



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0048743
(43) 공개일자 2018년05월10일

- (51) 국제특허분류(Int. Cl.)
C12N 15/86 (2006.01) A61K 48/00 (2006.01)
C07K 14/57 (2006.01) C07K 16/28 (2006.01)
C12N 15/113 (2010.01) C12N 9/78 (2006.01)
- (52) CPC특허분류
C12N 15/86 (2013.01)
A61K 48/005 (2013.01)
- (21) 출원번호 10-2018-7008093
- (22) 출원일자(국제) 2016년09월01일
심사청구일자 없음
- (85) 번역문제출일자 2017년03월21일
- (86) 국제출원번호 PCT/US2016/049947
- (87) 국제공개번호 WO 2017/040815
국제공개일자 2017년03월09일
- (30) 우선권주장
62/214,884 2015년09월04일 미국(US)

- (71) 출원인
토카겐 인크.
미국, 캘리포니아 92109, 샌 디에고, 스위트 230,
번커 힐 스트리트 3030
- (72) 발명자
줄리 더글러스 체이.
미국, 캘리포니아 92024, 엔시니타스, 382 선셋
드라이브
린 에이미 에이치.
미국, 캘리포니아 92130, 샌 디에고, 5488 소노마
플레이스
(뒷면에 계속)
- (74) 대리인
손민

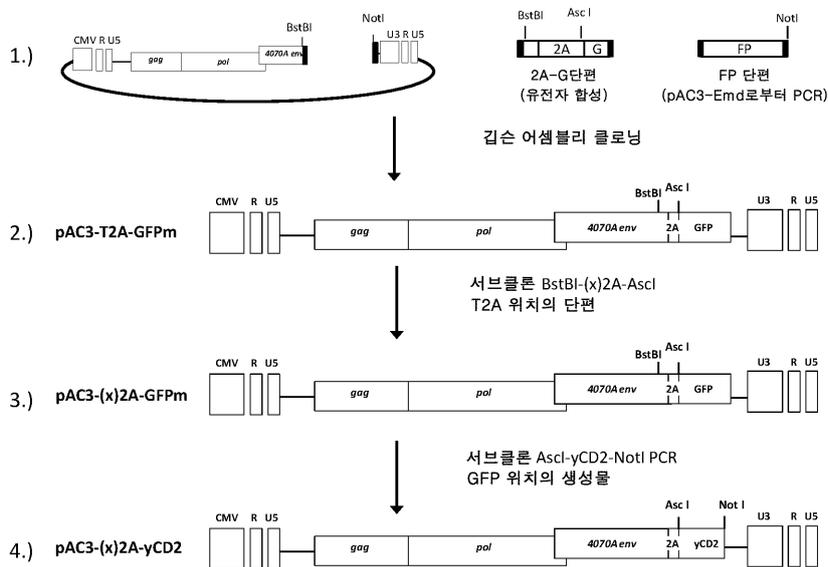
전체 청구항 수 : 총 58 항

(54) 발명의 명칭 **2A 펩타이드를 포함하는 재조합 벡터**

(57) 요약

본 발명은 이중성 폴리뉴클레오타이드에 작동가능하게 연결된 2A-펩타이드 또는 펩타이드-유사 코딩 서열을 함유하는 변형된 재조합 레트로바이러스를 제공한다. 본 발명은 또한 이러한 벡터를 발현 또는 포함하는 세포 및 벡터, 및 질병 및 장애의 치료에 상기 변형된 벡터를 사용하는 방법에 관한 것이다.

대표도 - 도3



(52) CPC특허분류

C07K 14/57 (2013.01)
C07K 16/2827 (2013.01)
C12N 15/1138 (2013.01)
C12N 9/78 (2013.01)
C07K 2319/50 (2013.01)
C12N 2310/141 (2013.01)
C12N 2740/13022 (2013.01)
C12N 2740/13043 (2013.01)
C12N 2840/20 (2013.01)

(72) 발명자

호파크레 앤드류

미국, 캘리포니아 92677, 라구나 니구엘, 24942 란
초 클레멘테

호간 다니엘 제이.

미국, 캘리포니아 92126, 샌 디에고, 7712 파르마
레인

오즈데르타그 데릭 지.

미국, 캘리포니아 92130, 샌 디에고, 6156 블루 다
운 트레일

명세서

청구범위

청구항 1

레트로바이러스 GAG 단백질;

레트로바이러스 POL 단백질;

레트로바이러스 외피(envelope);

레트로바이러스 폴리뉴클레오타이드 서열의 3' 말단에 긴-말단 반복(Long-term repeat, LTR) 서열을, 레트로바이러스 폴리뉴클레오타이드의 5' 말단에 프로모터 서열을, gag 핵산 도메인, pol 핵산 도메인 및 env 핵산 도메인을 포함하며, 상기 프로모터는 포유류 세포에서의 발현에 적합한 레트로바이러스 폴리뉴클레오타이드;

이종성 폴리뉴클레오타이드에 작동가능하게 연결된 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열을 포함하는 카세트로서, 상기 카세트가 5' 내지 3' LTR에 위치하고, 작동가능하게 연결되고, 3' 내지 레트로바이러스 외피를 코딩하는 env 핵산 도메인에 위치한 카세트; 및

표적 세포에서 역전사, 패키징 및 통합에 필요한 시스-작용 서열

을 포함하는 재조합 복제가능 레트로바이러스.

청구항 2

청구항 1에 있어서, 상기 외피는 양친매성(amphotropic), 다중지향성(polytropic), 이종향성(xenotropic), 10A1, GALV, 비비원숭이 내인성 바이러스, RD114, 람다바이러스, 알파바이러스, 홍역 또는 인플루엔자 바이러스 외피 중 하나로부터 선택되는, 재조합 복제가능 레트로바이러스.

청구항 3

청구항 1에 있어서, 상기 레트로바이러스 폴리뉴클레오타이드 서열은 무린 백혈병 바이러스(MLV), 몰로니 무린 백혈병 바이러스(MoMLV), 고양이 백혈병 바이러스(FeLV), 비비원숭이 내인성 레트로바이러스(BEV), 돼지 내인성 바이러스(PERV), 고양이 유래 레트로바이러스 RD114, 다람쥐 원숭이 레트로바이러스, 이종향성 무린 백혈병 바이러스-관련 바이러스(XMRV), 조류 망상내피증 바이러스(REV) 또는 긴팔원숭이 백혈병 바이러스(GALV)로 구성된 그룹으로부터 선택되는 바이러스로부터 유래되는, 재조합 복제가능 레트로바이러스.

청구항 4

청구항 1에 있어서, 상기 레트로바이러스는 감마레트로바이러스인, 재조합 복제가능 레트로바이러스.

청구항 5

청구항 1에 있어서, 상기 표적 세포는 포유류 세포인, 재조합 복제가능 레트로바이러스.

청구항 6

청구항 1에 있어서, 상기 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열은 서열식별번호: 1의 서열을 함유하는 펩타이드를 코딩하는, 재조합 복제가능 레트로바이러스.

청구항 7

청구항 1에 있어서, 상기 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열은 서열식별번호: 55 내지 125 중 어느 하나에 개시된 펩타이드를 코딩하는, 재조합 복제가능 레트로바이러스.

청구항 8

청구항 1에 있어서, 상기 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열은 서열식별번호: 8 내지 19 중 어느 하

나에 개시된 서열을 포함하는, 재조합 복제가능 레트로바이러스.

청구항 9

청구항 1 내지 8 중 어느 한 항에 있어서, 상기 이중성 폴리뉴클레오타이드는 > 500 bp인, 재조합 복제가능 레트로바이러스.

청구항 10

청구항 1에 있어서, 상기 이중성 폴리뉴클레오타이드는 2개 이상의 코딩 서열을 포함하는, 재조합 복제가능 레트로바이러스.

청구항 11

청구항 1에 있어서, 상기 카세트의 하류에 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열을 포함하는 제2 카세트를 추가로 포함하는, 재조합 복제가능 레트로바이러스.

청구항 12

청구항 1에 있어서, 상기 카세트의 하류에 제2 카세트를 추가로 포함하고, 제2 카세트는 내부 리보솜 진입 부위(IRES) 또는, 제2 이중성 폴리뉴클레오타이드에 작동가능하게 연결된 미니프로모터 또는 polIII 프로모터를 포함하는, 재조합 복제가능 레트로바이러스.

청구항 13

레트로바이러스 GAG 단백질;

레트로바이러스 POL 단백질;

레트로바이러스 외피;

레트로바이러스 폴리뉴클레오타이드 서열의 3' 말단에 긴-말단 반복(LTR) 서열을, 레트로바이러스 폴리뉴클레오타이드의 5' 말단에 프로모터 서열을, gag 핵산 도메인, pol 핵산 도메인 및 env 핵산 도메인을 포함하며, 상기 프로모터는 포유류 세포에서의 발현에 적합한, 레트로바이러스 폴리뉴클레오타이드;

제2 이중성 폴리뉴클레오타이드에 작동가능하게 연결된 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열을 포함하는 적어도 하나의 2A 카세트에 작동가능하게 연결된 제1 이중성 폴리뉴클레오타이드에 작동가능하게 연결된 조절 도메인을 포함하는 카세트로서, 상기 카세트가 5' 내지 3' LTR 및 3' 내지 레트로바이러스 외피를 코딩하는 env 핵산 도메인에 위치하고, 2A 카세트가 하류에 있고, 제1 이중성 폴리뉴클레오타이드에 작동가능하게 연결되는 카세트; 및

표적 세포에서 역전사, 패키징 및 통합에 필요한 시스-작용 서열(cis-acting integration)을 포함하는, 재조합 복제가능 레트로바이러스.

청구항 14

청구항 13에 있어서, 상기 외피는 양친매성, 다중지향성, 이중향성, 10A1, GALV, 비비원숭이 내인성 바이러스, RD114, 랫도바이러스, 알파바이러스, 홍역 또는 인플루엔자 바이러스 외피 중 하나로부터 선택되는, 재조합 복제가능 레트로바이러스.

청구항 15

청구항 13에 있어서, 상기 레트로바이러스 폴리뉴클레오타이드 서열은 뮤린 백혈병 바이러스(MLV), 몰로니 뮤린 백혈병 바이러스(MoMLV), 고양이 백혈병 바이러스(FeLV), 비비원숭이 내인성 레트로바이러스(BEV), 돼지 내인성 바이러스(PERV), 고양이 유래 레트로바이러스 RD114, 다람쥐 원숭이 레트로바이러스, 이중향성 뮤린 백혈병 바이러스-관련 바이러스(XMRV), 조류 망상내피증 바이러스(REV) 또는 긴팔원숭이 백혈병 바이러스(GALV)로 구성된 그룹으로부터 선택되는 바이러스로부터 유래되는, 재조합 복제가능 레트로바이러스.

청구항 16

청구항 13에 있어서, 상기 레트로바이러스는 감마레트로바이러스인, 재조합 복제가능 레트로바이러스.

청구항 17

청구항 13에 있어서, 상기 표적 세포는 포유류 세포인, 재조합 복제가능 레트로바이러스.

청구항 18

청구항 13에 있어서, 상기 2A 펩타이드 또는 펩타이드-유사 코딩 서열은 서열식별번호: 1의 서열을 함유하는 펩타이드를 코딩하는, 재조합 복제가능 레트로바이러스.

청구항 19

청구항 13에 있어서, 상기 2A 펩타이드 또는 펩타이드-유사 코딩 서열은 도 1 또는 2에 개시된 펩타이드를 코딩하는, 재조합 복제가능 레트로바이러스.

청구항 20

청구항 13에 있어서, 상기 2A 펩타이드 또는 펩타이드-유사 코딩 서열은 서열식별번호: 8 내지 19 중 어느 하나에 개시된 서열을 포함하는, 재조합 복제가능 레트로바이러스.

청구항 21

청구항 1 내지 20 중 어느 한 항에 있어서, 상기 표적 세포는 폐암 세포, 결장-직장암 세포, 유방암 세포, 전립선 암 세포, 요로 암 세포, 자궁암 세포, 뇌암 세포, 두경부암 세포, 췌장암 세포, 흑색종 세포, 위암 및 난소암 세포로 구성된 그룹으로부터 선택되는, 재조합 복제가능 레트로바이러스.

청구항 22

청구항 1 내지 21 중 어느 한 항에 있어서, 상기 프로모터 서열은 성장 조절 유전자와 관련된, 재조합 복제가능 레트로바이러스.

청구항 23

청구항 1 내지 21 중 어느 한 항에 있어서, 상기 프로모터 서열은 조직-특이적인 프로모터 서열을 포함하는, 재조합 복제가능 레트로바이러스.

청구항 24

청구항 23에 있어서, 상기 조직-특이적 프로모터 서열은 하나 이상의 안드로겐 반응 요소(ARE)를 포함하는, 재조합 복제가능 레트로바이러스.

청구항 25

청구항 1 또는 13에 있어서, 상기 프로모터는 뉴클레오타이드 1 내지 약 뉴클레오타이드 582에서 서열식별번호: 2에 개시된 서열을 갖는 CMV 프로모터를 포함하고, 하나 이상의 핵산 염기에 대한 변형을 포함할 수 있으며, 전사를 지시 및 개시할 수 있는, 재조합 복제가능 레트로바이러스.

청구항 26

청구항 1 또는 13에 있어서, 상기 프로모터는 CMV-R-U5 도메인 폴리뉴클레오타이드를 포함하는, 재조합 복제가능 레트로바이러스.

청구항 27

청구항 26에 있어서, 상기 CMV-R-U5 도메인은 MLV R-U5 영역에 연결된 인간 사이토메갈로바이러스로부터의 즉시 초기(immediately early) 프로모터를 포함하는, 재조합 복제가능 레트로바이러스.

청구항 28

청구항 27에 있어서, 상기 CMV-R-U5 도메인 폴리뉴클레오타이드는 약 뉴클레오타이드 1 내지 약 뉴클레오타이드

1202의 서열식별번호: 2에 개시된 서열, 또는 1 내지 약 1202의 서열식별번호: 2에 개시된 서열과 적어도 95% 동일한 서열을 포함하며, 상기 폴리뉴클레오타이드는 이에 작동가능하게 연결된 핵산 분자의 전사를 촉진하는, 재조합 복제가능 레트로바이러스.

청구항 29

청구항 1 또는 13에 있어서, 상기 gag 폴리뉴클레오타이드는 감마레트로바이러스로부터 유래되는, 재조합 복제가능 레트로바이러스.

청구항 30

청구항 29에 있어서, 상기 gag 핵산 도메인은 서열식별번호: 2의 약 뉴클레오타이드 번호 1203 내지 약 뉴클레오타이드 2819의 서열, 또는 이 서열과 적어도 95%, 98%, 99% 또는 99.8%의 동일성을 갖는 서열을 포함하는, 재조합 복제가능 레트로바이러스.

청구항 31

청구항 1 또는 13에 있어서, 상기 폴리뉴클레오타이드의 pol 도메인은 감마레트로바이러스로부터 유래되는, 재조합 복제가능 레트로바이러스.

청구항 32

청구항 31에 있어서, 상기 pol 도메인은 서열식별번호: 2의 약 뉴클레오타이드 번호 2820 내지 약 뉴클레오타이드 6358의 서열, 또는 이 서열과 적어도 95%, 98%, 99% 또는 99.9%의 동일성을 갖는 서열을 포함하는, 재조합 복제가능 레트로바이러스.

청구항 33

청구항 1 또는 13에 있어서, 상기 env 도메인은 서열식별번호: 2의 약 뉴클레오타이드 번호 6359 내지 약 뉴클레오타이드 8323의 서열 또는 이들과 적어도 95%, 98%, 99% 또는 99.8%의 동일성을 갖는 서열을 포함하는, 재조합 복제가능 레트로바이러스.

청구항 34

청구항 1 또는 13에 있어서, 상기 3' LTR은 감마레트로바이러스로부터 유래되는, 재조합 복제가능 레트로바이러스.

청구항 35

청구항 34에 있어서, 상기 3' LTR은 U3-R-U5 도메인을 포함하는, 재조합 복제가능 레트로바이러스.

청구항 36

청구항 35에 있어서, 상기 3' LTR은 약 뉴클레오타이드 9111 내지 약 11654의 서열식별번호: 2에 개시된 서열, 또는 이들과 적어도 95%, 98% 또는 99.5% 동일한 서열을 포함하는, 재조합 복제가능 레트로바이러스.

청구항 37

청구항 1 또는 13에 있어서, 상기 이중 핵산 서열은 생물학적 반응 조절제 또는 면역기능강화 사이토카인을 코딩하는, 재조합 복제가능 레트로바이러스.

청구항 38

청구항 37에 있어서, 상기 면역기능강화 사이토카인은 인터루킨 1 내지 38, 인터페론, 종양 괴사 인자(TNF) 및 과립구-대식세포-콜로니 자극 인자(GM-CSF)로 구성된 그룹으로부터 선택되는, 재조합 복제가능 레트로바이러스.

청구항 39

청구항 37에 있어서, 상기 면역기능강화 사이토카인은 인터페론 감마인, 재조합 복제가능 레트로바이러스.

청구항 40

청구항 1 또는 13에 있어서, 상기 이중 핵산은 무독성 프로드러을 독성 약물로 전환시키는 폴리펩타이드를 코딩하는, 재조합 복제가능 레트로바이러스.

청구항 41

청구항 40에 있어서, 상기 무독성 프로드러을 독성 약물로 전환시키는 폴리펩타이드는 티미딘 키나아제, 퓨린 뉴클레오사이드 포스포릴라제(PNP) 또는 시토신 데아미나제인, 재조합 복제가능 레트로바이러스.

