



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0012266  
(43) 공개일자 2022년02월03일

- |                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                                                                |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>(51) 국제특허분류(Int. Cl.)<br/>C12N 15/86 (2006.01) A61K 48/00 (2006.01)<br/>A61P 19/00 (2006.01) C07K 14/47 (2006.01)</p> <p>(52) CPC특허분류<br/>C12N 15/86 (2013.01)<br/>A61K 48/00 (2013.01)</p> <p>(21) 출원번호 10-2021-7040611</p> <p>(22) 출원일자(국제) 2020년05월22일<br/>심사청구일자 없음</p> <p>(85) 번역문제출일자 2021년12월10일</p> <p>(86) 국제출원번호 PCT/US2020/034394</p> <p>(87) 국제공개번호 WO 2020/237219<br/>국제공개일자 2020년11월26일</p> <p>(30) 우선권주장<br/>62/852,216 2019년05월23일 미국(US)</p> | <p>(71) 출원인<br/>스페이스크래프트 세븐, 엘엘씨<br/>미국, 뉴저지 08512, 크랜베리, 9 세다 브룩 드라이브</p> <p>(72) 발명자<br/>비어드 브라이언<br/>미국, 뉴저지 08512, 크랜베리, 9 세다 브룩 드라이브, 씨/오 스페이스크래프트 세븐, 엘엘씨<br/>릭스 데이비드<br/>미국, 뉴저지 08512, 크랜베리, 9 세다 브룩 드라이브, 씨/오 스페이스크래프트 세븐, 엘엘씨<br/>프라바카 라이<br/>미국, 뉴저지 08512, 크랜베리, 9 세다 브룩 드라이브, 씨/오 스페이스크래프트 세븐, 엘엘씨</p> <p>(74) 대리인<br/>특허법인한얼</p> |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

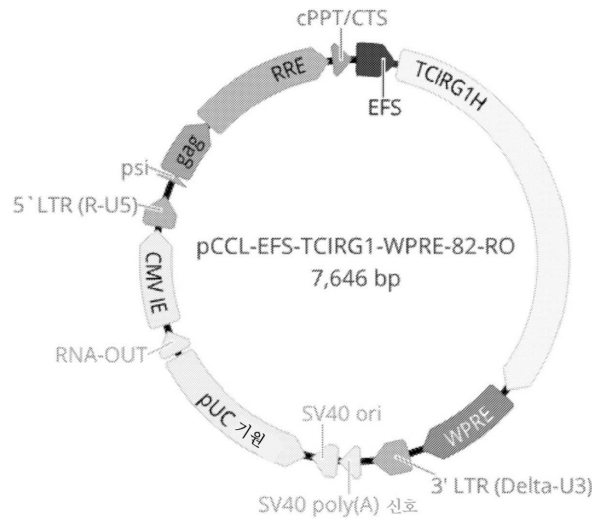
전체 청구항 수 : 총 70 항

(54) 발명의 명칭 영아 악성 골화석증을 위한 유전자 요법 벡터

(57) 요약

본 개시는 TCIRG1 폴리펩티드 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드 서열을 포함하는 개선된 유전자 요법 벡터, 이의 사용 방법, 약학적 조성물 등을 제공한다. 특히, 본 개시는 영아 악성 골화석증(IMO)의 치료를 위한 렌티바이러스 벡터를 제공한다.

대표도



(52) CPC특허분류

*A61P 19/00* (2018.01)

*C07K 14/4702* (2013.01)

*C12N 2740/16043* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

T 세포 면역 조절인자 1(TCIRG1)의 이소형 또는 이의 기능적 변이체를 암호화하는 코딩 폴리뉴클레오티드 및 프로모터를 포함하는 발현 카세트를 포함하는 전달 플라스미드로서, 여기서 폴리뉴클레오티드는 프로모터에 작동 가능하게 연결되고, 전달 플라스미드는 RNA-OUT 억제인자 및 CMV IE 프로모터를 포함하는, 전달 플라스미드.

#### 청구항 2

제1항에 있어서, RNA-OUT 억제인자는 서열번호 32와 적어도 95%의 동일성 또는 적어도 99%의 동일성을 공유하는, 전달 플라스미드.

#### 청구항 3

제1항 또는 제2항에 있어서, CMV IE 프로모터는 서열번호 33과 적어도 95%의 동일성 또는 적어도 99%의 동일성을 공유하는, 전달 플라스미드.

#### 청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 전달 플라스미드는 pCCL 백본을 포함하는, 전달 플라스미드.

#### 청구항 5

제4항에 있어서, pCCL 백본은 RNA-OUT 억제인자를 포함하는, 전달 플라스미드.

#### 청구항 6

제5항에 있어서, 전달 플라스미드는 서열번호 39와 적어도 95% 또는 100% 동일성을 공유하는, 전달 플라스미드.

#### 청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, 프로모터는 EFS 프로모터인, 전달 플라스미드.

#### 청구항 8

제7항에 있어서, EFS 프로모터는 서열번호 2와 적어도 95% 동일성을 공유하는, 전달 플라스미드.

#### 청구항 9

제8항에 있어서, EFS 프로모터는 서열번호 2인, 전달 플라스미드.

#### 청구항 10

제1항 내지 제9항 중 어느 한 항에 있어서, 코딩 폴리뉴클레오티드는 서열번호 3과 적어도 95% 동일성을 공유하는, 전달 플라스미드.

#### 청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서, 코딩 폴리뉴클레오티드는 서열번호 3과 적어도 99% 동일성을 공유하는, 전달 플라스미드.

#### 청구항 12

제11항에 있어서, 코딩 폴리뉴클레오티드는 서열번호 3인, 전달 플라스미드.

#### 청구항 13

제1항 내지 제12항 중 어느 한 항에 있어서, 발현 카세트는 마멋 감염 바이러스(WHP) 전사후 조절 요소(WPRE)를

포함하는, 전달 플라스미드.

**청구항 14**

제13항에 있어서, WPRE는 서열번호 4인, 전달 플라스미드.

**청구항 15**

제1항 내지 제14항 중 어느 한 항에 있어서, 발현 카세트는 서열번호 1과 적어도 95% 동일성을 공유하는, 전달 플라스미드.

**청구항 16**

제1항 내지 제15항 중 어느 한 항에 있어서, 발현 카세트는 5' 장 말단 반복(LTR) 및 3' LTR이 측면에 위치하는, 전달 플라스미드.

**청구항 17**

제16항에 있어서, 5' LTR은 서열번호 34이고/이거나 3' LTR은 서열번호 28인, 전달 플라스미드.

**청구항 18**

제1항 내지 제17항 중 어느 한 항에 있어서, 발현 카세트는 서열번호 1과 적어도 95% 동일성을 공유하는, 전달 플라스미드.

**청구항 19**

제1항 내지 제17항 중 어느 한 항에 있어서, 발현 카세트는 서열번호 1인, 전달 플라스미드.

**청구항 20**

제1항 내지 제20항 중 어느 한 항의 전달 플라스미드로 숙주 세포를 형질감염시킴으로써 생산된 렌티바이러스 입자.

**청구항 21**

T 세포 면역 조절인자 1(TCIRG1)의 이소형 또는 이의 기능적 변이체를 암호화하는 코딩 폴리뉴클레오티드 및 EFS 프로모터를 포함하는 발현 카세트로서, 여기서 폴리뉴클레오티드는 EFS 프로모터에 작동 가능하게 연결되는, 발현 카세트.

**청구항 22**

제1항에 있어서, 코딩 폴리뉴클레오티드는 서열번호 3과 적어도 95% 동일성을 공유하는, 발현 카세트.

**청구항 23**

제2항에 있어서, 코딩 폴리뉴클레오티드는 서열번호 3과 적어도 99% 동일성을 공유하는, 발현 카세트.

**청구항 24**

제3항에 있어서, 코딩 폴리뉴클레오티드는 서열번호 3인, 발현 카세트.

**청구항 25**

제1항 내지 제4항 중 어느 한 항에 있어서, EFS 프로모터는 서열번호 2와 적어도 95% 동일성을 공유하는, 발현 카세트.

**청구항 26**

제25항에 있어서, EFS 프로모터는 서열번호 2인, 발현 카세트.

**청구항 27**

제21항 내지 제26항 중 어느 한 항에 있어서, 마멋 간염 바이러스(WHP) 전사후 조절 요소(WPRE)를 포함하는, 발현 카세트.

**청구항 28**

제27항에 있어서, WPRE는 서열번호 4인, 발현 카세트.

**청구항 29**

제21항 내지 제28항 중 어느 한 항에 있어서, 발현 카세트는 서열번호 1과 적어도 95% 동일성을 공유하는, 발현 카세트.

**청구항 30**

제29항에 있어서, 발현 카세트는 서열번호 1인, 발현 카세트.

**청구항 31**

제조합 렌티바이러스 계놈으로서, 5'에서 3'의 순서로:

- (a) 렌티바이러스 5' 장 말단 반복(LTR);
- (b) 제21항 내지 제30항 중 어느 한 항의 발현 카세트; 및
- (c) 렌티바이러스 3' LTR을 포함하되,

제조합 렌티바이러스 계놈은 복제 불능인, 방법.

**청구항 32**

제31항의 제조합 렌티바이러스 계놈을 포함하는, 전달 플라스미드.

**청구항 33**

제31항의 제조합 렌티바이러스 계놈을 포함하는, 렌티바이러스 입자.

**청구항 34**

제33항의 렌티바이러스 입자를 포함하는 약학적 조성물.

**청구항 35**

제21항 내지 제30항 중 어느 한 항의 발현 카세트를 포함하는, 변형된 세포.

**청구항 36**

제31항의 제조합 렌티바이러스 계놈을 포함하는, 변형된 세포.

**청구항 37**

제36항에 있어서, 변형된 세포는 내인성 기능적 *TCIRG1* 유전자가 결여되어 있는, 변형된 세포.

**청구항 38**

제36항 또는 제37항에 있어서, 변형된 세포는 영아 악성 골석화증(IMO)을 앓고 있거나 앓을 것이 의심되는 대상 체로부터 유래되는, 변형된 세포.

**청구항 39**

제36항 내지 제38항 중 어느 한 항에 있어서, 변형된 세포는 기능적 *TCIRG1* 유전자를 갖는 파골세포에서 관찰된 *TCIRG1*의 발현 수준과 유사한 수준으로 *TCIRG1* 또는 이의 기능적 변이체를 발현하는, 변형된 세포.

**청구항 40**

제36항 내지 제39항 중 어느 한 항에 있어서, 변형된 세포는 IMO를 앓고 있지 않거나 앓을 것이 의심되지 않는 대상체로부터 유래된 파골세포에서 관찰된 TCIRG1의 발현 수준과 유사한 수준으로 TCIRG1 또는 이의 기능적 변이체를 발현하는, 변형된 세포.

**청구항 41**

제36항 내지 제40항 중 어느 한 항에 있어서, 변형된 세포는 조혈 줄기 세포(HSC)인, 변형된 세포.

**청구항 42**

제36항 내지 제41항 중 어느 한 항에 있어서, 변형된 세포는 CD34+ 전구세포(progenitor cell)인, 변형된 세포.

**청구항 43**

제41항 또는 제42항 중 어느 한 항에 있어서, 변형된 세포는 IMO를 앓고 있거나 앓을 것이 의심되는 대상체에게 임의로 G-CSF, 플레리팍서(plerifaxor), 또는 G-CSF와 플레리팍서의 조합을 투여하여 HSC를 동원한 후 단리한 HSC로부터 성분채집술에 의해 유래되는, 변형된 세포.

**청구항 44**

제35항 내지 제43항 중 어느 한 항에 있어서, 변형된 세포는 CD34+ 세포에 대해 농축된 세포 모집단으로부터 자기 포획에 의해 유래되는, 변형된 세포.

**청구항 45**

제35항 내지 제44항 중 어느 한 항의 변형된 세포를 포함하는 약학적 조성물.

**청구항 46**

IMO를 앓고 있거나 앓을 것이 의심되는 대상체의 하나 이상의 세포를 변형시키는 시험관 내 방법으로서,

(a) G-CSF, 플레리팍서, 또는 G-CSF와 플레리팍서의 조합을 포함하는 조성물을 대상체에게 투여함으로써 대상체로부터 동원한 말초 혈액 단핵 세포(PBMC)를 제공하는 단계;

(b) 자기 분리에 의해 CD34+ 세포에 대한 PBMC를 농축시켜 CD34-농축 세포의 모집단을 생성하는 단계; 및

(c) 재조합 렌티바이러스 게놈을 포함하는 렌티바이러스 입자와 CD34-농축 세포를 접촉시키는 단계를 포함하되, 재조합 렌티바이러스 게놈은 5'에서 3' 순서로:

(i) 렌티바이러스 5' 장 말단 반복(LTR);

(ii) 제21항 내지 제30항 중 어느 한 항의 발현 카세트; 및

(iii) 렌티바이러스 3' LTR을 포함하며,

재조합 렌티바이러스 게놈은 복제 불능인, 방법.

**청구항 47**

IMO를 앓고 있거나 앓을 것이 의심되는 대상체에서 영아 악성 골화석증(IMO)을 치료하는 방법으로서, 제35항 내지 제44항 중 어느 한 항의 변형된 세포 또는 제45항의 약학적 조성물을 대상체에게 투여하는 단계를 포함하는, 방법.

**청구항 48**

제47항에 있어서, 상기 방법은 TCIRG1 또는 이의 기능적 변이체를 발현하는 변형된 세포로 HSC 서식지를 재충진하는, 방법.

**청구항 49**

제47항 또는 제48항에 있어서, 상기 방법은 TCIRG1 또는 이의 기능적 변이체를 발현하는 변형된 세포로 파골세포 서식지를 재충진하는, 방법.

**청구항 50**

제47항 내지 제49항 중 어느 한 항에 있어서, 상기 방법은 IMO를 치료, 개선, 예방, 감소, 억제 또는 완화시키는, 방법.

**청구항 51**

제47항 내지 제50항 중 어느 한 항에 있어서, 상기 방법은 치료된 대상체의 평균 전체 생존기간을 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10년 또는 그 이상 연장시키는, 방법.

**청구항 52**

제47항 내지 제51항 중 어느 한 항에 있어서, 상기 방법은 IMO로 인한 대상체의 사망을 예방하는, 방법.

**청구항 53**

제47항 내지 제52항 중 어느 한 항에 있어서, 대상체는 인간인, 방법.

**청구항 54**

제47항 내지 제53항 중 어느 한 항에 있어서, 대상체는 치료 전에 IMO의 증상을 나타낸, 방법.

**청구항 55**

제47항 내지 제54항 중 어느 한 항에 있어서, 대상체는 치료 전에 TCIRG1의 발현이 감소되었거나 발현이 검출되지 않는 것으로 식별된, 방법.

**청구항 56**

제47항 내지 제55항 중 어느 한 항에 있어서, 대상체는 돌연변이된 TCIRG1 유전자를 갖는 것으로 식별된, 방법.

**청구항 57**

제47항 내지 제56항 중 어느 한 항에 있어서, 대상체는 영아인, 방법.

**청구항 58**

제47항 내지 제57항 중 어느 한 항에 있어서, 상기 방법은 자가 치료를 포함하는, 방법.

**청구항 59**

제47항 내지 제58항 중 어느 한 항에 있어서, 투여는 정맥내 주입을 통해 수행되는, 방법.

**청구항 60**

영아 악성 골석화증(IMO)의 치료 또는 예방을 위한 의약의 제조에 사용하기 위한 재조합 렌티바이러스 계놈으로서, 상기 렌티바이러스 계놈은 5'에서 3'의 순서로:

- (i) 렌티바이러스 5' 장 말단 반복(LTR),
- (ii) 제21항 내지 제30항 중 어느 한 항의 발현 카세트, 및
- (iii) 렌티바이러스 3' LTR을 포함하며,

재조합 렌티바이러스 계놈은 복제 불능인, 재조합 렌티바이러스 계놈.

**청구항 61**

영아 악성 골석화증(IMO)의 치료 또는 예방을 위한 의약의 제조에 사용하기 위한, 재조합 렌티바이러스 계놈을 포함하는 렌티바이러스 입자로서,

렌티바이러스 계놈은 5'에서 3' 순서로:

- (i) 렌티바이러스 5' 장 말단 반복(LTR),

- (ii) 제21항 내지 제50항 중 어느 한 항의 발현 카세트, 및
  - (iii) 렌티바이러스 3' LTR을 포함하며,
- 재조합 렌티바이러스 계통은 복제 불능인, 렌티바이러스 입자.

**청구항 62**

제24항 내지 제30항 중 어느 한 항의 발현 카세트를 포함하는 전달 플라스미드.

**청구항 63**

제62항에 있어서, RNA-OUT 서열을 추가로 포함하는, 전달 플라스미드.

**청구항 64**

제63항에 있어서, RNA-OUT 서열은 서열번호 22인, 전달 플라스미드.

**청구항 65**

제62항 또는 제63항에 있어서, RNA-OUT 서열은 전달 플라스미드가 패키징 세포주에서 안정하게 전파될 수 있도록 구성되는, 전달 플라스미드.

**청구항 66**

렌티바이러스 입자를 생산하는 방법으로서, 제1항 내지 제19항 또는 제62항 내지 제65항 중 어느 한 항의 전달 플라스미드가 복제되도록 상기 전달 플라스미드로 박테리아 세포를 형질전환시키는 단계, 복제된 전달 플라스미드를 단리하는 단계, 복제된 전달 플라스미드 및 임의로 하나 이상의 추가 플라스미드로 패키징 세포주를 형질 도입하여 렌티바이러스 입자를 생산하는 단계를 포함하는, 방법.

**청구항 67**

렌티바이러스 입자를 생산하는 방법으로서, 제1항 내지 제19항 또는 제62항 내지 제65항 중 어느 한 항의 전달 플라스미드 및 임의로 하나 이상의 추가 플라스미드로 패키징 세포주를 형질감염시키는 단계, 및 상기 패키징 세포주를 배양하는 단계를 포함하는, 방법.

**청구항 68**

제67항에 있어서, 전달 플라스미드는 안정적으로 전파되는, 방법.

**청구항 69**

제68항에 있어서, 전달 플라스미드는 진탕 플라스크를 사용하거나 적어도 1, 2, 3, 4, 5, 6 또는 7일 동안 발효 시킴으로써 30~37°C의 박테리아 숙주에서 안정적으로 전파되는, 방법.

**청구항 70**

제66항 내지 제69항 중 어느 한 항의 방법에 따라 생산된 렌티바이러스 입자.

**발명의 설명**

**기술 분야**

[0001] **관련 출원에 대한 상호참조**

[0002] 본 출원은 2019년 5월 23일에 출원된 미국 특허 가출원 제62/852,216호의 우선권 이익을 주장하며, 본 출원의 내용은 그 전체가 참조로서 본원에 통합된다.

[0003] **서열 목록에 관한 진술**

[0004] 본 출원과 연관된 서열 목록은 종이 사본 대신에 텍스트 형식으로 제공되며, 참조로서 본 명세서에 통합된다. 서열 목록을 포함하는 텍스트 파일의 명칭은 ROPA\_003\_01WQ\_ST25.txt이다. 텍스트 파일은 2020년 5월 22일에 생

성된 약 62 KB이며, EFS-Web을 통해 전자적으로 제출 중이다.

**[0005] 기술분야**

**[0006]** 본 개시는 일반적으로 T 세포 면역 조절인자 1, ATPase H+ 수송 V0 서브유닛 a3 유전자(TCIRG1)에서의 돌연변이와 관련된 질환을 위한 유전자 요법에 관한 것이다. 특히, 본 개시는 TCIRG1 단백질(TCIRG1)을 암호화하는 발현 카세트를 포함하는 플라스미드 및 유전자 요법을 제공한다.

**배경 기술**

**[0007]** 영아 악성 골화석증(IMO)은 기능장애성 파골세포에 의해 야기되는 골량 증가를 특징으로 하는 희귀한 열성 장애이다. 상기 질환은 가장 흔하게는 T 세포 면역 조절인자 1, ATPase H+ 수송 V0 서브유닛 a3(TCIRG1)에서의 돌연변이에 의해 발생한다. TCIRG1은 뼈를 재흡수하는 파골세포의 능력에 관여한다.

**[0008]** 파골세포 기능은 TCIRG1의 렌티바이러스 벡터 매개 발현에 의해 회복될 수 있다. Moscatelli 등의 문헌[Bone 57:1-9 (2013)] 참조. 추가의 연구는, TCIRG1의 렌티바이러스 매개 발현이 구성적 생리학적 프로모터가 포함된 렌티바이러스 벡터에 의해 발현됨에도 불구하고, 내인성 유전자 산물과 동일한 방식으로 조절된다는 것을 보여준다. Thudium 등의 문헌[Calcif Tissue Int. 99:638-648 (2016)] 참조. 또한 이들은, 유전자의 코돈 최적화된 cDNA에서의 mRNA 수준이 상당히 높았지만, 천연 TCIRG1 유전자 서열이 유전자의 코돈 최적화된 cDNA보다 파골세포에서 더 높은 수준의 단백질 발현 및 기능적 구제를 유도한다는 것을 확립하였다. 또한, 데이터는 시험관 내에서 재흡수 기능을 상당히 증가시키는 데에는 기능적 TCIRG1이 포함된 인간 전-골종의 낮은 분획만이 필요하다는 것을 보여주는데(재흡수성 파골세포와 비-재흡수성 파골세포의 융합으로 인한 것일 가능성이 높음), 이는 골석화증의 oc/oc 마우스 모델에서의 이전 결과와 일치한다. 효능 및 안전성 측면 모두에서, 상기 소견은 골석화증에 대한 유전자 요법의 추가적인 개발에 고무적이다.

**[0009]** TCIRG1에 대한 유전자 요법 벡터 및 이러한 벡터를 사용하는 치료 방법에 대한 필요성이 당업계에 남아 있다. 또한, 이러한 유전자 요법 벡터를 생산하는 신뢰할 수 있는 방법이 필요하다. 본 개시는 이러한 유전자 요법 벡터, 이의 제조 방법, 이의 사용 방법, 약학적 조성물 등을 제공한다.

**발명의 내용**

**[0010]** 본 개시는 TCIRG1 폴리펩티드 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드 서열을 포함하는 개선된 유전자 요법 벡터, 이의 사용 방법, 약학적 조성물 등을 제공한다.

**[0011]** 일 양태에서, 본 개시는 T 세포 면역 조절인자 1(TCIRG1)의 이소형 또는 이의 기능적 변이체를 암호화하는 코딩 폴리뉴클레오티드, 및 프로모터를 포함하는 발현 카세트를 포함하는 전달 플라스미드를 제공하며, 여기서 폴리뉴클레오티드는 프로모터에 작동 가능하게 연결되고, 전달 플라스미드는 RNA-OUT 억제인자 및 CMV IE 프로모터를 포함한다.

**[0012]** 일부 구현예에서, RNA-OUT 억제인자는 서열번호 32와 적어도 95%의 동일성 또는 적어도 99%의 동일성을 공유한다.

**[0013]** 일부 구현예에서, CMV IE 프로모터는 서열번호 33과 적어도 95%의 동일성 또는 적어도 99%의 동일성을 공유한다.

**[0014]** 일부 구현예에서, 전달 플라스미드는 pCCL 골격을 포함한다.

**[0015]** 일부 구현예에서, pCCL 골격은 RNA-OUT 억제인자를 포함한다.

**[0016]** 일부 구현예에서, 전달 플라스미드는 서열번호 39와 적어도 95%의 동일성을 공유한다.

**[0017]** 일부 구현예에서, 전달 플라스미드는 서열번호 39를 포함한다.

**[0018]** 일부 구현예에서, 프로모터는 EFS 프로모터이다.

**[0019]** 일부 구현예에서, EFS 프로모터는 서열번호 2와 적어도 95%의 동일성을 공유한다.

**[0020]** 일부 구현예에서, EFS 프로모터는 서열번호 2이다.

**[0021]** 일부 구현예에서, 코딩 폴리뉴클레오티드는 서열번호 3과 적어도 95%의 동일성을 공유한다.

- [0022] 일부 구현예에서, 코딩 폴리뉴클레오티드는 서열번호 3과 적어도 99%의 동일성을 공유한다.
- [0023] 일부 구현예에서, 코딩 폴리뉴클레오티드는 서열번호 3이다.
- [0024] 일부 구현예에서, 발현 카세트는 마멋 간염 바이러스(Woodchuck Hepatitis Virus, WHP) 전사후 조절 요소(WPRE)를 포함한다.
- [0025] 일부 구현예에서, WPRE는 서열번호 4이다.
- [0026] 일부 구현예에서, 발현 카세트는 서열번호 1과 적어도 95%의 동일성을 공유한다.
- [0027] 일부 구현예에서, 발현 카세트는 5' 긴 말단 반복(LTR) 및 3' LTR이 측면에 위치한다.
- [0028] 일부 구현예에서, 5' LTR은 서열번호 34이고/이거나 3' LTR은 서열번호 28이다.
- [0029] 일부 구현예에서, 발현 카세트는 서열번호 1과 적어도 95%의 동일성을 공유한다.
- [0030] 일부 구현예에서, 발현 카세트는 서열번호 1이다.
- [0031] 또 다른 양태에서, 본 개시는 T 세포 면역 조절인자 1(TCIRG1)의 이소형 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드, 및 EFS 프로모터를 포함하는 발현 카세트를 제공하며, 여기서 폴리뉴클레오티드는 임의로 EFS 프로모터에 작동 가능하게 연결된다.
- [0032] 일부 구현예에서, 코딩 폴리뉴클레오티드는 서열번호 3과 적어도 95%의 동일성을 공유한다. 일부 구현예에서, 코딩 폴리뉴클레오티드는 서열번호 3과 적어도 99%의 동일성을 공유한다. 일부 구현예에서, 코딩 폴리뉴클레오티드는 서열번호 3이다. 일부 구현예에서, EFS 프로모터는 서열번호 2와 적어도 95%의 동일성을 공유한다. 일부 구현예에서, EFS 프로모터는 서열번호 2이다.
- [0033] 일부 구현예에서, 발현 카세트는 마멋 간염 바이러스(Woodchuck Hepatitis Virus, WHP) 전사후 조절 요소(WPRE)를 포함한다. 일부 구현예에서, WPRE는 서열번호 4이다.
- [0034] 일부 구현예에서, 발현 카세트는 서열번호 1과 적어도 95%의 동일성을 공유한다. 일부 구현예에서, 발현 카세트는 서열번호 1이다.
- [0035] 또 다른 양태에서, 본 개시는 재조합 렌티바이러스 계놈을 제공하며, 상기 재조합 렌티바이러스 계놈은 5'에서 3'의 순서로 렌티바이러스 5' 장 말단 반복(LTR); 본원에 개시된 발현 카세트; 및 렌티바이러스 3' LTR을 포함하고, 여기서 재조합 렌티바이러스 계놈은 복제 불능(replication incompetent)이다.
- [0036] 또 다른 양태에서, 본 개시는 이러한 재조합 렌티바이러스 계놈을 포함하는 렌티바이러스 입자를 제공한다.
- [0037] 또 다른 양태에서, 본 개시는 이러한 재조합 렌티바이러스 계놈을 포함하는 전달 플라스미드를 제공한다. 특정 구현예에서, 전달 플라스미드는 RNA-OUT 서열을 포함한다. 일부 구현예에서, RNA-OUT 서열은 서열번호 22이다. 일부 구현예에서, RNA-OUT 서열은 패키징 세포주에서 전달 플라스미드가 안정하게 전파할 수 있도록 구성된다.
- [0038] 특정 구현예에서, 전달 플라스미드는 항생제 내성 유전자를 포함하지 않거나, 암피실린 내성 유전자, 예컨대 AmpR을 포함하지 않는다.
- [0039] 특정 구현예에서, 전달 플라스미드는 서열번호 23에 제시된 서열을 포함한다.
- [0040] 또 다른 양태에서, 본 개시는 렌티바이러스 입자를 생성하는 방법을 제공하며, 상기 방법은 본 개시의 임의의 전달 플라스미드 및 임의로 하나 이상의 추가 플라스미드로 패키징 세포주를 형질감염시키는 단계, 및 상기 패키징 세포주를 배양하는 단계를 포함한다. 일부 구현예에서, 전달 플라스미드는 진탕 플라스미드를 사용하거나 적어도 1, 2, 3, 4, 5, 6 또는 7일 동안 발효시킴으로써 30~37°C의 박테리아 숙주에서 안정적으로 전파된다.
- [0041] 관련 양태에서, 본 개시는 본원에 개시된 전달 플라스미드를 사용하여 생산된 렌티바이러스 입자를 제공한다.
- [0042] 또 다른 양태에서, 본 개시는 본 개시의 임의의 렌티바이러스 입자를 포함하는 약학적 조성물을 제공한다.
- [0043] 또 다른 양태에서, 본 개시는 본 개시의 임의의 발현 카세트를 포함하는 변형된 세포를 제공한다.
- [0044] 또 다른 양태에서, 본 개시는 본 개시의 임의의 재조합 렌티바이러스 계놈을 포함하는 변형된 세포를 제공한다.
- [0045] 일부 구현예에서, 변형된 세포는 내인성 기능성 *TCIRG1* 유전자가 결여되어 있다.
- [0046] 일부 구현예에서, 변형된 세포는 영아 악성 골석화증(IMO)을 앓고 있거나 앓을 것이 의심되는 대상체로부터 유

래된다.

