguuguggcu	1080
auugcaucua acaauggugg caaacaagcc uuggagacag ugcaacgauu acugccuguc	1140
uuaugucagg aucauggccu gacgcccggau cagguagugg caaucgcauc uaacauugga	1200
gguaaagcaag cacuggagac uguccagaga uuguuacccg uacuauguca agaucauggu	1260
uugacgcccug aucagguugu ugcgauagcc agcaauaacg gaggaaaca ggcucuugaa	1320
accguacagc gacuucuccc agucuugugc caagauacg ggcuuacucc ugaucaaguc	1380
guagcuaucg ccagcauaaa cgguggaaaa caggcccugg aaaccguaca acgucuccuc	1440
ccaguacuuu guaagacca cggguugacu ccggaucaag ucgucgcgau cgcgagcaau	1500
ggugggggga agcaggcgcu gaaacuguu cagagacugc ugccuguacu uugucaggac	1560
cauggucuga caccugacca aguuguggcg auagccagua auaucggggg aaaacaggca	1620
cuagagagca uuguggccca gcugagccgg ccugauccgg cguuggccgc guugaccaac	1680
gaccaccucg ucgcccuggc cugccucggc ggacguccug ccauggaugc agugaaaaag	1740
ggauugccgc acgcgcggaa auugaucaga agagucaauc gccguauugg cgaacgcacg	1800
ucccaucgcg uugcgauauc uagaguggga ggaagcucgc gcagagaguc caucaaccca	1860
uggauucuga cugguuucgc ugaugccgaa ggaugcucc gacuagacau ccacaacgca	1920
aacguauuga ggucugguuu cagaacuaga cuguccuucg aaaucguacu gcacaacaag	1980
gacaaaucga uucuggagaa uauccagucg acuuggaagg ucggcacaau cuacaacgcg	2040
ggcgacaaacg cagucagacu gcaagucaca cguuucgaaag auuugaaagu gauuaucgac	2100
cacuucgaga aaauucgcu gauaacacag aaauuggcg auuacaaguu guuuuacag	2160
gcauucagcc ucauggagaa caaagaacau cuuaaggaga augggauuaa ggagcucgua	2220
cgaaucaaag cuaagaugaa uuggggucuc aaugacgaa ugaaaaaaagc auuuuccagag	2280
aacauuagca aagagcgccc ccuuaucaau aagaacauuc cgaauucuaa auggcuggcu	2340
ggauucacau cuggugacgg cucguucgug guggaacuaa agaagagaag aagccccguc	2400
aagguaaggag ugcggcugcg auucggcauc acccagcaca ucagagacaa gaaccugaug	2460

aauucauuga uaacauaccu aggcuguggu cguaucguug agaauaaca	aucugagcac	2520
aguuggcucg aauucauugu aacaaaauuc agcgauauca acgacaagau cauuccggua		2580
uuccaggaaa auacucugau ugugcguaaa cucgaggacu uugaagauug gugcaagguu		2640
gccaaauuga ucgaagagaa gaaacaccug accgaaucgg guuuggauga gauuaagaaa		2700
aucaagcuga acaugaacaa aggucguguc uucagcggcc gcugauaa		2748
<210> 36		
<211> 2748		
<212> RNA		
<213> Artificial Sequence		
<220><223> Made in Lab - synthesized megaTAL CBLB.F6 mRNA		
<400> 36		
augggaauccg cggcaccuaa gaagaaacgc aaagucgugg aucuacgcac gcucggcuac		60
agucagcagc agcaagagaa gaucaaaccg aaggugcguu cgacaguggc gcagcaccac		120
gaggcacugg ugugcccaugg guuuuacacac ggcacaucg uugcgucucag ccaacacccg		180
gcagcguuag ggaccgucgc uguacguau cagcacauaa ucacggcguu gccagaggcg		240
acacacgaag acaucguugg cgucggcaaa cagugguccg ggcacgcgc ccuggaggcc		300
uugcucacgg augcggggga guugagaggu ccgcccguuac aguuggacac aggccaacuu		360
gugaagauug caaaaacgugg cggcgugacc gcaauggagg cagugcaugc aucgcgcaau		420
gcacugacgg gugcccccua gaacuuaaca cccgaucaag ugguagcgu agcgucaaac		480
ggcgaaaaaa aacaggcuuu ggagacggua cagcgguuau ugccgguacu cugccaggac		540
cacggauuga cacggacca agugguggcg auugcgucca acaacggagg caagcaggca		600
cuagagacug uccaacggcu ucuuuccguu cuuugucagg aucauggggcu aaccccugau		660
cagguagucg cuauagcucuuc aaauggaggg ggcaagcaag cacuggagac uguuacaacga		720
cuccugccag ugcucugcca agaccacgga cuuacaccag aucaaguggu ugcuaauugcc		780
uccaauggug gcgaaaaaca agcguuggaa acugugcaga gacuguuacc ugcuuugugu		840
caagaccacg gccucacgccc agaucaggug guagccauag cgucgaacau ugugguuag		900
caagccuug aaacggucca ggcguuucug ccgguguguu gccaggacca cggacuaacg		960
ccggauacagg uguagccau ugcguucaauu gggggcggca aacaggcgcu agagacaguc		1020
cagcgccucu ugcguuguguu augccaggau cacggcuuua ccccagacca aguuguggcu		1080
auugcaucua acaauggugg caaacaagcc uuggagacag ugcaacgauu acugccuguc		1140
uuauugucagg aucauggccu gacgcccggau cagguagugg caaucgcauc uaacauugga		1200