청구항 42

청구항 1 또는 13에 있어서, 상기 이중 핵산 서열은 수용체 도메인, 항체 또는 항체 단편을 코딩하는, 재조합 복제가능 레트로바이러스.

청구항 43

청구항 12에 있어서, 상기 제2 카세트는 억제성 폴리뉴클레오타이드를 포함하는, 재조합 복제가능 레트로바이러스.

청구항 44

청구항 43에 있어서, 상기 억제성 폴리뉴클레오타이드는 miRNA, RNAi 또는 siRNA 서열을 포함하는, 재조합 복제가능 레트로바이러스.

청구항 45

청구항 1 또는 13의 레트로바이러스를 생산하기 위한 재조합 레트로바이러스 폴리뉴클레오타이드 게놈.

청구항 46

청구항 45에 있어서, GSG 링커 코딩 서열을 갖거나 갖지 않는 2A 펩타이드 또는 펩타이드-유사 코딩 서열을 인-프레임으로 갖는 MLV 4070A 외피 단백질 유전자, 및 2A 펩타이드 또는 2A-유사 코딩 서열을 인-프레임으로 갖는 제2 유전자를 포함하는, 재조합 레트로바이러스 폴리뉴클레오타이드 게놈.

청구항 47

청구항 45에 있어서, GSG 링커 코딩 서열을 갖거나 갖지 않는 2A 펩타이드 또는 펩타이드-유사 코딩 서열을 인-프레임으로 갖는 MLV 10A1 외피 단백질 유전자, 및 2A 펩타이드 또는 2A-유사 코딩 서열을 인-프레임으로 갖는 제2 유전자를 포함하는, 재조합 레트로바이러스 폴리뉴클레오타이드 게놈.

청구항 48

청구항 45에 있어서, GSG 링커 코딩 서열을 갖거나 갖지 않는 2A 펩타이드 또는 펩타이드-유사 코딩 서열을 인-프레임으로 갖는 XMRV 외피 단백질 유전자, 및 2A 펩타이드 또는 2A-유사 코딩 서열을 인-프레임으로 갖는 제2 유전자를 포함하는, 재조합 레트로바이러스 폴리뉴클레오타이드 게놈.

청구항 49

청구항 45에 있어서, GSG 링커 코딩 서열을 갖거나 갖지 않는 2A 펩타이드 또는 펩타이드-유사 코딩 서열을 인-프레임으로 갖는 비-프렌드(non-Friend) MLV 외피 단백질 유전자, 및 2A 펩타이드 또는 2A-유사 코딩 서열을 인-프레임으로 갖는 제2 유전자를 포함하는, 재조합 레트로바이러스 폴리뉴클레오타이드 게놈.

청구항 50

청구항 45 내지 49 중 어느 한 항에 있어서, 상기 이중성은 분비성, 막, 세포질, 핵 또는 세포-구획-특이적 단백질인, 재조합 레트로바이러스 폴리뉴클레오타이드 게놈.

청구항 51

청구항 1에 있어서, 상기 재조합 레트로바이러스 및/또는 이중성 폴리뉴클레오타이드는 인간 APOBEC 과돌연변이

(hypermutation)를 일으키기 쉬운 트립토판 코돈을 제거하도록 조작되는, 재조합 복제가능 레트로바이러스.

청구항 52

청구항 13에 있어서, 상기 재조합 레트로바이러스 및/또는 제1 및/또는 제2 이중성 폴리뉴클레오타이드는 인간 APOBEC 과돌연변이를 일으키기 쉬운 트립토판 코돈을 제거하도록 조작되는, 재조합 복제가능 레트로바이러스.

청구항 53

청구항 52에 있어서, 상기 이중성 폴리뉴클레오타이드는 시토신 데아미나제 활성을 갖는 폴리펩타이드를 코딩하는, 재조합 복제가능 레트로바이러스.

청구항 54

청구항 53에 있어서, 상기 시토신 데아미나제 활성을 갖는 폴리펩타이드는 서열식별번호: 29의 폴리펩타이드를 코딩하며, 여기서 X는 트립토판을 제외한 임의의 아미노산인, 재조합 복제가능 레트로바이러스.

청구항 55

비-감수성 코돈에 APOBEC 과돌연변이가 일어나기 쉬운 레트로바이러스 폴리뉴클레오타이드에서 코돈을 조작함으로써 인간 APOBEC에 의한 불활성화에 저항성이 있는, 재조합 복제가능 레트로바이러스.

청구항 56

청구항 55에 있어서, 상기 APOBEC 과돌연변이되기 쉬운 코돈은 트립토판 아미노산을 코딩하는, 재조합 복제가능 레트로바이러스.

청구항 57

청구항 55 또는 56에 있어서, 상기 레트로바이러스는 env 유전자의 하류에 IRES 카세트, 프로모터 카세트 및/또는 2A 펩타이드 카세트를 포함하는, 재조합 복제가능 레트로바이러스.

청구항 58

이중성 폴리뉴클레오타이드가 발현되고 이중성 폴리뉴클레오타이드가 프로드럭을 세포독성 약물로 전환시키는 단백질을 코딩하는 조건하에, 청구항 1 또는 13의 레트로바이러스와 개체를 접촉시키는 단계를 포함하는 세포 증식성 장애의 치료 방법으로서, 상기 이중성 폴리뉴클레오타이드가 프로드럭을 세포독성 약물로 전환시키는 단백질을 코딩하는, 방법.

발명의 설명

기술 분야

[0001] 관련 출원에 대한 상호 참조

[0002] 본 출원은 2015년 9월 4일자로 출원된 미국 가출원 제62/214,884호의 우선권을 주장하며, 이들의 개시 내용은 본 명세서에 참고로 포함된다.

[0003] 서열 목록에 대한 참조

[0004] 본 출원은 전자 형식으로 서열 목록과 함께 출원되고 있다. 서열 목록은 2016년 9월 1일에 생성된 Sequence-Listing_ST25.txt라는 파일로 제공되며, 크기는 245,703 바이트(239KB)이다. 서열 목록의 전자 포맷 정보는 그 전체가 본원에 참고로 포함된다.

[0005] 기술 분야

[0006] 본 발명은 복제가능 레트로바이러스 벡터에 관한 것이다. 본 발명은 또한 세포내 이중 핵산(heterologous nucleic acid)의 전달 및 발현을 위한 상기 복제가능 레트로바이러스 벡터의 용도에 관한 것이다.

배경 기술

[0007] 유전자 및 이중 핵산을 세포 및 개체에 전달하는 효과적인 방법은 과학적 발달 및 질병 및 장애의 가능한 치료

를 위해 연구자들의 목표가 되어왔다.

[0008] 요약

[0009] 본원은 레트로바이러스 GAG 단백질; 레트로바이러스 POL 단백질; 레트로바이러스 외피(envelope); 레트로바이러스 폴리뉴클레오타이드 서열의 3' 말단에 긴-말단 반복(Long-term repeat, LTR) 서열을, 레트로바이러스 폴리뉴클레오타이드의 5' 말단에 프로모터 서열을, gag 핵산 도메인, pol 핵산 도메인 및 env 핵산 도메인을 포함하며, 상기 프로모터는 포유류 세포에서의 발현에 적합한 레트로바이러스 폴리뉴클레오타이드; 이중성 폴리뉴클레오타이드에 작동가능하게 연결된 2A 펩타이드 또는 펩타이드-유사 코딩 서열을 포함하는 카세트로서, 상기 카세트가 5' 내지 3' LTR에 위치하고, 작동가능하게 연결되고, 3' 내지 레트로바이러스 외피를 코딩하는 env 핵산 도메인에 위치한 카세트; 및 표적 세포에서 역전사, 패키징 및 통합에 필요한 시스-작용 서열을 포함하는 재조합 복제가능 레트로바이러스를 제공한다. 일 구현예에서, 외피는 양친매성(amphotropic), 다중지향성(polytropic), 이중향성(xenotropic), 10A1, GALV, 비비원숭이 내인성 바이러스, RD114, 랫도바이러스, 알파바이러스, 홍역 또는 인플루엔자 바이러스 외피 중 하나로부터 선택된다. 또다른 구현예에서, 레트로바이러스 폴리뉴클레오타이드 서열은 무린 백혈병 바이러스(MLV), 몰로니 무린 백혈병 바이러스(MoMLV), 고양이 백혈병 바이러스(FeLV), 비비원숭이 내인성 레트로바이러스(BEV), 돼지 내인성 바이러스(PERV), 고양이 유래 레트로바이러스 RD114, 다람쥐 원숭이 레트로바이러스, 이중향성 무린 백혈병 바이러스-관련 바이러스(XMRV), 조류 망상내피증 바이러스(REV) 또는 긴팔원숭이 백혈병 바이러스(GALV)로 구성된 그룹으로부터 선택되는 바이러스로부터 유래된다. 또다른 구현예에서, 레트로바이러스는 감마레트로바이러스이다. 또다른 구현예에서, 표적 세포는 포유류 세포이다. 또다른 구현예에서, 2A 펩타이드 또는 펩타이드-유사 코딩 서열은 서열식별번호: 1의 서열을 함유하는 펩타이드를 코딩한다. 또다른 구현예에서, 2A 펩타이드 또는 펩타이드-유사 코딩 서열은 서열식별번호: 55 내지 125 중 어느 하나에 개시된 펩타이드를 코딩한다. 또다른 구현예에서, 2A 펩타이드 또는 펩타이드-유사 코딩 서열은 서열식별번호: 8 내지 19 중 어느 하나에 개시된 서열을 포함한다. 상기 구현예 중 임의의 또다른 구현예에서, 이중성 폴리뉴클레오타이드는 > 500 bp, > 1000 bp, > 1100 bp, > 1200 bp, > 1300 bp, > 1400 bp 또는 > 1500 bp 길이이다. 또다른 구현예에서, 이중성 폴리뉴클레오타이드는 2개 이상의 코딩 서열을 포함한다. 또다른 구현예에서, 레트로바이러스는 카세트의 하류에 2A 펩타이드 또는 펩타이드-유사 코딩 서열을 포함하는 제2 카세트를 추가로 포함한다. 또다른 구현예에서, 레트로바이러스는 카세트의 하류에 제2 카세트를 추가로 포함하고, 제2 카세트는 내부 리보솜 진입 부위(IRES) 또는, 제2 이중성 폴리뉴클레오타이드에 작동가능하게 연결된 미니프로모터 또는 polIII 프로모터를 포함한다.

[0010] 본원은 또한: 레트로바이러스 GAG 단백질; 레트로바이러스 POL 단백질; 레트로바이러스 외피; 레트로바이러스 폴리뉴클레오타이드 서열의 3' 말단에 긴-말단 반복(LTR) 서열을, 레트로바이러스 폴리뉴클레오타이드의 5' 말단에 프로모터 서열을, gag 핵산 도메인, pol 핵산 도메인 및 env 핵산 도메인을 포함하며, 상기 프로모터는 포유류 세포에서의 발현에 적합한, 레트로바이러스 폴리뉴클레오타이드; 제2 이중성 폴리뉴클레오타이드에 작동가능하게 연결된 2A 펩타이드 또는 펩타이드-유사 코딩 서열을 포함하는 적어도 하나의 2A 카세트에 작동가능하게 연결된 제1 이중성 폴리뉴클레오타이드에 작동가능하게 연결된 조절 도메인을 포함하는 카세트로서, 상기 카세트가 5' 내지 3' LTR 및 3' 내지 레트로바이러스 외피를 코딩하는 env 핵산 도메인에 위치하고, 2A 카세트가 하류에 있고, 제1 이중성 폴리뉴클레오타이드에 작동가능하게 연결되는 카세트; 및 표적 세포에서 역전사, 패키징 및 통합에 필요한 시스-작용 서열(cis-acting integration)을 포함하는 재조합 복제가능 레트로바이러스를 제공한다. 일 구현예에서, 외피는 양친매성, 다중지향성, 이중향성, 10A1, GALV, 비비원숭이 내인성 바이러스, RD114, 랫도바이러스, 알파바이러스, 홍역 또는 인플루엔자 바이러스 외피 중 하나로부터 선택된다. 또다른 구현예에서, 레트로바이러스 폴리뉴클레오타이드 서열은 무린 백혈병 바이러스(MLV), 몰로니 무린 백혈병 바이러스(MoMLV), 고양이 백혈병 바이러스(FeLV), 비비원숭이 내인성 레트로바이러스(BEV), 돼지 내인성 바이러스(PERV), 고양이 유래 레트로바이러스 RD114, 다람쥐 원숭이 레트로바이러스, 이중향성 무린 백혈병 바이러스-관련 바이러스(XMRV), 조류 망상내피증 바이러스(REV) 또는 긴팔원숭이 백혈병 바이러스(GALV)로 구성된 그룹으로부터 선택되는 바이러스로부터 유래된다. 또다른 구현예에서, 레트로바이러스는 감마레트로바이러스이다. 또다른 구현예에서, 표적 세포는 포유류 세포이다. 또다른 구현예에서, 2A 펩타이드 또는 펩타이드-유사 코딩 서열은 서열식별번호: 1의 서열을 함유하는 펩타이드를 코딩한다. 또다른 구현예에서, 2A 펩타이드 또는 펩타이드-유사 코딩 서열은 서열식별번호: 55 내지 125 중 어느 하나에 개시된 펩타이드를 코딩한다. 또다른 구현예에서, 2A 펩타이드 또는 펩타이드-유사 코딩 서열은 서열식별번호: 8 내지 19 중 어느 하나에 개시된 서열을 포함한다. 상기 구현예 중 어느 하나에서, 표적 세포는 폐암 세포, 결장-직장암 세포, 유방암 세포, 전립선 암 세포, 요로 암 세포, 자궁암 세포, 뇌암 세포, 두경부암 세포, 췌장암 세포, 흑색종 세포, 위암 및 난소 암 세포로 구성된 그룹으로부터 선택된다. 상기 임의의 구현예에서, 프로모터 서열은 성장 조절 유전자와 관련된다.

상기 구현예 중 임의의 또다른 구현예에서, 프로모터 서열은 조직-특이적인 프로모터 서열을 포함한다. 추가의 구현예에서, 조직-특이적 프로모터 서열은 하나 이상의 안드로겐 반응 요소(ARE)를 포함한다. 상기 임의의 추가의 구현예에서, 프로모터는 뉴클레오타이드 1 내지 약 뉴클레오타이드 582에서 서열식별번호: 2에 개시된 서열을 갖는 CMV 프로모터를 포함하고, 하나 이상의 핵산 염기에 대한 변형을 포함할 수 있으며, 전사를 지시 및 개시할 수 있다. 상기 것들 중 임의의 또다른 구현예에서, 프로모터는 CMV-R-U5 도메인 폴리뉴클레오타이드를 포함한다. 또다른 구현예에서, CMV-R-U5 도메인은 MLV R-U5 영역에 연결된 인간 사이토메갈로바이러스로부터의 즉시 초기(immediately early) 프로모터를 포함한다. 또다른 구현예에서, CMV-R-U5 도메인 폴리뉴클레오타이드는 약 뉴클레오타이드 1 내지 약 뉴클레오타이드 1202의 서열식별번호: 2에 개시된 서열, 또는 뉴클레오타이드 1 내지 약 1202의 서열식별번호: 2에 개시된 서열과 적어도 95% 동일한 서열을 포함하며, 상기 폴리뉴클레오타이드는 이에 작동가능하게 연결된 핵산 분자의 전사를 촉진한다. 상기 것들 중 임의의 또다른 구현예에서, gag 폴리뉴클레오타이드는 감마레트로바이러스로부터 유래된다. 추가의 구현예에서, gag 핵산 도메인은 서열식별번호: 2의 약 뉴클레오타이드 번호 1203 내지 약 뉴클레오타이드 2819의 서열, 또는 이 서열과 적어도 95%, 98%, 99% 또는 99.8%의 동일성을 갖는 서열을 포함한다. 상기 임의의 또다른 구현예에서, 폴리뉴클레오타이드의 pol 도메인은 감마레트로바이러스로부터 유래된다. 추가의 구현예에서, pol 도메인은 서열식별번호: 2의 약 뉴클레오타이드 번호 2820 내지 약 뉴클레오타이드 6358의 서열, 또는 이 서열과 적어도 95%, 98%, 99% 또는 99.9%의 동일성을 갖는 서열을 포함한다. 상기 것들 중 임의의 다른 구현예에서, env 도메인은 서열식별번호: 2의 약 뉴클레오타이드 번호 6359 내지 약 뉴클레오타이드 8323의 서열 또는 이들과 적어도 95%, 98%, 99% 또는 99.8%의 동일성을 갖는 서열을 포함한다. 상기 것들 중 임의의 추가의 구현예에서, 3' LTR은 감마레트로바이러스로부터 유래된다. 추가의 구현예에서, 3' LTR은 U3-R-U5 도메인을 포함한다. 또다른 구현예에서, 3' LTR은 약 뉴클레오타이드 9111 내지 약 11654의 서열식별번호: 2에 개시된 서열, 또는 이들과 적어도 95%, 98% 또는 99.5% 동일한 서열을 포함한다. 상기 임의의 또다른 구현예에서, 이중 핵산 서열은 생물학적 반응 조절제 또는 면역기능강화 사이토카인을 코딩한다. 추가의 구현예에서, 면역기능강화 사이토카인은 인터루킨 1 내지 38, 인터페론, 종양 괴사 인자(TNF) 및 과립구-대식세포-콜로니 자극 인자(GM-CSF)로 구성된 그룹으로부터 선택된다. 추가의 구현예에서, 면역기능강화 사이토카인은 인터페론 감마이다. 상기 것들 중 임의의 또다른 구현예에서, 이중 핵산은 무독성 프로드럭을 독성 약물로 전환시키는 폴리펩타이드를 코딩한다. 추가의 구현예에서, 무독성 프로드럭을 독성 약물로 전환시키는 폴리펩타이드는 티미딘 키나아제, 퓨린 뉴클레오사이드 포스포릴라제(PNP) 또는 시토신 데아미나제이다. 상기 것들 중 임의의 또다른 구현예에서, 이중 핵산 서열은 수용체 도메인, 항체 또는 항체 단편을 코딩한다. 일 구현예에서, 제2 카세트는 억제성 폴리뉴클레오타이드를 포함한다. 추가의 구현예에서, 억제성 폴리뉴클레오타이드는 miRNA, RNAi 또는 siRNA 서열을 포함한다.