- [0047] 일부 구현예에서, 변형된 세포는 기능적 *TCIRG1* 유전자를 갖는 파골세포에서 관찰된 *TCIRG1*의 발현 수준과 유사한 수준으로 *TCIRG1* 또는 이의 기능적 변이체를 발현한다.
- [0048] 일부 구현예에서, 변형된 세포는 IMO를 앓고 있지 않거나 앓을 것이 의심되지 않는 대상체로부터 유래된 파골세포에서 관찰된 *TCIRG1*의 발현 수준과 유사한 수준으로 *TCIRG1* 또는 이의 기능적 변이체를 발현한다.
- [0049] 일부 구현예에서, 변형된 세포는 조혈 줄기 세포(HSC)이다.
- [0050] 일부 구현예에서, 변형된 세포는 CD34+ 전구 세포(progenitor cell)이다.
- [0051] 일부 구현예에서, 변형된 세포는 IMO를 앓고 있거나 앓을 것이 의심되는 대상체로부터 단리한 HSC로부터 성분채집술에 의해 유래된다.
- [0052] 일부 구현예에서, 변형된 세포는 IMO를 앓고 있거나 앓을 것이 의심되는 대상체에게 G-CSF, 플레리팍서(plerifaxor), 또는 G-CSF와 플레리팍서의 조합을 투여하여 HSC를 동원한 후 단리한 HSC로부터 성분채집술에 의해 유래된다.
- [0053] 일부 구현예에서, 변형된 세포는 CD34+ 세포가 풍부한 세포 모집단으로부터 자기 포획에 의해 유래된다.
- [0054] 또 다른 양태에서, 본 개시는 본 개시의 임의의 변형된 세포를 포함하는 약학적 조성물을 제공한다.
- [0055] 또 다른 양태에서, 본 개시는 IMO를 앓고 있거나 앓을 것이 의심되는 대상체의 하나 이상의 세포를 변형시키는 시험관 내 방법을 제공하며, 상기 방법은 G-CSF, 플레리팍서, 또는 G-CSF와 플레리팍서의 조합을 포함하는 조성물을 대상체에게 투여함으로써 대상체로부터 동원한 말초 혈액 단핵 세포(PBMC)를 제공하는 단계; 자기 분리에 의해 CD34+ 세포에 대한 PBMC를 농축시켜 CD34-농축 세포의 모집단을 생성하는 단계; 및 5'에서 3' 순서로 렌티바이러스 5' 장 말단 반복(LTR); 본 개시의 임의의 발현 카세트; 및 렌티바이러스 3' LTR을 포함하는 재조합 렌티바이러스 게놈을 포함하는 렌티바이러스 입자와 CD34-농축 세포를 접촉시키는 단계를 포함하며, 여기서 재조합 렌티바이러스 게놈은 복제 불능이다.
- [0056] 또 다른 양태에서, 본 개시는 IMO를 앓고 있거나 앓을 것이 의심되는 대상체에서 영아 악성 골화석증(IMO)을 치료하는 방법을 제공하며, 상기 방법은 본 개시의 임의의 변형된 세포 또는 본 개시의 임의의 약학적 조성물을 대상체에게 투여하는 단계를 포함한다.
- [0057] 일부 구현예에서, 상기 방법은 *TCIRG1* 또는 이의 기능적 변이체를 발현하는 변형된 세포로 HSC 서식지를 재충진한다.
- [0058] 일부 구현예에서, 상기 방법은 *TCIRG1* 또는 이의 기능적 변이체를 발현하는 변형된 세포로 파골세포 서식지를 재충진한다.
- [0059] 일부 구현예에서, 상기 방법은 IMO를 치료, 개선, 예방, 감소, 억제, 또는 완화시킨다.
- [0060] 일부 구현예에서, 상기 방법은 치료된 대상체의 평균 전체 생존기간을 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10년 또는 그 이상 연장시킨다.
- [0061] 일부 구현예에서, 상기 방법은 IMO로 인한 대상체의 사망을 예방한다.
- [0062] 일부 구현예에서, 대상체는 인간이다.
- [0063] 일부 구현예에서, 대상체는 치료 전에 IMO의 증상을 나타냈다.
- [0064] 일부 구현예에서, 대상체는 치료 전에 *TCIRG1*의 발현이 감소되었거나 발현이 검출되지 않는 것으로 식별되었다.
- [0065] 일부 구현예에서, 대상체는 돌연변이된 *TCIRG1* 유전자를 갖는 것으로 식별되었다.
- [0066] 일부 구현예에서, 대상체는 영아이다.
- [0067] 일부 구현예에서, 상기 방법은 자가 치료를 포함한다.
- [0068] 일부 구현예에서, 투여는 정맥내 주입을 통해 수행된다.
- [0069] 또 다른 양태에서, 본 개시는 영아 악성 골화석증(IMO)을 치료하거나 예방하기 위한 의약의 제조에 사용하기 위한 재조합 렌티바이러스 게놈을 제공하며, 여기서 렌티바이러스 게놈은 5'에서 3'의 순서로 렌티바이러스 5' 장

말단 반복(LTR), 본 개시의 임의의 발현 카세트, 및 렌티바이러스 3' LTR을 포함하며; 재조합 렌티바이러스 게놈은 복제 불능이다.

[0070] 또 다른 양태에서, 본 개시는 영아 악성 골석화증(IMO)을 치료하거나 예방하기 위한 의학의 제조에 사용하기 위한 렌티바이러스 벡터를 제공하며, 렌티바이러스 벡터는 렌티 바이러스 게놈을 포함하고, 렌티바이러스 게놈은 5'에서 3'의 순서로 렌티바이러스 5' 장 말단 반복(LTR), 본 개시의 임의의 발현 카세트, 및 렌티바이러스 3' LTR을 포함하며; 재조합 렌티바이러스 게놈은 복제 불능이다.

[0071] 본 발명의 다른 특징 및 이점은 다음의 상세한 설명 및 청구범위에 포함되며 이로부터 명백해질 것이다.

**도면의 간단한 설명**

[0072] 도 1은 TCIRG1을 암호화하는 렌티바이러스 유전자 요법 벡터를 생산하기 위한 전달 플라스미드 (pCCL.PPT.EFS.tcirg1h.wpre)의 다이어그램을 제공한다.

도 2는 발현 카세트(서열번호 1)의 유전자 서열을 보여주며, 상기 유전자 서열은 5'에서 3'의 순서로 신장 인자 1-? 짧은(EFS) 프로모터(밑줄 친 것; 서열번호 2), TCIRG1을 암호화하는 폴리뉴클레오티드(검은 바탕에 흰 글자; 서열번호 3), 및 마멋 간염 바이러스(WHP) 전사후 조절 요소(WPRE)(밑줄 친 굵은 글씨체; 서열번호 4)를 포함한다.

도 3a-3b는 2개의 상이한 렌티바이러스 플라스미드의 안정성을 비교한 것이다. 도 3a는 브롬화 에티듬으로 염색한 아가로오스 겔의 사진을 도시한 것으로서, 제한 효소로 분해되지 않거나("미절단"), Af1III("Af1III") 또는 Af1III 및 NarI("Af1III/NarI")로 소화된 플라스미드 pRRL.PPT.EFS.tcirg1h.wpre를 보여준다. 도 3b는 브롬화 에티듬으로 염색한 아가로오스 겔의 사진을 도시한 것으로서, 제한 효소로 분해되지 않거나("미절단"), Af1III("Af1III") 또는 Af1III 및 NarI("Af1III/NarI")로 소화된 플라스미드 pCCL.PPT.EFS.tcirg1h.wpre를 보여준다. 도 3c는 pRRL.PPT.EFS.tcirg1h.wpre 및 pCCL.PPT.EFS.tcirg1h.wpre 플라스미드의 개략도를 보여준다.

도 4는 렌티바이러스 입자 제조를 위한 예시적인 공정을 도시한다.

도 5a-5b는 형질도입 6일 후(도 5a) 및 12일 후(도 5b) 벌크 CD34+ 세포 액상 배양물 중 벡터 카피 수(VCN)를 보여준다. VCN은 형질도입한 CD34+ 세포를 SCGM 완전 배지에서 배양한 후 추출된 gDNA의 qPCR에 의해 평가하였다. 각각의 공여자에 대한 VCN 및 평균은 각각의 형질도입 조건별로 표시되어 있다.

**발명을 실시하기 위한 구체적인 내용**

[0073] 본 발명자들은, TCIRG1을 암호화하는 렌티바이러스 벡터를 형질도입한 자가 세포의 이식이 영아 악성 골석화증(IMO)의 치료에 효과적임을 보여주었다. 또한, TCIRG1을 암호화하는 유전자 요법 벡터의 발현 카세트 서열에 특정 서열 요소를 포함시키면 IMO에 대한 안전하고 효과적인 유전자 요법이 생성된다. 본 개시는 렌티바이러스 벡터를 생산하는 데 유리한 안정한 전달 플라스미드를 포함하는 렌티바이러스 벡터 및 TCIRG1을 암호화하는 플라스미드를 제공한다.

[0074] 벡터 및 플라스미드

[0075] 본 발명자들은 놀랍게도, TCIRG1 유전자 요법을 위한 렌티바이러스 벡터의 대규모 생산이 원하는 발현 카세트를 함유하는 pRRL 플라스미드를 2가지 방식으로 변형시킴으로써 개선된다는 것을 발견하였다: (i) pRRL 벡터 백본을 pCCL 벡터 백본으로 치환하는 것; 및 (ii) pCCL 백본 내의 종래의 항생제 내성 카세트를 RNA-OUT 선택성 마커로 치환하는 것. 그런 다음, 개선된 플라스미드를 헬퍼 플라스미드와 함께 렌티바이러스 입자 생산 시스템 내로 형질감염시켜 원하는 렌티바이러스 벡터를 생산한다.

[0076] 생성된 TCIRG1 유전자 요법용 pCCL/RNA-OUT 벡터(예를 들어, 도 1에도시된 벡터)는 개선된 안정성을 가지는데, 이는 대장균(*E. coli*) 기반 플라스미드 생산 대비 더 높은 플라스미드 수율; 및 정제된 플라스미드에서 바람직하지 않은 재조합 산물의 수준 감소에 반영되어 있다(실시예 1 및 도 3a-3c에 나타나 있음). 전달 플라스미드에 대한 이러한 개선은 TCIRG1 발현 카세트를 포함하는 렌티바이러스 입자를 임상 시험 및 사용에 충분한 수율로 제조할 수 있게 한다. 본원에 제공된 추가 데이터는 본원에 개시된 방법 및 조성물을 사용하여 생산된 렌티바이러스 입자가 임상적으로 유의미한 벡터 카피 수(VCN) 수준에 도달할만큼 충분히 효율적으로 CD34<sup>+</sup> 세포를 형질도입한다는 것을 입증한다.

[0077] 일부 구현예에서, 본 개시는 3세대 렌티바이러스 벡터 시스템에 사용된 pCCL 전달 플라스미드에 기초한 렌티바

이러스 벡터인 전달 플라스미드를 제공한다. pCCL 전달 플라스미드는 HIV 통합 부위에서 남은 대부분의 인간 서열이 치환된 키메라 사이토메갈로바이러스(CMV)-HIV 5' LTR 및 벡터 백본을 함유하는데, 여기서 시미안 바이러스 40 폴리아데닐화 서열 및 (인헨서가 없는) 복제 기원 서열은 HIV 3' LTR의 하류에 포함되어 있다. CCL 5' 하이브리드 장 말단 반복(LTR)은 HIV-1 LTR의 R 영역에 결합된 사이토메갈로바이러스(CMV)의 인헨서이자 프로모터이다(전사 개시 부위에 대한 뉴클레오티드 -673 내지 -1; GenBank 수탁 번호 K03104). 일부 구현예에서, 전달 플라스미드는 상류의 RRE 및 cPPT/CTS 요소와 함께 TCIRG1 유전에 연결된 EFS 프로모터 및 하류의 WPRE 요소를 포함한다(도 1). 일부 구현예에서, 전달 플라스미드는 상류의 RRE 및 cPPT/CTS 요소와 함께 TCIRG1 유전에 연결된 PGK 프로모터(서열번호 24) 및 하류의 WPRE 요소를 포함한다. 일부 구현예에서, 전달 플라스미드는 RNA-OUT 요소를 포함한다. 유리하게는, RNA-OUT 서열은 패키징 세포주에서 전달 플라스미드의 안정적인 전파에 기여한다. 일부 구현예에서, 전달 플라스미드는 항생제 내성 유전자, 예를 들어, AmpR을 포함하지 않는다.

[0078] 일부 구현예에서, PGK 프로모터는 서열번호 24와 적어도 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% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, PGK 프로모터는 서열번호 24와 적어도 80%, 85%, 90%, 95%, 99%, 또는 100% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, PGK 프로모터는 서열번호 24의 서열을 갖는다.

[0079] GGGGTGGGGTTGCGCCTTTTCCAAGGCAGCCCTGGGTTTTCGCGAGGGACGCGGCTGCTCTGGGCGTGGTTCGGGAAACGCAGCGCGCCGACCTGGGTC  
TCGCACATTTTCACGTCGGTTCGACGCGTACCCGGATCTTCGCCGTACCCCTTGTGGGCCCCCGGCGACGCTTCCTGCTCCGCCCTAAGTCGGGAAGG  
TTCCTGCGGTTTCGCGCGTTCGCGGACGTGACAAACGGAAGCCGACGCTCTACTAGTACCTCGCAGACGGACAGCGCCAGGAGCAATGGCAGCGCGCCG  
ACCGCATGGGCTGTGGCCAATAGCGGCTGCTCAGCAGGGCGCGCCGAGAGCAGCGCGCGGAAGGGGCGGTGCGGGAGGCGGGGTGTGGGGCGGTAGTGTG  
GGCCCTGTTCTGCCCGCGGGTGTTCGCGATTCTGCAAGCCTCCGGAGCGCAGCTCGGCGAGTTCCTCGTTGACCGAATCACCGACCTCTCTCCCA  
G

[0080] (서열번호 24)

[0081] 특정 구현예에서, 전달 플라스미드는 대장균에서 배양하거나 전파시킬 때 동일한 발현 카세트를 포함하는 다른 플라스미드보다 더 안정적이어서, 더 높은 수율의 플라스미드를 생성하며, 이는 벡터 생산에 사용하기에 유리하다. 일부 구현예에서, 전달 플라스미드는 pRRL.PPT.EFS.tcirg1h.wpre 전달 플라스미드보다 더 안정적이다. 특정 구현예에서, 동일한 배양 조건 하에서 생산된 pRRL.PPT.EFS.tcirg1h.wpre 전달 플라스미드의 양과 비교했을 때 적어도 2배, 적어도 5배, 또는 적어도 10배 더 많은 전달 플라스미드가 생산된다.

[0082] 특정 구현예에서, 전달 플라스미드는 pCCL.PPT.EFS.tcirg1h.wpre 또는 이의 기능적 변이체, 예를 들어, 본원에 개시된 것들이다. 본 개시는 특정 구현예에서, 전달 플라스미드 pCCL.PPT.EFS.tcirg1h.wpre 또는 이의 기능적 변이체를 제공한다. 전달 플라스미드 pCCL.PPT.EFS.tcirg1h.wpre는 서열번호 23을 가질 수 있다.

[0083] 대안적으로, 전달 플라스미드 pCCL.PPT.EFS.tcirg1h.wpre는 다음 서열 GATCAGAGACTAGCCTCGAGAAGCTTGATCGATTGGCTCCGGTGCC(서열번호 26)이 결실된 서열번호 25를 가질 수 있다.

[0084] 서열번호 25의 서열은 원형 플라스미드를 나타낸다. 염기쌍 1에서 EFS 프로모터로 시작하도록 재배치된 동일한 서열이 서열번호 27로서 제공된다.

[0085] 일부 구현예에서, 전달 플라스미드는 서열번호 23과 적어도 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% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 전달 플라스미드는 서열번호 23과 적어도 80%, 85%, 90%, 95%, 99%, 또는 100% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 전달 플라스미드는 서열번호 23의 서열을 포함한다.

[0086] 일부 구현예에서, 전달 플라스미드는 서열번호 25와 적어도 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% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 전달 플라스미드는 서열번호 25와 적어도 80%, 85%, 90%, 95%, 99%, 또는 100% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 전달 플라스미드는 서열번호 25의 서열을 포함한다.

[0087] 일부 구현예에서, 전달 플라스미드는 서열번호 27과 적어도 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% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 전달 플라스미드는 서열번호 27과 적어도 80%, 85%,

90%, 95%, 99%, 또는 100% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 전달 플라스미드는 서열번호 27의 서열을 포함한다.

[0088] 일부 구현예에서, 전달 플라스미드는 표 1에 열거된 벡터 요소 중 하나 이상을 포함한다.

표 1

[0089]

pCCL.PPT.EFS.tcirg1h.wpre 벡터 요소				
명칭	유형	최소 ~ 최대 기준 서열 내 위치 (뉴클레오티드)	길이 (염기 쌍)	서열번호
EFS	프로모터	1~243	243	서열번호 2
TCIRG1	CDS	257~2,749	2,493	서열번호 3
WPRE	조절	2,782~3,384	603	서열번호 4
3' LTR	LTR	3,471~3,704	234	서열번호 28
SV40 폴리(A)	폴리A 신호	3,776~3,907	132	서열번호 29
SV40 ori	복제 기원	3,917~4,076	160	서열번호 30
pUC 기원	복제 기원	4,115~5,129	1,015	서열번호 31
RNA-OUT	억제인자	5,146~5,284	139	서열번호 32
CMV IE	프로모터	5,334~5,910	577	서열번호 33
5-LTR	LTR	5,933~6,120	188	서열번호 34
psi	패키징	6,222~6,266	45	서열번호 35
gag	CDS	6,267~6,628	362	서열번호 36
RRE	조절	6,629~7,486	858	서열번호 37
cPPT/CTS	폴리 퓨린 구역	7,505~7,622	118	서열번호 38
-	백본	3,776~7,622	3,847	서열번호 39

[0090] 일부 구현예에서, 렌티바이러스 입자는 pCCL.PPT.EFS.tcirg1h.wpre를 포함하는 3세대 렌티바이러스 벡터 시스템의 일시적 형질감염에 의해 생성된다.

[0091] 일부 구현예에서, 발현 카세트는 서열번호 1과 적어도 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% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 발현 카세트는 서열번호 1과 적어도 80%, 85%, 90%, 95%, 99%, 또는 100% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 발현 카세트는 서열번호 1의 서열을 갖는다.

[0092] 일부 구현예에서, 발현 카세트는 5'에서 3'의 순서로 EFS 프로모터; T 세포 면역 조절인자 1, ATPase H+ 수송 V0 서브유닛 a3(TCIRG1) 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드; 및 마멋 간염 바이러스(WHP) 전사후 조절 요소(WPRE)를 포함한다. 일부 구현예에서, EFS 프로모터는 TCIRG1의 이소형 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드에 작동 가능하게 연결된다. 관련 구현예는 발현 카세트를 포함하는 전달 플라스미드, 및 전달 플라스미드를 사용하여 생산된 벡터를 포함한다.

[0093] 일부 구현예에서, EFS 프로모터는 서열번호 2와 적어도 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% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, EFS 프로모터는 서열번호 2와 적어도 80%, 85%, 90%, 95%, 99%, 또는 100% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, EFS 프로모터는 서열번호 2의 서열을 갖는다.

[0094] GGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGCAATTGAACCGGTGCCTAGAGAAGGTGGCGC  
GGGGTAAACTGGGAAAGTGATGTCGTACTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCGGTGAACGTTCTTTT  
CGCAACGGGTTTGCCGCCAGAACACAGGTGTCGTGACGC

[0095] (서열번호 2)

[0096] 일부 구현예에서, TCIRG1의 이소형 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드는 서열번호 3과 적어도 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% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서,

TCIRG1의 이소형 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드는 서열번호 3과 적어도 80%, 85%, 90%, 95%, 99%, 또는 100% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, TCIRG1의 이소형 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드는 서열번호 3의 서열을 갖는다.

[0097] 일부 구현예에서, WPRE는 서열번호 4와 적어도 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% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, WPRE는 서열번호 4와 적어도 80%, 85%, 90%, 95%, 99%, 또는 100% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, WPRE는 서열번호 4의 서열을 갖는다.

[0098] ATTCGAGCATCTTACCGCCATTATACCCATATTTGTTCTGTTTTTCTGTATTGGGTATACATTTAAATGTTAATAAAACAAAATGGTGGGGCAATCATTACATTTTTAGGGATATGTAATTACTAGTTCAGGTGATTGCCACAAGACAAACATGTTAAGAACTTCCCGTTATTTACGCTCTGTCTCTGTTAATCAACC TCTGGATTACAAAATTTGTGAAAGATTGACTGATATTCTTAACATATGTTGCTCCTTTTACGCTGTGTGGATATGCTGCTTTAATGCCTCTGTATCATGCTAT TGCTTCCCGTACGGCTTTCGTTTTCTCTCCTGTATAAATCCTGGTTGCTGTCTTTATGAGGAGTTGTGGCCGTGTCCGTCACAGTGGCGTGGTGTG CTCTGTGTTTGTGACGCAACCCCACTGGCTGGGGCATTGCCACCACCTGTCAACTCCTTTCTGGGACTTTCGCTTTCCCTCCCGATGCCACGGCAGA ACTCATCGCCGCTGCTGTGCCGCTGTGGACAGGGGCTAGGTTGCTGGGCACTGATAATTCGGTGTGTGCGGGGAAGCTGACGCTCCTTTTCG

[0099] (서열번호 4)

[0100] 일부 구현예에서, TCIRG1의 이소형 또는 이의 기능적 변이체는 서열번호 5와 적어도 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% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, TCIRG1의 이소형 또는 이의 기능적 변이체는 서열번호 5와 적어도 80%, 85%, 90%, 95%, 99%, 또는 100% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, TCIRG1의 이소형 또는 이의 기능적 변이체는 서열번호 5의 서열을 갖는다.

[0101] MGSFMRSEVALVQLFLPTAAAYTCVSRLGELGLVEFRDLNASVSAFQRRFVVDVRRCEELEKTFNFLQEEVRRAGLVLPKGRPLPAPPRDLLRIQEETE RLAQELRDVRGNQALRAQLHQLHAAVLRQGHEPQLAAHTDGASERTPLLQAPGGPHQDLRVNFVAGAVEPHKAPALERLLWRACRGLIASFRELEQP LEHPVTGEPATWMTFLISYWEQIGQKIRKI TDCFHCHVFPFLQEEEARLQALQQLQQSQELQEVLGETERFLSQVLGRVLQLLPPGQVQVHKMKAVYLAL NQCSVSTTHKCLIAEAWCSVRDLPALQEALRDSSMEEGVSVAHRIPCRDMPPTLIRTNRTASFQGI VDAYGVGRYQEVNPAPYTIITFPFLFAVMFGDVG HGLLMFLFALAMVLAENRPAVKAAQNEIWQTFFRGRYLLLLMGLFSIYTGFIYNECFSRATSIFFPSGWSVAAMANQSGWSDAFLAQHTMLTLDPNVTGVFLG PYPFGIDPIWLSAANHL SFLNSFKMKMSVILGVVHMAFGVVLGVFNHVFHGQRHRLLELPELTFLLGLFGYLVFLVIYKWLVCVWAARAASAPSIILHFIN MFLFSHSPSNRLLYPRQEVVQATLVVLALAMVPILLGTPHLHLHRHRRRLRRRPPADRQEENKAGLLDLPDASVNGWSDEEKAGLDDEEEAELVPSEVLM HQAIHTIEFCLGCVSNTASYRLWALSLAHAQLSEVLWAMVMRIGLGLGREVGVAAVLVP IFAAFVMTVA ILLVMEGLSAFLHALRLHWVEFQNKFYSGT GYKLSPTFAATDD

[0102] (서열번호 5)

[0103] 일 구현예에서, TCIRG1의 이소형 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드는 인간 숙주 세포에서의 발현을 위해 코돈 최적화된다. 일 구현예에서, TCIRG1 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드는 드물게 표시되는 코돈을 보다 빈번하게 표시되는 코돈으로 치환함으로써 발현을 향상시키도록 "코돈 최적화"되거나 변형된다. 일 구현예에서, TCIRG1의 이소형 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드는 코돈 최적화되지 않는다. 일 구현예에서, TCIRG1의 이소형 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드는 변형되지 않는다. 일 구현예에서, TCIRG1의 이소형 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드는 코돈 최적화되지 않는다. 일 구현예에서, TCIRG1의 이소형 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드는 천연 폴리뉴클레오티드 서열이다.