gguaagcaag cacuggagac uguccagaga uuguuacccg uacuauguca agaucauggu	1260
uugacgccug aucagguugu ugcgauagcc agcaauaacg gagggaaaca ggcucuugaa	1320
accguacagc gaciuucuccc agucuugugc caagaucacg ggcuuacucc ugaucuaguc	1380
guagcuaucg ccagcaauaa cgugggaaaa caggcccugg aaaccguaca acgucuccuc	1440
ccaguacuuu gucaagacca cggguugacu ccggaucaag ugcugcgcgau cgcgagcaau	1500
ggugggggga agcaggcgcu ggaaacuguu cagagacugc ugccuguacu uugucaggac	1560
cauggucuga caccugacca aguuguggcg auagccagua auaucggggg aaaacaggca	1620
cuagagagca uuguggccca gcugagccgg ccugauccgg cguuggccgc guugaccaac	1680
gaccaccucg ucgcuuggc cugccucggc ggacguccug ccauggaugc agugaaaaag	1740
ggaugccgc acgcgcccga auugaucaga agagucaauc gccguauugg cgaacgcacg	1800
ucccaucgcg uugcgauauc uagaguggga ggaagcucgc gcagagaguc caucaaccca	1860
uggauucuga cugguuuucgc ugaugccgaa ggaugcucc gacuagacau ccgcaacgca	1920
aacgauuuua gagccggaua cagaacuaga cuggccuucg aaaucguacu gcacaacaag	1980
gacaaaucga uucuggagaa uauccagucg acuuggaagg ucggcacaau cuacaacgca	2040
ggcgacaacg cagucagacu gcaagucaca cguuucgaag auuugaaagu gauuaucgac	2100
cacuucgaga aauauccgcu gauaacacag aaauuggcg auuacaaguu guuuuaacag	2160
gcauucagcg ucauggagaa caaagaacau cuuaaggaga augggauua ggagcucgua	2220
cgaaucaaag cuaagaugaa uuggggucuc aaugacgaaau ugaaaaaaagc auuuccagag	2280
aacauuagca aagagcgccc cciuaucaaau aagaacauuc cgaaucucaa auggcuggcu	2340
ggauucacau cuggugacgg cucguucaug guggaacuaa ugaagaauua gaauaacguu	2400
auuguaucug ugcgucugag auucucaauc ucccagcaca ucagagacaa gaaccugaug	2460
aauucauuga uaacauaccc aggcuguggu cguaucguug agaauaaacaa aucugagcac	2520
aguuggcucg aauucauugu aacaaaauuc agcgauauca acgacaagau cauuccggua	2580
uuccaggaaa auacucugau uggegucaaa cucgaggacu uugaagauug gugcaaggua	2640
gccaaauuga ucgaagagaa gaaacaccug accgaauccg guuuggauga gauuaagaaa	2700
aucaagcuga acaugaacaa aggcuguguc uucagcggcc gcugauaa	2748