[0011] 본원은 또한 상기 구현예 중 어느 하나에서 기술된 바와 같은 레트로바이러스를 생산하기 위한 재조합 레트로바이러스 폴리뉴클레오타이드 계통을 제공한다. 일 구현예에서, 폴리뉴클레오타이드는 GSG 링커 코딩 서열을 갖거나 갖지 않는 2A 펩타이드 또는 펩타이드-유사 코딩 서열을 인-프레임으로 갖는 MLV 4070A 외피 단백질 유전자, 및 2A 펩타이드 또는 2A-유사 코딩 서열을 인-프레임으로 갖는 제2 유전자를 포함한다. 또다른 구현예에서, 폴리뉴클레오타이드는 GSG 링커 코딩 서열을 갖거나 갖지 않는 2A 펩타이드 또는 펩타이드-유사 코딩 서열을 인-프레임으로 갖는 MLV 10A1 외피 단백질 유전자, 및 2A 펩타이드 또는 2A-유사 코딩 서열을 인-프레임으로 갖는 제2 유전자를 포함한다. 또다른 구현예에서, 폴리뉴클레오타이드는 GSG 링커 코딩 서열을 갖거나 갖지 않는 2A 펩타이드 또는 펩타이드-유사 코딩 서열을 인-프레임으로 갖는 XMRV 외피 단백질 유전자, 및 2A 펩타이드 또는 2A-유사 코딩 서열을 인-프레임으로 갖는 제2 유전자를 포함한다. 또다른 구현예에서, 폴리뉴클레오타이드는 GSG 링커 코딩 서열을 갖거나 갖지 않는 2A 펩타이드 또는 펩타이드-유사 코딩 서열을 인-프레임으로 갖는 비-프렌드(non-Friend) MLV 외피 단백질 유전자, 및 2A 펩타이드 또는 2A-유사 코딩 서열을 인-프레임으로 갖는 제2 유전자를 포함한다. 상기 중 임의의 다른 구현예에서, 이중성 폴리뉴클레오타이드는 분비성, 막, 세포질, 핵 또는 세포-구획-특이적 단백질이다.

[0012] 상기 구현예 중 임의의 구현예에서, 재조합 레트로바이러스 및/또는 이중성 폴리뉴클레오타이드는 인간 APOBEC 과돌연변이(hypermutation)를 일으키기 쉬운 트립토판 코돈을 제거하도록 조작된다. 일 구현예에서, 이중성 폴리뉴클레오타이드는 시토신 데아미나제 활성을 갖는 폴리펩타이드를 코딩한다. 추가의 구현예에서, 시토신 데아미나제 활성을 갖는 폴리펩타이드는 서열식별번호: 29의 폴리펩타이드를 코딩하며, 여기서 X는 트립토판을 제외한 임의의 아미노산이다.

[0013] 또한, 본 발명은 비-감수성 코돈에 APOBEC 과돌연변이가 일어나기 쉬운 레트로바이러스 폴리뉴클레오타이드에서 코돈을 조작함으로써 인간 APOBEC에 의한 불활성화에 저항성이 있는 재조합 복제가능 레트로바이러스를 제공한다. 일 구현예에서, APOBEC 과돌연변이되기 쉬운 코돈은 트립토판 아미노산을 코딩한다. 또다른 구현예에서, 재

조합 레트로바이러스는 env 유전자의 하류에 IRES 카세트, 프로모터 카세트 및/또는 2A 펩타이드 카세트를 포함한다.

[0014] 본원은 또한, 이중성 폴리뉴클레오타이드가 발현되고 이중성 폴리뉴클레오타이드가 프로드럭을 세포독성 약물로 전환시키는 단백질을 코딩하는 조건하에, 상기 구현예 중 어느 하나에 기술된 바와 같은 레트로바이러스와 개체를 접촉시키는 단계를 포함하는 세포 증식성 장애의 치료 방법을 제공한다.

[0015] 본 발명의 하나 이상의 구현예의 상세한 설명은 첨부된 도면 및 이하의 설명에서 개시된다. 다른 특징, 목적 및 이점은 상세한 설명 및 도면 및 청구범위로부터 명백해질 것이다.

도면의 간단한 설명

[0016] 도 1은 구제역 바이러스(F2A), 마비엄 A 바이러스(E2A), 토세아 아시그나(*Thosea asigna*) 바이러스(T2A) 및 돼지 테스코바이러스(porcine teschovirus)-1(P2A)의 2A 영역의 아미노산 서열의 서열 정렬(서열식별번호: 55 내지 58)을 나타낸다.

도 2는 바이러스의 상이한 부류에 존재하는 2A 펩타이드 서열의 서열 정렬(서열식별번호: 59 내지 125)을 나타낸다.

도 3은 pAC3-(x)2A-GFPm 및 pAC3-(x)2A-yCD2 벡터 세트의 클로닝 계획을 나타낸다. 검정색 박스는 깊은 어셈블리 클로닝(Gibson Assembly Cloning)에서 사용되는 중첩 서열을 나타내며; (x)는 마비엄 A(E), 구제역 바이러스(F), 돼지 테스코바이러스-1(P) 또는 토세아 아시그나 바이러스(T)로부터의 2A 펩타이드를 나타낸다.

도 4는 U87-MG 세포에서 일시적으로 형질감염된 HEK293T 세포로부터 생산된 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 벡터의 복제 동역학을 나타낸다.

도 5는 U87-MG 세포에서 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 벡터의 평균 형광 강도(MFI)로 나타난 GFP 발현 수준을 나타낸다. 표시된 백분율은 RRV-IRES-GFP의 MFI에 비례한다.

도 6은 U87-MG 세포에서의 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 벡터의 벡터 안정성을 나타낸다.

도 7은 최대 감염된 U87-MG 세포로부터 생산된 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 벡터의 역가를 나타낸다. 개시된 역가 값은 2회의 독립적인 실험에서 나온 값이다.

도 8은 최대 감염된 U87-MG 세포로부터 생산된 후, 나이브 U87-MG 세포에서의 감염 사이클에서 생산된 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 벡터의 복제 동역학을 나타낸다.

도 9는 제2 감염 사이클에서 최대 감염된 U87-MG 세포내 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 벡터에서의 GFPm 발현의 MFI를 나타낸다. 표시된 백분율은 RRV-IRES-GFP의 MFI에 비례한다.

도 10은 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 감염된 U87-MG 세포로부터의 세포 용해물의 항-GFP 면역블롯을 나타낸다. 약 120 KDa에서 검출된 단백질 밴드는 바이러스 외피-GFPm 융합 다단백질을 나타낸다. 약 27 KDa에서 검출된 단백질 밴드는 Env-GFPm 융합 다단백질로부터 분리된 천연 GFP 또는 2A-GFPm 단백질을 나타낸다.

도 11은 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 감염된 U87-MG 세포로부터의 세포 용해물의 항-gp70 면역블롯을 나타낸다. 약 120 KDa에서 검출된 단백질 밴드는 바이러스 외피-GFPm 융합 다단백질을 나타낸다. 약 75 KDa에서 검출된 단백질 밴드는 융합 다단백질로부터 분리된 Pr85/gp70 바이러스 외피 단백질을 나타낸다.

도 12는 각각 항-gp70 및 항-TM 항체에 의해 검출된 gp70 및 p15E 서브유닛을 함유하는 바이론(viron)-관련, 적절히 가공된 바이러스 외피 단백질의 면역블롯을 나타낸다. 항-p15E 항체는 전구체 TM 서브유닛 p15E 및 R-펩타이드 분열 TM 서브유닛 p12E 모두를 검출한다.

도 13은 최대 감염된 U87-MG 세포로부터 생산된 RRV-P2A-yCD2, RRV-T2A-yCD2, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 벡터의 역가를 나타낸다.

도 14는 RRV-P2A-yCD2, RRV-T2A-yCD2, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 감염된 U87-MG 세포로부터의 세포 용해물의 항-yCD2 면역블롯을 나타낸다. 약 110 KDa에서 검출된 단백질 밴드는 바이러스 외피-GFPm 융합 다단백질을 나타낸다. 약 15 KDa에서 검출된 단백질 밴드는 Env-yCD2 융합 다단백질로부터 분리된 yCD2 또는 2A-yCD2 단백질을 나타낸다.

도 15는 RRV-P2A-yCD2, RRV-T2A-yCD2, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 감염된 U87-MG 세포로부터의 세

포 용해물의 항-gp70 면역블롯을 나타낸다. 약 110 KDa에서 검출된 단백질 밴드는 바이러스 외피-yCD2 융합 다단백질을 나타낸다. 약 75 KDa에서 검출된 단백질 밴드는 융합 다단백질로부터 분리된 Pr85/gp70 바이러스 외피 단백질을 나타낸다.

도 16은 16개의 감염 사이클에 걸친 U87-MG 세포내 RRV-P2A-yCD2 및 RRV-T2A-yCD2, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 벡터의 장기간 벡터 안정성을 나타낸다. pDNA는 양성 대조군으로 포함된 pAC3-P2A-yCD2 및 pAC3-T2A-yCD2, pAC3-GSG-P2A-yCD2 및 pAC3-GSG-T2A-yCD2의 플라스미드 DNA이다.

도 17은 각각 항-gp70 및 항-TM 항체에 의해 검출된 gp70 및 p15E 서브유닛을 함유하는 바이론-관련, 적절히 가공된 바이러스 외피 단백질의 면역블롯을 나타낸다. 항-p15E 항체는 전구체 TM 서브유닛 p15E 및 R-펩타이드 분열된 TM 서브유닛 p12E 모두를 검출한다.

도 18은 5-FC 치료 후 7일째에 RRV-IRES-GFP, RRV-P2A-yCD2 및 RRV-T2A-yCD2, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 감염된 U87-MG 세포의 5-FC-매개된 살해를 나타낸다. 세포 생존율은 5-FC 처리되지 않았지만, RRV-감염된 세포에 비례하여 계산되었다. 5-FC의 비(non)-5-FU 매개된 세포독성 효과의 농도를 결정하기 위한 대조군으로서 나이브(Naive) U87-MG 세포가 포함되었다.

도 19는 비리온 조립 및 성숙 동안의 MLV 바이러스 외피 단백질 프로세싱을 나타낸다. 일반적으로, 천연 MLV 외피 단백질의 프로세싱은 감염된 숙주 세포에서 발생하는, 전구체 단백질 Pr85가 gp70(SU) 및 p15E(TM) 서브유닛으로 분열되는 단계를 포함한다. Pr85의 분열은 숙주 세포로부터의 출아(budding) 동안 바이러스 외피 단백질을 바이론(viron)에 효과적으로 혼입시키기 위해 필요하다. 비리온이 숙주 세포 막으로부터 떨어져 나올 때, 비리온은 감염성을 갖기 위해 성숙 과정을 거친다. MLV 비리온 돌연변이의 과정 중 하나는 바이러스 프로테아제에 의해 외피 단백질의 TM 서브유닛의 C-말단에 위치한 R-펩타이드의 제거를 포함한다. 2A 펩타이드는 R-펩타이드의 C-말단에 인-프레임으로 발현되어, 2A 펩타이드의 서열에 따라 R 펩타이드의 길이를 16개 아미노산에서 적어도 32개의 아미노산으로 증가시킨다. R-펩타이드의 길이가 2A 펩타이드 서열의 첨가에 의해 길어지지만, 2A 펩타이드는 R 펩타이드의 분열과 동시에 제거되어 기능성 외피 단백질을 생성할 것이다.

도 20a 내지 20b는 최대 감염된 Tu-2449 세포에서 RRV-P2A-yCD2, RRV-GSG-P2A-yCD2, RRV-T2A-yCD2 및 RRV-GSG-yCD2의 (A) yCD2 단백질 발현 및 (B) 5-FC 민감도를 나타낸다.

도 21은 RRV-IRES-yCD2 + 5-FC로 처리한 종양과 비교하여, RRV-GSG-T2A-yCD2 + 5-FC로 처리한 종양의 종양 성장 지연을 나타낸다.

도 22a 내지 22b는 (A) RRV-GSG-T2A-GMCSF-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2-GSG-P2A-GMCSF 일시적으로 형질감염된 HEK293T 세포로부터의 세포 용해물의 항-yCD2 면역블롯을 나타낸다. 약 110 KDa에서 검출된 단백질 밴드는 바이러스 외피-GFPm 융합 다단백질을 나타낸다. 약 15 KDa에서 검출된 단백질 밴드는 융합 다단백질로부터 분리된 yCD2 또는 2A-yCD2 단백질을 나타낸다. (B) GMCSF는 RRV-GSG-T2A-GMCSF-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2-GSG-P2A-GMCSF 일시적으로 형질감염된 HEK293T 세포로부터 배양 배지에 분비되었다.

도 23은 최대 감염된 U87-MG 세포로부터 생산된 RRV-P2A-TKO, RRV-T2A-TKO, RRV-GSG-P2A-TKO 및 RRV-GSG-T2A-TKO 벡터의 역할을 나타낸다.

도 24는 RRV-P2A-TKO, RRV-T2A-TKO, RRV-GSG-P2A-TKO 및 RRV-GSG-T2A-TKO 감염된 U87-MG 세포로부터의 세포 용해물의 항-HSV-tk 면역블롯을 나타낸다.

도 25는 16개의 감염 사이클 동안 U87-MG 세포내 RRV-P2A-TKO, RRV-T2A-TKO, RRV-GSG-P2A-TKO 및 RRV-GSG-T2A-TKO 벡터의 벡터 안정성을 나타낸다. pAC3-P2A-TKO 플라스미드 DNA를 양성 대조군으로 사용한다.

도 26a 내지 26b는 상이한 투여량의 RRV-P2A-TKO, RRV-T2A-TKO, RRV-GSG-P2A-TKO 및 RRV-GSG-T2A-TKO 감염된 U87-MG 세포의 GCV 민감도를 나타낸다.

도 27a 내지 27b는 일시적으로 형질감염된 293T 세포에서 PDL1scFv 및 PDL1scFvFc 단백질 발현 및 Env-scFv 및 Env-ScFvFc 다단백질의 분리 효율을 나타낸다. (A) pAC3-GSG-T2A-PDL1scFv, pAC3-GSG-T2A-PDL1scFvFc, pAC3-GSG-T2A-PDL1scFv-태그, pAC3-GSG-T2A-PDL1scFvFc-태그로 일시적으로 형질감염된 HEK293T 세포로부터의 scFv-태그(~30 KDa) 및 scFvFc-태그(~55 Kd) 단백질 발현. (B) 일시적으로 형질감염된 293T 세포로부터의 세포 용해물의 항-2A 면역블롯. 약 110 KDa 이상에서 검출된 단백질 밴드는 Env-scFv 및 Env-ScFvFc 융합 다단백질을 나타낸다. 약 85 KDa에서 검출된 단백질 밴드는 융합 다단백질로부터 분리된 Pr85 바이러스 외피 단백질을 나타낸다.

고, 약 15 KDa에서 검출된 단백질 밴드는 Pr85 바이러스 외피 단백질로부터 가공된 p15E-2A 단백질을 나타낸다.

도 28은 혈액 및 종양으로부터 채취된 20명의 환자 샘플로부터의 RRV(Toca 511, Tocagen Inc.) 점 돌연변이의 메타-분석을 나타낸다. 최소 1%의 검출 빈도를 갖는 품질 필터를 통과하는 모든 점 돌연변이를 각각의 샘플에 대해 준수시키고, 가능한 쌍 방식의 점 돌연변이의 각각을 계산하여 상대 빈도로 변환하였다. 이 플롯은 샘플 유형에 따라 그룹핑된, 상이한 점 돌연변이들의 상대적 빈도를 보여준다: 혈액 = 환자의 혈액 샘플, 종양 = Toca 511의 IV 투여 후 절제된 종양, 재발_종양(re_tumor) = 초기 절제시에 Toca 511로 치료한 환자의 재-절제된 종양.