[0104] 본원에서 사용되는 바와 같이, 용어 "이식 유전자(transgene)"는 TCIRG1의 이소형 또는 이의 기능적 변이체를 암호화하는 폴리뉴클레오티드를 지칭한다.

[0105] 코딩 서열은 번역을 위해 아미노산을 암호화하는 mRNA 서열의 일부이다. 번역하는 동안, 61개의 트리뉴클레오티드 코돈 각각은 20개의 아미노산 중 하나로 번역되어, 유전자 코드에서 퇴화성 또는 중복성을 야기한다. 그러나, 상이한 세포 유형, 및 상이한 동물 종은 상이한 빈도로 동일한 아미노산을 암호화하는 tRNA(각각이 항 코돈을 가짐)를 이용한다. 유전자 서열이 해당 tRNA에 의해 드물게 표시되는 코돈을 함유하는 경우, 리보솜 번역 기계가 느려져 효율적인 번역을 방해할 수 있다. 발현은 특정 종에 대한 "코돈 최적화"를 통해 개선될 수 있으며, 여기에서 암호화 서열은 동일한 단백질 서열을 암호화하도록 변경되지만, 고도로 발현되고/되거나 고도로 발현된 인간 단백질에 의해 활용되는 코돈을 활용한다(Cid-Arregui 등의 문헌[2003; J. Virol. 77:4928]).

[0106] 일부 구현예에서, 이식 유전자의 암호화 서열은 포유류 또는 영장류에서 드물게 발현되는 코돈을 영장류에서 빈

번하게 발현되는 코돈으로 대체하도록 변형된다. 예를 들어, 일부 구현예에서, 이식 유전자는 기준 폴리펩티드 (예를 들어, 야생형 TCIRG1; 서열번호 3)에 대해 적어도 85%의 서열 동일성-예를 들어, 기준 폴리펩티드에 대해 적어도 90%의 서열 동일성, 적어도 95%의 서열 동일성, 적어도 98%의 동일성, 또는 적어도 99%의 동일성-을 갖는 폴리펩티드를 암호화하되, 암호화 서열의 적어도 하나의 코돈은 상술하거나 본원에 개시된 서열의 상응하는 코돈보다 인간에서 더 높은 tRNA 빈도를 갖는다.

- [0107] 일 구현예에서, 이식 유전자는 서열번호 3보다 적은 수의 대안적인 개방 해독 프레임(ORF)을 포함한다. 일 구현예에서, 이식 유전자는 목적하는 이식 유전자를 암호화하지 않는 개방 해독 프레임(ORF)의 종결 또는 제거에 의해 발현을 향상시키도록 변형된다. 개방 해독 프레임(ORF)은 개시 코돈을 따르고 정지 코돈을 함유하지 않는 핵산 서열이다. ORF는 순방향 또는 역방향 배향일 수 있고, 관심 유전자와 비교하여 "프레임 내" 또는 "프레임 외"일 수 있다. 이러한 개방 해독 프레임은 관심 유전자와 함께 발현 카세트에서 발현될 가능성이 있으며, 목적하지 않는 부작용을 초래할 수 있다. 일부 구현예에서, 이식 유전자는 코돈 사용을 추가로 변경함으로써 개방 해독 프레임을 제거하도록 변형되었다. 이는 하나 이상의 개시 코돈(ATG)을 제거하고/하거나 하나 이상의 정지 코돈(TAG, TAA, 또는 TGA)을 목적하는 ORF에 역방향 또는 프레임 외로 도입하면서, 암호화된 아미노산 서열을 보존하고, 선택적으로, 관심 유전자에서 고도로 이용되는 코돈을 유지함으로써(즉, 빈도 < 20%인 코돈을 포함) 수행된다.
- [0108] 본 개시의 변형예에서, 이식 유전자 암호화 서열은 코돈 최적화 및 비-이식 유전자 ORF의 제거 중 하나에 의해 또는 두 가지 기술 모두를 사용하여 최적화될 수 있다. 일부 경우, 코돈 최적화 동안 도입된 ORF를 제거하기 위해 코돈 최적화 후 비-이식 유전자 ORF를 제거하거나 최소화한다.
- [0109] 일 구현예에서, 이식 유전자는 서열번호 3보다 더 적은 CpG 부위를 함유한다. 이론에 구속되지 않고, 폴리뉴클레오티드 서열에서 CpG 부위의 존재는 폴리뉴클레오티드 서열을 포함하는 바이러스 백터에 대한 숙주의 바람직하지 않은 면역학적 반응과 연관된 것으로 여겨진다. 일부 구현예에서, 이식 유전자는 CpG 부위의 수를 감소시키도록 설계된다. 예시적인 방법은 미국 특허 출원 공개 제US20020065236A1호에 제공되어 있다.
- [0110] 일 구현예에서, 이식 유전자는 서열번호 3보다 더 적은 암호화 스플라이스 부위를 함유한다. 최적화를 위해, GeneArt® 소프트웨어를 사용하여, 예를 들어 전사 발현 억제를 방지하기 위해 GC 함량을 증가시키고/증가시키거나 암호화 스플라이스 부위를 제거하여, 이식 유전자 발현을 증가시킬 수 있다. 대안적으로, 당업계에 공지된 임의의 최적화 방법이 사용될 수 있다. 암호화 스플라이스 부위의 제거는, 예를 들어, 국제 특허 출원 공개 제W02004015106A1호에 기술되어 있다.
- [0111] 또한, 공통 최적 코작 서열, 전장 폴리아데닐화(폴리아) 서열(또는 절단된 폴리A를 위한 전장 폴리A의 치환), 및 최소 개시 코돈 또는 상류(즉, 5')가 아닌 개시 코돈(즉, ATG 부위)을 포함하는 TCIRG1, 예를 들어, 본원에 개시된 TCIRG1 서열을 암호화하는 유전자 요법 백터 및 발현 카세트가 본원에 개시된다.
- [0112] 일부 구현예에서, 발현 카세트는 5' 장 말단 반복(LTR), 인핸서/프로모터 영역, 공통 최적 코작 서열, 이식 유전자(예를 들어, 본원에 개시된 TCIRG1을 암호화하는 이식 유전자), 전장 폴리A 서열을 포함하는 3' 미번역 영역, 및 3' LTR 중 2개 이상을 함유한다.
- [0113] 일 구현예에서, 발현 카세트는 이식 유전자에 작동 가능하게 연결된 코작 서열을 포함한다. 일 구현예에서, 코작 서열은 서열번호 6을 포함하거나 이로 이루어진 공통 최적 코작 서열이다.
- [0114] GCCGCCACCATGG (서열번호 6)
- [0115] 다양한 구현예에서, 발현 카세트는 이식 유전자에 작동 가능하게 연결된 대안적 코작 서열을 포함한다. 일 구현예에서, 코작 서열은 서열번호 14~18 중 어느 하나를 포함하거나 이로 이루어진 대안적 코작 서열이다.
- [0116] (gcc)gccRccAUGG (서열번호 14)
- [0117] AGNNAUGN (서열번호 15)
- [0118] ANNAUGG (서열번호 16)
- [0119] ACCAUGG (서열번호 17)
- [0120] GACACCAUGG (서열번호 18)
- [0121] 서열번호 14에서, 소문자는 염기가 어떻게든지 변경될 수 있는 위치에서 가장 일반적인 염기를 나타내고; 때문

자는 고도로 보존된 염기를 나타내며; 'R'은 아데닌 또는 구아닌을 나타낸다. 서열번호 14에서, 괄호 안의 서열 (gcc)는 선택적이다. 서열번호 15~17에서, 'N'은 임의의 염기를 나타낸다.

- [0122] 이러한 공통 최적 코작 서열 대신에 다양한 서열이 번역-개시 부위로서 사용될 수 있으며, 다른 서열을 식별하고 시험하는 것은 당업자의 기술 범위에 속한다. Kozak M.의 문헌[An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.* 115 (4): 887-903 (1991)] 참조.
- [0123] 일 구현예에서, 발현 카세트는 이식 유전자에 작동 가능하게 연결된 전장 폴리A 서열을 포함한다. 일 구현예에서, 전장 폴리A 서열은 서열번호 7을 포함한다.
- [0124] TGGCTAATAAAGGAAATTTATTTTCATTGCAATAGTGTGTGGAATTTTTTGTGTCTCTCACTCGGAAGGACATATGGGAGGGCAAATCATTTAAACATCA GAATGAGTATTTGGTTTAGAGTTTGGCAACATATGCCATATGTCTGGCTGCCATGAACAAAGGTTGGCTATAAAGAGGTCATCAGTATATGAAACAGCCCC TGCTGTCCATTCCCTATTCCATAGAAAAGCCTTGACTTGAGGTTAGATTTTTTTATATTTTGTGTTATTTTTTCTTTAACATCCCTAAAATTTTC CTTACATGTTTTACTAGCCAGATTTTCTCTCTCTGACTACTCCCAGTCATAGCTGTCCCTCTTCTTATGGAGATC
- [0125] (서열번호 7)
- [0126] 소 성장 호르몬 폴리아데닐화 신호(bGHPA) (서열번호 19), SV40 초기/후기 폴리아데닐화 신호(서열번호 20), 및 인간 성장 호르몬(HGH) 폴리아데닐화 신호(서열번호 21)를 포함하되 이에 제한되지 않고, 다양한 대안적 폴리A 서열이 본 개시의 발현 카세트에 사용될 수 있다.
- [0127] TCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCTTCTAATAAAAAAT GAGGAAATTCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGTGGGGTGGGGCAGGACACGAAGGGGAGGATTGGGAGGACAATAGCAGGCAT GCTGGGGATGCGGTGGGCTCTATGGCTTCTG
- [0128] (서열번호 19)
- [0129] CAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTG TAACCATTATAAGCTGCAATAACAAGTTAACAACAACAATTGCATTCTTTATGTTTCAGGTTTCAGGGGAGATGTGGGAGGTTTTTAAAGCAAGTAAA ACCTCTACAAATGTGGTA
- [0130] (서열번호 20)
- [0131] CTGCCGGGTGGCATCCCTGTGACCCCTCCCAGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTGCCACCAGCCTTGTCTAATAAAAAATTAAGTTGC ATCATTTTGTCTGACTAGGTGCTCCTTCTATAATATTATGGGGTGGAGGGGGTGGTATGGAGCAAGGGGCCAAGTTGGGAAGAACTGTAGGGCCTGC
- [0132] (서열번호 21)
- [0133] 일부 구현예에서, 발현 카세트는 폴리A 서열의 활성 단편을 포함한다. 특정 구현예에서, 폴리A 서열의 활성 단편은, 예를 들어, 본원에 개시된 폴리A 서열 중 어느 하나의 20 염기쌍(bp) 미만, 50 bp 미만, 100 bp 미만, 또는 150 bp 미만을 포함하거나 이로 이루어진다.
- [0134] 일부 경우, 발현 카세트가 경쟁 ORF를 함유하지 않도록 함으로써 이식 유전자의 발현이 증가된다. 일 구현예에서, 발현 카세트는 이식 유전자의 개시 코돈의 5'으로부터 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 또는 500개 염기쌍 이내에 개시 코돈을 포함하지 않는다. 일 구현예에서, 발현 카세트는 이식 유전자의 개시 코돈의 5'에 개시 코돈을 포함하지 않는다.
- [0135] 일 구현예에서, 발현 카세트는 5'에서 3' 방향으로 작동 가능하게 연결된 제1 역위 말단 반복, 인핸서/프로모터 영역, 인트론, 공통 최적 코작 서열, 이식 유전자, 전장 폴리A 서열을 포함하는 3' 미번역 영역, 및 제2 역위 말단 반복을 포함하며, 여기서 발현 카세트는 이식 유전자의 개시 코돈에 대한 5'에 개시 코돈을 포함하지 않는다.
- [0136] 일 구현예에서, 인핸서/프로모터 영역은 5'에서 3' 방향으로 CMV IE 인핸서 및 닭 베타-액틴 프로모터를 포함한다. 일 구현예에서, 인핸서/프로모터 영역은 CAG 프로모터를 포함한다. 본원에서 사용되는 "CAG 프로모터"는 CMV 초기 인핸서 요소, 닭 베타-액틴 프로모터, 닭 베타-액틴 유전자의 제1 엑손과 제1 인트론, 및 토끼 베타-글로빈 유전자로부터의 스플라이스 수용체를 포함하는 폴리뉴클레오티드 서열을 지칭한다.
- [0137] 일 구현예에서, 인핸서/프로모터 영역은 신장 인자 1? 짧은 프로모터(EFS 프로모터)를 포함하며, 이는 인트론이 없는 신장 인자 1? 프로모터의 더 짧은 버전이다. 본원에서 사용되는 바와 같이, "EFS 프로모터"는 짧고, 인트론이 없는 형태의 EF1알파를 포함하는 폴리뉴클레오티드 서열을 지칭한다. EFS 프로모터는 최근까지 많은 임상

시험에서 사용되어 왔다. 이는 인근 프로모터의 교차 활성화가 감소된 세포 유래 인헨서/프로모터이므로, 가설적으로 유전독성의 위험을 감소시킨다.

[0138] 일 구현예에서, 발현 카세트는 서열번호 1로부터 선택된 하나의 서열과 적어도 95%의 동일성을 공유한다. 일 구현예에서, 발현 카세트는 서열번호 1로부터 선택된 하나의 서열에 대해 완전한 동일성을 공유하거나, 서열번호 1로부터 선택된 하나의 서열에 대해 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 98%, 또는 적어도 99%의 동일성을 공유한다. 특정 구현예에서, 발현 카세트는 서열번호 1로부터 선택된 하나의 서열과 비교하여 하나 이상의 변형을 포함한다. 특정 구현예에서, 하나 이상의 변형은 다음 중 하나 이상을 포함한다: 하나 이상의(예를 들어, 모든) 상류 ATG 서열을 제거하는 것; 최적화된 공통 코작 서열 또는 다른 코작 서열(본원에 개시된 것들 중 어느 하나를 포함하지만 이에 한정되지는 않음)로 코작 서열을 치환하는 것; 및/또는 전장 폴리 아데닐화 서열 또는 다른 폴리 아데닐화 서열(본원에 개시된 것들 중 어느 하나를 포함하지만 이에 한정되지는 않음)로 폴리 아데닐화 서열을 치환하는 것. 이들 예시적인 발현 카세트 내의 유전자 요소의 예시적인 구성은 도 1에 도시되어 있다.

[0139] 관련 구현예에서, 본 개시는 본원에 개시된 발현 카세트를 포함하는 유전자 요법 벡터를 제공한다. 일반적으로, 본원에 기술된 유전자 요법 벡터는 TCIRG1의 하나 이상의 이소형을 암호화하는 폴리뉴클레오티드를 포함하는 발현 카세트를 포함하며, 상기 발현 카세트는 TCIRG1의 발현을 가능하게 하여, TCIRG1 단백질 발현 수준 결핍의 교정 및/또는 과골세포 형성 결핍의 교정을 필요로 하는 대상체(예를 들어, 적어도 부분적으로 TCIRG1 발현 결핍으로 인한 과골세포 형성 결핍을 특징으로 하는 영아 악성 골석화증 또는 다른 장애를 가진 대상체)에서 이를 부분적으로 또는 전체적으로 교정할 수 있게 한다. 특정 구현예에서, 발현 카세트는 본원에 개시된 TCIRG1을 암호화하는 폴리뉴클레오티드 서열(예: 서열번호 3) 또는 서열번호 3에 대해 적어도 90%, 적어도 95%, 적어도 98%, 또는 적어도 99%의 동일성을 갖는 서열을 포함한다. 유전자 요법 벡터는 바이러스 벡터 또는 비-바이러스 벡터일 수 있다. 예시적인 비-바이러스 벡터는, 예를 들어, 네이키드 DNA, 양이온성 리포솜 복합체, 양이온성 중합체 복합체, 양이온성 리포솜-중합체 복합체, 및 엑소솜을 포함한다. 바이러스 벡터의 예는 아데노바이러스, 레트로바이러스, 렌티바이러스, 헤르페스바이러스 및 아데노-관련 바이러스(AAV) 벡터를 포함하되, 이에 한정되지 않는다.

[0140] 본 발명의 실시예에 유용한 유전자 전달 바이러스 벡터는 분자 생물학 분야에 공지된 방법론을 사용하여 작제될 수 있다. 일반적으로, 이식 유전자를 운반하는 바이러스 벡터는 이식 유전자를 암호화하는 폴리뉴클레오티드, 적절한 조절 요소 및 세포 형질도입을 매개하는 바이러스 단백질의 생산에 필요한 요소로부터 조립된다. 이러한 재조합 바이러스는 당업계에 공지된 기술, 예를 들어, 포장 세포의 형질감염, 또는 헬퍼 플라스미드 또는 바이러스에 의한 일시적 형질감염으로 생산될 수 있다. 바이러스 포장 세포의 일반적인 예는 HeLa 세포, SF9 세포(선택적으로 바큘로바이러스 헬퍼 벡터 포함), 293 세포 등을 포함하지만 이에 한정되지 않는다. US20170218395A1 에 기술된 바와 같이, 헤르페스바이러스 기반 시스템을 사용하여 AAV 벡터를 생산할 수 있다. 이러한 복제-결합 재조합 바이러스를 생산하기 위한 상세한 프로토콜은, 예를 들어, W095/14785, W096/22378, 미국 특허 제5,882,877호, 미국 특허 제6,013,516호, 미국 특허 제4,861,719호, 미국 특허 제5,278,056호 및 W094/19478호에 기술되어 있으며, 이들 각각의 전체 내용은 본원에 참조로서 통합된다.

[0141] 일부 구현예에서, 벡터는 레트로바이러스 벡터, 또는 보다 구체적으로는 렌티바이러스 벡터이다. 본원에서 사용되는 바와 같이, 용어 "레트로바이러스" 또는 "레트로바이러스"는 이의 게놈 RNA를 선형 이중-가닥 DNA 카피로 역전사시키고, 이어서 이의 게놈 DNA를 숙주 게놈 내로 공유적으로 통합시키는 RNA 바이러스를 지칭한다. 레트로바이러스 벡터는 유전자 전달을 위한 일반적인 도구이다(Miller의 문헌[2000, *Nature*. 357: 455-460]). 바이러스가 숙주 게놈에 통합되면, 이는 "프로바이러스(provirus)"로서 지칭된다. 프로바이러스는 RNA 중합효소 II 에 대한 템플릿의 역할을 하고, 바이러스에 의해 암호화된 RNA 분자의 발현을 유도한다.

[0142] 예시적인 레트로바이러스(레트로바이러스과)는 (1) 몰로니 췌관 백혈병 바이러스(M-MuLV), 몰로니 췌관 육종 바이러스(MoMSV), 췌관 유방 종양 바이러스(MuMTV), 긴팔원숭이 백혈병 바이러스(GaLV), 및 고양이 백혈병 바이러스(FLV)와 같은 감마레트로바이러스속(genus gammaretrovirus), (2) 유인원 거품 바이러스와 같은 스푸마바이러스속(genus spumavirus), (3) 인간 면역결핍 바이러스-1 및 유인원 면역결핍 바이러스와 같은 렌티바이러스속(genus lentivirus)을 포함하지만 이들로 한정되지는 않는다.

[0143] 본원에서 사용되는 바와 같이, 용어 "렌티바이러스(lentiviral 또는 lentivirus)"는 복합 레트로바이러스들의 군(또는 속)을 지칭한다. 예시적인 렌티바이러스는, HIV(HIV 1형 및 HIV 2형을 포함하는 인간 면역결핍 바이러스; 비스나-메디 바이러스(Visna-maedi virus, VMV); 산양 관절염-뇌염 바이러스(caprine arthritis-

encephalitis virus, CAVAV); 말 전염성 빈혈 바이러스(equine infectious anemia virus, EIAV); 고양이 면역 결핍 바이러스(feline immunodeficiency virus, FIV); 소 면역 결핍 바이러스(bovine immune deficiency virus, BIV); 및 유인원 면역결핍 바이러스(simian immunodeficiency virus, SIV)를 포함하지만 이들로 한정되지는 않는다. 일 구현예에서, HIV-기반 벡터 백본(즉, HIV 시스-작용 서열 요소)이 바람직하다.

- [0144] 레트로바이러스 벡터, 보다 구체적으로는 렌티바이러스 벡터가 본 발명을 실시하는 데 사용될 수 있다. 따라서, 본원에서 사용되는 용어 "레트로바이러스 벡터"는 "렌티바이러스 벡터"를 포함하는 것을 의미하고; 본원에서 사용되는 용어 "레트로바이러스"는 "렌티바이러스"를 포함하는 것을 의미한다.
- [0145] 바이러스 벡터라는 용어는 핵산을 세포 내로 또는 전달된 핵산 자체에게 전달할 수 있는 벡터 또는 바이러스 입자를 지칭할 수 있다. 바이러스 벡터는 주로 바이러스로부터 유래되는 구조적 및/또는 기능적 유전 요소를 함유한다. 용어 "레트로바이러스 벡터"는 주로 레트로바이러스로부터 유래되는 구조적 및 기능적 유전 요소 또는 이의 일부를 함유하는 바이러스 벡터를 지칭한다. 용어 "렌티바이러스 벡터"는 LTR을 포함하여, 주로 렌티바이러스로부터 유래되는 구조적 및 기능적 유전 요소 또는 이의 일부를 함유하는 바이러스 벡터를 지칭한다. 용어 "하이브리드(hybrid)"는 레트로바이러스(예: 렌티바이러스) 서열 및 비-렌티바이러스 서열 둘 다를 함유하는 벡터, LTR, 또는 다른 핵산을 지칭한다. 일 구현예에서, 하이브리드 벡터는 역전사, 복제, 통합, 및/또는 패키징을 위한 레트로바이러스(예: 렌티바이러스) 서열을 포함하는 벡터 또는 전달 플라스미드를 지칭한다.
- [0146] 특정 구현예에서, 용어 "렌티바이러스 벡터" 및 "렌티바이러스 발현 벡터"는 렌티바이러스 전달 플라스미드 및/또는 감염성 렌티바이러스 입자를 지칭하도록 사용될 수 있다. 본원에서 클로닝 부위, 프로모터, 조절 요소, 이중 핵산 등과 같은 요소가 언급되는 경우, 이들 요소의 서열이 본 발명의 렌티바이러스 입자에서는 RNA 형태로 존재하고 본 발명의 DNA 플라스미드에서는 DNA 형태로 존재한다는 것을 이해해야 한다.
- [0147] 특정 구체적인 구현예에 따르면, 바이러스 벡터 백본 서열의 대부분 또는 전부는 렌티바이러스(예: HIV-1)로부터 유래된다. 그러나, 렌티바이러스 서열의 많은 상이한 공급원이 사용될 수 있고, 렌티바이러스 서열 중 일부에서의 많은 치환 및 변경이 본원에 기술된 기능을 수행하는 전달 벡터의 능력을 손상시키지 않고도 수용될 수 있다는 것을 이해해야 한다. 또한, 다양한 렌티바이러스 벡터가 당업계에 공지되어 있으며(Naldini 등의 [1996a, 1996b, 및 1998 문헌]; Zufferey 등의 [1997 문헌]; Dull 등의 [1998 문헌], 미국 특허 제6,013,516호; 및 제5,994,136호 참조), 이들 중 다수가 본 발명의 바이러스 벡터 또는 전달 플라스미드를 생산하는 데 적합할 수 있다.
- [0148] 렌티바이러스 벡터를 제조함에 있어서, 예를 들어 포유류 세포(예를 들어, HEK 293T 세포)를 포함하여, 렌티바이러스 벡터를 생산하기 위한 임의의 숙주 세포가 사용될 수 있다. 숙주 세포는 렌티바이러스 *gag/pol* 및 *rev* 유전자가 숙주 세포에서 안정적으로 유지되는 패키징 세포이거나, 렌티바이러스 벡터 게놈이 안정적으로 유지되고 포장되는 패키징 세포일 수도 있다. 렌티바이러스 벡터는 당업계에 공지된 표준 기술을 사용해 정제하고 제형화한다.
- [0149] 특정 구현예에서, 본 발명은 본원에 개시된 유전자 발현 카세트, 유전자 전달 카세트, 또는 제조함 렌티바이러스 벡터를 포함하는 세포를 포함한다. 관련 구현예에서, 세포는 본원에 개시된 발현 카세트를 포함하는 제조함 렌티바이러스 벡터로 형질도입되거나, 본원에 개시된 발현 카세트를 세포 게놈에 통합된 상태로 갖는다. 특정 구현예에서, 세포는 제조함 레트로바이러스 벡터(예: 패키징 세포)를 생산하는 데 사용되는 세포이다.
- [0150] 일부 구현예에서, 렌티바이러스 벡터는 위형화된 것이다. 예를 들어, 이중 *env* 유전자를 포함하는 플라스미드가 위형화에 사용될 수 있다. 적절한 *env* 유전자는 VSV-G를 포함하지만 이에 한정되지는 않는다.
- [0151] 일부 구현예에서, 전달 플라스미드는 백본은 RNA-OUT 서열을 포함한다. RNA-OUT은, 예를 들어 본원에 참조로서 통합된 미국 특허 제9,109,012호 및 제9,737,620호에 기술된 바와 같이, 항생제의 사용 내에서 전달 플라스미드를 보유하는 세포의 선택을 용이하게 하는 선택성 마커 시스템이다. 일부 구현예에서, RNA-OUT 서열은 다음과 같다:
- [0152] GTAGAATTGGTAAAGAGAGTCGTGATAAAATATCGAGTTCGCACATCTTGTGTCTGATTATTGATTTTTGGCGAAACCATTTGATCATATGACAAGATGTGTATCTACCTTAACTTAATGATTTTTGATAAAAATCATTAGG
- [0153] (서열번호 22)
- [0154] 유리하게는, RNA-OUT 서열은 패키징 세포주에서 전달 플라스미드의 안정적인 전파에 기여한다.
- [0155] 일부 구현예에서, 본 개시는 발현 카세트를 포함하는 전달 플라스미드를 제공하며, 발현 카세트는 서열번호 1과

적어도 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% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 발현 카세트는 서열번호 1과 적어도 80%, 85%, 90%, 95%, 99%, 또는 100% 동일성을 공유하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 발현 카세트는 서열번호 1의 서열을 갖는다.