<210> 37  
 <211> 711  
 <212> RNA  
 <213> *Mus musculus*  
 <400> 37

augucugagc caccucgggc ugagaccuuu guauuccugg accuagaagc cacugggcuc 60  
 ccaaacaugg acccugagau ugcagagaua uccuuuuug cuguucaccg cucuuccug 120  
 gagaacccag aacgggauga uucugguucc uuggugcugc cccguguucu ggacaagcuc 180

acacugugca ugugcccgga gcgccccuuu acugccaagg ccagugagau uacugguuug 240  
 agcagcgaaa gccugaugca cugcgggaag gcugguuuca auggcgcugu gguaggaca 300  
 cugcagggcu uccuaagccg ccaggaggc cccaucugcc uuguggccca caauggcuuc 360  
 gauuaugacu ucccacugcu gugcacggag cuacaacguc uggugccca ucugcccaa 420  
 gacacugucu gccuggacac acugccugca uugcggggcc uggaccgugc ucacagccac 480  
 ggcaccaggc cucaaggccg caaaagcuac agccuggcca gucucuucca ccgcuacuuc 540  
 caggcugaac ccagugcugc ccauucagca gaaggugaug ugcacacccu gcuucugauc 600

uuccugcauc gugcuccuga gcugcucgccc ugggcagaug agcaggcccg cagcuggggcu 660  
 cauauugagc ccauguacgu gccaccugau gguccaagcc ucgaagccug a 711

<210> 38

<211> 236

<212> PRT

<213> Mus musculus

<400> 38

Met Ser Glu Pro Pro Arg Ala Glu Thr Phe Val Phe Leu Asp Leu Glu

1 5 10 15

Ala Thr Gly Leu Pro Asn Met Asp Pro Glu Ile Ala Glu Ile Ser Leu

20 25 30

Phe Ala Val His Arg Ser Ser Leu Glu Asn Pro Glu Arg Asp Asp Ser

35 40 45

Gly Ser Leu Val Leu Pro Arg Val Leu Asp Lys Leu Thr Leu Cys Met

50 55 60

Cys Pro Glu Arg Pro Phe Thr Ala Lys Ala Ser Glu Ile Thr Gly Leu

65 70 75 80

Ser Ser Glu Ser Leu Met His Cys Gly Lys Ala Gly Phe Asn Gly Ala

85 90 95

Val Val Arg Thr Leu Gln Gly Phe Leu Ser Arg Gln Glu Gly Pro Ile

100 105 110

Cys Leu Val Ala His Asn Gly Phe Asp Tyr Asp Phe Pro Leu Leu Cys

115 120 125

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

130 135 140

Leu Asp Thr Leu Pro Ala Leu Arg Gly Leu Asp Arg Ala His Ser His

145 150 155 160

Gly Thr Arg Ala Gln Gly Arg Lys Ser Tyr Ser Leu Ala Ser Leu Phe

165 170 175

His Arg Tyr Phe Gln Ala Glu Pro Ser Ala Ala His Ser Ala Glu Gly

180 185 190

Asp Val His Thr Leu Leu Leu Ile Phe Leu His Arg Ala Pro Glu Leu

195 200 205

Leu Ala Trp Ala Asp Glu Gln Ala Arg Ser Trp Ala His Ile Glu Pro

210 215 220

Met Tyr Val Pro Pro Asp Gly Pro Ser Leu Glu Ala

225 230 235

<210> 39

<211> 3

<212> PRT

<213> Artificial Sequence

<220><223> Exemplary linker sequence

<400

> 39

Gly Gly Gly

1

<210> 40

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> Exemplary linker sequence

<400> 40

Asp Gly Gly Gly Ser

1 5

<210> 41

<211> 5  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Exemplary linker sequence  
 <400> 41

Thr Gly Glu Lys Pro

1 5

<210> 42  
 <211> 4  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Exemplary linker sequence  
 <400> 42

Gly Gly Arg Arg

1  
 <210> 43  
 <211> 5  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Exemplary linker sequence  
 <400> 43

Gly Gly Gly Gly Ser

1 5  
 <210> 44  
 <211> 14  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Exemplary linker sequence  
 <400> 44