발명을 실시하기 위한 구체적인 내용

- [0017] 본 명세서 및 첨부된 청구범위에서 사용된 단수 형태("a", "and" 및 "the")는 문맥에 따라 다르게 지시하지 않는 한 복수 대상을 포함한다. 따라서, 예를 들어, "하나의 세포"에 대한 언급은 복수의 상기 세포를 포함하고, "백터"에 대한 언급은 하나 이상의 백터에 대한 언급을 포함한다.
- [0018] 또한, " 또는"의 사용은 달리 언급되지 않는 한 "및/또는"을 의미한다. 마찬가지로, "포함한다(comprise)", "포함하다(comprises)", "포함하는(comprising)", "포함하다(include)", "포함하다(includes)" 및 "포함하는(including)"은 서로 교환가능하며, 제한하려는 것은 아니다.
- [0019] 다양한 구현예들의 설명이 "포함하는"이라는 용어를 사용하는 경우, 당업자들은 일부 특정 예들에서, 구현예들은 "본질적으로 구성된" 또는 "구성된"이라는 용어를 사용하여 대안적으로 기술될 수 있음을 이해할 것이다.
- [0020] 달리 정의되지 않는 한, 본원에서 사용되는 모든 기술 및 과학 용어는 본 발명이 속하는 기술 분야의 통상의 지식을 가진 자에게 일반적으로 이해되는 것과 동일한 의미를 갖는다. 본원에 기술된 것과 유사하거나 균등한 방법 및 물질이 개시된 방법의 실시 및 조성물에 사용될 수 있지만, 예시적인 방법, 장치 및 재료들이 본원에 기술되어 있다.
- [0021] 백터, 프로모터 및 많은 다른 관련 주제의 사용을 포함하여, 본원에서 유용한 분자 생물학적 기술을 기술하는 일반 텍스트는 다음을 포함한다: Berger and Kimmel, *Guideto Molecular Cloning Techniques, Methods in Enzymology Volume 152*, (Academic Press, Inc., San Diego, Calif.) ("Berger"); Sambrook 등, *Molecular Cloning--A Laboratory Manual*, 2d ed., Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989 ("Sambrook"); *Current Protocols in Molecular Biology*, F. M. Ausubel 등, eds., CurrentProtocols, a joint venture between Greene PublishingAssociates, Inc. and John Wiley & Sons, Inc., (supplemented through 1999) ("Ausubel"); and S.Carson, H.B.Miller & D.S.Witherow and *Molecular Biology Techniques: A Classroom Laboratory Manual*, Third Edition, Elsevier, San Diego (2012). 중합효소 연쇄 반응(PCR), 리가아제 연쇄 반응(LCR), Qβ-레플리카제 증폭 및, 예를 들어 본 발명의 상동성 핵산을 생성하기 위한 다른 RNA 중합효소 매개 기술(예를 들어, NASBA)을 포함하는 시험관내 증폭 방법을 통해 숙련된 기술자를 지시하기에 충분한 프로토콜의 예는 Berger, Sambrook 및 Ausubel 뿐만 아니라 Mullis 등(1987) 미국 특허 제4,683,202호; Innis 등, eds. (1990) *PCR Protocols: A Guideto Methods and Applications* (Academic Press Inc. San Diego, Calif.) ("Innis"); Arnheim & Levinson (Oct. 1, 1990) *C&E* 36-47; *The Journal Of NIH Research* (1991) 3: 81-94; Kwoh 등 (1989) *Proc. Natl. Acad. Sci. USA* 86: 1173; Guatelli 등 (1990) *Proc. Nat'l. Acad. Sci. USA* 87: 1874; Lomell 등 (1989) *J. Clin. Chem* 35: 1826; Landegren 등 (1988) *Science* 241: 1077-1080; Van Brunt (1990) *Biotechnology* 8: 291-294; Wuand Wallace (1989) *Gene* 4:560; Barringer 등 (1990) *Gene* 89:117; and Sooknanan and Malek (1995) *Biotechnology* 13: 563-564에서 발견된다. 시험관내 증폭된 핵산을 클로닝하기 위한 개선된 방법은 문헌 Wallace 등, 미국 특허 제5,426,039호에 기술되어 있다. PCR에 의해 큰 핵산을 증폭시키기 위한 개선된 방법은 Cheng 등(1994) *Nature* 369: 684-685 및 본원에 인용된 참고 문헌에 요약되어 있으며, 여기서는 최대 40kb의 PCR 앰플리콘(amplicon)이 생성된다. 당업자는 본질적으로 임의의 RNA가 역전사 효소 및 중합효소를 사용하여 제한효소 소화, PCR 팽창 및 시퀀싱에 적합한 이중 가닥 DNA로 전환될 수 있다는 것을 이해할 것이다. 예를 들어, 위의 Ausubel, Sambrook 및 Berger를 참조한다.
- [0022] 본 명세서 전반에 걸쳐 개시된 공보들은 본원의 출원일 이전에 그들의 개시를 위해서만 제공된다. 본 명세서의 어떠한 내용도 발명자가 선행 공개에 의해 상기 개시보다 시기적절할 자격이 없다는 것을 인정하는 것으로 해석되어서는 안된다.
- [0023] 본원은 세포 또는 개체의 유전자 또는 단백질 전달에 유용한 방법 및 조성물을 제공한다. 상기 방법 및 조성

물은 암 및 기타 세포 증식성 질병 및 장애를 포함하는 개체의 다양한 질병 및 장애를 치료하는데 사용될 수 있다. 본원은 세포로의 유전자 전달을 위한 복제가능 레트로바이러스 벡터를 제공한다.

- [0024] 용어 "벡터", "벡터 구조체" 및 "발현 벡터"는 DNA 또는 RNA 서열(예를 들어, 외래 유전자)이 숙주 세포로 도입되어, 숙주를 형질전환시키고 도입된 서열의 발현(예를 들어, 전사 및 번역)을 촉진시키는 비핵질을 의미한다. 벡터는 전형적으로 단백질을 코딩하는 외래 DNA가 제한 효소 기술에 의해 삽입되는 DNA 또는 RNA를 포함한다. 공통 유형의 벡터는 일반적으로 추가(외래) DNA를 쉽게 받아들일 수 있는 이중 가닥 DNA의 자체 함유 분자이고, 적당한 숙주 세포에 쉽게 도입될 수 있는 "플라스미드"이다. 플라스미드 및 균류 벡터(fungal vector)를 포함하는 다수의 벡터가 다양한 진핵 및 원핵 숙주에서 복제 및/또는 발현용으로 기술되어 있다. 비-제한적인 예로는 본원에 개시되거나 인용되거나, 그렇지 않으면 관련 분야의 당업자에게 공지된 방법을 사용하여, pKK 플라스미드(Clonetech), pUC 플라스미드, pET 플라스미드(Novagen, Inc., Madison, Wis.), pRSET 또는 pREP 플라스미드(Invitrogen, San Diego, Calif.), 또는 pMAL 플라스미드(New England Biolabs, Beverly, Mass.) 및 다수의 적절한 숙주 세포를 포함한다. 재조합 클로닝 벡터는 종종 클로닝 또는 발현을 위한 하나 이상의 복제 시스템, 숙주에서의 선택을 위한 하나 이상의 마커, 예를 들어 항생제 내성 및 하나 이상의 발현 카세트를 포함할 것이다.
- [0025] 재조합 복제가능 레트로바이러스 벡터 또는 레트로바이러스 복제 벡터(RRV)는 레트로바이러스 계열의 바이러스의 구성원을 기반으로 하는 벡터를 지칭한다. 레트로바이러스의 구조는 하기에서 보다 완전하게 기술된 바와 같이 특징화된다. 상기 벡터는 이중 유전자 또는 서열을 삽입함으로써 모(parent) 바이러스가 비-자연발생 RRV로 변형되도록 재조합 유전자 기술을 사용하여 조작될 수 있다. 상기 변형은 시험관내 또는 생체내에서 숙주 세포로 발현되는 유전자를 전달할 수 있도록 벡터에 속성을 제공할 수 있다.
- [0026] 레트로바이러스는 다양한 방법으로 분류되었지만, 지난 10년 동안 그 명명법이 표준화되었다(ICTVdb - The Universal Virus Database, v 4 on the World Wide Web (www) at ncbi.nlm.nih.gov/ICTVdb/ICTVdb/ 및 교과서 "Retroviruses" Eds Coffin, Hughsand Varmus, Cold Spring Harbor Press 1997 참조; 이들의 개시 내용은 본원에 참고로 포함됨). 일 구현예에서, 복제가능 레트로바이러스 벡터는 오르토투트로바이러스 또는 보다 전형적으로 감마 레트로바이러스 벡터를 포함할 수 있다.
- [0027] 레트로바이러스는 그들의 유전 물질을 복제하는 방법에 의해 정의된다. 복제하는 동안 RNA는 DNA로 변환된다. 세포 감염 후 DNA의 이중-가닥 분자는 역전사로 알려진 분자 과정에 의해 바이러스 입자에서 운반되는 두 RNA 분자로부터 생성된다. DNA 형태는 프로바이러스(provirus)로서 숙주 세포 게놈에서 공유적으로 통합되며, 이로부터 바이러스 RNA가 세포 및/또는 바이러스 인자의 도움을 받아 발현된다. 발현된 바이러스성 RNA는 입자로 패키징되어 감염성 비리온으로서 방출된다.
- [0028] 레트로바이러스 입자는 2개의 동일한 RNA 분자로 구성된다. 각각의 야생형 게놈은 5' 말단에서 캡핑되고 3' 꼬리에서 폴리아데닐화된 양성 센스, 단일-가닥 RNA 분자를 갖는다. 이배체 바이러스 입자는 gag 단백질, 바이러스 효소(pol 유전자 산물) 및 gag 단백질의 '코어' 구조내 숙주 tRNA 분자와 착화된 2개의 RNA 가닥을 함유한다. 이 캡시드를 둘러싸고 보호하는 것은 지질 이중층이며, 숙주 세포 막으로부터 유래하고, 바이러스 외피(env) 단백질을 함유한다. env 단백질은 바이러스에 대한 세포 수용체에 결합하며, 입자는 전형적으로 수용체-매개 엔도시토시스 및/또는 막 융합을 통해 숙주 세포에 유입된다.
- [0029] 표적화된 세포 내로 바이러스 입자가 방출된 후, 외측 외피가 흘러지고(shed), 역전사에 의해 바이러스 RNA가 DNA로 복사된다. 이것은 pol 영역에 의해 코딩된 역전사 효소에 의해 촉매되고, DNA 합성을 위한 프라이머로서 비리온으로 패키징된 숙주 세포 tRNA를 사용한다. 이런 방식으로 RNA 게놈은 보다 복잡한 DNA 게놈으로 변환된다.
- [0030] 역전사에 의해 생성된 이중-가닥의 선형 DNA는 핵에서 환형되어야 할 수도 있고 그렇지 않을 수도 있다. 프로바이러스는 이제 긴 말단 반복(LTR)으로 알려진 어느 한 끝에 2개의 동일한 반복을 가지고 있다. 두개의 LTR 서열의 말단은 프로바이러스가 LTR의 말단으로부터 숙주 DNA 두 염기쌍(bp)에 항상 결합되도록 통합을 촉매하는 pol 생성물인, 인테그라아제(integrase) 단백질에 의해 인식되는 부위를 생성한다. 세포 서열의 복제는 수송가능한 유전 요소의 통합 패턴을 연상시키는, 양쪽 LTR의 말단에서 보인다. 레트로바이러스는 숙주 DNA의 많은 부위에서 DNA를 통합할 수 있지만, 다른 레트로바이러스는 통합 부위 선호도가 다르다. HIV-1과 유인원 면역결핍 바이러스(simian immunodeficiency virus) DNA는 발현된 유전자에 우선적으로 통합되며, 뮤린 백혈병 바이러스(MLV) DNA는 전사 개시 부위(TSS) 근처에 우선적으로 통합되며, 조류 육종 백혈증 바이러스(ASLV)와 인간 T 세포 백혈병 바이러스(HTLV) DNA는 거의 무작위로 통합하며, 이는 유전자에 대한 약한 선호도를 나타낸다(Derse D, 등(2007), J Virol 81: 6731-6741, Lewinski MK, 등 (2006), PLoS Pathog 2:e601).

- [0031] 통합된 바이러스 DNA의 전사, RNA 스플라이싱 및 번역은 숙주 세포 단백질에 의해 매개된다. 다양하게 스플라이싱된 전사물이 생성된다. 인간 레트로바이러스의 경우, HIV-1/2 및 HTLV-I/II 바이러스 단백질도 유전자 발현을 조절하는데 사용된다. 세포 및 바이러스 인자 사이의 상호작용은 바이러스 잠복성 조절 및 바이러스 유전자가 발현되는 시간적 서열의 요소이다.
- [0032] 레트로바이러스는 수평 및 수직으로 전달될 수 있다. 레트로바이러스의 효과적인 전염성 전달에는 바이러스 외피 단백질을 특이적으로 인식하는 수용체의 표적 세포에서의 발현이 필요하지만, 바이러스는 낮은 효율로 수용체-독립적인, 비특이적 도입경로를 사용할 수 있다. 일반적으로, 바이러스 감염은 수용체 마스킹 또는 하향-조절 때문에 세포 당 바이러스 계놈의 단일 또는 수개의 복제를 유도하며, 이는 차례로 중복감염에 대한 내성을 일으킨다(Ch3 p104 in "Retroviruses", JM Coffin, SH Hughes, & HE Varmus 1997 Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY; Fan 등 J.Virol 28:802, 1978). 조직 배양에서 상황을 조작함으로써, 어느 정도의 다중 감염을 얻는 것이 가능하지만, 전형적으로 5개 미만의 복제본/배수체 계놈이다. 또한, 표적 세포 유형은 바이러스가 결합되어 침투한 후 복제주기의 모든 단계를 지원할 수 있어야한다. 종적인 전염은 바이러스 계놈이 숙주의 생식계열내에 통합될 때 발생한다. 그후, 프로바이러스는 마치 세포 유전자 인 것처럼 대대로 이어질 것이다. 따라서, 내인성 프로바이러스는 빈번하게 잠복 해 있지만, 숙주가 적절한 약제에 노출되었을 때 활성화될 수 있다.
- [0033] "렌티바이러스"라는 용어는 역전사 효소를 함유하는 바이러스류를 기술하기 위한 통상적인 의미로 사용된다. 렌티바이러스에는 인간 면역결핍 바이러스(HIV) 1형 및 2형(HIV-1 및 HIV-2)과 유인원 면역결핍 바이러스(SIV)를 포함하는 "면역결핍 바이러스"가 포함된다.
- [0034] 온코바이러스는 역사적으로 바이러스 성숙 동안 전자 현미경 하에서 관찰되는 바와 같이, 입자 형태에 기초하여 그룹 A, B, C 및 D로 더욱 세분되었다. A형 입자는 감염된 세포의 세포질에서 볼 수 있는 B형 및 D형 바이러스의 미성숙 입자를 나타낸다. 이 입자는 전염성이 없다. B형 입자는 세포질내 A형 입자의 포위에 의해 원형질막으로부터 성숙한 비리온으로 성장한다. 막에서 그들은 긴 당단백질 스파이크가 돌출하는 75 nm의 토로이드 코어를 가지고 있다. 출아후, B-형 입자는 편심되어 위치하고 전자 밀도가 높은 코어를 함유하고 있다. 프로토타입 B형 바이러스는 마우스 유방 종양 바이러스(MMTV)이다. C형 바이러스에 감염된 세포에서는 세포질내 입자가 관찰되지 않을 수 있다. 대신, 성숙한 입자는 초승달 모양의 'C'-모양의 응축을 통해 세포 표면에서 직접 출아되고, 그 다음 자체적으로 닫히고, 원형질막으로 둘러싸인다. 외피 당단백질 스파이크가 균일하게 전자 밀도가 높은 코어와 함께 보여질 수 있다. 표면 원형질 막에서 출아가 일어날 수 있거나, 세포내 액포로 직접 출아될 수 있다. C형 바이러스는 가장 일반적으로 연구되었으며, 조류 및 무린 백혈병 바이러스(MLV)를 많이 포함한다. 소백혈병 바이러스(BLV) 및 인간 T-세포 백혈병 바이러스 I형과 II형(HTLV-I/II)은 세포 표면으로부터의 그들의 출아 형태로 인해 C형 입자로 유사하게 분류된다. 그러나 이들은 또한 무린 백혈병 바이러스(MLV)와 같은 프로토타입 C형 바이러스보다 복잡한 계놈 구조와 정육각형 형태를 가지고 있다. D형 입자는 감염된 세포 세포질에서 고리 모양의 구조를 나타낸다는 점에서 B형 입자와 유사하며, 세포 표면에서 출아하지만, 비리온은 짧은 표면 당단백질 스파이크를 포함하고 있다. 전자 밀도가 높은 코어는 또한 입자 내에 편심되어 위치한다. 메이슨 화이저 원숭이 바이러스(Mason Pfizer monkey virus, MPMV)는 프로토타입 D형 바이러스이다.
- [0035] 재조합 복제가능 레트로바이러스를 치료적으로 사용하는 많은 경우에, 재조합 복제가능 레트로바이러스에 의해 코딩되는 이식유전자의 높은 수준의 발현을 갖는 것이 유리하다. 예를 들어, 시토신 데아미나제 유전자와 같은 프로드럭 활성화 유전자의 경우, 프로드럭 5-FU의 5-FU로의 전환이 보다 효율적이 되도록 세포에서 CD 단백질의 높은 수준의 발현을 갖는 것이 유리하다. 마찬가지로 siRNA 또는 shRNA의 발현 수준이 높으면 표적 유전자 발현을 보다 효과적으로 억제할 수 있다. 또한 사이토카인 또는 단일 사슬 항체(scAb)의 경우, 높은 수준의 사이토카인 또는 scAb를 발현하는 것이 유리하다. 또한, 벡터 또는 이식유전자의 활성을 불활성시키거나 손상시키는 벡터의 일부 복제물에 돌연변이가 있는 경우, 이는 손상되지 않은 이식유전자의 효율적인 발현의 높은 확률을 제공하기 때문에 표적 세포내에 벡터의 다수의 복제물을 갖는 것이 유리하다 .
- [0036] 상기 언급된 바와 같이, 통합된 DNA 중간체는 프로바이러스(provirus)로 지칭된다. 이전의 유전자 치료 또는 유전자 전달 시스템은 적절한 헬퍼 바이러스의 존재하에, 또는 오염된 헬퍼 바이러스의 동시 생성없이 캡시드형성(encapsidation)을 가능하게하는 적당한 서열을 함유하는 세포주에서, 프로바이러스 및 조립체를 감염성 바이러스로 전사하는 것을 요구하는 방법 및 레트로바이러스를 사용한다. 이하에 설명되는 바와 같이, 캡시드형성을 위한 서열이 계놈내에 제공되어 유전자 전달 또는 치료를 위한 복제가능 레트로바이러스 벡터를 제공하기 때문에, 헬퍼 바이러스는 본 발명의 재조합 레트로바이러스의 생산에 요구되지 않는다.