[0156] AAV는 4.7 kb, 단일 가닥 DNA 바이러스이다. 야생형 AAV는 비병원성이며 알려진 질환과 병인적 연관성이 없기 때문에, AAV에 기초한 재조합 벡터는 우수한 임상 안전성과 관련이 있다. 또한, AAV는 다수의 조직에서 매우 효율적인 유전자 전달 및 지속적인 이식 유전자 발현 능력을 제공한다. "AAV 벡터"는 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAVrh.10, AAVrh.74 등을 포함하되 이들로 한정되지 않는, 아테노-관련 바이러스 혈청형으로부터 유래된 벡터를 의미한다. AAV 벡터는 전체적으로 또는 부분적으로, 예를 들어, *rep* 및/또는 *cap* 유전자가 결실된 하나 이상의 AAV 야생형 유전자를 가질 수 있지만, 기능성 측면 역위 말단 반복(ITR) 서열은 유지한다. 기능성 ITR 서열은 AAV 비리온의 구조, 복제 및 포장에 필요하다. 따라서, AAV 벡터는 적어도 바이러스의 복제 및 포장(예를 들어, 기능성 ITR)을 위해 시스(cis)에서 필요한 이들 서열을 포함하도록 본원에서 정의된다. ITR은 야생형 뉴클레오티드 서열일 필요는 없으며, 서열이 기능적 구조, 복제 및 포장을 제공하는 한, 예를 들어 뉴클레오티드의 삽입, 결실 또는 치환에 의해 변경될 수 있다. AAV 벡터는 하나 이상의 변형된 캡시드 단백질(예를 들어, VP1, VP2, 및/또는 VP3)을 포함하지만 이에 한정되지 않는 다른 변형을 포함할 수 있다. 예를 들어, 캡시드 단백질은 항성을 변화시키고/시키거나 면역원성을 감소시키기 위해 변형될 수 있다. AAV 발현 벡터는, 적어도 전사 방향으로 작동 가능하게 연결된 성분으로서, 전사 개시 영역, 관심 DNA(즉, TCIRG1 유전자) 및 전사 종결 영역을 포함하는 제어 요소를 제공하도록 공지된 기술을 사용하여 작제된다.

[0157] 약학적 조성물 및 사용 방법

[0158] 본 개시는 또한 본원에 개시된 발현 카세트 또는 벡터(예를 들어, 유전자 요법 벡터) 및 하나 이상의 약학적으로 허용 가능한 담체, 희석제 또는 부형제를 포함하는 약학적 조성물을 제공한다. 일부 구현예에서, 약학적 조성물은 본원에 개시된 발현 카세트를 포함하는 AAV 벡터를 포함하며, 예를 들어, 발현 카세트는 TCIRG1을 암호화하는 코돈-이식 유전자(예: 서열번호 3)를 포함한다. 예를 들어, 파골세포 형성 결핍(예: 영아 악성 골화석증)을 특징으로 하는 장애를 예방하거나 치료하는 데 사용하기 위한 약학적 조성물이 제공되며, 상기 조성물은, TCIRG1의 하나 이상의 이소형을 암호화하는 폴리뉴클레오티드의 핵산 서열을 포함하는 렌티바이러스 벡터의 치료적 유효량을 포함한다.

[0159] 특정 구현예에서, 본원에 개시된 렌티바이러스 입자는 대상체로부터 유래된 자가 CD34+ 조혈 줄기 세포(HSC)를 형질도입하여 유전적 결함을 보완하는 데 사용된다. 형질도입은 생체 내에서 또는 생체 외에서 발생할 수 있다. CD34+ 농축 세포 모집단은, 일부 구현예에서, CellGenix 줄기 세포 성장 배지(SCGM)에서 재조합 인간 사이토카인과 함께 배양하고, 37°C에서 5% CO<sub>2</sub> 및 5% O<sub>2</sub> 하에 인큐베이션한다. CD34+ 농축 세포 모집단은, 일부 구현예에서, 사전 자극에 사용된 것과 동일한 첨가제와 함께 인큐베이션하되, 임의로 형질도입 인핸서 및 (예를 들어, MOI 50으로) 발현 카세트 EFS-TCIRG1-WPRE를 포함하는 렌티바이러스 입자를 첨가하여 인큐베이션한다. 일부 구현예에서, 형질도입 후 방출 시험을 위해 세포 현탁액을 세척하여 세포의 일부와 상청액을 제거하고, 약물 생성물을 주입용 제제로서 냉동시킨다. 일부 구현예에서, G-CSF, 플레리팍서, 또는 G-CSF와 플레리팍서의 조합으로 환자를 치료하여 HSC를 동원한다. 그런 다음, 성분채집술에 의해 환자의 말초 혈액으로부터 HSC를 수집한다. 예를 들어, 자기 포획(예를 들어, Miltenyi Biotec CliniMACs 시스템을 이용해)을 사용해 CD34+ 세포를 농축시키고, CD34+ 농축 세포에 렌티바이러스 입자를 생체외에서 형질도입한다. 일부 구현예에서, 형질도입 공정은, 폴리락사머 및 프로스타글란딘 E2(PGE2)와 같은 형질도입 인핸서의 사용을 포함하지만, 이에 한정되지는 않는다.

[0160] 그런 다음에, 일부 구현예에서, 체중 1 kg당 적어도 2.0x10<sup>6</sup>개의 CD34+ 세포를 대상체(예: 대상체)에게 주입함으로써, 형질도입된 HSC를 이식한다. 일부 구현예에서, 이들은 TCIRG1-발현 세포를 HSC 서식지에 재충진한다. 일부 구현예에서, 이들은 TCIRG1-발현 세포를 파골세포 서식지에 재충진한다.

[0161] 예를 들어, 파골세포 형성 결핍(예: 영아 악성 골화석증)을 특징으로 하는 장애를 예방하거나 치료하는 데 사용하기 위한 약학적 조성물이 또한 제공되며, 상기 조성물은, TCIRG1의 하나 이상의 이소형을 암호화하는 폴리뉴클레오티드의 핵산 서열을 포함하는 변형된 세포의 치료적 유효량을 포함한다. 일부 구현예에서, 변형된 세포는 기능적 TCIRG1 유전자를 갖는 파골세포에서 관찰된 TCIRG1의 발현 수준과 유사한 수준으로 TCIRG1 또는 이의 기능적 변이체를 발현한다. 일부 구현예에서, 변형된 세포는 IMO를 앓고 있지 않거나 앓을 것이 의심되지 않는 대상체로부터 유래된 파골세포에서 관찰된 TCIRG1의 발현 수준과 유사한 수준으로 TCIRG1 또는 이의 기능적 변이

체를 발현한다. 일부 구현예에서, 변형된 세포는 조혈 줄기 세포(HSC)이다. 일부 구현예에서, 변형된 세포는 CD34+ 전구 세포(progenitor cell)이다. 일부 구현예에서, 변형된 세포는 IMO를 앓고 있거나 앓을 것이 의심되는 대상체로부터 단리한 HSC로부터 성분채집술에 의해 유래된다. 일부 구현예에서, 변형된 세포는 대상체에 대한 자가 세포이다. 일부 구현예에서, 변형된 세포는 IMO를 앓고 있거나 앓을 것이 의심되는 대상체에게 G-CSF, 플레리팍서(plerifaxor), 또는 G-CSF와 플레리팍서의 조합을 투여하여 HSC를 동원한 후 단리한 HSC로부터 성분채집술에 의해 유래된다. 일부 구현예에서, 변형된 세포는 CD34+ 세포가 풍부한 세포 모집단으로부터 자기 포획에 의해 유래된다. 일부 구현예에서, 변형된 세포는 본원에 개시된 벡터, 예를 들어, 본원에 개시된 전달 플라스미드를 사용하여 생산된 렌티바이러스 벡터를 사용하여 형질도입하였다.

[0162] 발현 카세트 또는 렌티바이러스 벡터 또는 변형된 세포를 함유하는 약학적 조성물은 선택된 투여 방식, 예를 들어, 심실 내, 심근 내, 관상 내, 정맥 내, 동맥 내, 신장 내, 요도 내, 경막 외 또는 근육 내 투여에 적합한 임의의 형태일 수 있다. 하나 이상의 TCIRG1 이소형을 암호화하는 폴리뉴클레오티드를 포함하는 유전자 조작 세포는, 단일 활성제로서, 또는 종래의 약학적 보조제와의 혼합물로서 다른 활성제와 조합되어 단위 투여 형태로, 동물 및 인간에게 투여될 수 있다. 일부 구현예에서, 약학적 조성물은 본 개시의 유전자 요법 벡터 중 어느 하나가 생체외에서 형질도입된 세포를 포함한다.

[0163] 다양한 구현예에서, 약학적 조성물은 주사 가능한 제형을 위해 약학적으로 허용 가능한 비히클(예를 들어, 담체, 희석제, 및 부형제)을 함유한다. 이들은 특히 등장성 멸균 식염수 용액(인산나트륨 또는 인산2나트륨, 나트륨, 칼륨, 칼슘, 또는 염화마그네슘 등 또는 이러한 염들의 혼합물)이거나, 경우에 따라, 멸균수 또는 생리 식염수를 첨가하면 주사 용액을 구성할 수 있게하는 건조 조성물, 특히 동결 건조 조성물일 수 있다. 주사용으로 적합한 예시적인 약학적 형태는, 예를 들어, 멸균 수용액 또는 분산액; 참기름, 땅콩유, 또는 수성 프로필렌 글리콜을 포함하는 제형; 및 멸균 주사 용액 또는 분산액의 즉석 제조를 위한 멸균 분말을 포함한다.

[0164] 또 다른 양태에서, 본 개시는 영아 악성 골석화증(IMO) 또는 다른 장애의 하나 이상의 증상의 예방, 완화, 경감, 감소, 억제, 제거 및/또는 역전을 필요로 하는 대상체에게서 이러한 증상을 예방, 완화, 경감, 감소, 억제, 제거 및/또는 역전시키는 방법을 제공하며, 상기 방법은 본 개시의 유전자 요법 벡터를 대상체에게 투여하는 단계를 포함한다. 용어 "영아 악성 골석화증" 또는 "악성 영아 골석화증" 또는 "영아 상염색체 열성 골석화증" 또는 "영아 골석화증" 또는 "IMO"는 일반적으로 영아기에 나타나고 독특한 방사선학적 발현인 전신성 골과 다증(뼈의 과도한 성장)을 특징으로 하는 희귀한 골경화증 유형의 골격 이형성증을 지칭한다. 골밀도의 전신성 증가는 수질 부분에 관여하되 피질은 상대적으로 보존하는 특별한 호발성을 갖는다. 골수 공간의 폐색에 이어지는 세포 기능 저하는 심각한 혈액학적 합병증을 초래할 수 있다. 골 팽창에 따른 시신경 위축과 뇌신경 손상은 현저한 이환율을 초래할 수 있다. 치료하지 않은 경우의 예후는 매우 나쁘다. 단순 방사선 촬영은 진단에 대한 주요 정보를 제공한다. 임상적 및 방사선학적 상관관계도 진단 과정에 기초적인 것이며, 유전자 검사를 추가하면 확실하게 알 수 있다.

[0165] 일 구현예에서, 변형된 세포, 예를 들어, 본 개시의 렌티바이러스 입자를 형질도입한 자가 세포는 다음으로 이루어진 군으로부터 선택된 경로를 통해 투여된다: 비경구, 정맥 내, 동맥 내, 심장 내, 관상동맥 내, 심근 내, 신장 내, 요도 내, 경막 외, 및 근육 내. 일부 구현예에서, 변형된 세포는 주입(예: 정맥내 주입)에 의해 투여된다. 일 구현예에서, 변형된 세포가 여러번 투여된다. 일 구현예에서, 변형된 세포는 주입에 의해 투여된다.

[0166] 일 구현예에서, 본 개시는 질환 또는 장애, 임의로 IMO의 치료를 필요로 하는 대상체에게서 이를 치료하는 방법을 제공하며, 상기 방법은 본 개시에 따른 유전자 요법 벡터와 세포를 접촉시키는 단계, 및 상기 세포를 대상체에게 투여하는 단계를 포함한다. 일 구현예에서, 세포는 줄기 세포, 임의로 다능성 줄기 세포이다. 일 구현예에서, 줄기 세포는 골 세포로 분화될 수 있다. 일 구현예에서, 줄기 세포는 파골세포로 분화될 수 있다. 일 구현예에서, 줄기 세포는 자가 조직이다. 일 구현예에서, 줄기 세포는 CD34+ 줄기 세포이다.

[0167] 일 구현예에서, 대상체는 IMO 또는 다른 장애의 증상을 나타낸다. 일 구현예에서, 대상체는 TCIRG1 발현이 감소되었거나 검출 불가능한 것으로 식별되었다. 일 구현예에서, 대상체는 돌연변이된 TCIRG1 유전자를 갖는 것으로 식별되었다.

[0168] 본원에 기술된 방법을 사용하는 치료를 받을 수 있는 대상체/환자는 불충분한 파골세포를 특징으로 하는 질환 또는 장애(예: IMO)뿐만 아니라 다른 알려진 파골세포 형성 장애의 위험이 있는 개체를 포함한다. 일부 구현예에서, 대상체는 증상을 나타내지 않는다. 일부 구현예에서, 대상체는 현재 증상을 나타내는 중이다. 이러한 대상체는 돌연변이된 TCIRG1 유전자를 갖거나, TCIRG1 발현 수준이 감소되었거나 검출 불가능한 것으로 식별될 수 있다. 증상은 적극적으로 나타날 수 있거나, (예를 들어, 약물에 의해) 억제 또는 조절될 수 있거나, 관

해 상태일 수 있다. 대상체는, 예를 들어, 자격을 갖춘 의사에 의해 장애로 진단을 받았거나 그렇지 않을 수 있다.

[0169] 정의

[0170] 용어 "T 세포 면역 조절인자 1, ATPase H<sup>+</sup> 수송 V0 서브유닛 a3" 및 "TCIRG1"은 (1) TCIRG1 핵산에 의해 암호화된 아미노산 서열(예를 들어, GenBank 수탁 번호 NM\_006019.4 (변이체 1), NM\_006053.3 (변이체 2), NM\_001351059.1 (변이체 3) 참조) 또는 TCIRG1 폴리펩티드의 아미노산 서열(예를 들어, GenBank 수탁 번호 NP\_006044.1 (이소형 A), NP\_006044.1 (이소형 B), NP\_001337988.1 (이소형 C) 참조)에 대해, 바람직하게는 이들 아미노산 서열 중 적어도 약 25, 50, 100, 200, 300, 400개 또는 그 이상의 아미노산 영역에 걸쳐, 또는 이들 아미노산 서열의 전장에 걸쳐 약 90%를 초과하는 아미노산 서열 동일성(예: 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 이상의 아미노산 서열 동일성)을 갖고; (2) TCIRG1 폴리펩티드(예를 들어, 본원에 기술된 TCIRG1 폴리펩티드)의 아미노산 서열 또는 TCIRG1 핵산(예를 들어, 본원에 기술된 TCIRG1 폴리뉴클레오티드)에 의해 암호화된 아미노산 서열을 포함하는 면역원에 대해 성장시킨 항체(예를 들어, 다클론 항체) 및 이의 보존적으로 변형된 변이체에 결합하고; (3) TCIRG1 단백질을 암호화하는 핵산 서열에 상응하는 안티-센스 가닥 및 이의 보존적으로 변형된 변이체와 엄격한 혼성화 조건 하에서 특이적으로 혼성화되고; (4) TCIRG1 핵산(예: 본원에 기술된 TCIRG1 폴리뉴클레오티드, 및 본원에 기술된 것과 같은 TCIRG1 폴리펩티드를 암호화하는 TCIRG1 폴리뉴클레오티드)에 대해, 바람직하게는 상기 핵산의 25, 50, 100, 200, 500, 1000, 2000개 또는 그 이상의 뉴클레오티드에 걸쳐, 또는 상기 핵산의 전장에 걸쳐 약 90%를 초과하는, 바람직하게는 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 그 이상의 뉴클레오티드 서열 동일성을 갖는 핵산 서열을 갖는 핵산 및 폴리펩티드 다형성 변이체, 대립 유전자, 돌연변이체, 및 중간 동족체를 지칭하도록 상호 교환적으로 사용된다.

[0171] *TCIRG1* 유전자는 여러 개의 단백질 이소형을 암호화하는데, 이중 주된 이소형은 2가지이다. 전장 이소형 a(OC116)는, 세포내 구획 및 파골세포의 소기관의 pH를 포함하여, 진행 세포의 세포내 구획 및 소기관의 pH 조절에 관여하는, 공포성 H<sup>(+)</sup>-ATPase의 A3 서브유닛을 암호화한다. 더 짧은 이소형 b(TIRC7)는 T 림프구 활성화 및 면역 반응에서 필수적인 역할을 하는 T 세포-특이적 막 단백질을 암호화한다.

[0172] 2개 이상의 핵산 또는 폴리펩티드 서열의 맥락에서의 "동일한" 또는 "동일성" 백분율이란 용어는, 비교 윈도우 또는 지정된 영역에 걸쳐 최대 상응도에 대해 비교하고 정렬할 때, 다음의 서열 비교 알고리즘 중 하나를 사용하거나 수동 정렬 또는 육안 검사를 통해 측정했을 때, 2개 이상의 서열 또는 하위 서열이 동일하거나, 이들의 아미노산 잔기 또는 뉴클레오티드의 명시된 백분율이 명시된 영역에 걸쳐 기준 서열(예를 들어, 본원에 기술된 것과 같은 TCIRG1 폴리뉴클레오티드 또는 폴리펩티드 서열)과 동일한 것, 즉 기준 서열과 적어도 약 80% 동일성, 예를 들어, 적어도 약 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 동일성을 공유하는 것을 지칭한다. 이러한 서열은 "실질적으로 동일한" 것으로 언급된다. 이 정의는 또한 시험 서열의 보안을 지칭한다. 바람직하게는, 동일성은 길이가 적어도 약 25개 아미노산 또는 뉴클레오티드인 영역에 걸쳐, 예를 들어, 길이가 50, 100, 200, 300, 400개 아미노산 또는 뉴클레오티드인 영역에 걸쳐, 또는 기준 서열의 전장에 걸쳐 존재한다.

[0173] 서열 비교를 위해, 통상적으로 하나의 서열이 시험 서열이 비교되는 기준 서열로서 작용한다. 서열 비교 알고리즘을 사용할 때, 시험 및 기준 서열은 컴퓨터에 입력되고, 필요한 경우 하위 서열 좌표가 지정되고, 서열 알고리즘 프로그램 파라미터가 지정된다. 기본 프로그램 파라미터가 사용될 수 있거나, 대체 파라미터가 지정될 수 있다. 이어서, 서열 비교 알고리즘은 프로그램 파라미터에 기초하여 기준 서열에 대한 시험 서열의 서열 동일성 백분율을 계산한다. 핵산 및 단백질의 TCIRG1 핵산 및 단백질에 대한 서열 비교를 위해, BLAST 및 BLAST 2.0 알고리즘 및 기본 파라미터가 사용된다.

[0174] 본원에서 사용되는 "비교 윈도우"는 20 내지 600, 일반적으로 약 50 내지 약 200, 보다 일반적으로 약 100 내지 약 150으로 이루어진 군으로부터 선택된 연속 위치의 수 중 임의의 하나의 세그먼트에 대한 참고를 포함하며, 여기에서 서열은 2개의 서열이 최적으로 정렬된 후 연속 위치의 동일한 수의 기준 서열과 비교될 수 있다. 비교를 위한 서열의 정렬 방법은 당업계에 잘 알려져 있다. 비교를 위한 서열의 최적 정렬은, 예를 들어, Smith 및 Waterman의 국소 상동성 알고리즘(the local homology algorithm)[*Adv. Appl. Math.* 2:482(1981)], Needleman 및 Wunsch의 상동성 정렬 알고리즘(the homology alignment algorithm)[*J. Mol. Biol.* 48:443(1970)], Pearson 및 Lipman의 유사도 검색 방법(the search for similarity method)[*Proc. Nat'l. Acad. Sci. USA* 85:2444(1988)], 이들 알고리즘의 컴퓨터화된 구현(Wisconsin Genetics 소프트웨어 패키지의 GAP, BESTFIT, FASTA, 및 TFASTA; Genetics Computer Group, 575 Science Dr., Madison, WI), 또는 수동 정렬 및 육안 검사

(예를 들어, Ausubel(eds) 등의 문헌[Current Protocols in Molecular Biology(1995 supplement)] 참조)에 의해 수행될 수 있다. 서열 동일성 백분율 및 서열 유사성을 결정하는 데 적합한 알고리즘의 예는 BLAST 및 BLAST 2.0 알고리즘이며, 이들은 Altschul 등의 문헌[*Nucleic Acids Res.* 25:3389-3402 (1977)] 및 Altschul 등의 문헌[*J. Mol. Biol.* 215:403-410 (1990)]에 각각 기술되어 있다. BLAST 분석을 수행하기 위한 소프트웨어는 National Center for Biotechnology Information(www.ncbi.nlm.nih.gov)을 통해 공개적으로 이용할 수 있다.

[0175] 2개의 핵산 서열 또는 폴리펩티드가 실질적으로 동일하다는 표시는, 후술하는 바와 같이, 제1 핵산에 의해 암호화된 폴리펩티드가 제2 핵산에 의해 암호화된 폴리펩티드에 대해 생성된 항체와 면역학적으로 교차 반응한다는 것이다. 따라서, 폴리펩티드는 전형적으로 제2 폴리펩티드와 실질적으로 동일하다(예를 들어, 2개의 펩티드는 보존적 치환에 의해서만 상이하다). 2개의 핵산 서열이 실질적으로 동일하다는 다른 표시는 2개의 분자 또는 이들의 보체가 엄격한 조건 하에서 서로 하이브리드화된다는 것이다. 2개의 핵산 서열이 실질적으로 동일하다는 또 다른 표시는 서열을 증폭하는 데 동일한 프라이머가 사용될 수 있다는 것이다.

[0176] 본원에서 사용되는 "투여"는 국소 및 전신 투여를 지칭하며, 예를 들어, 장관, 비경구, 폐, 및 국소/경피 투여를 포함한다. 본원에 기술된 방법에서 유용한 화합물(예를 들어, 하나 이상의 TCIRG1 이소형을 암호화하는 폴리뉴클레오티드)의 투여 경로는, 예를 들어, 대상체에 대한 경구(os 당(P.O.)) 투여, 비강 또는 흡입 투여, 좌약 투여, 국소 접촉, 경피 전달(예를 들어, 경피 패치를 통한 전달), 경막 내(IT) 투여, 정맥 내("iv") 투여, 복강 내("ip") 투여, 근육 내("im") 투여, 병변 내 투여, 또는 피하("sc") 투여, 또는 서방출 장치(예를 들어, 소형-삼투압 펌프, 데포 제형 등)의 이식을 포함한다. 투여는 비경구 및 경점막(예를 들어, 경구, 비강, 질, 직장 또는 경피)을 포함하는 임의의 경로에 의해 이루어질 수 있다. 비경구 투여는, 예를 들어, 정맥 내, 근육 내, 동맥 내, 신장 내, 요도 내, 심장 내, 관상 내, 심근 내, 피부 내, 경막 외, 피하, 복강 내, 심실 내, 이온 영동(ionophoretic) 및 두개 내를 포함한다. 다른 전달 모드는 리포솜 제형, 정맥 내 주입, 경피 패치 등의 사용을 포함하지만, 이에 한정되지 않는다.

[0177] 용어 "전신 투여" 및 "전신적 투여"는 화합물 또는 조성물이 순환계를 통해, 약학적 작용의 표적화된 부위를 포함하는 신체 내 부위에 전달되도록 화합물 또는 조성물을 포유동물에게 투여하는 방법을 지칭한다. 전신 투여는 경구, 비강 내, 직장 및 비경구(예를 들어, 근육 내, 정맥 내, 동맥 내, 경피 및 피하와 같은, 소화관을 통하는 것 외) 투여를 포함하나 이에 한정되지 않는다.

[0178] 예를 들어 화합물(예를 들어, TCIRG1 폴리뉴클레오티드) 및/또는 이의 유사체 및 또 다른 활성제와 관련하여 사용될 때의 용어 "공동 투여" 또는 "동시 투여"는, 화합물 및/또는 유사체 및 활성제 모두가 동시에 생리학적 효과를 달성할 수 있도록 투여하는 것을 지칭한다. 그러나, 두 가지 제제가 반드시 함께 투여되어야 하는 것은 아니다. 특정 구현예에서, 하나의 제제의 투여는 다른 하나의 투여에 선행할 수 있다. 동시 생리학적 효과는 반드시 동시에 순환하는 두 제제의 존재를 필요로 하지 않는다. 그러나, 특정 구현예에서, 공동 투여는 일반적으로 임의의 투여량에 대한 최대 혈청 농도의 상당한 분율(예를 들어, 20% 이상, 예를 들어, 30% 또는 40% 이상, 예를 들어, 50% 또는 60% 이상, 예를 들어, 70% 또는 80% 또는 90% 이상)로 신체 내에(예를 들어, 혈장 내에) 동시에 존재하는 두 가지 제제를 생성한다.

[0179] 용어 "유효량" 또는 "약학적 유효량"은, 손상되거나 결핍된 자가포식(예를 들어, IMO)을 특징으로 하는 질환의 궁극적인 중증도를 감소시키기에 충분한 양으로 원하는 결과(예를 들어, 하나 이상의 TCIRG1 이소형의 발현 증가)를 달성하는 데 필요한 하나 이상의 화합물(예를 들어, 유전자 요법 벡터, 변형된 세포)의 양 및/또는 투여량, 및/또는 투여 요법을 지칭한다.

[0180] "투여 사유"라는 어구는, 문제가 되는 제제(들)/화합물(들)을 대상체에게 투여하는 것을 통제하고/하거나 허용하는 의료 전문가(예를 들어, 의사) 또는 대상체의 의학적 치료를 통제하는 사람이 취한 조치를 지칭한다. 투여 사유는 적절한 치료적 또는 예방적 처방, 및/또는 대상체에 대한 특정 제제(들)/화합물 처방의 진단 및/또는 결정을 포함할 수 있다. 이러한 처방은, 예를 들어, 처방 양식 초안 작성, 의료 기록 주석 달기 등을 포함할 수 있다.

[0181] 본원에서 사용되는 용어 "치료하는" 및 "치료"는 이 용어가 적용되는 질환 또는 병태, 또는 이러한 질환 또는 병태의 하나 이상의 증상의 개시를 지연시키거나, 진행을 지연 또는 경감시키거나, 중증도를 감소시키거나, 완화시키거나 예방하는 것을 지칭한다. 용어 "치료하는" 및 "치료"는 또한 질환 또는 병태의 하나 이상의 증상을 예방, 완화, 경감, 감소, 억제, 제거 및/또는 역전시키는 것을 포함한다.

[0182] 용어 "완화"는 그 병리 또는 질환의 하나 이상의 증상의 감소 또는 제거, 및/또는 그 병리 또는 질환의 하나 이

상의 증상의 속도의 감소, 또는 발병 또는 중증도의 지연 및/또는 그 병리 또는 질환의 예방을 지칭한다. 소정의 구현예에서, 병리 또는 질환의 하나 이상의 증상의 감소 또는 제거는, 예를 들어, TCIRG1의 하나 이상의 이소형의 발현 수준의 측정 가능하고 지속적인 증가를 포함할 수 있다.

[0183] 본원에서 사용되는 문구 "본질적으로 구성되는"은 방법 또는 조성물에 인용된 활성 약학적 제제의 유전자 또는 종을 지칭하고, 그 자체로 인용된 지시 또는 목적에 대해 실질적인 활성을 갖지 않는 다른 제제를 추가로 포함할 수 있다.