Glu Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Val Asp

1 5 10  
 <210> 45  
 <211> 18  
 <212> PRT

<213> Artificial Sequence

<220><223> Exemplary linker sequence

<400> 45

Lys Glu Ser Gly Ser Val Ser Ser Glu Gln Leu Ala Gln Phe Arg Ser

1 5 10 15

Leu Asp

<210> 46

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> Exemplary linker sequence

<400> 46

Gly Gly Arg Arg Gly Gly Gly Ser

1 5

<210> 47

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> Exemplary linker sequence

<400> 47

Leu Arg Gln Arg Asp Gly Glu Arg Pro

1 5

<210> 48

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> Exemplary linker sequence

<400> 48

Leu Arg Gln Lys Asp Gly Gly Ser Glu Arg Pro

1 5 10

<210> 49

<211> 16

<212> PRT

<213> Artificial Sequence

<220><223> Exemplary linker sequence

<400> 49

Leu Arg Gln Lys Asp Gly Gly Ser Gly Gly Ser Glu Arg Pro

1 5 10 15

<210> 50

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Cleavage sequence by TEV protease

<220><221> MISC\_FEATURE

<222> (2)..(3)

<223> Xaa is any amino acid

<220><221> MISC\_FEATURE

<222> (5)

<223> Xaa is any amino acid

<220><221> MISC\_FEATURE

<222> (7)

<223> Xaa = Gly or Ser

<400> 50

Glu Xaa Xaa Tyr Xaa Gln Xaa

1 5

<210> 51

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Cleavage sequence by TEV protease

<400> 51

Glu Asn Leu Tyr Phe Gln Gly

1 5

<210> 52

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Cleavage sequence by TEV protease

<400> 52

Glu Asn Leu Tyr Phe Gln Ser

1 5

<210> 53

<211> 22

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 53

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> 54

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 54

Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn

1 5 10 15

Pro Gly Pro

<210> 55

<211> 14

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 55

Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro

1 5 10

<210> 56

<211> 21

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 56

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> 57

<211> 18

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 57

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

1 5 10 15

Gly Pro

<210> 58

<211> 13

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 58

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

1 5 10

<210> 59

<211> 23

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 59

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> 60

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 60

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> 61

<211> 14

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 61

Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro Gly Pro

1 5 10

<210> 62

<211> 25

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 62

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> 63

<211> 22

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 63

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> 64  
<211> 14  
<212> PRT  
<213> Artificial Sequence  
<220><223> Self-cleaving polypeptide comprising 2A site  
<400> 64  
Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro Gly Pro

1 5 10

<210> 65

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 65

Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn

1 5 10 15  
Pro Gly Pro

<210>	66
<211>	19
<212>	PRT
<213>	Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 66

Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn

1 5 10 15

Pro Gly Pro

<210> 67

<211> 14

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 67

Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro Gly Pro

1 5 10

<210> 68

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 68

Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro Gly

1 5 10 15

Pro

<210> 69

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 69

Gln Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser

1 5 10 15

Asn Pro Gly Pro

20

&lt;210&gt; 70

&lt;211&gt; 24

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Self-cleaving polypeptide comprising 2A site

&lt;400&gt; 70

Ala Pro Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly

1 5 10 15

Asp Val Glu Ser Asn Pro Gly Pro

20

&lt;210&gt; 71

&lt;211&gt; 40

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Self-cleaving polypeptide comprising 2A site

&lt;400&gt; 71

Val Thr Glu Leu Leu Tyr Arg Met Lys Arg Ala Glu Thr Tyr Cys Pro

1 5 10 15

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

20 25 30

Ile Val Ala Pro Val Lys Gln Thr

35 40

&lt;210&gt; 72

&lt;211&gt; 18

&lt;212&gt; PRT

&lt;

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Self-cleaving polypeptide comprising 2A site

&lt;400&gt; 72

Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro

1 5 10 15

Gly Pro

<210> 73  
 <211> 40  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Self-cleaving polypeptide comprising 2A site  
 <400> 73

Leu Leu Ala Ile His Pro Thr Glu Ala Arg His Lys Gln Lys Ile Val

1 5 10 15

Ala Pro Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly  
 20 25 30

Asp Val Glu Ser Asn Pro Gly Pro  
 35 40

<210> 74  
 <211> 33  
 <212> PRT  
 <213> Artificial Sequence

<220><223> Self-cleaving polypeptide comprising 2A site

<400> 74

Glu Ala Arg His Lys Gln Lys Ile Val Ala Pro Val Lys Gln Thr Leu

1 5 10 15

Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro Gly  
 20 25 30

Pro

<210> 75  
 <211> 10  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Consensus Kozak sequence  
 <400> 75