- [0037] 본 발명의 레트로바이러스 게놈 및 프로바이러스성 DNA는 적어도 3개의 유전자: *gag*, *pol* 및 *env*를 가지며, 이들 유전자는 하나 또는 2개의 긴 말단(LTR) 반복에 의해 플랭크될(flank) 수 있거나, 프로바이러스에서 2개의 긴 말단 반복(LTR) 및 *psi*와 같은 시스-작용 서열을 함유하는 서열에 의해 플랭크된다. *gag* 유전자는 내부 구조체(매트릭스, 캡시드 및 뉴클레오캡시드) 단백질을 코딩하며; *pol* 유전자는 RNA-지향 DNA 중합효소(역전사 효소), 프로테아제 및 인테그라제를 코딩하며; *env* 유전자는 바이러스 외피 당단백질을 코딩한다. 5' 및/또는 3' LTR은 비리온 RNA의 전사 및 폴리아데닐화를 촉진시키는 역할을 한다. LTR은 바이러스 복제에 필요한 모든 다른 시스-작용 서열을 함유한다. 렌티바이러스는 *vif*, *vpr*, *tat*, *rev*, *vpu*, *nef* 및 *vpx*(HIV-1, HIV-2 및/또는 SIV에서)를 포함하는 추가 유전자를 가지고 있다.
- [0038] 5' LTR에 인접하여 게놈의 역전사(tRNA 프라이머 결합 부위) 및 바이러스 RNA의 입자 내로의 효과적인 캡시드형성(Psi 부위)에 필요한 서열이 있다. 캡시드형성(또는 레트로바이러스 RNA를 감염성 비리온으로 패키징)하는데 필요한 서열이 바이러스 게놈에서 없으면, 결과적으로 게놈 바이러스 RNA의 캡시드형성을 방해하는 *cis* 결핍이 생긴다. 이러한 유형의 변형된 벡터는 전형적으로 이전의 유전자 전달 시스템(즉, 비리온의 캡시드형성에 필요한 요소가 결여된 시스템)에서, 복제되지 않은 패키징가능한 RNA 게놈을 포장하는 바이러스 단백질을 트랜스로 제공하는 '헬퍼' 요소로서 사용되어왔다.
- [0039] "발현하다" 및 "발현"이란 용어는 유전자 또는 DNA 서열내 정보가 나타나게하거나, 예를 들어 상응하는 유전자 또는 DNA 서열의 전사 및 번역에 관여하는 세포 기능을 활성화시킴으로써, 또는 억제제 RNA(RNAi)의 경우, 처리되어 표적 유전자의 발현을 억제할 수 있도록 RNAi 분자를 전사함으로써 단백질을 생성시키는 것을 의미한다.
- [0040] DNA 서열은 세포 내에서 또는 세포에 의해 발현되어 단백질과 같은 "발현 생성물"을 형성한다. 발현 생성물 자체, 예를 들어, 결과 단백질은 또한 세포에 의해 "발현"된다고 일컬어질 수 있다. 폴리뉴클레오타이드 또는 폴리펩타이드는 예를 들어, 외래 또는 천연 프로모터의 조절하에 외래 숙주 세포에서 발현되거나 생산될 때, 또는 외래 프로모터의 조절하에 천연 숙주 세포에서 발현되거나 생산될 때, 재조합적으로 발현된다.
- [0041] 상기한 바와 같이, 일부 경우, 용어 "발현"은 억제성 RNA 분자(RNAi)의 생산을 포함한다. 상기 분자의 발현은 세포의 번역 기계를 포함하지 않지만, 숙주 세포의 유전자 발현을 변형시키기 위해 세포내 기계를 이용한다. 일부 구현예에서, 본 발명의 재조합 바이러스 벡터는 코딩 서열(예를 들어, 단백질)을 발현시키거나, RNAi 분자를 발현시키거나, 코딩 서열(예를 들어, 단백질을 발현) 및 발현 및 RNAi 분자 모두를 발현하도록 변형될 수 있다.
- [0042] 전형적으로, 재조합 복제 바이러스 벡터는 전형적으로 효과적인 발현을 허용하는 요소에 조작가능하게 결합된 발현될 이종 유전자 또는 서열을 함유하는 "카세트"(예를 들어, 프로모터, IRES 또는, 이종 서열의 전사 및 번역을 허용하는 관통-통과 요소(read-through element))를 포함하도록 변형된다.
- [0043] 이식유전자(예를 들어, 발현될 이종 서열)는 긴-말단 반복(LTR's), 외피의 삽입 하류, 스플라이스 수용체 후, 바이러스 *gag* 또는 *pol* 단백질, 내부 IRES 서열 또는 외피 코딩 서열의 하류의 작은 내부 프로모터와의 융합을 포함하는 다수의 위치에서 레트로바이러스 게놈으로 삽입될 수 있다. LTR에 이식유전자를 삽입하고, 여분의 스플라이스 수용체를 도입하면, 벡터 게놈이 급속히 불안정해지는 반면, IRES 및 기타 방법은 더 많은 가능성을 보여준다. 크립틱(cryptic) 스플라이스 수용체의 제거 및 이식유전자 서열의 인간화와 같은, 중요한 서열의 적절한 변화에 의해 이식유전자의 발현 및 구성이 적어도 부분적으로 영향을 받을 수 있다(예를 들어, 미국 특허 제8,722,867호 참조, 이는 이후에 참조로 본원에 통합됨). 이식유전자의 크기는 또한 벡터 안정성(stability)에 영향을 줄 수 있다. 예를 들어, 특정 벡터에서 이식유전자의 크기가 증가함에 따라 바이러스는 불안정해지고, 이종 유전자 또는 서열의 적어도 일부를 신속하게 결실시킨다. 이러한 한계는 IRES(보통 약 600bp, 예를 들어 미국 특허 제8,722,867호 참조) 또는 소형 프로모터(보통 약 250 내지 300bp, 예를 들어 국제 출원 공보 WO 2014/066700 참조, 이는 본원에 참고로 포함됨), 잠재적으로 예를 들어, MLV에서 이종 유전자 또는 서열의, 잠재적으로 남는 900 내지 1200bp 삽입물과 같은 발현가능 서열을 포함할 필요에 의해 악화된다. 따라서, 보다 많은 이식유전자 또는 다중 이식유전자를 포함하도록 이용가능한 이식유전자 크기를 최대화할 수 있는 것이 매우 유용할 것이다.
- [0044] 인간 세포에서 효율적으로 복제하는 레트로바이러스의 몇 가지 예는 무린 백혈병 바이러스(MLV) 뿐만 아니라 긴 팔원숭이 백혈병 바이러스(GALV), 비비원숭이 내인성 바이러스, 고양이 유래 바이러스 RD114의 양친매성, 다중 지향성, 이중향성, 및 10A1 균주를 포함한다. 마찬가지로, 양친매성-의사형(pseudotyped) RRV와 같은 비-에코트로픽(ecototropic) 외피 유전자를 포함하도록 변형된 MLV의 에코트로픽 균주는 또한 치료될 다양한 종과 세포 유형에서 효율적으로 복제될 수 있다. 그러나 레트로바이러스 외피는 람다바이러스, 알파바이러스, 홍역 또는

인플루엔자 바이러스 외피와 같은 비-레트로바이러스 외피로 대체될 수도 있다.

- [0045] 피코르나바이러스 및 뇌심근염 바이러스를 비롯한 여러 바이러스는 단일 ORF로부터 다중 단백질 발현을 매개하기 위해 그들의 게놈에 2A 또는 2A-유사 펩타이드를 코딩한다. 2A 펩타이드는 전형적으로 서열내 약 16 내지 18 개의 아미노산이고, 컨센서스 모티프(D[V/I]EXNPGP(서열식별번호: 1), 여기서 X는 임의의 아미노산임)를 공유한다. 2A 펩타이드가 인공 다중시스트론(multicistronic) mRNA에서 ORF 사이에 코딩되는 경우, 번역 폴리펩타이드에서 2A 펩타이드의 C-말단에서 리보솜이 정지하여 각 ORF에서 유래된 폴리펩타이드가 분리된다(Doronina 등, 2008). 분리 지점은 2A의 C-말단에 있으며, 하류 ORF의 첫번째 아미노산은 프롤린이다(예를 들어, 도 1 참조). 2A 펩타이드의 독특한 특징은 단일 다중시스트론 mRNA 구성에서 다중-단백질 발현을 위한 분자 도구로서의 활용을 이끌어 냈다.
- [0046] 2A 펩타이드는 구제역 바이러스 및 마비염 A 바이러스와 같은 피코르나바이러스 바이러스 계열 및, 돼지 테스코 바이러스-1 및 곤충 바이러스 토세아 아시그나 바이러스와 같은 기타 바이러스의 바이러스 게놈내에 존재한다(도 1). 2A 펩타이드는 원래의 상황에서 100%에 가까운 "분리" 효율을 가지며, 비-천연 서열에 도입될 때 종종 더 낮은 "분리" 효율을 갖는다. 다른 종류의 바이러스에서 발견되는 다른 2A-유사 서열은 또한 비-천연 서열에서 ~ 85% "분리" 효율을 달성하는 것으로 나타났다(Donnelly 등, 1997). 이식유전자를 발현하기 위한 본원의 방법 및 조성물에 사용될 수 있는 다수의 2A-유사 서열(도 2)이 있다.
- [0047] 2A 서열이 약 20년 동안 존재하는 것으로 알려졌지만, 비-천연 환경에서 기능할 수 있는 그들의 능력은 의문시되고 있다. 특히, 2A 서열은 이전에 번역된 단백질의 C 말단에 약 17 내지 22개의 여분의 아미노산을 남기고, 하류 단백질의 N-말단에 프롤린을 부가하여, 선행 단백질의 기능에 영향을 줄 가능성이 있다. 많은 바이러스 외피 단백질(T. Murakami, Mol Biol Int. 2012)의 경우와 같이, 단백질이 내부원형질 및 골지체내에, 및/또는 비리온의 성숙동안, 단백질이 번역-후 변형을 필요로 하는 경우, 선행 단백질에 대한 기능적 무능력의 위험이 추가되었다.
- [0048] 도 19는 외피 단백질의 C-말단에서 2A 펩타이드를 갖는 MLV 외피 단백질의 처리를 나타낸다. 일반적으로, 천연 MLV 외피 단백질의 처리는 감염된 숙주 세포에서 발생하는 전구체 단백질 Pr85 내지 gp70(SU) 및 p15E(TM) 서브유닛의 분열을 수반한다. 숙주 세포로부터의 출아 중에 바이러스 외피 단백질을 비리온에 효과적으로 도입하기 위해 Pr85의 분열이 필요하다. 비리온이 숙주 세포 막으로부터 출아함에 따라, 비리온은 감염성을 갖기 위해 성숙 과정을 거친다. MLV 비리온 성숙의 과정들 중 하나는 바이러스 단백질 분해효소에 의해 외피 단백질의 TM 서브유닛의 C-말단에 위치한 R-펩타이드의 제거를 포함한다. 도 19에 도시된 시나리오에서, 최종 아미노산 잔기인 프롤린(Pro)을 제외한 2A 펩타이드는 R-펩타이드의 하류에서 발견되어, 2A 펩타이드의 서열에 따라, R 펩타이드의 길이를 16개 아미노산에서 적어도 32개의 아미노산으로 만든다. 이론적으로 2A 펩타이드 서열의 첨가에 의해 R-펩타이드의 길이가 길어지지만, 2A 펩타이드는 R 펩타이드의 분열과 함께 동시에 제거되어, 기능성 외피 단백질이 생성될 것이다.
- [0049] 따라서, 외피 서열이 비-기능적이거나 약화되는 경우, 바이러스 벡터는 유용하지 않을 수 있다. 마우스만 감염시키는 특정 외피를 갖는 레트로바이러스 구조체(에코트로픽)에서 특정 2A 서열(돼지 테스코바이러스-1, "P2A")을 사용하기 위해 여러 시도들이 있었다(S. Stavrou 등, PLoS Pathog 10(5):e1004145, 2014; and E.P. Browne, J.Virol. 89:155-64, 2015). 그러나, 상기 바이러스는 인간 세포를 감염시키지 않으며, 일반적인 단백질 처리 문제가 해결될 것이라는 기대는 없다. 더욱이, 이렇게 구성된 바이러스는 치료 효과를 얻기보다는 생체 내에서 바이러스 복제를 촉진하는 유전자를 발현하도록 고안되었다.
- [0050] 본원은 세포 또는 개체에 전달될 수 있는, 예를 들어 시토신 데아미나제 또는 이의 돌연변이체, miRNA 또는 siRNA, 사이토키인, 항원 결합 도메인, 또는 코딩 서열의 조합 등을 코딩하는 이중성 폴리뉴클레오타이드를 함유하는 복제가능 바이러스 벡터를 제공한다. 바이러스 벡터는 아데노바이러스 벡터, 홍역 벡터, 헤르페스 벡터, 레트로바이러스 벡터(렌티바이러스 벡터 포함), 람다바이러스 벡터, 에컨대 수포성 구내염 바이러스 벡터, 레오 바이러스 벡터, 세네카 밸리 바이러스 벡터, 폭스바이러스 벡터(동물성 수두 또는 백시니아 유래 벡터 포함), 파보바이러스 벡터(AAV 벡터 포함), 알파바이러스 벡터 또는 당업자에게 공지된 다른 바이러스 벡터일 수 있다(예를 들어, *Concepts in Genetic Medicine*, ed. Boro Dropulic and Barrie Carter, Wiley, 2008, Hoboken, NJ.; *The Development of Human Gene Therapy*, ed. Theodore Friedmann, Cold Springs Harbor Laboratory Press, Cold springs Harbor, New York, 1999; *Gene and Cell Therapy*, ed. Nancy Smyth Templeton, Marcel Dekker Inc., New York, New York, 2000 and *Gene & Cell Therapy: Therapeutic Mechanism and Strategies*, 3rd ed., ed. Nancy Smyth Templeton, CRC Press, Boca Raton, FL, 2008 참고; 이들은 참조로 본 명세서에

통합된다).

- [0051] 이하에 기술된 바와 같이, 본원의 RRV는 MLV, MoMLV, GALV, FELV 등으로부터 유래될 수 있으며(즉, 모 뉴클레오타이드 서열은 MLV, MoMLV, GALV, FELV 등으로부터 얻어짐), 이중 뉴클레오타이드 서열에 결합된 2A 펩타이드 또는 2A 유사-펩타이드(때로는 "2A-펩타이드 카세트"로 언급됨)를 함유하도록 조작된다.
- [0052] 본 발명의 RRV는 그의 안정성 및/또는 발현을 변형시키도록 조작될 수 있다. 예를 들어, 발현의 변화는 중앙 조직에서 점진적으로 복제될 때 복제하는 레트로바이러스 벡터에 불활성화 또는 감쇠 돌연변이가 축적되는 빈도로 인해 발생할 수 있다. 조사에 따르면 가장 빈번한 반응 중 하나가 G에서 A로의 돌연변이임이 나타났다(역전사 단계의 제1 복제 단계에서 음성 가닥 단일 가닥 DNA에서 C에서 T 특성 ApoBec 매개 돌연변이에 해당함). 이것은 RRV 단백질의 아미노산 조성 변화 및 TGG(트립토판)에서 정지코돈(TAG 또는 TGA)으로의 엄청난 변화를 일으킬 수 있다. 일 구현예에서, 이러한 불활성화 변화는 페닐알라닌 또는 티로신과 같은 유사한 화학적 또는 구조적 성질을 갖는 다른 아미노산의 치환 코돈에 의해, 상기 가능성 없이, 회피된다.
- [0053] 따라서, 2A-펩타이드 카세트에 추가하여, RRV는 숙주 세포에서 구조체의 발현 및/또는 안정성을 개선시키는 다수의 추가 돌연변이를 포함할 수 있다. 상기 돌연변이는 트립토판 코돈을 GAG, POL 및/또는 ENV 도메인의 생물학적 활성을 유지하는 허용 코돈으로 변경시키는, GAG, POL 및/또는 ENV 코딩 서열에서의 하나 이상의 코돈의 변형을 포함할 수 있다. 트립토판에 대한 코돈은 UGG(DNA에서 TGG)인 것으로 당 분야에 공지되어있다. 또한, "정지 코돈"이 UAA, UAG 또는 UGA(DNA의 TAA, TAG 또는 TGA)인 것이 당 분야에 공지되어있다. 트립토판 코돈에 단일 지점 돌연변이가 있어, 비정상적인 정지 코돈을 유발한다(예를 들어, UGG → UAG 또는 UGG → UGA). 또한 인간 APOBEC3GF(hA3G/F)가 G → A 과돌연변이를 통한 레트로바이러스 복제를 억제한다는 것도 알려져있다(Neogi 등, *J. Int. AIDS Soc.*, 16(1): 18472, Feb. 25, 2013). 더욱이, 하기 기술된 바와 같이, 장기간 발현 및 바이러스 안정성은 코딩 서열에서 트립토판 코돈의 사용을 피함으로써, 그후 hA3G/F에 의한 과돌연변이로 인한 비정상적인 정지 코돈의 혼입을 피함으로써 개선될 수 있다. 예를 들어, 일 구현예에서, MLV 유래 핵산 서열은 GAG, POL 및 ENV 코딩 도메인을 포함한다(예를 들어, gag 핵산 도메인은 서열식별번호: 2의 약 뉴클레오타이드 번호 1203 내지 약 뉴클레오타이드 2819의 서열을 포함하고, pol 도메인은 서열식별번호: 2의 약 뉴클레오타이드 번호 2820 내지 약 뉴클레오타이드 6358의 서열을 포함하고, env 도메인은 서열식별번호: 2의 약 뉴클레오타이드 번호 6359 내지 약 뉴클레오타이드 8323의 서열을 포함한다). 트립토판 코돈에 있는, 표 1에서 확인된 뉴클레오타이드를 함유하는 코돈(서열식별번호: 2에 대해 언급된 뉴클레오타이드 번호)을 변경시킴으로써, hA3G/F 내성 RRV를 제공할 수 있다.