[0184] 용어 "대상체", "개체" 및 "환자"는 상호 교환적으로 포유류, 바람직하게는 인간 또는 비인간 영장류를 지칭할 뿐 아니라, 가축 포유류(예를 들어, 개 또는 고양이), 실험실 포유류(예를 들어, 마우스, 랫트, 토끼, 햄스터, 기니피그) 및 농업 포유류(예를 들어, 말, 소, 돼지, 양) 또한 지칭한다. 다양한 구현예에서, 대상체는 인간(예를 들어, 성인 남성, 성인 여성, 청소년 남성, 청소년 여성, 남성 아동, 여성 아동)일 수 있다.

[0185] 용어 "유전자 전달" 또는 "유전자 이송"은 외래 DNA를 숙주 세포 내로 신뢰성 있게 삽입하기 위한 방법 또는 시스템을 지칭한다. 이러한 방법은 비통합으로 전달된 DNA의 일시적인 발현, 전달 레플리콘(예를 들어, 에피솜)의 염색체 외 복제 및 발현, 또는 숙주 세포의 게놈 DNA 내로의 전달 유전 물질의 통합을 초래할 수 있다.

[0186] 본원에서 용어 "백터"는 또 다른 핵산 분자를 전달 또는 수송할 수 있는 핵산 분자를 지칭하기 위해 사용된다. 전달된 핵산은 일반적으로 백터 핵산 분자에 연결, 예를 들어 백터 핵산 분자 내로 삽입된다. 백터는 세포에서의 자율 복제 또는 역전사를 유도하는 서열을 포함할 수 있거나, 숙주 세포 DNA로의 통합을 허용하기에 충분한 서열을 포함할 수 있다. "백터"는 유전자 요법 백터를 포함한다. 본원에서 사용되는 용어 "유전자 요법 백터"는 유전자 요법을 수행하는 데 사용할 수 있는 백터, 예를 들어, 치료 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열을 대상체에게 전달하는 백터를 지칭한다. 유전자 요법 백터는 대상체에게 도입시킬 경우, 치료적 활성 폴리펩티드, 예를 들어, TCIRG1 또는 대체 유전자 요법에 유용한 다른 유전자를 암호화하는 핵산 분자("이식 유전자")를 포함할 수 있다. 유용한 백터는 바이러스 백터를 포함하지만, 이에 한정되지 않는다.

[0187] 본원에서 사용되는 용어 "발현 카세트"는 전술한 발현 카세트에 혼입된 치료적 활성 폴리펩티드(예를 들어, TCIRG1)를 암호화하는 폴리뉴클레오티드("이식 유전자")의 발현을 유도하는 적절한 환경에서 능력이 있는 DNA 세그먼트를 지칭한다. 숙주 세포 내로 도입될 경우, 발현 카세트는 특히 세포의 기구가 이식 유전자를 RNA로 전사하도록 유도할 수 있고, 이어서 이는 일반적으로 추가로 가공되고 최종적으로 치료적 활성 폴리펩티드로 번역된다. 발현 카세트는 유전자 요법 백터에 포함될 수 있다. 일반적으로, 발현 카세트라는 용어에는 5' LTR에 대해 5'에 있는 폴리뉴클레오티드 서열 및 3' LTR에 대해 3'에 있는 폴리뉴클레오티드 서열이 포함되지 않는다.

[0188] 본 명세서에 참조되고 식별된 모든 특허, 특허 공개, 및 기타 공개문헌은 모든 목적을 위해 그 전체가 참조로서 본원에 개별적으로 및 명시적으로 통합된다.

[0189] 실시예

[0190] 실시예 1: 전달 플라스미드의 안정한 전파

[0191] 최소 TCIRG1 발현 카세트인 EFS-TCIRG1-WPRE(서열번호 1)를 포함하는 상이한 플라스미드들의 안정성을 조사하였다. 암피실린 저항성(AmpR)을 가진 플라스미드 작제물 pRRL.PPT.EFS.tcirg1h.wpre(도 3c)는, 플라스미드를 패키징 세포주 내로 형질감염시키기 전에 이를 전파하는 데 사용한 대장균 세포에서 배양 하는 동안 예상 외로 낮은 성장 및 불안정성을 나타냈으며, 이는 플라스미드의 낮은 수율, 및 불안정성을 나타내는 전체적으로 분해된 DNA 얼룩에 의해 나타났다(도 3a). 다양한 다른 플라스미드 백본도 불안정성을 나타냈다(데이터 미도시). 그러나, 최소 발현 카세트인 EFS-TCIRG1-WPRE(서열번호 1)를 pRRL 백터로부터 RNA-OUT 서열을 갖는 pCCL 백터로 클로닝한 경우(도 3c), 생성된 플라스미드 작제물 pCCL.PPT.EFS.tcirg1h.wpre(서열번호 27)는 대장균에서 전파될 때 예상 외로 양호한 성장 및 안정성을 나타냈다. 이는 플라스미드의 높은 수율 및 제한된 분해 패턴으로 나타났다(도 3b). 전체 백터 서열은 서열번호 27로서 제공되며, 각 백터 요소의 위치는 표 2에 제공되어 있다.

표 2

pCCL.PPT.EFS.tcirg1h.wpre 백터 요소			
명칭	유형	기준 서열 내 최소 ~ 최대위치 (뉴클레오티드)	길이 (염기 쌍)
EFS	프로모터	1~243	243

[0192]

TCIRG1	CDS	257~2,749	2,493
WPRE	조절	2,782~3,384	603
3' LTR	LTR	3,471~3,704	234
SV40 폴리(A)	폴리A 신호	3,776~3,907	132
SV40 ori	복제 기원	3,917~4,076	160
pUC 기원	복제 기원	4,115~5,129	1,015
RNA-OUT	억제인자	5,146~5,284	139
CMV IE	프로모터	5,334~5,910	577
5-LTR	LTR	5,933~6,120	188
psi	패키징	6,222~6,266	45
gag	CDS	6,267~6,628	362
RRE	조절	6,629~7,486	858
cPPT/CTS	폴리 퓨린 구역	7,505~7,622	118

[0193] 렌티바이러스 벡터는 pCCL/RNA-OUT 벡터를 패키징 플라스미드(pCMV ΔR8.91) 및 외피 플라스미드(VSV-G pMDG)와 함께 293T 세포 내로 일시적으로 형질감염시켜 생산하며, 도 4에 도시된 프로토콜에 따라 생산한다.

[0194] **실시예 2: pCCL.PPT.EFS.tcirg1h.wpre를 이용해 IMO 환자로부터 파골세포의 흡수 기능 복원**

[0195] 본 실시예는 환자 유래 HSC에서 렌티바이러스-매개 TCIRG1 유전자 전달을 위한 pCCL.PPT.EFS.tcirg1h.wpre의 용도를 입증한다. HSC를 획득하고, 증식시키고, 실시예 1에 기술된 pCCL.PPT.EFS.tcirg1h.wpre를 가진 렌티바이러스 입자를 형질도입하여 유전자-변형 HSC를 획득한다. 주입 후, 유전자-변형 HSC는 파골세포로 분화하게 된다. 사용된 방법은 Moscatelli 등의 문헌[Hum. Gene Therap. 29:938-949 (2017)]에 기술된 것과 본질적으로 같다.

[0196] IMO 환자의 말초 혈액 샘플 또는 정상 분만아의 체대혈(CB) 샘플을 구한다. 이들 공급원으로부터의 단핵 세포를 피콜(Ficoll)과 함께 밀도 구배 원심분리하여 단리하고, CD34+ 세포를 자석 활성화 세포 분류(MACS) 컬럼(Miltenyi Biotec, Bergisch Gladbach, Germany)을 사용하여 단핵 세포 분획으로부터 분리한다. 증식을 위해, 인간 재조합 사이토카인 M-CSF(50 ng/ml), GM-CSF(30 ng/ml), SCF(200 ng/ml), IL-6(10 ng/ml), 및 Flt3L(50 ng/ml)(R&D Systems, Minneapolis MN)이 포함된 SFEM StemSpan 배지(StemCell Technologies, Vancouver, BC)에서 세포를 배양한다. CD34+ 세포를 24-웰 세균학 플레이트를 사용하여 1 ml 배지에  $5 \times 10^4$ 개의 세포의 밀도로 도말하고, 37°C에서 1주일 동안 인큐베이션한 후 수집하고,  $1 \times 10^5$ /웰의 밀도로 재도말한다. 7일차부터, 2~3일마다 배지를 절반씩 고갈시켜 교환한다. 이식을 위해, CD34+ 세포를 다음의 인간 재조합 사이토카인과 함께 SFEM StemSpan 배지(StemCell Technologies, Vancouver, BC)에서 30시간 동안 배양한다: SCF(100 ng/ml), Flt3L(100 ng/ml), 및 TPO(100 ng/ml) (R&D Systems, Minneapolis MN).

[0197] 레트로벡터(Takara Bio, Otsu, Japan)으로 코팅된 24-웰 플레이트에서 형질도입을 수행한다. 시험관 내 실험을 위해, 3일차에 CD34+ 세포에 제1 히트(first hit)를 30의 감염다중도(MOI)로 6시간 동안 형질도입하고, 7일차에 제2 히트(second hit)를 30의 MOI로 6시간 동안 형질도입한 다음, 골수성 사이토카인 카테일과 함께 배양한 후 파골세포로 분화시켰다. 생체 내 실험을 위해, CD34+ 모집단의 줄기/전구 세포 성질을 유지하면서 효율적인 형질도입을 허용하기 위한 더 짧은 형질도입 프로토콜이 사용된다. 단핵 세포를 단리하고, 제1 히트(MOI 30 또는 100)를 밤새 형질도입한 다음, 다음 날 제2 히트(MOI 30 또는 100)를 6시간 동안 형질도입하면, 세포를 대상체(마우스 또는 인간 환자)에게 이식할 준비가 된다.

[0198] 파골세포형성은 50 ng/ml M-CSF 및 50 ng/ml RANKL의 존재 하에 약 10일 동안 분화시켜 평가할 수 있고, 이어서 추가 분석을 위해 4% 포르말데히드로 세포를 고정시키거나 웨스턴 블롯 분석을 위해 세포를 용해시킨다. 재흡수는, c-말단 I형 콜라겐 단편(CTX-I)의 방출 또는 배지 내로 Ca<sup>2+</sup>의 방출을 검정하고, 고정된 세포를 헤마톡실린으로 염색하여 재흡수 함요(resorption pit)의 형성을 시각화함으로써 평가한다.

[0199] 동물 연구를 위해, 형질도입된 파골세포를 NSG 마우스에게 이식한다. 8 내지 15주령의 NSG 마우스에게 300 cGy를 준치사량으로 조사하고, 6시간 후에 형질도입되지 않은  $1 \times 10^5$ 개의 CB CD34+ 세포 또는 pCCL.PPT.EFS.tcirg1h.wpre로부터 유래된 렌티바이러스 입자를 형질도입한  $1 \times 10^5$ 개의 IMO CD34+ 세포를 꼬리 정맥 주사에 의해 이식한다. 이식 후 감염을 피하기 위해, 마우스에게 2주 동안 식수를 통해 시프로플록사신을 투여한다. 말초 혈액은 상이한 시점에 채취하였고, 골수 세포는 마우스를 희생시킨 후 모르타르로 대퇴골을 분쇄

하여 채취하였다.

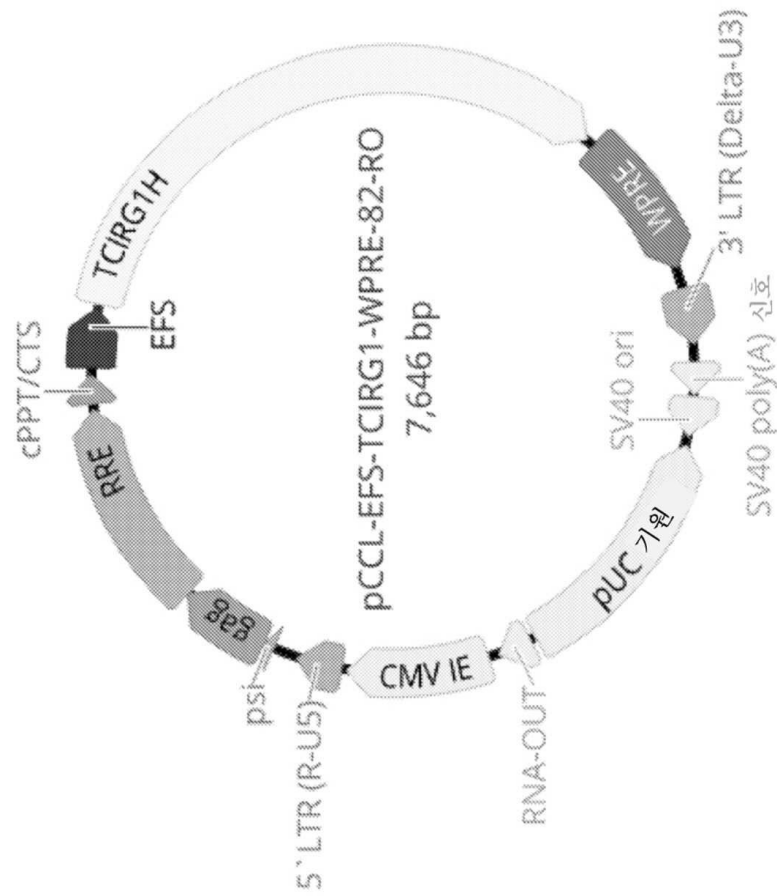
- [0200] 이식 후 9~19주차에 마우스에게서 채취한 샘플의 전체 골수 계층 DNA를 이용해 백터 카피 수 분석을 수행한다. 이식된 NSG 마우스의 말초 혈액 및 골수를, huCD45-APC에 대해 양성인 세포의 백분율을 결정함으로써 인간 재구성에 대해 분석한다. 계통 분석을 위해, CD33-PeCy7, CD15-PeCy7, CD19-BV605, 및 CD3-PE에 대해 유도된 항체로 세포를 염색하였다.
- [0201] 전술한 방법은 렌티바이러스 매개 TCIRG1 유전자 전달 및 형질도입된 CD34<sup>+</sup> 세포의 장기 생착 후 IMO 환자로부터 파골세포의 재흡수 기능의 복원을 확인하는 데 사용된다.
- [0202] **실시예 3: 임상적으로 확립된 형질도입 프로토콜을 사용한 인간 CD34<sup>+</sup> 농축 세포의 시험관 내 형질도입**
- [0203] 본 실시예는 실시예 1에 기술된 플라스미드를 사용하여 제조된 EFS-TCIRG1-WPRE(서열번호 25 또는 27)의 pre-GMP 배치의 적합성을 다음에 의해 입증한다: (1) 생존 가능한 CD34<sup>+</sup> 세포의 백분율(%) 및 다중 계통 분화 능력에 의한, 형질도입된 CD34<sup>+</sup> 세포의 표현형 특성 분석, 및 (2) 백터 카피 수에 의한 형질도입 효율(액체 배양물 및 콜로니 둘 다에서의 VCN 결정).
- [0204] *EFS-TCIRG1-WPRE*로 형질도입한 후 *mPB CD34<sup>+</sup>* 세포의 보존된 표현형 및 다중 계통 분화 능력
- [0205] pre-GMP EFS-TCIRG1-WPRE 배치의 성능을 다른 장애의 치료를 위해 생산된 LV(4개의 배치)와 비교하였다. 동원된 PB CD34<sup>+</sup> 세포를 계획한 IMO 임상시험에서와 유사하게 표적 세포로서 사용하였다. 다양한 MOI를 시험하였다. 형질도입 20시간 후에, 시험한 모든 백터 및 MOI에 걸쳐 높은 세포 생존력(95% 초과)을 얻었으며, 이는 단기 독성이 없음을 나타낸다. CD34<sup>+</sup> 세포의 백분율은 형질도입 직후 시험한 모든 조건에 대해 매우 높았고(97% 초과), 액체 배양물에서 2일간 배양한 후에는, 이러한 유형의 배양물에 대해 예상되는 바와 같이, 시간 경과에 따라 모든 조건에서 유사한 수준으로 점진적으로 줄어들었다.
- [0206] 형질도입된 CD34<sup>+</sup> 세포의 다계통 능력은 반-고형 메틸셀룰로오스 배지 배양물에서 분화된 CFU의 정량화에 의해 평가하였다. 적혈구 계통 및 골수 계통을 의미하는 총 CFU뿐만 아니라 BFU-E, CFU-GM, 및 CFU-GEMM도 평가하였다. EFS-TCIRG1-WPRE의 존재는 CFU 성장에 영향을 미치지 않았는데, 이는 Mock 대조군과 비교했을 때, 실험 조건들 간에 CFU의 총 수에서 유의한 차이가 관찰되지 않았기 때문이다. Mock과 비교했을 때 어떤 콜로니 유형에서도 차이가 관찰되지 않았으므로, 높은 MOI 값으로 EFS-TCIRG1-WPRE로 형질도입한 후에도 다계통 능력이 시험관 내에서 유지됨을 확인하였다.
- [0207] *EFS-TCIRG1-WPRE*는 *mPB CD34<sup>+</sup>* 세포의 높은 형질도입 효율을 보여준다
- [0208] 치료적으로 사용하기에 적합한 형질도입 효율을 제공하는 데 필요한 EFS-TCIRG1-WPRE의 백터 투여량을 결정하기 위해, 확립된 임상 형질도입 프로토콜을 사용하여 렌티바이러스 백터를 시험하였다. VCN 평가에 앞서, 형질도입된 CD34<sup>+</sup> 세포를 최대 12일 동안 액체 배양물 중에 유지시켜 에피솜 LV 계층 카피를 세포 제거하였다. IMO 백터로 높은 VCN/세포 값을 획득하였는데, 이는 투여량 의존적이었다. 투여량 증가의 효과는 시험된 모든 백터에 걸쳐 일관되었으며, 동일한 MOI로 형질도입한 결과 LAD-I 및 FA 백터에 비해 IMO 백터에 대한 VCN/세포 값이 더 높았는데(도 5a~5b), 이는 EFS-TCIRG1-WPRE의 높은 형질도입 효율을 나타내는 것이다.
- [0209] VCN/세포를 세포의 표현형에 따라 단리된 CFU, 즉 BFU-E, CFU-GM, 또는 CFU-GEMM에서도 평가하여, 상이한 전구 세포에서의 형질도입을 확인하였다. EFS-TCIRG1-WPRE는 두 공여자 모두의 세포에서 더 높은 콜로니 VCN 값을 나타냈다. 액체 배양물에서의 결과와 유사하게, IMO 백터를 이용한 형질도입의 결과 형질도입 효율이 더 높았고, 메틸셀룰로오스 배지에서 배양된 콜로니의 VCN/세포가 더 높았다.
- [0210] 상이한 CFU 유형(BFU-E, CFU-GM, 및 CFU-GEMM)에서의 VCN 패턴은 다른 백터와 유사한 것으로 밝혀졌으며, 그 값은 이전에 Charrier 등의 문헌[*Gene Therapy* 18:479-487(2011)]에 기술된 것과 같이 적혈구 콜로니에서 흔히 발견되는 가장 높은 값이었다.
- [0211] IMO 백터 EFS-TCIRG1-WPRE는 렌티바이러스의 임상 로트와 유사한 수준으로 인간 CD34<sup>+</sup> 세포를 형질도입한다. CD34<sup>+</sup> 세포의 표현형 및 다중계통 능력은 높은 형질도입 효율을 달성하면서도 보존되었다. IMO 백터는 대조군 백터보다 낮은 MOI에서 성공적으로 작용하여, IMO 환자의 유전자 요법에 사용하기에 적합함을 입증하였다.

[0212]

이들 연구는, VCN 및 형질도입 효율이 생체 내 유전자 변형 조절 세포의 교정 수준과 유사하게 달성되어, *TCIRG1*의 돌연변이로 인한 영아 악성 골화석증의 치료에 이러한 유전자 요법을 사용할 수 있음을 입증하였다.

도면

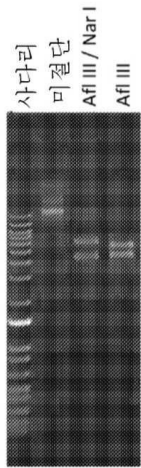
도면1



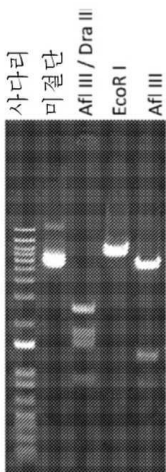
도면2

GGCTCCGGTGCCTCAGTGGGAGAGCGGCACATCGCCCAAGTCCCGGAAATTTGGGGGAGGGGTGGCAATTTGAACCGGTCCTAGAGAGGTGGCGGC  
GGTAAACTGGGAAAGTATGTCGTGTAAGTCTGGCTCGCCCTTTTCCGAGGGTGGGGGAGAACCTTATAAGTCAAGTATCGCCGTGAACGTTCTTTTTCG  
CAAGGGTTTGGCGCCAGAAACAGAGGTGTCGTGACGCGGGATCCCGCACATGGGGTCCATGTTTCGGAGGAGGTTGGCCCTGGTCCAGCTCTTTCTGGCC  
CACAGCGGCTGCCTACACTCGGTGAGTCGGCTGGCGAGCTGGGCTCFGGAGTTCAGAGACCTCAAACGCTCGGTGAGCGCCTCCAGAGAGCGCTTTGTGTG  
GTGATGTTGGCGCTGAGGAGCTGGAGAAGACCTTACCTTCAGAGGAGGCTGGCGGGTGGCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG  
CGGCACCCCAACCCCGGSACTGCTGGACATCCAGGAGGACGGAGCCCTGGCCAGGACTGGGGATGTGGGGGCAACACAGACGCCCTGCGGGCCCA  
GTGCACACAGCTCAGCTCCAGCCCGCTGTACGCCAGGCAAGAACCTAGCTGGACCCGCCACACAGATGGGGCTCAGAGAGGACCGCCCTGCTC  
CAGCCCCGGGGGGCCGACAGGACCTGAGGGTCAACTTTGTTGGAGGTTGCTGGGCCCAAGGCCCCAAGGCCCTTAGAGGCCCTGCTGGAGGCGCTT  
GCAGAGGCTTCCATTTGCCACTTCAGGGAGCTGGAGCAGCCGCTGAGCCCGGTGAGCGGGGAGCCAGGAGTCAACCCCGCTCCCTACCATCACTCACTG  
GGGTGACAGATCGGACAGAACTCCGAAAGATCACGGACTGCTTCCCTGCTCCAGTCCAGGAGGAGGGCCCGCTGGGGCCCTGGGGCCCTGGCAG  
CAGCTGACAGAGCAGAGCAGAGTCCAGGATCCGGSAGACACGCGGTTCCTGAGCCAGGTGTAGGCCGGGTGTGAGCGCTGCTGGCCGCTGGC  
AGGTGACGTTCCAAAGATGAGAGCCCTGTACTGGCCCTGAACCAAGTGCAGCTGAGCACACCAAGTGCCTCATTTGCCAGGCTGGTGTCTGTGGG  
AGACTGGCCCGCTGAGAGGCCCTGGGGGACAGTGCAGTGGAGAGGGAGTGGCTGGCTACCGCACTCCCTGGCCGGGACATGCCCCCAGACTC  
ATCCGCACCAACCGCTTACGGCTCCAGGGCATCTGGATGCTCTAGCCCTGGGGGCTACAGGAGTCAACCCCGCTCCCTACCATCACTCACTCACT  
TCCCTTCCCTGTTGTGTGATGTTGGGGATGGGCCAGGGCTGCTCATGTTCTTCCCTGGCCATGGCTATGGTCTTGGGAGAACCGACCGGCTGTGAA  
GGCCCGCAGAGCAGATGTCGACACTTTCTCAGGGCCGCTACTCTGCTCTGTTATGGGGCTGTTCTCCATCTACACCCGGTTCATCTACAACGAGTGC  
TTCAGTCCGCGCCACAGCATCTCCCTCGGGCTGGAGTGGCCCAAGTGGCCCAACAGTCTGGCTGGAGTGTGCACTTCCCTGGCCACACAGCATGCTTA  
CCCTGGACCCCAAGTCAACGGTGTCTCTGGGACCTTACCCCTTTGGGATCGATCTATTTGGAGCTGGCTGGCCACCACTTGGAGCTTCTCAACTCTCT  
CAAGATGAAGATGCTCCCTCACTCTGGGGTCTGCACTGGCCCTTTGGGGTGTCTCGGATCTTCAAACACCTGCACTTTGGCCAGAGGCAACCGGCTGCTG  
CTGGAGCTGCCGGACTCACTTCTGCTGGGACTCTTCGGTTACCTGTTTCCCTAGTCACTACAAAGTGGCTGTGTCTGGGGCTGCCAGGGCCGCT  
CGGCCCCAGCATCTCACTCAATGTTCTTCTTCCACAGCCCAAGCAACAGGCTGCTTACCCCGGCAAGGAGTGGTTCAGGGCCAGCT  
GTTGGTCTGGCTGGCCATGGCCATGCTGCTGCTGGCACACCCCTGCACTCCCGCCACCGCCCTGGGAGGAGGGCCCGCTGACCGG  
CAGAGGAACAAGGGCCGGTGTGGACCTGCTGAACTGTGAAATGGCTGGAGCTCCGATGAGGAAGAGGCAAGGCAAGGGGGCTGGATGATGAAGAGGAGG  
CCGAGCTGCTCCCTCGAGGTCTCATGCCACAGGCAATCCACACTCAGTCTGCTGGGGTGTCTCCAAACCGCCCTCTACTGCGCCCTGTGGG  
CCTGAGCCTGGCCACGCCCAAGTCTCGAGGTCTGTGGCCATGGTGTGAGTGGCCATGAGGCTGGCCCTGGCCCGGGAGGAGGGCTGGCGGCTGTGGTGTG  
GTCCCACTTTGGCCCTTTGCCGTGATGACCGTGGCTACTCTGCTGGTGTGATGAGAGGAGCTCTCAGCCTTCTGCAAGCCCTGGGGCTGGGCTGGGCTGGG  
TCCAGAACAAAGTCTACTCAGGCACGGGCTCAAGCTGAGTCCCTTCACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG  
CGAGCATCTTAACGCCAATTAACCCAAATTTGTTCTGTTTCTGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG  
TTTTAGGGATATGATACTAGTTTCAAGGTGATGCCACAAAGATGTAAGAACTTCCCGTTATTTACGCTCTGTTCTGTTTAAATCAACCTCTGG  
ATTACAAATTTGTGAAGATGACTGATACTTAACTATGTTGCTTCTTTTACCTGTGGATGCTGTTTAAATGCTCTGTATCATGCTATGCTATTGCTTC  
CCGTACGGCTTTCTCTCTCTGTAATAATCCTGGTGTGCTCTCTTTTAGAGGATTTGGCCGTTGCTCCGTTCAACGTTGGGCTGGTGTGCTGTG  
TTTGTGACGCAACCCCACTGGCTGGGCAATGCCACCACTGTCAACTCTTTCTGGGACTTTGCTTCCCTTCCGATCCGCAACGGCAGAACTCATCG  
CCGCCCTGGCCCTGGCTGGACAGGGGCTAGGTTGTGCTGGGCACTGATAATTCGCTGGTGTGTTGCTGGGGAAAGCTGACGCTCTTTTCG