gccrccatgg 10

<210> 76  
 <211> 22

<212> DNA

<213> Homo sapiens

<400> 76

tggggatggg aatatcttac ag 22

<210> 77

<211> 293

<212> PRT

<213> Artificial Sequence

<220><223> Synthesized I-OnuI variant CBLB.E3

<400> 77

Ser Ile Asn Pro Trp Ile Leu Thr Gly Phe Ala Asp Ala Glu Gly Cys

1 5 10 15

Phe Arg Leu Asp Ile Arg Asn Ala Asn Asp Leu Arg Ala Gly Tyr Arg

20 25 30

Thr Arg Leu Ala Phe Glu Ile Val Leu His Asn Lys Asp Lys Ser Ile

35 40 45

Leu Glu Asn Ile Gln Ser Thr Trp Lys Val Gly Thr Ile Tyr Asn Ala

50 55 60

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

65 70 75 80

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

85 90 95

Gly Asp Tyr Lys Leu Phe Lys Gln Ala Phe Ser Val Met Glu Asn Lys

100 105 110

Glu His Leu Lys Glu Asn Gly Ile Lys Glu Leu Val Arg Ile Lys Ala

115 120 125

Lys Met Asn Trp Gly Leu Asn Asp Glu Leu Lys Lys Ala Phe Pro Glu

130 135 140

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

145 150 155 160

Lys Trp Leu Ala Gly Phe Thr Ser Gly Asp Gly Ser Phe Val Val Glu

165 170 175

Leu Lys Lys Arg Arg Ser Pro Val Lys Val Gly Val Arg Leu Arg Phe  
 180 185 190  
 Ser Ile Thr Gln His Ile Arg Asp Lys Asn Leu Met Asn Ser Leu Ile  
 195 200 205

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

His Leu Thr Glu Ser Gly Leu Asp Glu Ile Lys Lys Ile Lys Leu Asn  
 275 280 285  
 Met Asn Lys Gly Arg

290  
 <210> 78  
 <211> 293  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Synthesized I-OnuI variant CBLB.F6

<400> 78  
 Ser Ile Asn Pro Trp Ile Leu Thr Gly Phe Ala Asp Ala Glu Gly Cys  
 1 5 10 15  
 Phe Arg Leu Asp Ile Arg Asn Ala Asn Asp Leu Arg Ala Gly Tyr Arg  
 20 25 30

Thr Arg Leu Ala Phe Glu Ile Val Leu His Asn Lys Asp Lys Ser Ile  
 35 40 45

Leu Glu Asn Ile Gln Ser Thr Trp Lys Val Gly Thr Ile Tyr Asn Ala  
 50 55 60  
 Gly Asp Asn Ala Val Arg Leu Gln Val Thr Arg Phe Glu Asp Leu Lys  
 65 70 75 80