【표 1】

코돈 변화를 정지시키는 트립토판을 유도하는 재발성 G에서 A로의 돌연변이의 요약. 뉴클레오타이드는 서열식별번호: 2 RRV 게놈내 위치이며, "유전자"는 뉴클레오타이드가 위치한 유전자이고, AA는 폴리펩타이드내 아미노산 위치이다.

뉴클레오타이드	유전자	AA
1306	GAG	35
5299	POL	718
5557	POL	804
5806	POL	887
6193	POL	1016
6232	POL	1029
6298	POL	1051
6801	ENV	148
6978	ENV	207
7578	ENV	407

[0054]

[0055]

따라서, 본 발명의 일 구현예에서, 트립토판에 대한 코돈에 하나 이상의 돌연변이를 포함하고, 상기 돌연변이가 그 코돈을 트립토판 이외의 아미노산에 대한 코돈으로 변화시키고, 생체적합한 코돈(즉, 백터의 기능을 방해하지 않는 코돈)을 제공하는 경우, 재조합 복제가능 레트로바이러스가 제공된다. 이 백터는 "ApoBec 불활성화 내성"이다. 재조합 ApoBec 불활성화 내성 백터는 IRES 카세트, 프로모터 카세트 및/또는 2A 펩타이드 카세트를 포함할 수 있다. 본원에 사용된 바와 같이, IRES 카세트는 원하는 생물학적 활성 분자를 코딩하는 이중성 폴리뉴클레오타이드에 작동가능하게 연결된 내부 리보솜 진입 부위를 포함한다(예를 들어, 미국 특허 제8,829,173호 참조, 본원에 참고로 포함됨). 본원에 사용된 바와 같이, 프로모터 카세트는 원하는 생물학적 활성 분자를 코딩하는 하류 이중성 폴리뉴클레오타이드의 전사를 개시하는 조절 도메인을 포함한다(예를 들어, 미국 특허 제8,829,173호 및 미국 공개특허 공보 제2015/0273029A1호 참조, 본원에 참고로 포함됨). 프로모터는 조직특이적 프로모터, polIII 프로모터 또는 미니-프로모터일 수 있다. 2A 펩타이드 카세트는 본 명세서의 다른 곳에서 기술된다.

[0056]

일 구현예에서, 바이러스 백터는 분열 포유류 세포만을 감염시킬 수 있는 복제가능 레트로바이러스 백터일 수 있다. 일 구현예에서, 복제가능 레트로바이러스 백터는 레트로바이러스 외피 즉시 하류에 및 작동가능하게 연결된, 및 발현될 이중 핵산 서열의 즉시 상류에 2A 펩타이드 또는 2A 펩타이드-유사 서열을 포함한다. 특정 구현예에서, 백터는 IRES 카세트 또는 polIII(또는 미니프로모터) 카세트를 추가로 포함할 수 있다. 이중성 폴리뉴클레오타이드는 예를 들어, 시토신 데아미나제, 티미딘 키나아제, 사이토카인, 수용체, 항체 등을 코딩할 수 있다. polIII 프로모터 또는 미니프로모터가 포함된 경우, 백터는 miRNA, siRNA 또는 기타 RNAi 서열을 추가로 발현할 수 있다.

[0057]

또다른 구현예에서, 본원은 ENV-2A-이중 유전자 카세트를 제공한다. 카세트는 양친매성, 다중지향성, 이중향성, 10A1, GALV, 비비원숭이 내인성 바이러스, RD114, 랫도바이러스, 알파바이러스, 홍역 또는 인플루엔자 바이러스 외피 중 하나로부터 선택된 외피를 포함할 수 있다. 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열은 외피 코딩 서열의 C-말단에 작동가능하게 연결된 도 1 또는 2에 개시된 임의의 서열일 수 있다. 또다른 구현예에서, 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열은 GSG 링커 서열(예를 들어, ggaagcgga(서열식별번호: 3))을 통해 연결된다. 이중 유전자는 2A 펩타이드 또는 2A 펩타이드-유사 서열의 C-말단에 작동가능하게 연결된다. 이중 유전자는 표적 세포에서 전달되고 발현되는 임의의 원하는 유전자일 수 있다. 일 구현예에서, 이중 유전자는 길이가

500 내지 1500 bp 또는 그 사이의 임의의 수치(예를 들어, 1000 bp, 1100 bp, 1200 bp, 1300 bp, 1400 bp 등)를 포함한다. 또다른 일 구현예에서, 이중 유전자는 > 1500bp의 길이를 포함한다. 또다른 구현예에서, 카세트는 2A 펩타이드 또는 2A 펩타이드-유사 서열에 의해 분리된 2개의 이중 유전자를 포함한다. 또다른 구현예에서, 카세트는 이중 유전자의 env의 C-말단과 및 N-말단 사이에 작동가능하게 연결된 2A 펩타이드 또는 2A 펩타이드-유사 서열을 포함할 수 있으며, 상기 이중 유전자 다음에 IRES 또는 제2 이중성 서열에 연결된 프로모터를 포함하는 제2 카세트가 뒤따른다.

[0058] 본원은 변형된 레트로바이러스 벡터를 제공한다. 상기 변형된 레트로바이러스 벡터는 레트로바이러스과 계열의 구성원으로부터 유래될 수 있으며, ENV-2A-이식유전자 카세트를 함유하도록 조작될 수 있다. 위에서 언급했듯이, 레트로바이러스과 계열은 3개의 그룹으로 구성된다: 인간의 포미 바이러스(HFV)와 같은 스푸마바이러스(spumaviruses)-(또는 포미 바이러스(foamy virus)); 양의 비스나 바이러스 뿐만 아니라 렌티바이러스; 및 온코바이러스(비록 이 그룹 내의 모든 바이러스가 종양유발성은 아님).

[0059] 일 구현예에서, 본원은 비분열 세포, 분열 세포 또는 세포 증식 장애를 갖는 세포를 감염시킬 수 있는 재조합 레트로바이러스를 제공한다. 본 발명의 재조합 복제가능 레트로바이러스는 바이러스 GAG, 바이러스 POL, 바이러스 ENV, 바이러스 ENV 서열의 즉시 하류의 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열이 앞에 있고 비리온내에서 캡슐화된 이중성 폴리뉴클레오타이드(예를 들어, 1 내지 50개의 뉴클레오타이드(1 내지 10, 10 내지 15, 15 내지 20, 20 내지 25, 25 내지 30, 30 내지 35, 35 내지 40, 40 내지 45, 45 내지 50 또는 그 사이의 임의의 정수))를 포함한다.

[0060] "비분열" 세포라는 구절은 유사 분열을 거치지 않는 세포를 의미한다. 비분할 세포는 세포가 활발히 분할하지 않는 한 세포주기의 임의의 지점(예를 들어, G₀/G₁, G_{1/S}, G_{2/M})에서 차단될 수 있다. 생체의 감염을 위해, 분할 세포는 방사선 조사, 아피도콜린 치료, 혈청 기아 및 접촉 저해를 포함하는, 당업자에 의해 사용된 표준 기술에 의해 세포 분열을 차단하도록 처리될 수 있다. 그러나, 많은 세포가 이미 저지되어 있기 때문에(예를 들어, 줄기 세포), 생체의 감염은 종종 세포를 차단하지 않고 수행되는 것으로 이해해야한다. 예를 들어, 재조합 렌티바이러스 벡터는 비-분열 세포를 감염시킬 수 있다. 체내 기존의 비분열 세포의 예로는 신경 세포, 근육 세포, 간 세포, 피부 세포, 심장 세포, 폐 세포 및 골수 세포 및 이들의 유도물이 있다. 세포를 분열시키기 위해, 온코레트로바이러스 벡터가 사용될 수 있다.

[0061] "분열" 세포는 활성 유사분열 또는 감수분열을 겪는 세포를 의미한다. 상기 분열 세포는 줄기 세포, 피부 세포(예를 들어, 섬유아세포 및 케라틴세포), 생식 세포 및 당 분야에 공지된 다른 분열 세포를 포함한다. 특히 관심의 대상이 되고 분열 세포라는 용어는 신생 세포와 같은, 세포 증식 장애를 갖는 세포이다. 용어 "세포 증식성 장애"는 세포 분열의 비정상적인 수를 특징으로 하는 상태를 나타낸다. 상기 상태는 비대(조직 내의 세포 집단의 과증식을 초래하는 세포의 지속적인 증식) 및 영양장애증(조직 내 세포의 결핍 또는 부족) 세포 성장 또는 신체 영역으로의 세포의 과도한 유입 또는 이동을 모두 포함할 수 있다. 세포 집단은 반드시 형질전환되거나, 종양성 또는 악성 세포는 아니지만, 정상 세포도 포함할 수 있다. 세포증식성 장애는 결합 조직의 과다 성장과 관련된 질환, 예컨대 피부경화증, 관절염 및 간경화증을 비롯한 다양한 섬유증 질환을 포함한다. 세포 증식 장애는 두경부암과 같은 종양성 질병을 포함한다. 두경부암은 구강암, 식도암, 인후암, 후두암, 갑상선암, 설암, 구순암, 침샘암, 비암(nose carcinoma), 부비강암(paranasal sinuses carcinoma), 비인두암, 상비 천장암(superior nasal vault carcinoma) 및 동종양(sinus tumor), 감각신경모세포종(esthesioneuroblastoma), 편평세포암(squamous cell cancer), 악성 흑색종(malignant melanoma), 비강상악동 미분화암종(sinonasal undifferentiated carcinoma, SNUC), 뇌종양(교아세포종(glioblastomas), 예컨대 다형성 교아종(glioblastoma multiforme) 포함) 또는 혈액 종양(blood neoplasia)을 포함한다. 또한, 경부 림프절, 후두앞 림프절, 허파 식도옆 림프절 및 턱밑 림프절을 포함하는 국소 림프절의 암종도 포함된다(Harrison's Principles of Internal Medicine (eds., Isselbacher, 등, McGraw-Hill, Inc., 13th Edition, pp1850-1853, 1994). 다른 암 유형은 폐암, 결장 직장암, 유방암, 전립선 암, 요로 암, 자궁암 림프종, 구강암, 췌장암, 백혈병, 흑색종, 위암, 피부암 및 난소암을 포함하지만, 이에 한정되지는 않는다. 세포증식성 질병은 또한, 류마티스 관절염(O'Dell NEJM 350:2591 2004) 및 기타 자가면역 질환(Mackay *et al* NEJM 345:340 2001)을 포함하며, 이는 종종 면역 체계 세포의 부적절한 증식을 특징으로 한다.

[0062] 이중 핵산 서열은 2A 펩타이드 또는 2A 펩타이드-유사 서열을 코딩하는 서열에 작동가능하게 연결된다. 본원에 사용된 용어 "이중" 핵산 서열 또는 이식유전자는 (i) 야생형 레트로바이러스에 정상적으로 존재하지 않는 서열, (ii) 외래종으로부터 유래된 서열, 또는 (iii) 동일한 종으로부터인 경우, 그것은 원래의 형태로부터 실

질적으로 변형될 수 있는 서열을 의미한다. 대안적으로, 세포에서 정상적으로 발현되지 않는 변하지않는 핵산 서열은 이중 핵산 서열이다.

[0063] 본 발명의 레트로바이러스 벡터의 의도된 용도에 따라, 임의의 수의 이중성 폴리뉴클레오타이드 또는 핵산 서열이 레트로바이러스 벡터에 삽입될 수 있다. 예를 들어, *생체외* 연구를 위해 항생제 내성 및 형광 분자(예를 들어, GFP) 또는 발광 분자를 포함하여 일반적으로 사용되는 마커 유전자 또는 리포터 유전자가 사용될 수 있다. 임의의 목적하는 폴리펩타이드 서열을 코딩하는 추가의 폴리뉴클레오타이드 서열은 또한 본 발명의 벡터에 삽입될 수 있다.

[0064] 이중 핵산 서열의 생체내 전달이 요구되는 경우, 치료적 및 비치료적 서열 모두가 사용될 수 있다. 예를 들어, 일부 구현예에서, ENV-2A-이식유전자 카세트 다음에는 polIII-RNAi 카세트 또는 IRES-카세트가 뒤따를 수 있다. 예를 들어, 미니프로모터 또는 polIII 카세트가 사용되는 경우, 카세트는 세포 증식성 장애 또는 기타 유전자-관련 질환 또는 장애와 관련된 특정 유전자에 대한 miRNA, siRNA 등을 포함하는 이중 서열을 포함할 수 있다. 다른 구현예에서, 2A 펩타이드 또는 2A 펩타이드-유사 서열 또는 IRES 하류의 이중 유전자는 자살 유전자(예를 들어, HSV-tk 또는 PNP 또는 시토신 데아미나제 활성을 갖는 폴리펩타이드; 변형된 또는 변형되지않음), 성장인자 또는 치료용 단백질(예를 들어, 인자 IX, IL2 등)일 수 있다. 본원에 적용 가능한 다른 치료용 단백질은 당 기술 분야에서 쉽게 확인된다.

[0065] 일 구현예에서, 벡터 내의 이중성 폴리뉴클레오타이드는 인간 세포에서의 발현을 위해 최적화된 시토신 데아미나제 또는 티미딘 키나아제를 포함한다. 추가의 구현예에서, 시토신 데아미나제는 인간 코돈 최적화된 서열을 포함하며, 시토신 데아미나제의 안정성을 증가시키고(예를 들어, 감소된 분해 또는 증가된 열-안정성), 및/또는 야생형 시토신 데아미나제와 비교하여 트립토판 코돈을 비-트립토판 코딩 코돈으로 변경시키는 돌연변이를 포함하는 서열을 포함한다. 또다른 구현예에서, 이중성 폴리뉴클레오타이드는 UPRT 또는 OPRT 활성을 갖는 폴리펩타이드를 코딩하는 폴리뉴클레오타이드에 작동가능하게 연결된 시토신 데아미나제 활성을 갖는 폴리펩타이드(인간 코돈 최적화되거나 최적화되지않음, 돌연변이되거나 돌연변이되지 않음)를 포함하는 융합 구조체를 코딩한다.

[0066] 상기한 바와 같이, 인간 APOBEC3g는 G → A를 변환하는 바이러스 벡터 서열에서 과돌연변이를 일으킨다. 따라서, 2A 펩타이드 카세트에 함유된 이중성 폴리뉴클레오타이드의 트립토판 코돈은 hAPOBEC3에 의해 정지 코돈으로 변환되기 쉽다. 상기 돌연변이를 피하기 위해, 트립토판 코돈은 다른 아미노산의 생물학적으로 허용되는 코돈으로 대체될 수 있다. 예를 들어, 일 구현예에서, 본 발명의 2A-카세트는 시토신 데아미나제 활성을 갖는 폴리펩타이드를 코딩하는 폴리뉴클레오타이드를 포함할 수 있으며, 여기서 폴리뉴클레오타이드는 하기 서열을 포함한다:

```
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat atc
gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct atc ggc
gct tgt ctg atc aac aac aag gac gcc agt gtg ctg ggc agg ggc cac aac
atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag atc tcc acc ctg
gag aac tgt ggc agg ctg gag ggc aag gtg tac aag gac acc acc ctg tac
acc acc ctg tcc cct tgt gac atg tgt acc ggc gct atc atc atg tac ggc
atc cct agg tgt gtg atc ggc gag aac gtg aac ttc aag tcc aag ggc gag
aag tac ctg caa acc agg ggc cac gag gtg gtg gtt gtt gac gat gag agg
tgt aag aag ctg atg aag cag ttc atc gac gag agg cct cag gac tgg ttc
gag gat atc ggc gag taa (SEQ ID NO:28)
```

[0067]

[0068] 이 서열은 2개의 트립토판 코돈(볼드체/밑줄)을 포함한다. 본원의 일 구현예에서, 상기 코돈은 D, M, T, E, S, Q, N, F, Y, A, K, H, P, R, V, L, G, I 및 C로 구성된 그룹으로부터 선택된 아미노산을 제공하는 코돈으로 독립적으로 변경된다. 결과적으로 얻은 폴리펩타이드는 하기 서열을 포함한다:

```
M V T G G M A S K X D Q K G M D I A Y E E A L L G Y K E G G V P I G
G C L I N N K D G S V L G R G H N M R F Q K G S A T L H G E I S T L
E N C G R L E G K V Y K D T T L Y T T L S P C D M C T G A I I M Y G
I P R C V I G E N V N F K S K G E K Y L Q T R G H E V V V V D D E R
```

[0069]

[0070] C K K L M K Q F I D E R P Q D X F E D I G E (서열식별번호:29)

[0071] 상기 폴리펩타이드는 시토신 데아미나제 활성을 포함하며, X는 트립토판을 제외한 임의의 아미노산이다. 일 구현예에서, 서열식별번호: 29의 X는 각각 독립적으로 F, D, M, L, S 또는 R로 구성된 그룹으로부터 선택된다.