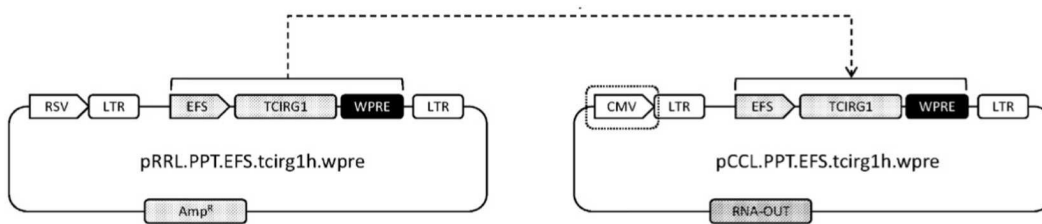
도면3a



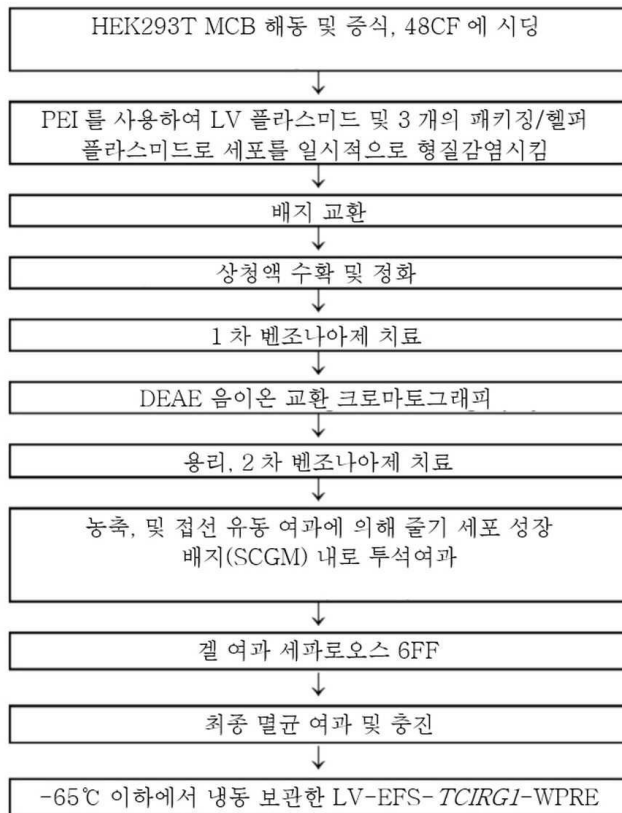
도면3b



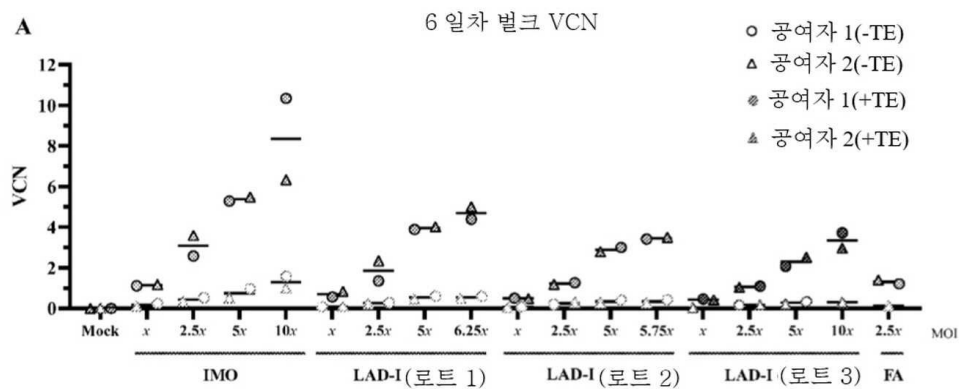
도면3c



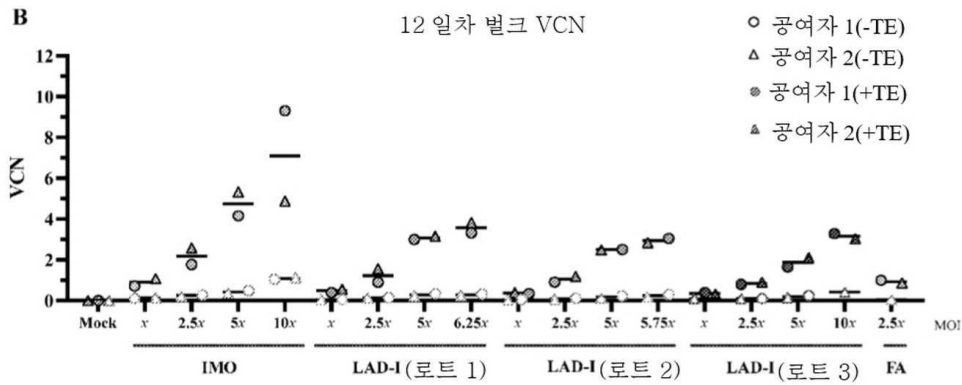
도면4



도면5a



도면5b



서열목록

SEQUENCE LISTING

<110> Rocket Pharmaceuticals, LTD

Beard, Brian

Ricks, David

Prabhakar, Raj

<120> GENE THERAPY VECTORS FOR INFANTILE MALIGNANT OSTEOPETROSIS

<130> ROPA-003/01WO 329592-2098

<150> US 62/852,216

<151> 2019-05-23

<160> 39

<170> PatentIn version 3.5

<210> 1

<211> 3384

<212> DNA

<213> Artificial Sequence

<220><223> Made in Lab - Expression cassette

<400> 1

ggctccggtg cccgtcagtg ggcagagcgc acatcgcca cagtccccga gaagttgggg 60

ggaggggtcg gcaattgaac cggtgcctag agaaggtggc gcggggtaaa ctgggaaagt 120

gatgtcgtgt actggctccg cttttttccc gagggtgggg gagaaccgta tataagtgca 180

gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca ggtgtcgtga 240

cgcgggatcc gccaccatgg gctecatgtt tcggagcgag gaggtggccc tggtecagct 300

ctttctgcc acagcggctg cctacacctg cgtgagtcgg ctgggcgagc tgggcctcgt 360

ggagttcaga gacctcaacg cctcggtag cgccttcag agacgcttg tggttgatg 420  
 tcggcgtgt gaggagctgg agaagacctt caccttctg caggaggagg tgcggcgggc 480  
  
 tgggctggtc ctccccgc caaaggggag gctgccggca cccccaccc gggacctgct 540  
 gcgcatccag gaggagacgg agcgctggc ccaggagctg cgggatgtgc ggggcaacca 600  
 gcaggcctg cggggccagc tgcaccagct gcagctccac gccccgtgc tacgccaggg 660  
 ccatgaacct cagctggcag ccgcccacac agatggggcc tcagagagga cgcccctgct 720  
 ccaggcccc gggggccgc accaggacct gagggtcaac tttgtggcag gtgccgtgga 780  
 gccccacaag gccctgccc tagagcgct gctctggagg gcctgcagag gcttctcat 840  
 tgccagcttc agggagctgg agcagccgct ggagcacccc gtgacggcg agccagccac 900  
  
 gttgatgacc ttctcatct cctactgggg tagcagatc ggacagaaga tccgcaagat 960  
 cacggactgc ttccactgcc acgtcttccc gtttctgcag caggaggagg cccgctcgg 1020  
 ggccctgcag cagctgcaac agcagagcca ggagctgcag gaggctctcg gggagacaga 1080  
 gcggttctg agccaggtgc taggccgggt gctgcagctg ctgccgccag ggcaggtgca 1140  
 ggtccacaag atgaaggccg tgtacctggc cctgaaccag tgcagcgtga gcaccacgca 1200  
 caagtgcctc attgccgagg cctggtgctc tgtgcgagac ctgcccgcc tgcaggaggc 1260  
 cctgcgggac agctcgatgg aggagggagt gagtgccgtg gctcaccgca tcccctgccg 1320  
  
 ggacatgccc cccacactca tccgaccaa ccgcttcacg gccagcttcc agggcatcgt 1380  
 ggatgcctac ggcgtgggcc gctaccagga ggtcaacccc gctccctaca ccatcatcac 1440  
 ctccccctc ctgtttgctg tgatgttcgg ggatgtgggc cacgggctgc tcatgttct 1500  
 ctteccctg gccatgttcc ttgcggagaa ccgaccgct gtgaaggccg cgcagaacga 1560  
 gatctggcag actttctca ggggccgcta cctgctctg cttatgggcc tgttctcat 1620  
 ctacaccgge ttcatctaca acgagtgtt cagtccgccc accagcatct tcccctcggg 1680  
 ctggagtgtg gccccatgg ccaaccagtc tgctggagt gatgcattcc tggcccagca 1740  
  
 cagatgctt accctggacc ccaacgtcac cgggtcttcc ctgggacct accctttgg 1800  
 catcgatcct atttgagcc tggctgcaa ccacttgagc ttcctcaact cttcaagat 1860  
 gaagatgtcc gtcacctgg gcgtcgtgca catggccttt ggggtggtcc tcggagtctt 1920  
 caaccacgtg cactttggcc agaggcaccg gctgctgctg gagacgtgc cggagctcac 1980  
 ctctctgctg ggactcttcg gttacctcgt gttcctagtc atctacaagt ggctgtgtgt 2040  
 ctgggctgcc agggccgct cggccccag catctcacc cacttcatca acatgttct 2100

cttctccac agccccagca acaggctgct ctacccccgg caggaggtgg tccaggccac 2160

gctggtggtc ctggccttgg ccatggtgcc catcctgctg cttggcacac ccctgcacct 2220

gctgcaccgc caccgccgcc gcctgcggag gaggccccct gaccgacagg aggaaaaaca 2280

ggccgggttg ctggacctgc ctgacgcac tgtgaatggc tggagctccg atgaggaaaa 2340

ggcagggggc ctggatgatg aagaggaggc cgagctcgtc ccctccgagg tgctcatgca 2400

ccaggccate cacaccatcg agttctgect gggctgcgtc tccaacaccg cctcctacct 2460

gcgctgtgg gccctgagcc tggcccacgc ccagctgtcc gaggttctgt gggccatggt 2520

gatgcgcata ggctgggcc tgggccggga ggtgggcgtg gcggctgtgg tgctggtccc 2580

catctttgcc gcctttgccg tgatgaccgt ggctatcctg ctggtgatgg agggactctc 2640

agccttctg cagccctgc ggctgcaactg ggtggaattc cagaacaagt tctactcagg 2700

cacgggctac aagctgagtc ccttcacctt cgtgccaca gatgactagt aagtcgacgg 2760

atccccggg ctgcaggaat tcgagcatct taccgccatt tataccata tttgttctgt 2820

ttttcttgat ttgggtatac atttaaatgt taataaaaca aaatggtggg gcaatcattt 2880

acatttttag ggatatgtaa ttactagtgc aggtgtattg ccacaagaca aacatgttaa 2940

gaaactttcc cgtatttac gctctgttcc tgtaataca cctctggatt acaaaatttg 3000

tgaaagattg actgatattc ttaactatgt tgctcctttt acgctgtgtg gatatgctgc 3060

tttaatgcct ctgtatcatg ctattgcttc ccgtacggct ttcgttttct cctccttgta 3120

taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtccgtc aacgtggcgt 3180

ggtgtgctct gtgtttgctg acgcaacccc cactggctgg ggcatgcca ccacctgtea 3240

actcctttct gggactttcg ctttccccct cccgatgcc acggcagaac tcatgccgcg 3300

ctgccttgcc cgtgctgga caggggctag gttgctgggc actgataatt ccgtggtgtt 3360

gtcggggaag ctgacgtcct ttcc 3384

<210> 2

<211> 243

<212> DNA

<213> Homo sapiens

<400> 2

ggctccgggtg cccgtcagtg ggcagagcgc acatcgcca cagtccccga gaagtgggg 60

ggaggggtcg gcaattgaac cggcgcctag agaaggtggc gcggggtaaa ctgggaaagt 120

gatgtcgtgt actggtccg ccttttccg gagggtgggg gagaaccgta tataagtgca 180

glagtgcgcc tgaacgttct ttttcgaac gggtttgccg ccagaacaca ggtgtcgtga 240

cgc 243

<210> 3

<211> 2493

<212> DNA

<213> Homo sapiens

<400> 3

atgggctcca tgtttcggag cgaggagggtg gccctgggtcc agctctttct gcccacagcg 60

gctgcctaca cctcgtgag tcggctgggc gagctgggcc tcgtggagtt cagagacctc 120

aacgcctcgg tgagcgcctt ccagagacgc ttgtgtgttg atgttcggcg ctgtgaggag 180

ctggagaaga ccttcacctt cctgcaggag gaggtgcggc gggctgggct ggtcctgccc 240

ccgcaaaagg ggaggctgcc ggcaacccca ccccgggacc tgcctgcgat ccaggaggag 300

acggagcgcc tggcccagga gctgcgggat gtgcggggca accagcaggc cctgcgggcc 360

cagctgcacc agctgcagct ccacgccgcc gtgctacgcc agggccatga acctcagctg 420

gcagccgccc acacagatgg ggcctcagag aggacgcccc tgctccaggc ccccgggggg 480

ccgcaccagg acctgagggt caactttgtg gcagggtccg tggagccca caaggccct 540

gccttagagc gcctgctctg gagggcctgc agaggcttcc tcattgccag cttcaggag 600

ctggagcagc cgctggagca ccccgtagc ggcgagccag ccacgtggat gaccttctc 660

atctcctact ggggtgagca gatcgacag aagatccga agatcacgga ctgcttcac 720

tgccacgtct tcccgtttct gcagcaggag gagccccgcc tcggggccct gcagcagctg 780

caacagcaga gccaggagct gcaggaggtc ctcggggaga cagagcggtt cctgagccag 840

gtgctaggcc ggggtgctgca gctgctgccg ccagggcagg tgcaggcca caagatgaag 900

gccgtgtacc tggccctgaa ccagtgcagc gtgagcacca cgcacaagtg cctcattgcc 960

gaggcctggt gctctgtgcg agacctgcc gcctgcagg aggcctgcg ggacagctcg 1020

atggaggagg gaggtagtgc cgtggctcac cgcateccct gccgggacat gccccca 1080

ctcatccga ccaaccgctt cacggccagc ttccaggga tcgtggatgc ctacggcgtg 1140

ggccgctacc aggaggtaa ccccgctccc tacacatca tcacctccc cttcctgttt 1200

gctgtgatgt tcgggatgt gggccacggg ctgctcatgt tctcttcgc cctggccatg 1260

gtccttgagg agaaccgacc ggctgtgaag gccgcgaga acgagatctg gcagacttc 1320

ttcagggcc gctacctgct cctgcttatg ggctgttct ccatctaac cggttcate 1380

tacaacgagt gcttcagtcg cgccaccagc atcttccct cgggctggag tgtggccgc 1440

atggccaacc agtctggctg gaggatgca ttctggccc agcacacgat gcttacctg 1500

gacccaacg tcaccgggtg cttcctggga ccctaccct ttggcatcga tcctatttgg 1560  
 agcctggctg ccaaccactt gagcttctc aactcctca agatgaagat gtccgtcatc 1620  
 ctgggcgtcg tgcacatggc ctttgggggtg gtccctcggag tcttcaacca cgtgcacttt 1680

ggccagaggc accggctgct gctggagacg ctgccggagc tcaccttct gctgggactc 1740  
 ttcggttacc tcgtgttctt agtcatctac aagtggctgt gtgtctgggc tgccagggcc 1800  
 gcctcggccc ccagcatcct catecacttc atcaacatgt tctcttctc ccacagcccc 1860  
 agcaacaggc tgctctacc cggcaggag gtggtccagg ccacgtggt ggtcctggcc 1920  
 ttggccatgg tgcccatcct gctgcttggc acaccctgc acctgctgca ccgccaccgc 1980  
 cgccgctgc ggaggaggcc cgctgaccga caggaggaaa acaaggccgg gttgctggac 2040  
 ctgcctgacg catctgtgaa tggctggagc tccgatgagg aaaaggcagg gggcctggat 2100

gatgaagagg aggccgagct cgteccctcc gaggtgetca tgcaccaggc catccacacc 2160  
 atcgagttct gcctgggctg cgtctccaac accgcctcct acctgcgcct gtgggccttg 2220  
 agcctggccc acgccagct gtccgaggtt ctgtgggcca tggatgatgc cataggcctg 2280  
 ggctgggccc gggagggtgg cgtagcggtt gtggtgetgg tcccactctt tgcccctttt 2340  
 gccgtgatga ccgtggctat cctgctggtg atggaggac tctcagcctt cctgcacgcc 2400  
 ctgcgctgc actgggtgga attccagaac aagtctact caggcacggg ctacaagctg 2460  
 agtccctca ccttcgctgc cacagatgac tag 2493

<210> 4

<211> 606

<212> DNA

<213> Woodchuck hepatitis virus

<400> 4

attcgagcat ctiaccgcca tttatacca ttttgttct gttttcttg atttgggtat 60  
 acatttaaat gtaataaaa caaaatggtg gggcaatcat ttacattttt agggatatgt 120  
 aattactagt tcagggtgat tgccacaaga caaacatggt aagaaacttt cccgttattt 180  
 acgctctggt ccigttaatc aacctctgga ttacaaaatt tgtgaaagat tgactgatat 240  
 tcttaactat gttgtcctt ttacgtctg tggatatgct gctttaatgc ctctgtatca 300  
 tgctattgct tcccgtagg ctttcgtttt ctctccttg tataaatcct ggttgetgtc 360  
 tctttatgag gatttggc ccgttgtccg tcaactggc gtggtgtgct ctgtgttgc 420  
 tgacgaace cccactggct ggggcattgc caccactgt caactcctt ctgggacttt 480  
 cgctttccc ctccgatcg ccacggcaga actcatgcc gcctgccttg cccgctgctg 540

gacaggggct aggttgctgg gcactgataa ttccgtggtg ttgtcgggga agctgacgtc 600  
ctttcg 606

<210> 5  
<211> 830  
<212> PRT  
<213> Homo sapiens  
<400> 5

Met Gly Ser Met Phe Arg Ser Glu Glu Val Ala Leu Val Gln Leu Phe

1 5 10 15  
Leu Pro Thr Ala Ala Ala Tyr Thr Cys Val Ser Arg Leu Gly Glu Leu  
20 25 30  
Gly Leu Val Glu Phe Arg Asp Leu Asn Ala Ser Val Ser Ala Phe Gln  
35 40 45  
Arg Arg Phe Val Val Asp Val Arg Arg Cys Glu Glu Leu Glu Lys Thr  
50 55 60  
Phe Thr Phe Leu Gln Glu Glu Val Arg Arg Ala Gly Leu Val Leu Pro

65 70 75 80  
Pro Pro Lys Gly Arg Leu Pro Ala Pro Pro Pro Arg Asp Leu Leu Arg  
85 90 95  
Ile Gln Glu Glu Thr Glu Arg Leu Ala Gln Glu Leu Arg Asp Val Arg  
100 105 110  
Gly Asn Gln Gln Ala Leu Arg Ala Gln Leu His Gln Leu Gln Leu His  
115 120 125  
Ala Ala Val Leu Arg Gln Gly His Glu Pro Gln Leu Ala Ala Ala His

130 135 140  
Thr Asp Gly Ala Ser Glu Arg Thr Pro Leu Leu Gln Ala Pro Gly Gly  
145 150 155 160  
Pro His Gln Asp Leu Arg Val Asn Phe Val Ala Gly Ala Val Glu Pro  
165 170 175  
His Lys Ala Pro Ala Leu Glu Arg Leu Leu Trp Arg Ala Cys Arg Gly  
180 185 190

Phe Leu Ile Ala Ser Phe Arg Glu Leu Glu Gln Pro Leu Glu His Pro  
 195 200 205  
 Val Thr Gly Glu Pro Ala Thr Trp Met Thr Phe Leu Ile Ser Tyr Trp  
 210 215 220  
 Gly Glu Gln Ile Gly Gln Lys Ile Arg Lys Ile Thr Asp Cys Phe His  
 225 230 235 240  
 Cys His Val Phe Pro Phe Leu Gln Gln Glu Glu Ala Arg Leu Gly Ala  
 245 250 255  
 Leu Gln Gln Leu Gln Gln Gln Ser Gln Glu Leu Gln Glu Val Leu Gly  
 260 265 270  
 Glu Thr Glu Arg Phe Leu Ser Gln Val Leu Gly Arg Val Leu Gln Leu  
 275 280 285  
 Leu Pro Pro Gly Gln Val Gln Val His Lys Met Lys Ala Val Tyr Leu  
 290 295 300  
 Ala Leu Asn Gln Cys Ser Val Ser Thr Thr His Lys Cys Leu Ile Ala  
 305 310 315 320  
 Glu Ala Trp Cys Ser Val Arg Asp Leu Pro Ala Leu Gln Glu Ala Leu  
 325 330 335  
 Arg Asp Ser Ser Met Glu Glu Gly Val Ser Ala Val Ala His Arg Ile  
 340 345 350  
 Pro Cys Arg Asp Met Pro Pro Thr Leu Ile Arg Thr Asn Arg Phe Thr  
 355 360 365  
 Ala Ser Phe Gln Gly Ile Val Asp Ala Tyr Gly Val Gly Arg Tyr Gln  
 370 375 380  
 Glu Val Asn Pro Ala Pro Tyr Thr Ile Ile Thr Phe Pro Phe Leu Phe  
 385 390 395 400  
 Ala Val Met Phe Gly Asp Val Gly His Gly Leu Leu Met Phe Leu Phe  
 405 410 415  
 Ala Leu Ala Met Val Leu Ala Glu Asn Arg Pro Ala Val Lys Ala Ala  
 420 425 430  
 Gln Asn Glu Ile Trp Gln Thr Phe Phe Arg Gly Arg Tyr Leu Leu Leu



Trp Ser Ser Asp Glu Glu Lys Ala Gly Gly Leu Asp Asp Glu Glu Glu  
 690 695 700  
 Ala Glu Leu Val Pro Ser Glu Val Leu Met His Gln Ala Ile His Thr  
 705 710 715 720  
 Ile Glu Phe Cys Leu Gly Cys Val Ser Asn Thr Ala Ser Tyr Leu Arg  
 725 730 735  
 Leu Trp Ala Leu Ser Leu Ala His Ala Gln Leu Ser Glu Val Leu Trp  
 740 745 750  
 Ala Met Val Met Arg Ile Gly Leu Gly Leu Gly Arg Glu Val Gly Val  
 755 760 765  
 Ala Ala Val Val Leu Val Pro Ile Phe Ala Ala Phe Ala Val Met Thr

770 775 780  
 Val Ala Ile Leu Leu Val Met Glu Gly Leu Ser Ala Phe Leu His Ala  
 785 790 795 800  
 Leu Arg Leu His Trp Val Glu Phe Gln Asn Lys Phe Tyr Ser Gly Thr  
 805 810 815  
 Gly Tyr Lys Leu Ser Pro Phe Thr Phe Ala Ala Thr Asp Asp  
 820 825 830

<210> 6

<211> 13

<212> DNA

<213> Artificial Sequence

<220><223> Consensus Kozak sequence

<400> 6

gccgccacca tgg

13

<210> 7

<211> 387

<212> DNA

<213> Mus musculus

<400> 7

tggctaataa aggaaattta ttttcattgc aatagtgtgt tggaaatntt tgtgtctctc

60

actcgaagg acatatggga gggcaaatca tttaaaacat cagaatgagt atttggttta

120

gagtttggca acatatgcc atagtctggc tgccatgaac aaagttggc tataaagagg 180  
 tcatcagtat atgaaacagc cccctgctgt ccattcctta ttccatagaa aagccttgac 240  
 ttgaggttag atttttttta tttttggtt tgtgttattt ttttctttaa catccctaaa 300

attdtctta catgttttac tagccagatt tttctctctc tctgactac tcccagtcac 360  
 agctgtccct cttctcttat ggagatc 387

<210> 8  
 <400> 8  
 000

<210> 9  
 <400> 9  
 000

<210> 10  
 <400> 10  
 000

<210> 11  
 <400> 11  
 000

<210> 12  
 <400> 12  
 000

<210> 13  
 <400> 13  
 000

<210> 14  
 <211> 13  
 <212> RNA

<213> Artificial Sequence

<220><223> Consensus Kozak sequence

<400> 14  
 gccgcrcca ugg 13

<210> 15  
 <211> 8  
 <212> RNA

<213> Artificial Sequence  
 <220><223> Kozak sequence  
 <220><221> misc\_feature  
 <222> (2)..(3)  
 <223> n is A, C, T, G or U  
 <220><221> misc\_feature  
 <222> (4)..(4)  
 <223> n is a, c, g, or u  
 <220><221> misc\_feature  
 <222> (8)..(8)  
 <223> n is A, C, T, G or U  
 <400> 15

agnnaugn 8

<210> 16  
 <211> 7  
 <212> RNA

<213> Artificial Sequence  
 <220><223> Kozak sequence  
 <220><221> misc\_feature