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

85 90 95

Gly Asp Tyr Lys Leu Phe Lys Gln Ala Phe Ser Val Met Glu Asn Lys

100 105 110

Glu His Leu Lys Glu Asn Gly Ile Lys Glu Leu Val Arg Ile Lys Ala

115 120 125

Lys Met Asn Trp Gly Leu Asn Asp Glu Leu Lys Lys Ala Phe Pro Glu

130 135 140

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

145 150 155 160

Lys Trp Leu Ala Gly Phe Thr Ser Gly Asp Gly Ser Phe Met Val Glu

165 170 175

Leu Met Lys Asn Lys Asn Asn Val Ile Val Arg Val Arg Leu Arg Phe

180 185 190

Ser Ile Ser Gln His Ile Arg Asp Lys Asn Leu Met Asn Ser Leu Ile

195 200 205

Thr Tyr Leu Gly Cys Gly Arg Ile Val Glu Asn Asn Lys Ser Glu His

210 215 220

Ser Trp Leu Glu Phe Ile Val Thr Lys Phe Ser Asp Ile Asn Asp Lys

225 230 235 240

Ile Ile Pro Val Phe Gln Glu Asn Thr Leu Ile Gly Val Lys Leu Glu

245 250 255

Asp Phe Glu Asp Trp Cys Lys Val Ala Lys Leu Ile Glu Glu Lys Lys

260 265 270

His Leu Thr Glu Ser Gly Leu Asp Glu Ile Lys Lys Ile Lys Leu Asn

275 280 285

Met Asn Lys Gly Arg

290

<210> 79

<211> 293

<212> PRT

<213> Artificial Sequence

&lt;220&gt;&lt;223&gt; Synthesized I-OnuI variant CBLB.E6

&lt;400&gt; 79

Ser Ile Asn Pro Trp Ile Leu Thr Gly Phe Ala Asp Ala Glu Gly Cys

1 5 10 15

Phe Arg Leu Asp Ile Arg Asn Ala Asn Asp Leu Arg Ala Gly Tyr Arg

20 25 30

Thr Arg Leu Ser Phe Glu Ile Ser Leu His Asn Lys Asp Lys Ser Ile

35 40 45

Leu Glu Asn Ile Gln Ser Thr Trp Lys Val Gly Thr Ile Tyr Asn Ala

50 55 60

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

65 70 75 80

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

85 90 95

Gly Asp Tyr Lys Leu Phe Lys Gln Ala Phe Ser Val Met Glu Asn Lys

100 105 110

Glu His Leu Lys Glu Asn Gly Ile Lys Glu Leu Val Arg Ile Lys Ala

115 120 125

Lys Met Asn Trp Gly Leu Asn Asp Glu Leu Lys Lys Ala Phe Pro Glu

130 135 140

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

145 150 155 160

Lys Trp Leu Ala Gly Phe Thr Ser Gly Asp Gly Ser Phe Met Val Glu

165 170 175

Leu Met Lys Asn Lys Asn Asn Val Ile Val Arg Val Arg Leu Arg Phe

180 185 190

Ser Ile Ser Gln His Ile Arg Asp Lys Asn Leu Met Asn Ser Leu Ile

195 200 205

Thr Tyr Leu Gly Cys Gly Arg Ile Val Glu Asn Asn Lys Ser Glu His

210 215 220

Ser Trp Leu Glu Phe Ile Val Thr Lys Phe Ser Asp Ile Asn Asp Lys

225 230 235 240  
 Ile Ile Pro Val Phe Gln Glu Asn Thr Leu Ile Gly Val Lys Leu Glu  
 245 250 255  
 Asp Phe Glu Asp Trp Cys Lys Val Ala Lys Leu Ile Glu Glu Lys Lys  
 260 265 270  
 His Leu Thr Glu Ser Gly Leu Asp Glu Ile Lys Lys Ile Lys Leu Asn  
 275 280 285  
 Met Asn Lys Gly Arg  
 290  
 <210> 80  
 <211> 293  
 <212  
 > PRT  
 <213> Artificial Sequence  
 <220><223> Synthesized I-OnuI variant CBLB.D2  
 <400> 80  
 Ser Ile Asn Pro Trp Ile Leu Thr Gly Phe Ala Asp Ala Glu Gly Cys  
 1 5 10 15  
 Phe Arg Leu Asp Ile His Asn Ala Asn Val Leu Arg Ser Gly Tyr Arg  
 20 25 30  
 Thr Arg Leu Ser Phe Glu Ile Val Leu His Asn Lys Asp Lys Ser Ile  
 35 40 45  
 Leu Glu Asn Ile Gln Ser Thr Trp Lys Val Gly Lys Ile Tyr Asn Ala  
 50 55 60  
 Gly Asp Asn Ala Val Arg Leu Gln Val Thr Arg Phe Glu Asp Leu Lys  
 65 70 75 80  
 Val Ile Ile Asp His Phe Glu Lys Tyr Pro Leu Ile Thr Gln Lys Leu  
 85 90 95  
 Gly Asp Tyr Lys Leu Phe Lys Gln Ala Phe Ser Val Met Glu Asn Lys  
 100 105 110  
 Glu His Leu Lys Glu Asn Gly Ile Lys Glu Leu Val Arg Ile Lys Ala  
 115 120 125  
 Lys Met Asn Trp Gly Leu Asn Asp Glu Leu Lys Lys Ala Phe Pro Glu

130	135	140	
Asn Ile Ser Lys Glu Arg Pro Leu Ile Asn Lys Asn Ile Pro Asn Leu			
145	150	155	160
Lys Trp Leu Ala Gly Phe Thr Ser Gly Asp Gly Ser Phe Met Val Glu			
165	170	175	
Leu Met Lys Asn Lys Asn Val Ile Val Arg Val Arg Leu Arg Phe			
180	185	190	
Ser Ile Ser Gln His Ile Arg Asp Lys Asn Leu Met Asn Ser Leu Ile			