[0072] 또다른 구현예에서, 복제가능 레트로바이러스 벡터는 (본원에 기술된 바와 같이) 시토신 데아미나제를 포함하는

폴리펩타이드를 코딩하는 이중성 폴리뉴클레오타이드를 포함할 수 있으며, 바이러스성 프로모터부터의 1차 전사물의 일부로서 또는 프로모터에 연결된 miRNA 또는 siRNA 분자를 포함하는 폴리뉴클레오타이드를 추가로 포함할 수 있으며, 이는 세포-유형 또는 조직 특이적일 수 있다. 또다른 구현예에서, miRNA 또는 siRNA는 pol III 프로모터에 의해 선행될 수 있다.

[0073] 마이크로RNA(miRNA)는 작은 비-코딩 RNA이다. 이들은 코딩 유전자 또는 비-코딩 유전자의 인트론, 비-코딩 유전자의 엑손 또는 게놈간 영역 내에 위치한다. miRNA 코딩 서열은 1차 전구체 miRNA(pri-miRNA)라고 불리는 전구체 폴리뉴클레오타이드를 생성하는 RNA 중합효소 II에 의해 전사된다. 핵 내의 pri-miRNA는 리보뉴클레아제 Drosha에 의해 처리되어 짧은 헤어핀 구조를 형성하는 miRNA 전구체(pre-miRNA)를 생성한다. 그 다음, pre-miRNA는 엑스포틴(Exportin) 5를 통해 세포질로 운반되고, Dicer라는 다른 리보뉴클레아제에 의해 처리되어 활성, 성숙 miRNA를 생성한다.

[0074] 성숙한 miRNA는 길이가 약 21개의 뉴클레오타이드이다. 그것은 표적 유전자의 mRNA의 3' 비번역 영역에 결합하고, 단백질 번역의 억제 또는 mRNA의 분해에 의해 단백질 발현을 억제함으로써 기능을 발휘한다. miRNA는 발달, 세포 증식, 분화 및 암 진행을 포함한 생물학적 과정에 관여한다. miRNA 프로파일링에 대한 연구는 일부 miRNA 발현이 조직 특이적이거나 특정 조직에 풍부해졌음을 보여준다. 예를 들어, miR-142-3p, miR-181 및 miR-223 발현은 인간 및 마우스의 조혈 조직에서 풍부해지는 것으로 입증되었다(Baskerville 등, 2005 *RNA* 11, 241-247; Chen 등, 2004 *Science* 303, 83-86).

[0075] 여러 종양에서 일부 miRNA는 상향 조절(발암 miRNA) 또는 하향 조절(억제자)하는 것으로 관찰되었다(Spizzo 등, 2009 *Cell* 137, 586e1). 예를 들어, miR-21은 교아세포종, 유방암, 폐암, 전립선 암, 대장 암, 위암, 식도암, 및 자궁 경부암, 자궁 평활근육종, DLBCL, 두경부암에서 과발현된다. 대조적으로, let-7의 구성원은 교아세포종, 폐암, 유방암, 위암, 난소암, 전립선 암 및 대장암에서 하향 조절되는 것으로 보고되었다. 암에서의 miRNA 발현의 항상성의 재구성은 암 진행을 억제하거나 역전시키는 필수 메커니즘이다.

[0076] 암에서 하향 조절되는 miRNA는 항암제로서 유용할 수 있다. 그 예로는 mir-128-1, let-7, miR-26, miR-124, 및 miR-137이 있다(Esquele-Kerscher 등, 2008 *Cell Cycle* 7, 759-764; Kumar 등, 2008 *Proc Natl Acad Sci USA* 105, 3903-3908; Kota 등, 2009 *Cell* 137, 1005-1017; Silber 등, 2008 *BMC Medicine* 6:14 1-17). miR-128 발현은 중추 신경계에 풍부한 것으로 보고되었으며, 교아세포종에서 하향 조절되는 것으로 관찰되었다(Sempere 등, 2004 *Genome Biology* 5:R13.5-11; Godlewski 등, 2008 *Cancer Res* 68: (22) 9125-9130). miR-128은 miR-128-1과 miR-128-2의 2개의 별개 유전자에 의해 코딩된다. 둘 다 동일한 성숙 서열로 처리된다. Bmi-1 및 E2F3a는 miR-128의 직접 표적인 것으로 보고되었다(Godlewski 등, 2008 *Cancer Res* 68: (22) 9125-9130; Zhang 등, 2009 *J. Mol Med* 87:43-51). 또한, Bmi-1 발현은 신경교종, 맨틀세포 림프종, 비-소세포 폐암 B-세포 비-호지킨 림프종, 유방암, 대장암 및 전립선 암을 포함한 다양한 인간 암에서 상향 조절되는 것으로 관찰되었다. 또한, Bmi-1은 신경 줄기 세포 뿐만 아니라 신경교종내 "줄기-유사" 세포 집단을 포함하는 다양한 조직의 줄기 세포의 자가-재생에 필요한 것으로 입증되었다.

[0077] 본원에 사용된 바와 같이, 용어 "RNA 간섭"(RNAi)은 짧은 간섭 핵산(siRNA 또는 마이크로RNA(miRNA))에 의해 매개되는 서열-특이적 후-전사 유전자 사일런싱의 프로세스를 지칭한다. "RNA 간섭을 매개할 수 있는 체제"라는 용어는 세포 내에서 전사될 때 siRNA를 코딩하는 DNA 및 RNA 벡터뿐만 아니라 siRNA를 의미한다. 용어 siRNA 또는 miRNA는 서열 특이적 RNA 간섭을 매개할 수 있는 임의의 핵산 분자, 예를 들어, 짧은 간섭 RNA(siRNA), 이중-가닥 RNA(dsRNA), 마이크로-RNA(miRNA), 짧은 헤어핀 RNA(shRNA), 짧은 간섭 올리고뉴클레오타이드, 짧은 간섭 핵산, 짧은 간섭 변형된 올리고뉴클레오타이드, 화학적으로 변형된 siRNA, 후-전사 유전자 사일런싱 RNA(ptgsRNA) 및 다른 것들을 포함하여 의미한다.

[0078] 헤어핀 이중가닥의 줄기 길이를 설계하기 위한 적합한 범위는 20 내지 30개의 뉴클레오타이드, 30 내지 50개의 뉴클레오타이드, 50 내지 100개의 뉴클레오타이드, 100 내지 150개의 뉴클레오타이드, 150 내지 200개의 뉴클레오타이드, 200 내지 300개의 뉴클레오타이드, 300 내지 400개의 뉴클레오타이드, 400 내지 500개의 뉴클레오타이드, 500 내지 600개의 뉴클레오타이드, 및 600 내지 700개의 뉴클레오타이드의 줄기 길이를 포함한다. 헤어핀 이중가닥의 루프 길이를 설계하기 위한 적합한 범위는 헤어핀 이중가닥의 줄기 길이가 상당할 경우, 4 내지 25개의 뉴클레오타이드, 25 내지 50개의 뉴클레오타이드 또는 그 이상의 루프 길이를 포함한다. 특정 상황에서, 21개 이상의 뉴클레오타이드인 이중가닥 영역을 갖는 헤어핀 구조는 루프 서열 및 길이에 관계없이 효과적인 siRNA-방향 사일런싱을 촉진할 수 있다.

[0079] 또다른 또는 추가적인 구현예에서, 이중성 폴리뉴클레오타이드는 인터루킨, 인터페론 감마 등과 같은 사이토카

인을 포함할 수 있다. 본 발명의 레트로바이러스 벡터로부터 발현될 수 있는 사이토카인은 IL-1 α , IL-1 β , IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IL-19, IL-20, IL-21, IL-22, IL-23, IL-24, IL-25, IL-26, IL-27, IL-28, IL-29, IL-30, IL-31, IL-32, IL-33, IL-34, IL-35, IL-36, IL-37, IL-38, 항-CD40, CD40L, IFN- γ 및 TNF- α , TNF- α 의 가용성 형태, 림프 독소- α (LT- α , TNF- β 로도 알려져 있음), LT- β (복합 헤테로트리머 LT- α 2- β 에서 발견됨), OPGL, FasL, CD27L, CD30L, CD40L, 4-1BBL, DcR3, OX40L, TNF- γ (국제 공개 번호 WO 96/14328), AIM-I (국제 공개 번호 WO 97/33899), 엔도킨- α (국제 공개 번호 WO 98/07880), OPG, 및 뉴트로카인- α (국제 공개 번호 WO 98/18921, OX40, 및 신경성장인자(NGF), 및 Fas, CD30, CD27, CD40 및 4-1BB의 가용성 형태, TR2(국제 공개 번호 WO 96/34095), DR3(국제 공개 번호 WO 97/33904), DR4(국제 공개 번호 WO 98/32856), TR5(국제 공개 번호 WO 98/30693), TRANK, TR9(국제 공개 번호 WO 98/56892), TR10(국제 공개 번호 WO 98/54202), 312C2(국제 공개 번호 WO 98/06842), 및 TR12, 및 가용성 형태 CD154, CD70, 및 CD153을 포함하지만, 이에 한정되지 않는다. 혈관 형성 단백질은 일부 구현예에서, 특히 세포주로부터의 단백질 생산에 유용할 수 있다. 상기 혈관형성 인자는 신경교종 유도 성장인자(Glioma Derived Growth Factor, GDGF), 혈소판 유도 성장인자-A(PDGF-A), 혈소판 유도 성장인자-B(PDGF-B), 태반 성장인자(PIGF), 태반 성장인자-2(PEGF-2), 혈관 내피 성장인자(VEGF), 혈관 내피 성장인자-A(VEGF-A), 혈관 내피 성장인자-2(VEGF-2), 혈관 내피 성장인자 B(VEGF-3), 혈관 내피 성장인자 B-186(VEGF-B186), 혈관 내피 성장인자-D(VEGF-D), 혈관 내피 성장인자-D(VEGF-D) 및 혈관 내피 성장인자-E(VEGF- VEGF-E)를 포함하지만, 이에 한정되지 않는다. 섬유아세포 성장인자는 FGF-1, FGF-2, FGF-3, FGF-4, FGF-5, FGF-6, FGF-7, FGF-8, FGF-9, FGF-10, FGF-11, FGF-12, FGF-13, FGF-14 및 FGF-15를 포함하지만, 이에 한정되지 않는다. 조혈 성장인자는 본원의 벡터들을 사용하여 전달될 수 있으며, 상기 성장인자들은 과립구 대식세포 콜로니 자극 인자(GM-CSF)(사르그라마스티(sargramostim)), 과립구 콜로니 자극 인자(G-CSF)(필그라스티(filgrastim)), 대식세포 콜로니 자극 인자(M-CSF, CSF-1), 에리스로포이에틴(에포에틴 알파), 줄기 세포 인자(SCF, c-kit 리간드, 스틸 인자), 거핵구 콜로니 자극 인자, PIXY321(GMCSF/IL-3) 융합 단백질 등을 포함하지만, 이에 한정되지 않는다.

[0080] 용어 "조절 핵산 서열"은 수용체 세포의 코딩 서열의 복제, 전사, 및 번역을 위해 전체적으로 제공하는 프로모터 서열/영역, 폴리아데닐화 시그널, 전사 종결 서열, 상류 조절 도메인, 복제 기원, 인핸서 등을 집합적으로 지칭한다. 선택된 코딩 서열이 적당한 숙주 세포에서 복제, 전사 및 번역될 수 있는 한 상기 제어 서열 모두가 항상 존재할 필요는 없다. 당업자는 공개 데이터베이스 및 물질로부터 조절 핵산 서열을 용이하게 확인할 수 있다. 또한, 당업자는 예를 들어, *생체내*, *생체외* 또는 *시험관내에서* 의도된 용도에 적용가능한 조절 서열을 확인할 수 있다.

[0081] 용어 "프로모터 영역"은 DNA 조절 서열을 포함하는 뉴클레오타이드 영역을 의미하는 것을 지칭하는, 통상적인 의미로 사용되며, 상기 조절 서열은 RNA 중합효소에 결합할 수 있고 하류 (3'-방향) 코딩 서열의 전사를 개시할 수 있는 유전자로부터 유래된다. 조절 서열은 원하는 유전자 서열과 동종이거나 이종일 수 있다. 예를 들어, 바이러스 또는 포유동물 프로모터를 포함하여, 광범위한 프로모터가 이용될 수 있다.

[0082] "2A 펩타이드 또는 2A 펩타이드-유사 서열"은 도 1 및 2의 서열 중 임의의 서열과 97% 동일하고, 서열식별번호: 1의 컨센서스 서열을 함유하는 서열인, 서열식별번호: 1의 컨센서스 서열을 갖는 펩타이드를 지칭한다. 2A 펩타이드 또는 2A 펩타이드-유사 서열을 "코딩"하는 서열은 예를 들어 서열식별번호: 1의 컨센서스 서열을 갖는 2A 펩타이드 또는 펩타이드-유사 서열을 코딩하는 폴리뉴클레오타이드 서열이다. 코딩 서열은 일 구현예에서, ENV 및 이중 서열 사이에 작동가능하게 연결되어 배치되고, 한번 서열이 전사되면 단일 전사물(예를 들어, poly(m)RNA)로 전사되고, 전사물이 번역될 때 2개의 폴리펩타이드(예를 들어, ENV 및 이중 폴리펩타이드)가 생성된다.

[0083] 내부 리보솜 진입 부위("IRES")는 통상적으로 3'에서 IRES의 코딩 서열의 번역동안 리보솜의 진입 또는 보유를 촉진시키는 핵산 분절을 의미한다. 일부 구현예에서, IRES는 스플라이스 수용체/공여체 부위를 포함할 수 있지만, 바람직한 IRES는 스플라이스 수용체/공여체 부위가 없다. 일반적으로 메신저 RNA에 리보솜을 넣는 것은 모든 진핵 세포 mRNA의 5' 말단에 위치한 캡(cap)을 통해 이루어진다. 그러나, 이 보편적 규칙에는 예외가 있다. 어떤 바이러스성 mRNA에 캡이 없다는 것은 상기 RNA의 내부 부위에 리보솜의 진입을 허용하는 대체 구조가 있음을 시사한다. 현재까지, 기능때문에 IRES로 명명된 많은 구조체가 피코나바이러스, 특히 소아마비 바이러스(poliomyelitis virus)(Pelletier 등, 1988, Mol. Cell. Biol., 8, 1103-1112) 및 EMCV 바이러스(뇌심근염 바이러스)(Jang 등, J. Virol., 62, 2636-2643 1988; B.T.Baranick 등, Proc Natl Acad Sci U S A. 105:4733-8, 2008)와 같은 캡핑되지않은 바이러스 mRNA의 5' 비코딩 영역에서 확인되었다. 본원은 복제가능 레트로바이러스

백터와 관련하여 IRES의 사용을 제공한다.