<222> (2)..(3)  
 <223> n is A, C, T, G or U  
 <400> 16

annaugg 7

<210> 17  
 <211> 7  
 <212> RNA

<213> Artificial Sequence  
 <220><223> Kozak sequence  
 <400> 17

accaugg 7

<210> 18  
 <211> 10  
 <212> RNA

<213> Artificial Sequence

<220><223> n is A, C, T, G or U

<400> 18

gacaccaugg 10

<210

> 19

<211> 235

<212> DNA

<213> Bos taurus

<400> 19

tcgactgtgc ctctagttg ccagecatct gttgtttgcc cctccccctg gccttccttg 60

accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcacgcac 120

tgtctgagta ggtgtcattc tattctgggg ggtgggtgg ggcaggacag caagggggag 180

gattgggagg acaatagcag gcatgctggg gatgcggtgg gctctatggc ttctg 235

<210> 20

<211> 222

<212> DNA

<213> simian virus 40

<400> 20

cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa 60

aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca 120

ataaacaagt taacaacaac aattgcattc attttatggt tcaggttcag ggggagatgt 180

gggaggtttt ttaaagcaag taaaacctct acaaatgtgg ta 222

<210> 21

<211> 202

<212> DNA

<213> Homo sapiens

<400> 21

ctgccccggt ggcacccctg tgaccctcc ccagtgcctc tcctggccct ggaagtggc 60

actccagtgc ccaccagcct tgtcctaata aaattaagtt gcatcatttt gctgactag 120

gtgtccttct ataataattat ggggtggagg ggggtggtat ggagcaaggg gcccaagttg 180

ggaagaaacc tgtaggcct gc 202

<210> 22

<211> 141

<212> DNA  
 <213> Escherichia coli  
 <400> 22  
 gtagaattgg taaagagagt cgtgtaaaat atcgagttcg cacatcttgt tgtctgatta 60  
 ttgatttttg gcgaaacat ttgatcatal gacaagatgt gtatctacct taacttaatg 120  
 attttgataa aatcattag g 141  
 <210> 23  
 <211> 7660  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Made in Lab - plasmid construct  
 <400> 23  
 acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc 60  
  
 atatatggag ttccgcgta cataacttac ggtaaatggc ccgcctggct gaccgccc 120  
 cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac 180  
 tttccattga cgtcaatggg tggagtattt acggtaaact gccacttgg cagtacatca 240  
 agtgtatcat atgccaagta cgccccctat tgacgtcaat gacggtaaat ggccccctg 300  
 gcattatgcc cagtacatga cttatggga ctttcctact tggcagtaca tctacgtatt 360  
 agtcatcgct attaccatgg tgatgcgggt ttggcagtac atcaatgggc gtggatagcg 420  
 gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga gtttgttttg 480  
  
 gcacaaaaat caacgggact ttcaaaaatg tcgtaacaac tccgccccat tgacgcaaat 540  
 gggcggtagg cgtgtacggt gggaggctta tataagcaga gctcgtttag tgaaccgggg 600  
 tctctctggt tagaccagat ctgagcctgg gagctctctg gctaactagg gaaccactg 660  
 cttaacctc aataaagctt gccttgagtg cttcaagtag tgtgtgcccg tctgttgtgt 720  
 gactctggta actagagatc cctcagacc ttttagtcag tgtggaaaat ctctagcagt 780  
 ggccccgaa cagggacttg aaagcgaag ggaaccaga ggagctctct cgacgcagga 840  
 ctcggttgc tgaagcgcgc acggcaagag gcgaggggcg gcgactggtg agtacgcaa 900  
  
 aaattttgac tagcggaggc tagaaggaga gagatgggtg cgagagcgtc agtattaagc 960  
 gggggagaat tagatcgca tgggaaaaa ttcggttaag gccaggggga aagaaaaat 1020  
 ataaatata acatatagta tgggcaagca gggagctaga acgattcgca gttaatcctg 1080  
 gcctgtaga aacatcagaa ggctgtagac aaatactggg acagctaca ccatccttc 1140  
 agacaggatc agaagaactt agatcattat ataatacagt agcaaccctc tattgtgtgc 1200

atcaaaggat agagataaaa gacaccaagg aagctttaga caagatagag gaagagcaaa 1260  
 acaaaagtaa gaccaccgca cagcaagcgg ccgctgatct tcagacctgg aggaggagat 1320  
  
 atgagggaca attggagaag tgaattatat aaatataaag tagtaaaaat tgaaccatta 1380  
 ggagtagcac ccaccaaggc aaagagaaga gtggtgcaga gagaanaaag agcagtggga 1440  
 ataggagctt tgttcttgg gttcttggga gcagcaggaa gcactatggg cgcagcgtca 1500  
 atgacgctga cggtagcggc cagacaatta ttgtctggta tagtgcagca gcagaacaat 1560  
 ttgctgaggg ctattgaggc gcaacagcat ctgttgcaac tcacagtctg gggcatcaag 1620  
 cagctccagg caagaatcct ggctgtggaa agatacctaa aggatcaaca gctcctgggg 1680  
 atttggggtt gccttgaaa actcatttgc accactgctg tgccttgaa tgctagtgg 1740  
  
 agtaataaat ctctggaaca gatttgaat cacacgacct ggatggagtg ggacagagaa 1800  
 attaacaatt acacaagctt aatacactcc ttaattgaag aatcgcaaaa ccagcaagaa 1860  
 aagaatgaac aagaattatt ggaattagat aaatgggcaa gtttgtggaa ttggttaac 1920  
 ataacaatt ggctgtggta tataaaatta tcataatga tagtaggagg cttagtaggt 1980  
 ttaagaatag ttttctgt actttctata gtgaatagag ttaggcaggg atattcacca 2040  
 ttatcgttcc agaccacct cccaaccccg aggggacccg acaggcccgagg agaatagaa 2100  
 gaagaaggtg gagagagaga cagagacaga tccattcgat tagtgaacgg atctcgacgg 2160  
  
 tatcggttaa cttttaaag aaaagggggg attggggggt acagtgcagg ggaaagaata 2220  
 gtagacataa tagcaacaga catacaact aaagaattac aaaaacaaat tacaaaaatt 2280  
 caaaatitta tcgatcacga gactagcctc gagaagcttg atcgattggc tccggtgcc 2340  
 gtcagtgggc agagcgaca tcgccacag tccccagaa gttgggggga ggggtcggca 2400  
 attgaaccgg tccttagaga aggtggcgcg gggtaaactg ggaaagtgat gtcgtgtact 2460  
 ggctccgct ttttccgag ggtgggggag aaccgtatat aagtgcagta gtcgccgtga 2520  
 acgttctttt tcgcaacggg tttgcccga gaacacaggt gtcgtgacgc gggatccgcc 2580  
  
 accatgggct ccatgttctg gagcaggag gtggccctgg tccagctctt tctgcccaca 2640  
 gcggctgcct acacctgctg gagtcggctg ggcgagctgg gcctcgtgga gttcagagac 2700  
 ctcaacgct cggtagcgc cttccagaga cgcttctgg ttgatgttcg gcgctgtgag 2760  
 gagctggaga agacctcac cttctgcag gaggaggtgc ggcgggctgg gctggtcctg 2820  
 ccccccaa aggggaggt gccgcacc ccacccggg acctgctgcg catccaggag 2880  
 gagacggagc gcctggccca ggagctgcgg gatgtcggg gcaaccagca ggcctgcgg 2940

gcccagctgc accagctgca gctccacgcc gccgtgctac gccagggcca tgaacctcag 3000

ctggcagccg cccacacaga tggggcctca gagaggacgc ccctgctcca ggccccggg 3060

gggccgcacc aggacctgag ggtcaacttt gtggcaggtg ccgtggagcc ccacaaggcc 3120

cctgccttag agcgcctgct ctggagggcc tgcagaggct tcctcattgc cagcttcagg 3180

gagctggagc agccgctgga gcaccccgctg acgggcgagc cagccacgtg gatgaccttc 3240

ctcatctcct actgggggtga gcagatcgga cagaagatcc gcaagatcac ggactgcttc 3300

cactgccacg tcttcccgtt tctgcagcag gaggaggccc gcctcggggc cctgcagcag 3360

ctgcaacagc agagccagga gctgcaggag gtctctgggg agacagagcg gttcttgagc 3420

caggtgctag gccgggtgct gcagctgctg ccccagggc aggtgcaggt ccacaagatg 3480

aaggccgtgt acctggccct gaaccagtgc agcgtgagca ccacgcacaa gtgcctcatt 3540

gccgaggcct ggtgctctgt gcgagacctg cccgccctgc aggaggccct gcgggacagc 3600

tcgatggagg agggagtgag tgccgtggct caccgcatcc cctgccggga catgcccccc 3660

acactcatcc gcaccaaccg cttcacggcc agcttcagg gcatcgtgga tgectacggc 3720

gtgggccgct accaggaggt caaccccgt ccttacacca tcatcacctt ccccttctg 3780

tttgctgtga tgttcgggga tgtgggccac gggctgctca tgttctctt cgccttggcc 3840

atggtccttg cggagaaccg accggctgtg aaggccgcgc agaacgagat ctggcagact 3900

ttcttcaggg gccgctacct gctcctgctt atgggcctgt tctccatcta caccggcttc 3960

atctacaacg agtgcttcag tcgcgccacc agcatcttcc cctcgggctg gagtgtggcc 4020

gccatggcca accagtctgg ctggagtgat gcattcctgg cccagcacac gatgcttacc 4080

ctggacceca acgtcaccgg tgtcttcttg ggaccctacc cctttggcat cgatectatt 4140

tggagcctgg ctgccaacca cttgagcttc ctcaactcct tcaagatgaa gatgtccgtc 4200

atcctgggcg tcgtgcatat ggcctttggg gtggtcctcg gagtcttcaa ccacgtgcac 4260

tttggccaga ggcaccggct gctgctggag acgctgccgg agctcacctt cctgctggga 4320

ctcttcgggt acctcgtgtt cctagtcate tacaagtggc tgtgtgtctg ggctgccagg 4380

gccgcctcgg cccccagcat cctcatccac ttcatacaaa tgttctctt ctcccacagc 4440

cccagcaaca ggtcgtctta cccccggcag gaggtgttcc aggccacgct ggtggtcctg 4500

gccttggcca tggtgcccat cctgctgctt ggcacacccc tgcacctgct gcaccgccac 4560

cgcccgccg cgcggaggag gcccgtgac cgacaggagg aaaacaaggc cgggttgctg 4620

gacctgctg acgcatctgt gaatggctgg agctccgatg aggaaaaggc agggggcctg 4680

gatgatgaag aggaggccga gctcgtcccc tccgaggtgc tcatgcacca ggccatccac 4740  
accatcgagt tctgcctggg ctgcgtctcc aacaccgct cctacctgcg cctgtgggcc 4800  
ctgagcctgg cccacgcccc gctgtccgag gttctgtggg ccatggtgat gcgcataggc 4860  
ctgggcctgg gccgggaggt gggcgtggcg gctgtggtgc tggccccat ctttgccgcc 4920  
tttgccgtga tgaccgtggc tatectgtg gtgatggagg gactctcagc cttcctgcac 4980  
gccctcgggc tgactgggt ggaattccag aacaagtct actcaggcac gggctacaag 5040  
ctgagtcctt tcaccttcgc tgccacagat gactagtaag tcgacggatc ccccgggctg 5100  
  
caggaattcg agcatcttac cgccatttat acccatatth gttctgtttt tcttgatttg 5160  
ggtatacatt taaatgttaa taaaacaaaa tggtagggca atcatttaca ttttaggga 5220  
tatgtaatta ctagtccagg tgtattgcca caagacaaac atgttaagaa actttcccgt 5280  
tatttacgt ctgttctgt taatcaacct ctggattaca aaattgtga aagattgact 5340  
gatattctta aciatgttgc tccttttacg ctgtgtggat atgctgcttt aatgcctctg 5400  
tatcatgcta ttgcttcccg tacggctttc gttttctct ccttgataa acctggttg 5460  
ctgtctcttt atgaggagt gtggcccgtt gtcctcaac gtggcgtggt gtgctctgtg 5520  
  
tttgctgacg caacccccac tggctggggc attgccacca cctgtcaact cttttctggg 5580  
actttcgctt tccccctccc gatcggcag gcagaactca tcgccgctg ctttggccgc 5640  
tgctggacag ggcctaggtt gctgggcact gataattccg tgggtttgtc ggggaagctg 5700  
acgtcctttc gaattcgata tcaagctgta cctttaagac caatgactta caaggcagct 5760  
gtagatctta gccacttttt aaaagaaaag gggggactgg aagggtaat tcaactccaa 5820  
cgaagacaag atctgctttt tgcttgact ggtctctct ggtagacca gatctgagcc 5880  
tgggagctct ctggctaact agggaaacca ctgcttaagc ctcaataaag cttgccttga 5940  
  
gtgcttcaag tagtgtgtgc ccgtctgttg tgtgactctg gtaactagag atccctcaga 6000  
cccttttagt cagtgtggaa aatctctagc agtagtagt catgtcatct tattattcag 6060  
tatttataac ttgcaaagaa atgaatatca gagagtgaga ggaacttgtt tattgcagct 6120  
tataatggtt acaataaag caatagcatc acaatttca caaataaagc attttttca 6180  
ctgacttcta gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctggctctag 6240  
ctatcccgc ctaactccg cccatcccgc ccctaactcc gccagttcc gccattctc 6300  
cgccccatgg ctgactaatt tttttattt atgcagaggc cgaggccgc tcggcctctg 6360  
  
agctattcca gaagtagiga ggagctttt ttggaggcct aggtagccc cctaatgagc 6420  
gggctttttt ttcttaggcc ttcttccgct tctcctca ctgactcgt gcgctcggtc 6480  
gttcggctgc ggcgagcgt atcagctcac tcaaaggcgg taatcgggt atccacagaa 6540

tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt 6600  
 aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa 6660  
 aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt 6720  
 ccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg 6780

tccgcctttc tccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc 6840  
 agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc 6900  
 gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggttaag acacgactta 6960  
 tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatgt aggcggtgct 7020  
 acagagtctt tgaagtggtg gcctaactac ggctacacta gaagaacagt atttggatc 7080  
 tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa 7140  
 caaacaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa 7200

aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa 7260  
 aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt 7320  
 ttaaattaaa aatgaagtgt taaatcaatc taaagtatat atgagtaaac ttggtctgac 7380  
 agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcaccc 7440  
 atagttgcct gactcctgca aaccacgttg tggtagaatt ggtaagaga gtcgtgtaaa 7500  
 atatcgagtt cgcacatctt gttgtctgat tattgatctt tggcgaaacc atttgcacat 7560  
 atgacaagat gtgtatctac cttacttaa tgattttgat aaaaatcatt aggtacctgt 7620

acatttatat tggtcatgt ccaacattac cgccatgttg 7660

<210> 24  
 <211> 511  
 <212> DNA  
 <213> Homo sapiens  
 <400> 24

ggggttgggg ttgcgccttt tccaagcag ccttgggttt gcgcagggac gcggtgctc 60  
 tggcggtggt tccgggaaac gcagcggcgc cgaccctggg tctcgacat tcttcacgtc 120  
 cgttcgcage gtcaccgga tcttcgccgc taccttgtg ggcccccg cgacgttcc 180  
 tgctccgcc ctaagtcggg aaggttcctt gcggttcgcg gcgtgccgga cgtgacaaac 240  
 ggaagccgca cgtctacta gtaccctcgc agacggacag cgccaggag caatggcagc 300

gcgccgaccg cgatgggctg tggccaatag cggctgctca gcagggcgcg ccgagagcag 360  
 cggccgggaa gggcggtg gcggagcggg gtgtggggcg gtagtgtggg ccctgttctc 420

gccccgcgg tgttccgat tctgcaagcc tccggagcgc acgtcggcag tccgctccct 480  
 cgttgaccga atcaccgacc tctctccca g 511  
 <210> 25  
 <211> 7613  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Made in Lab - plasmid construct  
 <400> 25  
 acattgatta ttgactagt attaatagta atcaattacg gggtcattag ttcatagccc 60  
 atatatggag ttccgcgta cataacttac ggtaaatggc cgcctggct gaccgcccaa 120  
  
 cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac 180  
 ttccattga cgicaatggg tggagtattt acggtaaact gccacttgg cagtacatca 240  
 agtgtatcat atccaagta cccccctat tgacgtcaat gacggtaaat gccccgctg 300  
 gcattatgcc cagtacatga cttatggga ctttctact tggcagtaca tctacgtatt 360  
 agtcatcgct attaccatgg tgatcgggtt ttggcagtac atcaatgggc gtggatagcg 420  
 gtttgactca cggggatttc caagtctca cccattgac gtcaatggga gtttgitttg 480  
 gcacaaaaat caacgggact ttcaaaaatg tcgtaacaac tccgccccat tgacgcaaat 540  
  
 gggcggtagg cgtgtacgt gggaggtcta tataagcaga gctcgtttag tgaaccgggg 600  
 tctctctggt tagaccagat ctgagcctgg gagctctctg gctaactagg gaaccactg 660  
 cttaacctc aataaagctt gccttgagtg cttcaagtag tgtgtgcccg tctgttgtgt 720  
 gactctgcta actagagatc cctcagacc ttttagtcag tgtggaaaat ctctagcagt 780  
 ggcgccccga cagggacttg aaagcgaag ggaaccaga ggagctctct cgacgcagga 840  
 ctcggcttgc tgaagcgcgc acggcaagag gcgaggggcg gcgactggtg agtacgcaa 900  
 aaatcttgac tagcggaggc tagaaggaga gagatgggtg cgagagcgtc agtattaagc 960  
  
 gggggagaat tagatcgcga tgggaaaaaa ttcggttaag gccaggggga aagaaaaat 1020  
 ataaattaa acatatagta tgggcaagca gggagctaga acgattcgca gtaaatcctg 1080  
 gcctgttaga aacatcagaa ggctgtagac aaatactggg acagctaca ccatcccttc 1140  
 agacaggatc agaagaactt agatcattat ataatacagt agcaaccctc tattgtgtgc 1200  
 atcaaaggat agagataaaa gacaccaagg aagctttaga caagatagag gaagagcaaa 1260  
 acaaaagtaa gaccaccga cagcaagcgg ccgctgatct tcagacctgg aggaggagat 1320  
 atgagggaca attggagaag tgaattatat aatatataag tagtaaaaat tgaaccatta 1380

ggagttagcac ccaccaaggc aaagagaaga gtggtgcaga gagaaaaaag agcagtggga 1440  
 ataggagctt tgttccttgg gttcttggga gcagcaggaa gcactatggg cgcagcgtca 1500  
 atgacgctga cggtagcaggc cagacaatta ttgtctggta tagtgcagca gcagaacaat 1560  
 ttgctgaggg ctattgaggg gcaacagcat ctgttgcaac tcacagtctg gggcatcaag 1620  
 cagctccagg caagaatcct ggctgtggaa agatacctaa aggatcaaca gctcctgggg 1680  
 atttggggtt gctctggaaa actcatttgc accactgctg tgccctggaa tgctagttag 1740  
 agtaataaat ctctggaaca gatttggaaat cacacgacct ggatggagtg ggacagagaa 1800  
  
 attaacaatt acacaagctt aatacactcc ttaattgaag aatcgcaaaa ccagcaagaa 1860  
 aagaatgaac aagaattatt ggaattagat aaatgggcaa gtttgtggaa ttggtttaac 1920  
 ataacaatt ggctgtggta tataaaatta ttcataatga tagtaggagg cttggtaggt 1980  
 ttaagaatag tttttgctgt actttctata gtgaatagag ttaggcaggg atattcacca 2040  
 ttatcgtttc agaccacct cccaaccccg aggggacccg acaggcccgga aggaatagaa 2100  
 gaagaaggtg gagagagaga cagagacaga tccattcgat tagtgaacgg atctcgacgg 2160  
 tatcggttaa cttttaaaag aaaagggggg attggggggt acagtgcagg ggaaagaata 2220  
  
 gtagacataa tagcaacaga catacaact aaagaattac aaaaacaaat tacaaaaatt 2280  
 caaaatttta tccgtcagtg ggcagagcgc acatcgccca cagtccccga gaagtgggg 2340  
 ggaggggtcg gcaattgaac cgggtcctag agaaggtggc gcggggtaaa ctgggaaagt 2400  
 gatgtcgtgt aciggtccg ctttttccc gaggggtggg gagaaccgta tataagtca 2460  
 gtagtcgccc tgaacgttct ttttcgcaac gggtttgccg ccagaacaca ggtgtcgtga 2520  
 cgcggtatcc gccaccatgg gctccatggt tcggagcgag gaggtggccc tgggtccagct 2580  
 ctttctgccc acagcggctg cctacacctg cgtgagtcgg ctgggcgagc tgggcctcgt 2640  
  
 ggagttcaga gacctcaacg cctcggtgag cgccttcag agacgctttg tggttgatgt 2700  
 tcggcgtgt gaggagctgg agaagacctt caccttctg caggaggagg tgcggcgggc 2760  
 tgggctggte ctgccccgc caaaggggag gctgcccgca cccccaccc gggacctgct 2820  
 gcgcatccag gaggagacgg agcgcctggc ccaggagctg cgggatgtgc ggggcaacca 2880  
 gcaggccctg cgggcccagc tgcaccagct gcagctccac gccgccgtgc tacgccaggg 2940  
 ccatgaacct cagctggcag cgcgccacac agatggggcc tcagagagga cgccctgct 3000  
 ccaggccccc ggggggcccgc accaggacct gaggtcaac tttgtggcag gtgccgtgga 3060  
  
 gccccacaag gccctgccc tagagcgcct gctctggagg gcctgcagag gcttctcat 3120  
 tgccagcttc agggagctgg agcagccgct ggagcacccc gtgacgggag agccagccac 3180  
 gtggatgacc ttctcatct cctactgggg tgagcagatc ggacagaaga tccgcaagat 3240

cacggactgc ttccactgcc acgtcttccc gtttctgcag caggaggagg cccgcctcgg 3300  
 ggccctgcag cagctgcaac agcagagcca ggagctgcag gaggtcctcg gggagacaga 3360  
 gcggttctctg agccaggtgc taggccgggt gctgcagctg ctgccccag ggcaggtgca 3420  
 ggtccacaag atgaaggccg tgtacctggc cctgaaccag tgcagcgtga gcaccacgca 3480  
  
 caagtgcctc attgccgagg cctggtgctc tgtgcgagac ctgcccgccc tgcaggaggc 3540  
 cctgcgggac agctcgatgg aggagggagt gagtgccctg gctcaccgca tcccctgccg 3600  
 ggacatgccc cccacactca tccgaccaa ccgcttcacg gccagcttcc agggcatcgt 3660  
 ggatgcctac ggcgtgggcc gctaccagga ggtcaacccc gctccctaca ccatcatcac 3720  
 cttecccttc ctgtttctg tgatgttcgg ggatgtgggc cacgggctgc tcatgttctt 3780  
 ctteccctg gccatggtcc ttgcggagaa ccgaccggct gtgaaggccg gcgagaacga 3840  
 gatctggcag actttcttca ggggccgcta cctgctcctg cttatgggcc tgttctccat 3900  
  
 ctacaccggc ttcatctaca acgagtgtt cagtcgcgcc accagcatct tcccctcggg 3960  
 ctggagtgtg gccgccatgg ccaaccagtc tgctggagt gatgcattcc tggcccagca 4020  
 cacgatgctt acctggacc ccaacgacac cgggtgtctt ctgggacct accccttgg 4080  
 catcgatcct atttggacc tggctgcaa ccacttgagc ttcctcaact cttcaagat 4140  
 gaagatgtcc gtcactctgg gcgtcgtgca catggccttt ggggtggtcc tggagtctt 4200  
 caaccacgtg cactttggcc agaggcaccg gctgctgctg gagacgctgc cggagctcac 4260  
 cttectgctg ggactcttcg gttacctcgt gttcctagtc atctacaagt ggctgtgtgt 4320  
  
 ctgggctgcc agggccgctt cggcccccag catctcacc cacttcatca acatgttctt 4380  
 cttctccac agccccagca acaggctgct ctacccccg caggaggtgg tccaggccac 4440  
 gctggtggtc ctggccttgg ccatggtgcc catcctgctg cttggcacac ccctgcacct 4500  
 gctgcaccgc caccgccgcc gcctcgggag gaggccccct gaccgacagg aggaaaaca 4560  
 ggccgggttg ctggacctgc ctgacgcac tgtgaatggc tggagctccg atgaggaaaa 4620  
 ggcagggggc ctggatgatg aagaggaggc cgagctcgtc ccctccgagg tgctcatgca 4680  
 ccaggccate cacaccatcg agttctgctt gggctgcgtc tccaacaccg cctcctacct 4740  
  
 gcgcctgtgg gcctgagcc tggcccacgc ccagctgtcc gaggttctgt gggccatggt 4800  
 gatgcgcata ggctgggcc tggccggga ggtgggcgtg gcggtctggt tgctggtccc 4860  
 catctttgcc gcctttgccg tgatgacctg ggctatcctg ctggtgatgg agggactctc 4920  
 agccttctg cacgccctgc ggctgcaact ggtggaattc cagaacaagt tctactcagg 4980  
 cacgggctac aagctgagtc ccttcacctt cgtgccaca gatgactagt aagtcgacgg 5040  
 atccccggg ctgcaggaat tcgagcatct taccgccatt tataccata tttgttctgt 5100

ttttcttgat ttgggtatac attttaatgt taataaaaca aaatgggtggg gcaatcattt 5160  
  
 acatttttag ggatatgtaa ttactagtgc aggtgtattg ccacaagaca aacatgtaa 5220  
 gaaactttcc cgttatttac gctctgttcc tgtaatacaa cctctggatt acaaaatttg 5280  
 taaaagattg acigatattc ttaactatgt tgctcctttt acgctgtgtg gatatgctgc 5340  
 tttaatgcct ctgtatcatg ctattgcttc ccgtacggct ttcgttttct cctccttgta 5400  
 taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtccgtc aacgtggcgt 5460  
 ggtgtgctct gtgtttgctg acgcaacccc cactggctgg ggcatgcca ccacctgtca 5520  
 actcctttct gggactttcg ctttcccct cccgatcgcc acggcagaac tcatcgccgc 5580  
  
 ctgccttgcc cgctgctgga caggggctag gttgctgggc actgataatt cegtgtgtt 5640  
 gtcggggaag ctgacgtcct ttcgaattcg atatcaagct gtaccttaa gaccaatgac 5700  
 ttacaaggca gctgtagatc ttagccactt tttaaaagaa aaggggggac tggagggtt 5760  
 aattcactcc caacgaagac aagatctgct ttttgcttgt actgggtctc tctggttaga 5820  
 ccagatctga gcctgggagc tctctggcta actagggaac ccactgctta agcctcaata 5880  
 aagcttgctt tgagtgttc aagtagtgtg tgcccgtctg ttgtgtgact ctggtacta 5940  
 gagatccctc agaccctttt agtcagtgtg gaaaatctct agcagtagta gttcatgtca 6000  
  
 tcttattatt cagtattat aacttgcaa gaaatgaata tcagagagtg agaggaactt 6060  
 gtttattgca gcttataatg gttacaaata aagcaatagc atcaciaatt tcaciaataa 6120  
 agcatttttt tcaactgact ctagtgtggg ttgtccaaa ctcatcaatg tatcttatca 6180  
 tgtctggctc tagctatccc gccctaaact ccgccatcc cggccctaac tccgccagct 6240  
 tccgccatt ctccgccca tggctgacta attttttta tttatgcaga ggccgaggcc 6300  
 gcctggcct ctgagctatt ccagaagtag tgaggaggct ttttggagg cctaggtagc 6360  
 ccgcctaatg agcgggcttt ttttcttag gccttcttc gcttctcgc tcaactgactc 6420  
  
 gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cgtaatacg 6480  
 gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 6540  
 gccaggaac cgiaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 6600  
 cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcacag gactataaag 6660  
 ataccaggcg tttcccctg gaagctccct cgtgcctct cctgttccga cctgccct 6720  
 taccggatac ctgtccgct tctcccttc gggaagcgtg gcgctttctc atagctcacg 6780  
 ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 6840

ccccgttcag cccgaccgct gcgccttata cggttaactat cgtcttgagt ccaaccgggt 6900  
 aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 6960  
 tgtaggcgggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 7020  
 agtatttgggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 7080  
 ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 7140  
 tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 7200  
 tcagtggaac gaaaactcac gttaagggat ttgggtcatg agattatcaa aaaggatctt 7260

cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 7320  
 aacttggctc gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 7380  
 atttcgttca tccatagttg cctgactcct gcaaccacg ttgtggtaga attggtaaag 7440  
 agagtcgtgt aaaaatcgca gttcgccat cttgttgtct gattattgat ttttggcgaa 7500  
 accatttgat catatgacaa gatgtgtatc taccttaact taatgatttt gataaaaatc 7560  
 attaggtacc tgtacattta tattggctca tgcacaacat taccgcatg ttg 7613