195	200	205	
Thr Tyr Leu Gly Cys Gly Arg Ile Val Glu Asn Asn Lys Ser Glu His			
210	215	220	
Ser Trp Leu Glu Phe Ile Val Thr Lys Phe Ser Asp Ile Asn Asp Lys			
225	230	235	240
Ile Ile Pro Val Phe Gln Glu Asn Thr Leu Ile Gly Val Lys Leu Glu			
245	250	255	
Asp Phe Glu Asp Trp Cys Lys Val Ala Lys Leu Ile Glu Glu Lys Lys			
260	265	270	

His Leu Thr Glu Ser Gly Leu Asp Glu Ile Lys Lys Ile Lys Leu Asn			
275	280	285	
Met Asn Lys Gly Arg			

290			
<210> 81			
<211> 293			
<212> PRT			
<213> Artificial Sequence			
<220><223> Synthesized I-OnuI variant CBLB.C2			
<400> 81			

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

Thr Arg Leu Ser Phe Glu Ile Ser Leu His Asn Lys Asp Lys Ser Ile

35	40	45
Leu Glu Asn Ile Gln Ser Thr Trp Lys Val Gly Thr Ile Tyr Asn Ala		
50	55	60
Gly Asp Asn Ala Val Arg Leu Gln Val Thr Arg Phe Glu Asp Leu Lys		
65	70	75
Val Ile Ile Asp His Phe Glu Lys Tyr Pro Leu Ile Thr Gln Lys Leu		
85	90	95
Gly Asp Tyr Lys Leu Phe Lys Gln Ala Phe Ser Leu Met Glu Asn Lys		
100	105	110
Glu His Leu Lys Glu Asn Gly Ile Lys Glu Leu Val Arg Ile Lys Ala		
115	120	125
Lys Met Asn Trp Gly Leu Asn Asp Glu Leu Lys Lys Ala Phe Pro Glu		
130	135	140
Asn Ile Ser Lys Glu Arg Pro Leu Ile Asn Lys Asn Ile Pro Asn Leu		
145	150	155
Lys Trp Leu Ala Gly Phe Thr Ser Gly Asp Gly Ser Phe Met Val Glu		
165	170	175
Leu Met Lys Asn Lys Asn Asn Val Ile Val Arg Val Arg Leu Arg Phe		
180	185	190
Ser Ile Ser Gln His Ile Arg Asp Lys Asn Leu Met Asn Ser Leu Ile		
195	200	205
Thr Tyr Leu Gly Cys Gly Arg Ile Val Glu Asn Asn Lys Ser Glu His		
210	215	220
Ser Trp Leu Glu Phe Ile Val Thr Lys Phe Ser Asp Ile Asn Asp Lys		
225	230	235
Ile Ile Pro Val Phe Gln Glu Asn Thr Leu Ile Gly Val Lys Leu Glu		
245	250	255
Asp Phe Glu Asp Trp Cys Lys Val Ala Lys Leu Ile Glu Glu Lys Lys		
260	265	270
His Leu Thr Glu Ser Gly Leu Asp Glu Ile Lys Lys Ile Lys Leu Asn		
275	280	285
Met Asn Lys Gly Arg		

290

<210> 82

<211> 293

<212> PRT

<213> Artificial Sequence

<220><223> Synthesized I-OnuI variant CBLB.B5

<400> 82

Ser Ile Asn Pro Trp Ile Leu Thr Gly Phe Ala Asp Ala Glu Gly Cys

1	5	10	15
Phe Arg Leu Asp Ile His Asn Ala Asn Val Leu Arg Ser Gly Tyr Arg			
20	25	30	
Thr Arg Leu Ser Phe Glu Ile Val Leu His Asn Lys Asp Lys Ser Ile			
35	40	45	
Leu Glu Asn Ile Gln Ser Thr Trp Lys Val Gly Thr Ile Tyr Asn Ala			
50	55	60	
Gly Asp Asn Ala Val Arg Leu Gln Val Thr Arg Phe Glu Asp Leu Lys			
65	70	75	80
Val Ile Ile Asp His Phe Glu Lys Tyr Pro Leu Ile Thr Gln Lys Leu			
85	90	95	
Gly Asp Tyr Lys Leu Phe Lys Gln Ala Phe Ser Leu Met Glu Asn Lys			
100	105	110	
Glu His Leu Lys Glu Asn Gly Ile Lys Glu Leu Val Arg Ile Lys Ala			
115	120	125	
Lys Met Asn Trp Gly Leu Asn Asp Glu Leu Lys Lys Ala Phe Pro Glu			
130	135	140	
Asn Ile Ser Lys Glu Arg Pro Leu Ile Asn Lys Asn Ile Pro Asn Leu			
145	150	155	160
Lys Trp Leu Ala Gly Phe Thr Ser Gly Asp Gly Ser Phe Val Val Glu			
165	170	175	
Leu Lys Lys Arg Arg Ser Pro Val Lys Val Gly Val Arg Leu Arg Phe			
180	185	190	
Gly Ile Thr Gln His Ile Arg Asp Lys Asn Leu Met Asn Ser Leu Ile			

195

200

205

Thr Tyr Leu Gly Cys Gly Arg Ile Val Glu Asn Asn Lys Ser Glu His

210

215

220

Ser Trp Leu Glu Phe Ile Val Thr Lys Phe Ser Asp Ile Asn Asp Lys

225 230 235 240

Ile Ile Pro Val Phe Gln Glu Asn Thr Leu Ile Gly Val Lys Leu Glu

245

250

255

Asp Phe Glu Asp Trp Cys Lys Val Ala Lys Leu Ile Glu Glu Lys Lys

260

265

270

His Leu Thr Glu Ser Gly Leu Asp Glu Ile Lys Lys Ile Lys Leu Asn

275

280

285

Met Asn Lys Gly Arg

290

&lt;210&gt; 83

&lt;211&gt; 293

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Synthesized I-OnuI variant CBLB.A8

&lt;400&gt; 83

Ser Ile Asn Pro Trp Ile Leu Thr Gly Phe Ala Asp Ala Glu Gly Cys

1

5

10

15

Phe Arg Leu Asp Ile Arg Asn Ala Asn Asp Leu Arg Ala Gly Tyr Arg

20

25

30

Thr Arg Leu Ala Phe Glu Ile Val Leu His Asn Lys Asp Lys Ser Ile

35

40

45

Leu Glu Asn Ile Gln Ser Thr Trp Lys Val Gly Thr Ile Tyr Asn Ala

50

55

60

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

65

70

75

80

Val Ile Ile Gly His Phe Glu Lys Tyr Pro Leu Ile Thr Gln Lys Leu

85

90

95

Gly Asp Tyr Lys Leu Phe Lys Gln Ala Phe Ser Val Met Glu Asn Lys

100	105	110
Glu His Leu Lys Glu Asn Gly Ile Lys Glu Leu Val Arg Ile Lys Ala		
115	120	125
Lys Met Asn Trp Gly Leu Asn Asp Glu Leu Lys Lys Ala Phe Pro Glu		
130	135	140
Asn Ile Ser Lys Glu Arg Pro Leu Ile Asn Lys Asn Ile Pro Asn Leu		
145	150	155
160		
Lys Trp Leu Ala Gly Phe Thr Ser Gly Asp Gly Ser Phe Met Val Glu		
165	170	175
Leu Met Lys Asn Lys Asn Asn Val Ile Val Arg Val Arg Leu Arg Phe		
180	185	190
Ser Ile Ser Gln His Ile Arg Asp Lys Asn Leu Met Asn Ser Leu Ile		
195	200	205
Thr Tyr Leu Gly Cys Gly Arg Ile Val Glu Asn Asn Lys Ser Glu His		
210	215	220
Ser Trp Leu Glu Phe Ile Val Thr Lys Phe Ser Asp Ile Asn Asp Lys		
225	230	235
240		
Ile Ile Pro Val Phe Gln Glu Asn Thr Leu Ile Gly Val Lys Leu Glu		
245	250	255
Asp Phe Glu Asp Trp Cys Lys Val Ala Lys Leu Ile Glu Glu Lys Lys		
260	265	270
His Leu Thr Glu Ser Gly Leu Asp Glu Ile Lys Lys Ile Lys Leu Asn		
275	280	285
Met Asn Lys Gly Arg		
290		
<210>	84	
<211>	39	
<212>	DNA	
<213>	Homo sapiens	
<400>	84	
tggggatggg aatatcttac agaccatacc tcataacaa		39