<210> 26

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> Made in Lab - section of plasmid construct

<400> 26

gatcacgaga ctacccctga gaagcttgat cgattggctc cggtgcc 47

<210> 27

<211> 7646

<212> DNA

<213> Artificial Sequence

<220><223> Made in lab - plasmid construct

<400> 27

ggctccgggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga gaagttgggg 60

ggaggggtcg gcaattgaac cggtccttag agaagggtggc gcggggtaaa ctgggaaagt 120

gatgtcgtgt actggctccg cctttttccc gaggggtgggg gagaaccgta tataagtgca 180

gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca ggtgtcgtga 240

cgcgggatcc gccacatgg gctccatggt tcggagcgag gaggtggccc tgggtccagct 300

ctttctgccc acagcggctg cctacacctg cgtgagtcgg ctgggcgagc tgggcctcgt 360

ggagttcaga gacctcaacg cctcggtag cgcttccag agacgctttg tggttgatgt 420  
 tcggcctgt gaggagctgg agaagacctt caccttctg caggaggagg tccggcgggc 480  
 tgggctggtc ctgccccgc caaaggggag gctgccggca cccccaccc gggacctgct 540  
 gcgcatccag gaggagacgg agcgcctggc ccaggagctg cgggatgtgc ggggcaacca 600  
 gcaggccctg cgggcccagc tgcaccagct gcagctccac gccgccgtgc tacgccaggg 660  
  
 ccatgaacct cagctggcag ccgcccacac agatggggcc tcagagagga cgcccctgct 720  
 ccaggcccc ggggggccc accaggacct gagggtcaac tttgtggcag gtgccgtgga 780  
 gccccacaag gccctgccc tagagcgcct gctctggagg gcctgcagag gcttctcat 840  
 tgccagcttc agggagctgg agcagccgct ggagcaccac gtgacgggcg agccagccac 900  
 gtggatgacc ttctcatct cctactgggg tagcagatc ggacagaaga tccgcaagat 960  
 cacggactgc ttccactgcc acgtttccc gtttctgcag caggaggagg cccgcctcgg 1020  
 ggccctgcag cagctgcaac agcagagcca ggagctgcag gagtcctcg gggagacaga 1080  
  
 gcggttctg agccaggctg taggccgggt gctgcagctg ctgccgccag ggcagggtgca 1140  
 ggtccacaag atgaaggccg tgtacctggc cctgaaccag tgcagcgtga gcaccacgca 1200  
 caagtgcctc attgccgagg cctggtgctc tgtgcgagac ctgcccgcc tgcaggaggc 1260  
 cctgcgggac agctcgatgg aggagggagt gagtgccgtg gctcaccgca tcccctgccg 1320  
 ggacatgccc ccacactca tccgaccaa ccgcttcacg gccagcttcc agggcatcgt 1380  
 ggatgcctac ggcgtgggcc gctaccagga ggtcaacccc gctccctaca ccatcatcac 1440  
 ctccccctc ctgtttgctg tgatgttcgg ggatgtgggc cacgggctgc tcatgttct 1500  
  
 ctteccctg gccatggtcc ttgcggagaa ccgaccgct gtgaaggccg cgcagaacga 1560  
 gatctggcag actttctca ggggccgcta cctgctctg cttatgggcc tgttctcat 1620  
 ctacaccgce ttcatctaca acgagtgtt cagtccgccc accagcatct tcccctcggg 1680  
 ctggagtgtg gccccatgg ccaaccagtc tgctggagt gatgattcc tggcccagca 1740  
 cacgatgctt acctggacc ccaacgac cgtgtcttc ctgggacct accccttgg 1800  
 catcgatcct atttgagcc ttgctgcaa ccacttgagc ttectcaact cctteaagat 1860  
 gaagatgtcc gtcacctgg gcgtcgtgca catggccttt ggggtggtcc tcggagtctt 1920  
  
 caaccacgtg cactttggcc agaggcaccg gctgctgctg gagacgctgc cggagctcac 1980  
 cttectgctg ggactcttcg gttacctcgt gttcctagtc atctacaagt ggctgtgtgt 2040  
 ctgggctgcc agggccgct cggccccag catctcacc cacttcatca acatgttct 2100  
 ctctcccaac agccccagca acaggctgct ctacccccg caggagggtg tccaggccac 2160  
 gctggtggtc ctggccttgg ccatggtgcc catcctgctg cttggcacac cctgcacct 2220

gctgcaccgc caccgccgcc gcctgctggag gaggccccct gaccgacagg aggaaaaaaa 2280  
 ggccccgttg ctggacctgc ctgacgcatc tgtgaatggc tggagctccg atgaggaaaa 2340  
  
 ggcagggggc ctggatgatg aagaggaggc cgagctcgtc ccctccgagg tgctcatgca 2400  
 ccaggccatc cacaccatcg agttctgcct gggctcgtc tccaacaccg cctcctacct 2460  
 gcgcctgtgg gccctgagcc tggcccacgc ccagctgtcc gaggttctgt gggccatggt 2520  
 gatgcgcata ggctgggcc tgggccggga ggtggcgtg gcgctgtgg tgctgtccc 2580  
 catctttgcc gcctttgccg tgatgaccgt ggctatcctg ctggatgagg agggactctc 2640  
 agccttctg cacgccctgc ggctgcaact ggtggaattc cagaacaagt tctactcagg 2700  
 cacgggctac aagctgagtc ccttcacctt cgtgccaca gatgactagt aagtcgacgg 2760  
  
 atccccggg ctgcaggaat tcgagcatct taccgccatt tatacccata tttgttctgt 2820  
 ttttcttgat ttgggtatac atttaaatgt taataaaaca aaatggtggg gcaatcattt 2880  
 acatTTTTag ggatatgtaa ttactagttc aggtgtattg ccacaagaca aacatgttaa 2940  
 gaaactttcc cgttatttac gctctgttcc tgttaatcaa cctctggatt acaaaatttg 3000  
 tgaaagattg actgatatc ttaactatgt tctcctttt acgctgtgtg gatatgctgc 3060  
 tttaatgcct ctgtatcatg ctattgcttc ccgtacggct ttcgttttct cctccttgta 3120  
 taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtccgtc aacgtggcgt 3180  
  
 ggtgtgctct gtgtttgctg acgcaacccc cactggctgg ggcatgcca ccacctgtca 3240  
 actcctttct gggactttcg ctttccccct ccgcatgcc acggcagaac tcacgccgc 3300  
 ctgccttgcc cgtgctgga caggggctag gttgctgggc actgataatt cegtgtgtt 3360  
 gtcggggaag ctgacgtcct ttcgaattcg atatcaagct gtaccttaa gaccaatgac 3420  
 ttacaaggca gctgtagatc ttagccactt tttaaaagaa aaggggggac tggaaaggct 3480  
 aattcactcc caacgaagac aagatctgct tttgcttgt actgggtctc tetggttaga 3540  
 ccagatctga gcctgggagc tctctggcta actaggaac cactgctta agcctcaata 3600  
  
 aagcttgctt tgagtgtctc aagtagtgtg tccccgtctg ttgtgtgact ctggttaacta 3660  
 gagatccctc agaccctttt agtcagtgtg gaaaatctct agcagtagta gttcatgtca 3720  
 tcttattatt cagtatttat aacttgcaaa gaaatgaata tcagagagtg agaggaactt 3780  
 gtttattgca gcttataatg gttacaata aagcaatagc atcacaatt tcacaaataa 3840  
 agcatttttt tactgcatt ctagtgtgg tttgtccaaa ctcatcaatg tatcttatca 3900  
 tgtctggctc tagctatccc gccctaact ccgccatcc cggccctaac tccgccagt 3960

tccgccatt ctccgccca tggctgacta attttttta tttatgcaga ggccgaggcc 4020

gcctcggcct ctgagctatt ccagaagtag tgaggaggct tttttggagg cctaggtagc 4080

ccgcctaag agcgggcttt tttttcttag gccttcttcc gcttcctcgc tcaactgactc 4140

gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cgtaataacg 4200

gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 4260

ggccaggaac cgtaaaaagg ccgcgttgcg ggcgtttttc cataggtcc gcccccctga 4320

cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagag gactataaag 4380

ataccaggcg tttcccctg gaagctccct cgtgcgctct cctgttccga cctgcccct 4440

taccggatac ctgtccgct tttcccttc gggaagcgtg gcgctttctc atagctcacg 4500

ctgtaggtat ctcaattcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 4560

ccccgttcag cccgaccgt gcgcttctc cgtaactat cgtcttgagt ccaaccgggt 4620

aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 4680

tgtaggcggg gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 4740

agtatttggg atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 4800

ttgatccggc aaacaaacca ccgctggtag cgggtggtttt tttgtttgca agcagcagat 4860

tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 4920

tcagtggaac gaaaactcac gtttaaggat tttggctatg agattatcaa aaaggatctt 4980

cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 5040

aacttggctc gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5100

atctcttca tccatagttg cctgactcct gcaaacaccg ttgtggtaga attggtaaag 5160

agagtctgt aaaaatcgca gttcgacat ctgtgtgtct gattattgat ttttggcgaa 5220

accatttgat catatgacaa gatgtgtatc taccttaact taatgatttt gataaaaatc 5280

attaggtacc tgiacattta tattggctca tgtccaacat taccgcatg ttgacattga 5340

ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg 5400

gagttccgcg ttacataact tacggtaaat ggccccctg gctgaccgcc caacgacccc 5460

cgccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 5520

tgacgtcaat ggggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 5580

catatgcaa gtacgcccc tattgacgtc aatgacggta aatggcccc ctggcattat 5640

gcccagtaca tgacctatg ggactttcct acttggcagt acatctacgt attagtcate 5700

gctattacca tggatgatcgc gttttggcag tacatcaatg ggctggata gcggtttgac 5760  
tcacgggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa 5820  
aatcaacggg actttccaaa atgtcgtaac aactccgcc cattgacgca aatggcggt 5880  
aggcgtgtac ggiggagggt ctatataagc agagctcgtt tagtgaaccg ggtctctct 5940  
ggttagacca gatctgacc tgggagctct ctggctaact agggaacca ctgcttaagc 6000  
ctcaataaag cttgccttga gtgcttcaag tagtgtgtgc ccgtctgttg tgtgactctg 6060  
gtaactagag atccctcaga ccccttttagt cagtgtggaa aatctctagc agtggcgccc 6120  
  
gaacaggac ttgaaagcga aagggaacc agaggagctc tctcgacgca ggactcggt 6180  
tgctgaagcg cgcacggcaa gaggcgaggg gcggcgactg gtgagtacgc caaaaat 6240  
gactagcgga ggctagaagg agagagatgg gtgagagac gtcagtatta agcgggggag 6300  
aattagatcg cgatgggaaa aaattcgggt aaggccaggg ggaaagaaaa aatataaatt 6360  
aaaacatata gtatgggcaa gcaggagct agaacgattc gcagttaatc ctggcctgtt 6420  
agaaacatca gaaggctgta gacaaact gggacagcta caaccatccc ttcagacagg 6480  
atcagaagaa cttagatcat tatataatac agtagcaacc ctctattgtg tgcataaag 6540  
  
gatagagata aaagacacca aggaagcttt agacaagata gaggaagagc aaaacaaaag 6600  
taagaccacc gcacagcaag cggccgctga tcttcagacc tggaggagga gatatgaggg 6660  
acaattggag aagtgaatta tataaatata aagtagtaaa aattgaacca ttaggagtag 6720  
caccaccaa ggcaaagaga agagtgggtc agagagaaaa aagagcagtg ggaataggag 6780  
ctttgttct tgggttcttg ggagcagcag gaagcactat gggcgagcg tcaatgacgc 6840  
tgacgttaca ggccagacaa ttattgtctg gtatagtca gcagcagaac aatttctga 6900  
gggctattga ggcgaacag catctgttc aactcacagt ctggggcatc aagcagctcc 6960  
  
aggcaagaat cctggctgtg gaaagatac taaaggatca acagctcctg gggatttggg 7020  
gttgctctgg aaaactcatt tgcaccactg ctgtgccttg gaatgctagt tggagtaata 7080  
aatctctgga acagatttgg aatcacacga cctggatgga gtgggacaga gaaattaaca 7140  
attacacaag cttataatc tccttaattg aagaatcgca aaaccagcaa gaaaagaatg 7200  
aacaagaatt attggaatta gataaatggg caagtttctg gaattggtt aacataacaa 7260  
attggctgtg gtatataaaa ttattcataa tgatagtagg aggccttgga ggtttaagaa 7320  
tagtttttgc tgiactttct atagtgaata gagttaggca gggatattca ccattatcgt 7380  
  
ttcagacca cctcccaacc ccgaggggac ccgacaggcc cgaaggaata gaagaagaag 7440  
gtggagagag agacagagac agatccattc gattagtga cggatctcga cggatcgggt 7500  
taacttttaa aagaaaagg gggattgggg ggtacagtgc aggggaaaga atagtagaca 7560

taatagcaac agacatacaa actaaagaat tacaaaaaca aattacaaaa attcaaaatt 7620  
 ttatcgatca cgagactagc ctcgag 7646  
 <210> 28  
 <211> 234  
 <212> DNA  
 <213> human immunodeficiency virus  
 <400> 28  
 tggaagggct aattcactcc caacgaagac aagatctgct ttttgcttgt actgggtctc 60  
  
 tctggttaga ccagatctga gcctgggagc tctctggcta actaggaac cactgctta 120  
 agcctcaata aagcttgctt tgagtgcttc aagtagtggtg tgcccgtctg ttgtgtgact 180  
 ctggtaacta gagatccctc agaccctttt agtcagtgtg gaaaatctct agca 234  
 <210> 29  
 <211> 132  
 <212> DNA  
 <213> Simian virus 40  
 <400> 29  
 aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca 60  
 aataaagcat ttttttact gcattctagt tggggttgt ccaaactcat caatgtatct 120  
 tatcatgtct gg 132  
  
 <210> 30  
 <211> 160  
 <212> DNA  
 <213> Simian virus 40  
 <400> 30  
 tcccgcacct aactcgcacc atcccgcacc taactcgcgc cagttcgcgc cattctcgc 60  
 cccatggctg actaattttt tttatttatg cagaggccga ggccgcctcg gcctctgagc 120  
 tattccagaa gtagtgagga ggcttttttg gaggcctagg 160  
 <210> 31  
 <211> 1015  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Made in Lab  
 <400> 31

tcttccgctt cctcgcctac tgactcgcig cgtcggctcg ttcggctgcg gcgagcggta 60  
 tcagctcact caaaggcggg aatacgggta tccacagaat caggggataa cgcaggaaag 120

aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg 180  
 tttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg 240  
 tggcgaaaac cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg 300  
 cgctctcctg ttccgacct gccgettac ggatacctgt ccgcctttct cccttcggga 360  
 agcgtggcgc tttctcatag ctacgctgt aggtatctca gttcgggtga ggtcgttcgc 420  
 tccaagctgg gctgtgtgca cgaaccccc gttcagcccg accgctgcgc cttatccggt 480  
 aactatcgtc ttgagtccaa cccgtaaga cagcacttat cgccactggc agcagccact 540

ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagtctt gaagtggagg 600  
 cctaactacg gctacactag aagaacagta tttggtatct gcgctctgct gaagccagtt 660  
 accttcggaa aaagagtgg tagctcttga tccggcaaac aaaccaccgc tggtagcggg 720  
 ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct 780  
 ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg 840  
 gtcattgagat tatcaaaaa gatcttcacc tagatccttt taaattaa atgaagtttt 900  
 aaatcaatct aaagtatata tgagtaaact tggcttgaca gttaccaatg cttaatcagt 960

gaggcaccta ttcagcgat ctgtctattt cgttcatcca tagttgcctg actcc 1015

<210> 32  
 <211> 139  
 <212> DNA  
 <213> Escherichia coli  
 <400> 32

gtagaattgg taaagagagt cgtgtaaaat atcgagttcg cacatcttgt tgtctgatta 60  
 ttgatttttg gcgaaacat ttgatcatat gacaagatgt gtatctacct taacttaatg 120  
 attttgataa aaatcatta 139

<210> 33  
 <211> 577  
 <212> DNA  
 <213> Human betaherpesvirus 5  
 <400> 33

acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc 60

atatatggag ttccgcgtta cataacttac ggtaaattggc ccgcctggct gaccgccc 120  
 cgacccccgc ccattgacgt caataatgac gtagtccc atagtaacgc caataggac 180  
 ttccattga cgtcaatggg tggagtattt acggtaaact gccacttgg cagtacatca 240  
 agtgtatcat atgccaagta cgccccctat tgacgtcaat gacggtaaat ggccccctg 300  
 gcattatgcc cagtacafta ctttatggga ctttctact tggcagtaca tctacgtatt 360  
 agtcatcgct attaccatgg tgatcgggtt ttggcagtac atcaatgggc gtggatagcg 420  
 gtttgactca cggggatttc caagtctcca ccccatgac gtcaatggga gtttgtttg 480

gcacccaaat caacgggact ttccaaatg tcgtaacaac tccgccccat tgacgcaat 540  
 gggcggtagg cgtgtacgtt gggaggtcta tataagc 577

<210> 34

<211> 188

<212> DNA

<213> human immunodeficiency virus

<400> 34

gtctctctgg ttagaccaga tctgagcctg ggagctctct ggctaactag ggaaccact 60  
 gcttaagcct caataaagct tgccttgagt gcttcaagta gtgtgtgcc gtctgtgtg 120  
 tgactctggt aactagat ccctcagacc cttttagtca gtgtggaaaa tctctagcag 180  
 tggcgccc 188

<210> 35

<211> 45

<212> DNA

<213> Human immunodeficiency virus 1

<400> 35

tgagtacgcc aaaaattttg actagcggag gctagaagga gagag 45

<210> 36

<211> 362

<212> DNA

<213> human immunodeficiency virus

<400> 36

atgggtgcga gagcgtcagt attaagcggg ggagaattag atcgcgatgg gaaaaattc 60  
 ggtaaggcc agggggaaag aaaaaatata aattaaaca tatagtatgg gcaagcagg 120  
 agctagaac attcgcagtt aatcctggcc tgttagaac atcagaagc ttagacaaa 180

tactgggaca gctacaacca tcccttcaga caggatcaga agaacttaga tcattatata 240

atacagtagc aaccctctat tgtgtgcatc aaaggataga gataaaagac accaaggaag 300

ctttagacaa gatagaggaa gagcaaaaca aaagtaagac caccgcacag caagcggcgcg 360

ct 362

<210> 37

<211> 858

<212> DNA

<213> Artificial Sequence

<220><223> Made in Lab - plasmid element

<400> 37

gatcttcaga cctggaggag gagatatgag ggacaattgg agaagtgaat tatataaata 60

taaagtagta aaaattgaac cattaggagt agcaccacc aaggcaaaga gaagagtgg 120

gcagagagaa aaaagagcag tgggaatagg agctttgttc cttgggttct tgggagcagc 180

aggaagcact atgggcgcag cgtcaatgac gctgacggta caggccagac aattattgtc 240

tggatatagt cagcagcaga acaatttgct gagggttatt gaggcgcaac agcatctgtt 300

gcaactcaca gtctggggca tcaagcagct ccaggcaaga atcctggctg tggaaagata 360

cctaaaggat caacagctcc tggggatttg gggttgcctt ggaaaactca tttgaccac 420

tgctgtgctt tggaatgcta gttggagtaa taaatctctg gaacagattt ggaatcacac 480

gacctggatg gagtgggaca gagaaattaa caattacaca agcttaatac actccttaat 540

tgaagaatcg caaaaccagc aagaaaagaa tgaacaagaa ttattggaat tagataaatg 600

ggcaagtttg tggaaattgg ttaacataac aaattggctg tggatatataa aattattcat 660

aatgatagta ggaggcttgg taggtttaag aatagttttt gctgtacttt ctatagttaa 720

tagagttagg cagggatatt caccattatc gtttcagacc cacctcccaa ccccagggg 780

acccgacagg cccgaaggaa tagaagaaga aggtggagag agagacagag acagatccat 840

tcgattagtg aacggatc 858

<210> 38

<211> 118

<212> DNA

<213> human immunodeficiency virus

<400> 38

ttttaaaaga aaagggggga ttggggggta cagtgcaggg gaaagaatag tagacataat 60

agcaacagac atacaaacta aagaattaca aaaacaaatt acaaaaattc aaaatttt 118  
 <210> 39  
 <211> 3847  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Made in lab - plasmid backbone construct  
 <400> 39  
 aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca 60  
 aataaagcat ttttttcaact gcattctagt tgtggtttgt ccaaactcat caatgtatct 120  
 tatcatgtct ggctctagct atcccgcccc taactccgcc catcccgcc ctaactccgc 180  
 ccagttccgc ccatttccgc ccccatggct gactaatttt ttttatttat gcagaggccg 240  
  
 aggccgcctc ggctctgag ctattccaga agtagtgagg aggctttttt ggaggcctag 300  
 gtagcccgcc taatgagcgg gctttttttt cttaggcctt cttccgcttc ctgctcact 360  
 gactcgtgc gctcggctgt tggctgagg cgagcggtat cagctcactc aaaggcggtg 420  
 atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag 480  
 caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc 540  
 cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccg gacaggacta 600  
 taaagatacc aggcgtttcc ccctggaagc tcctcgtgc gctctcctgt tccgaccctg 660  
  
 ccgcttaccg gatacctgtc cgcttttctc ccttcgggaa gcgtggcgct ttctcatagc 720  
 tcacgctgta ggtatctcag ttcgggtgtag gtcgttcgct ccaagctggg ctgtgtgcac 780  
 gaacccccg ttacgccga cgctgcgcc ttatccgta actatcgtct tgagtccaac 840  
 ccgtaagac acgacttacc gccactggca gcagccactg gtaacaggat tagcagagcg 900  
 aggtatgtag gcggtgctac agagtcttgg aagtggggc ctaactacgg ctacactaga 960  
 agaacagtat ttggtatctg cgctctgctg aagccagtta ccttcgaaa aagagtggg 1020  
 agctcttgat ccggcaaaaa aaccaccgct ggtagcgggtg gtttttttgt ttgcaagcag 1080  
  
 cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct 1140  
 gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg 1200  
 atcttcacct agatcctttt aaattaaana tgaagtttta aatcaatcta aagtatatat 1260  
 gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctgagcagtc 1320  
 tgtctatttc gttcatccat agttgctgca ctctgcaaaa ccacgttgtg gtagaattgg 1380  
 taaagagagt cgtgtaaaa atcaggttcg cacatcttgt tgtctgatta ttgatttttg 1440

gcgaaacat ttgatcatat gacaagatgt gtatctacct taacttaatg attttgataa 1500

aaatcattag gtacctgtac atttatattg gctcatgtcc aacattaccg ccatgttgac 1560

attgattatt gactagtat taatagtaat caattacggg gtcattagtt catagcccat 1620

atatggagtt ccgcgttaca taacttacgg taatggccc gcctggctga ccgccaacg 1680

acccccgcc attgacgtca ataatgacgt atgttcccat agtaacgcca atagggactt 1740

tccattgacg tcaatgggtg gagtatttac ggtaaactgc ccaactggca gtacatcaag 1800

tgtatcatat gccaaagtacg cccctattg acgtcaatga cggtaaatgg cccgcctggc 1860

attatgcca gtacatgacc ttatgggact ttctacttg gcagtacatc tacgtattag 1920

tcatcgtat taccatggtg atgcggtttt ggcagtacat caatggcgt ggatagcgg 1980

ttgactcacg gggatttcca agtctccacc ccattgacgt caatgggagt ttgttttgcc 2040

accaaaaatca acgggacttt ccaaaaatgtc gtaacaactc cgccccattg acgcaaatgg 2100

gcggtaggcg tgiacgggtg gaggtctata taagcagagc tcgtttagtg aaccggggtc 2160

tctctggtta gaccagatct gagcctggga gctctctggc taactagga acccactgct 2220

taagcctcaa taaagcttgc cttgagtgtc tcaagtagtg tgtgccctc tgttgtgtga 2280

ctctggtaac tagagatccc tcagaccctt ttagtcagtg tggaaaatct ctgacgtgg 2340

cgccccgaaca gggacttgaa agcgaaaggg aaaccagagg agctctctcg acgcaggact 2400

cggcttgctg aagcgcgcac ggcaagaggc gagggcgccg gactggtgag tacgccaaaa 2460

attttgacta gcggaggcta gaaggagaga gatgggtgag agagcgtcag tattaagcgg 2520

gggagaatta gatcgcgatg ggaaaaaatt cggttaaggc cagggggaaa gaaaaaatat 2580

aaattaaaac atatagtatg ggcaagcagg gagctagaac gattcgcagt taatcctggc 2640

ctgttagaaa catcagaagg ctgtagacaa atactgggac agctacaacc atccctcag 2700

acaggatcag aagaacttag atcattatat aatacagtag caacccteta ttgtgtgcat 2760

caaaggatag agataaaaga caccaaggaa gcttttagaca agatagagga agagcaaac 2820

aaaagtaaga ccaccgcaca gcaagcggcc gctgatcttc agacctggag gaggagatat 2880

gagggacaat tggagaagtg aattatataa atataaagta gtaaaaattg aaccattagg 2940

agtagcacc accaaggcaa agagaagagt ggtgcagaga gaaaaagag cagtgggaat 3000

aggagctttg ttctttgggt tcttgggagc agcaggaagc actatgggag cagcgtcaat 3060

gacgctgacg gtacaggcca gacaattatt gtctggtata gtcagcagc agaacaattt 3120

gctgagggct attgaggcgc aacagcatct gttgcaactc acagtctggg gcatcaagca 3180

gctccaggca agaatcctgg ctgtggaaag atacctaaag gatcaacagc tcctggggat	3240
ttggggttgc tctggaaaac tcatttgac cactgctgtg ccttggaaatg ctagtggag	3300
taataaatct ctggaacaga tttggaatca cacgacctgg atggagtggg acagagaaat	3360
taacaattac acaagcttaa tacactcctt aattgaagaa tgcaaaacc agcaagaaaa	3420
gaatgaacaa gaattattgg aattagataa atgggcaagt ttgtggaatt ggtttaacat	3480
aacaaattgg ctgtggtata taaaattatt cataatgata gtaggaggct tggtaggttt	3540
aagaatagtt ttigtgtac tttctatagt gaatagagtt aggcagggat attcaccatt	3600
atcgtttcag acccacctcc caaccccgag gggacccgac aggccccaag gaatagaaga	3660
agaaggtgga gagagagaca gagacagatc cattcgatta gtgaacggat ctcgacgta	3720
tcggttaact tttaaaagaa aaggggggat tgggggtac agtgcagggg aaagaatagt	3780
agacataata gcaacagaca taaaactaa agaattacaa aaacaaatta caaaaattca	3840
aaatttt	3847