- [0084] 이종 핵산 서열은 전형적으로 바이러스 LTR 프로모터-인핸서 요소 또는 내부 프로모터 중 어느 하나의 조절하에 있으며, 레트로바이러스 LTR 내의 보유된 요소는 백터의 숙주 세포 게놈 내로의 효율적인 통합을 가져올 수 있다. 따라서, 본원의 재조합 레트로바이러스 백터, 원하는 서열, 유전자 및/또는 유전자 단편은 여러 부위에, 및 상이한 조절 서열하에 삽입될 수 있다. 예를 들어, 삽입 부위는 바이러스 인핸서/프로모터 근위 부위(즉, 5' LTR-유도 유전자좌)일 수 있다.
- [0085] 일 구현예에서, 본 발명의 레트로바이러스 게놈은 원하는/이종성 폴리뉴클레오타이드의 삽입을 위한 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열의 하류에 클로닝 부위를 포함하는 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열을 함유한다. 일 구현예에서, 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열은 3'에서 레트로바이러스 백터에서 env 유전자에 위치하지만, 5'에서 원하는 이종성 폴리뉴클레오타이드에 위치한다. 따라서, 목적하는 폴리펩타이드를 코딩하는 이종성 폴리뉴클레오타이드는 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열에 작동가능하게 연결된다.
- [0086] 또다른 구현예에서, 표적화 폴리뉴클레오타이드 서열은 본원의 재조합 레트로바이러스 백터의 일부로서 포함된다. 표적 폴리뉴클레오타이드 서열은 표적 리간드(예를 들어, 펩타이드 호르몬, 예컨대 헤레굴린(hereregulin), 단일사슬 항체, 수용체 또는 수용체에 대한 리간드), 조직-특이적 또는 세포-유형 특이적 조절 요소(예를 들어, 조직-특이적 또는 세포-유형 특이적 프로모터 또는 인핸서), 또는 표적화 리간드와 조직-특이적/세포-유형 특이적 조절 요소의 조합이다. 바람직하게, 표적화 리간드는 레트로바이러스의 env 단백질에 작동가능하게 연결되어, 키메라 레트로바이러스성 env 단백질을 생성한다. 바이러스 GAG, 바이러스 POL 및 바이러스 ENV 단백질은 임의의 적합한 레트로바이러스(예를 들어, MLV 또는 렌티바이러스-유래)로부터 유래될 수 있다. 또다른 구현예에서, 바이러스성 ENV 단백질은 비-레트로바이러스-유래(예를 들어, CMV 또는 VSV)이다.
- [0087] 일 구현예에서, 본 발명의 재조합 레트로바이러스는 바이러스가 특정 세포 유형(예를 들어, 평활근 세포, 간세포, 신장 세포, 섬유아세포, 각질세포, 중간엽 줄기 세포, 골수 세포, 연골 세포, 상피 세포, 장 세포, 유방 세포, 종양 세포, 교모세포, 신경 세포 및 당 분야에 공지된 다른 세포들)에 표적화되는 방식으로 유전적으로 변형되어, 레트로바이러스 백터의 재조합 게놈이 표적 분열되지 않은 세포, 표적 분열 세포, 또는 세포 증식성 장애를 갖는 표적 세포로 전달되도록 한다.
- [0088] 일 구현예에서, 레트로바이러스 백터는 세포의 외부 표면에 분자를 갖는 세포에 결합함으로써 세포에 표적화된다. 레트로바이러스를 표적화하는 상기 방법은 레트로바이러스의 외막에 표적 리간드를 발현시켜 레트로바이러스 표면의 표적 리간드와 상호작용하는 수용체 또는 결합 분자를 갖는 세포 또는 조직에 바이러스를 표적화시키는 것을 돕는다. 바이러스에 의한 세포 감염 후, 바이러스는 핵산을 세포 내로 주입하고, 레트로바이러스 유전 물질은 숙주 세포 게놈에 통합될 수 있다.
- [0089] 예를 들어 특정 표적 세포상의 수용체에 대한 리간드를 코딩하는 또다른 유전자와 함께, 이종 폴리뉴클레오타이드를 본 발명의 바이러스 백터에 삽입함으로써, 백터가 이제 표적 특이적이 된다. 바이러스 백터는 예를 들어 당, 당지질 또는 단백질을 부착하여 표적을 특이적으로 만들 수 있다. 표적화는 바이러스 백터를 표적으로 하는 항체를 사용함으로써 달성될 수 있다. 당업자는 바이러스 게놈에 삽입될 수 있는 특정 폴리뉴클레오타이드 서열, 또는 바이러스 외피에 부착되어 핵산 서열을 함유하는 바이러스 백터의 표적 특이적 전달을 허용할 수 있는 단백질을 인지하거나 용이하게 확인할 수 있을 것이다.
- [0090] 따라서, 본원은 일 구현예에서, 표적 폴리펩타이드에 작동가능하게 연결된 레트로바이러스 ENV 단백질을 포함하는 키메라성 env 단백질을 포함한다. 표적 폴리펩타이드는 세포 특이적 수용체 분자, 세포 특이적인 수용체에 대한 리간드, 세포 특이적 항원성 에피토프에 대한 항체 또는 항체 단편, 또는 표적 세포와 결합 또는 상호작용할 수 있는 당 분야에서 쉽게 확인된 임의의 다른 리간드일 수 있다. 표적화 폴리펩타이드 또는 분자의 예는 링커로서 비오틴-스트렙타비딘을 사용하는 2가 항체(Etienne-Julan 등, J. Of General Virol., 73, 3251-3255 (1992); Roux 등, Proc. Natl. Acad. Sci USA 86, 9079-9083 (1989)), 그의 외피에 합텐에 대한 단일-사슬 항체 가변 영역을 코딩하는 서열을 함유하는 재조합 바이러스(Russell 등, NucleicAcids Research, 21, 1081-1085 (1993)), 펩타이드 호르몬 리간드의 레트로바이러스 외피로의 클로닝(Kasahara 등, Science, 266, 1373-1376 (1994)), 키메라 EPO/env 구조체(Kasahara 등, 1994), LDL 수용체를 발현하는 HeLa 세포의 특정 감염을 일으키는, 에코트로픽 MLV 외피내 지단백(LDL) 수용체내 저밀도에 대한 단일-사슬 항체(Somia 등, Proc. Natl. Acad. Sci USA, 92, 7570-7574 (1995)), 유사하게, 인테그린 리간드의 혼입에 의해 변형될 수 있어, 바이러스가 래트 교아세포종 세포들을 특이적으로 감염시키는 종을 통과할 수 있게 하는, 비장 피사 바이러스(SNV)의 조직-특이

적 표적을 보고한, ALV의 숙주 범위(Valsesia-Wittmann 등, J. Virol. 68, 4609-4619 (1994)), and Dornberg and co-workers (Chu and Dornburg, J. Virol 69, 2659-2663 (1995); M. Engelstadter 등 Gene Therapy 8, 1202-1206 (2001)), 중앙 표지자에 대한 단일-사슬 항체를 함유하는 외피를 사용한, 조류 레트로바이러스를 포함한다

[0091] 본 발명은 표적 세포를 감염시킬 수 있는 재조합 레트로바이러스를 생산하는 방법으로서, 적합한 숙주 세포를 하기에 의해 형질감염시키는 단계 및 재조합 바이러스를 회수하는 단계를 포함하는 방법을 제공한다: 바이러스 gag, 바이러스 pol 및 바이러스 env를 코딩하는 폴리뉴클레오타이드 서열을 포함하는 벡터, 2A 펩타이드 또는, env와 이중성 폴리뉴클레오타이드 사이에 작동가능하게 연결된 2A 펩타이드-유사 코딩 서열.

[0092] 본 발명의 레트로바이러스 및 방법은 비리온을 증식 및 생성시키기 위해 헬퍼 바이러스 또는 추가의 핵산 서열 또는 단백질을 필요로 하지 않는 복제가능 레트로바이러스를 제공한다. 예를 들어, 본 발명의 레트로바이러스의 핵산 서열은 상기 바와 같이 각각 그룹 특이적 항원 및 역전사 효소 (및 성숙 및 역전사에 필요한 인테그라아제 및 프로테아제-효소)를 코딩한다. 바이러스 gag 및 pol은 렌티바이러스, 예컨대 HIV 또는 온코바이러스 또는 감마레트로바이러스, 예컨대 MoMLV로부터 유래될 수 있다. 또한, 본 발명의 레트로바이러스의 핵산 계놈은 바이러스 외피(ENV) 단백질을 코딩하는 서열을 포함한다. env 유전자는 임의의 레트로바이러스에서 유래될 수 있다. env는 인간 및 다른 종의 세포의 형질도입을 허용하는 양친매성 외피 단백질일 수 있거나, 마우스 및 래트 세포만을 형질도입할 수 있는 에코트로픽 외피 단백질일 수 있다. 또한, 특정 세포 유형의 수용체를 표적화하기 위한 특정 리간드 또는 항체와 외피 단백질을 결합시켜 재조합 바이러스를 표적화하는 것이 바람직할 수 있다. 상기 언급한 바와 같이, 레트로바이러스 벡터는 예를 들어, 당지질 또는 단백질을 삽입함으로써 표적 특이적으로 될 수 있다. 표적화는 특정 세포 유형(예를 들어, 특정 조직에서 발견되는 세포 유형 또는 암 세포 유형)의 항원에 레트로바이러스 벡터를 표적화하기 위한 항체를 사용함으로써 종종 달성된다. 당업자는 과도한 실험없이 특정 표적에 레트로바이러스 벡터를 전달시키기 위한 특정 방법을 알거나 쉽게 확인할 수 있다. 일 구현예에서, env 유전자는 비-레트로바이러스(예를 들어, CMV 또는 VSV)로부터 유도된다. 레트로바이러스 유래 env 유전자의 예는 몰로니 뮤린 백혈병 바이러스(MoMuLV), 하비 뮤린 육종 바이러스(HaMuSV), 뮤린 유방 중앙 바이러스(MuMTV), 긴팔원숭이 백혈병 바이러스(GaLV), 인간 면역결핍 바이러스(HIV) 및 라우스 육종 바이러스(RSV)를 포함하지만, 이에 한정되지 않는다. 수포성 구내염 바이러스 벡터(VSV)(단백질 G), 시토크롬 바이러스 외피(CMV) 또는 인플루엔자 바이러스 혈구응집소(HA)와 같은 다른 env 유전자도 사용될 수 있다.

[0093] 일 구현예에서, 레트로바이러스 계놈은 온코-레트로바이러스, 보다 구체적으로는 포유동물 온코-레트로바이러스로부터 유래된다. 추가의 구현예에서, 레트로바이러스 계놈은 감마 레트로바이러스, 보다 구체적으로는 포유동물 감마 레트로바이러스로부터 유래된다. "유래된"은 모 폴리뉴클레오타이드 서열이 자연 발생 서열의 삽입(예를 들어, 2A 펩타이드 또는 2A 펩타이드 유사 코딩 서열 및 폴리펩타이드를 코딩하는 이중성 폴리뉴클레오타이드 및 IRES, 또는 또다른 이중성 폴리뉴클레오타이드 또는 관심의 역제성 핵산에 연결된 polIII 프로모터 중 임의로 하나 이상의 삽입) 또는 제거에 의해 변형된 야생형 온코바이러스인 것을 의미한다.

[0094] 또다른 구현예에서, 본원은 조절 서열을 이용하여 표적화된 레트로바이러스 벡터를 제공한다. 세포- 또는 조직-특이적 조절 서열(예를 들어, 프로모터)은 특정 세포 집단에서 유전자 서열의 발현을 표적화하는데 이용될 수 있다. 본원에 적합한 포유동물 및 바이러스 프로모터는 본 명세서의 다른 곳에서 설명된다. 따라서, 일 구현예에서, 본원은 레트로바이러스 계놈의 5' 말단에 조직-특이적인 프로모터 요소를 갖는 레트로바이러스를 제공한다. 전형적으로, 조직-특이적 조절 요소/서열은 신생 세포(예를 들어, 중앙 세포-특이적 인핸서 및 프로모터), 및 유도성 프로모터(예를 들어, 테트라사이클린)에 예를 들어 세포-특이적 또는 중앙-특이적 프로모터 및 인핸서를 포함하는, 레트로바이러스 계놈의 LTR의 U3 영역에 존재한다.

[0095] 본 발명의 전사 제어 서열은 또한 초항원, 사이토카인 또는 케모카인을 코딩하는 유전자와 자연적으로 결합된 자연발생적인 전사 제어 서열을 포함할 수 있다.

[0096] 어떤 경우에는, 발현을 조절하는 것이 바람직할 수 있다. 예를 들어, 다양한 발현 강도를 갖는 상이한 바이러스 프로모터가 원하는 발현 수준에 따라 사용될 수 있다. 포유동물 세포에서 CMV의 초기 프로모터는 종종 강한 전사 활성화를 제공하는데 사용된다. 효능이 적은 CMV 프로모터의 변형된 버전은 또한 이식유전자의 감소된 수준의 발현이 요구될 때 사용되었다. 조혈 세포에서 이식유전자의 발현이 요구되는 경우, MLV 또는 MMTV로부터의 LTR과 같은 레트로바이러스 프로모터가 사용될 수 있다. 사용될 수 있는 다른 바이러스 프로모터는 SV40, RSV LTR, HIV-1 및 HIV-2 LTR, 아데노 바이러스 프로모터, 예컨대 E1A, E2A 또는 MLP 영역, AAV LTR, 콜리플라워 모자이크 바이러스, HSV-TK 및 조류 육종 바이러스를 포함한다.

- [0097] 유사하게, 조직-특이적 또는 선택적 프로모터가 사용되어, 비-표적 조직에 대한 잠재적인 독성 또는 바람직하지 않은 영향을 감소시키기 위해 특정 조직 또는 세포에서 전사를 수행할 수 있다. 예를 들어, PSA, 프로바신 (probasin), 전립선 산 포스파타제 또는 전립선-특이성 선 칼리크레인(hK2)과 같은 프로모터를 사용하여 전립선에서 유전자 발현을 표적화할 수 있다. 유청 보조 단백질(WAP)은 유방 조직 발현에 사용될 수 있다(Andres 등, PNAS 84:1299-1303, 1987). 사용될 수 있는 다른 프로모터/조절 도메인은 하기에 기술되어 있다.
- [0098] "조직-특이적 조절 요소"는 한 조직에서 유전자의 전사를 유도할 수 있는 반면, 다른 조직 유형에서는 크게 "사일런트"를 유지할 수 있는 조절 요소(예를 들어, 프로모터)이다. 그러나, 조직-특이적 프로모터는 사일런트할 것으로 예상되는 조직에서 검출가능한 양의 "배경" 또는 "기본" 활성을 가질 수 있음을 이해할 것이다. 프로모터가 표적 조직에서 선택적으로 활성화되는 정도는 선택도 비율(표적 조직에서의 활성/대조군 조직에서의 활성)로 나타낼 수 있다. 이와 관련하여, 본 발명의 실시예에 유용한 조직 특이적 프로모터는 전형적으로 약 5보다 큰 선택도 비율을 갖는다. 바람직하게는, 상기 선택도 비율은 약 15보다 크다.
- [0099] 특정 징후에서, 본 발명의 재조합 복제가능 레트로바이러스(RRV)의 투여 후 특정 시간에 전사를 활성화시키는 것이 바람직할 수 있다.
- [0100] 이것은 호르몬 또는 사이토카인 조절가능한 프로모터로 수행될 수 있다. 예를 들어, 특정 스테로이드가 생성되거나 전달되는 생식소 조직을 나타내는 경우, 치료적 용도로 안드로겐 또는 에스트로겐 조절 프로모터의 사용이 유리할 수 있다. 호르몬 조절가능한 상기 프로모터는 MMTV, MT-1, 엑디손(ecdysone) 및 루비스코(RuBisco)를 포함한다. 갑상선, 뇌하수체 및 부신 호르몬에 반응하는 것과 같은 다른 호르몬 조절 프로모터가 사용될 수 있다. 사이토카인 및 염증 단백질 반응 촉진제로는 K 및 T 키니노겐(kininogen)(Kageyama 등, 1987), c-fos, TNF- α , C-반응성 단백질(Arcane 등, 1988), 합토클로빈(Oliviero 등, 1987), 혈청 아밀로이드 A2, C/EBP 알파, IL-1, IL-6(Poli and Cortese, 1989), 보체 C3(Wilson 등, 1990), IL-8, 알파-1 산 당단백질(Prowse and Baumann, 1988), 알파-1 안티티핀(antitypsin), 지단백질 리파아제(Zechner 등, 1988), 안지오텐시노젠(Ron 등, 1990), 피브리노겐, c-jun(포르볼 에스테르, TNF- α , UV 방사선, 레티노산 및 과산화수소에 의해 유도됨), 콜라게나아제(포르볼 에스테르 및 레티노산에 의해 유도됨), 메탈로티오네인(중금속 및 글루코코르티코이드 유도성), 스트로멜리신(포르볼 에스테르, 인터루킨-1 및 EGF에 의해 유도됨), 알파-2 매크로글로블린 및 알파-1 안티키모트립신이 포함된다. 오스테오칼신, 저산소증-반응 요소(HRE), MAGE-4, CEA, α -태아단백질, GRP78/BiP 및 티로시나제와 같은 중앙 특이적 프로모터는 또한 중앙 세포에서의 유전자 발현을 조절하는데 사용될 수 있다.
- [0101] 또한, 이러한 프로모터 목록은 완전하거나 제한적으로 해석되어서는 안되며, 당업자는 본원에 개시된 프로모터 및 방법과 관련하여 사용될 수 있는 다른 프로모터를 알 것이다.

【표 2】

조직 특이적 프로모터	
조직	프로모터
췌장	인슐린 엘라스틴 아밀라아제 pdr-1 pdx-1 글루코키나아제
간	알부민 PEPCK HBV 인헨서 α 태아단백질 아포지질단백질 C α -1 항트립신 비텔로제닌, NF-AB 트랜스타이레틴
골격근	미오신 H 사슬 근육 크레아틴 키나아제 디스트로핀 칼페인 p94 골격 알파-액틴 빠른 트로포닌 1
피부	케라틴 K6 케라틴 K1
폐	CFTR 인간 사이토케라틴 18(K18) 폐 계면활성제 단백질 A, B 및 C CC-10 P1
평활근	sm22 α SM- α -액틴
내피	엔도텔린-1 E-셀렉틴 폰 빌레브란트 인자 TIE(Korhonen et al., 1995) KDR/flk-1 멜라노사이트 티로시나아제
지방 조직	지단백질 리파아제(Zechner et al., 1988) 아딤신(Spiegelman et al., 1989) 아세틸-CoA 카르복실라제(Pape and Kim, 1989) 글리세로포스페이트 탈수소효소(Dani et al., 1989) 지방 세포 P2(Hunt et al., 1986)
유방	유청 산성 단백질(WAP)(Andres et al., PNAS 84:1299-1303 1987
혈액	β -글로빈

[0102]

- [0103] 특정 프로모터는 단일 조직 유형에 대한 활성이 제한되지는 않지만, 그럼에도 불구하고 한 그룹의 조직에서 활성일 수 있고, 다른 그룹에서는 덜 활성이거나 사일런트할 수 있다는 점에서 선택성을 나타낼 수 있다는 것이 또한 이해될 것이다. 상기 프로모터는 또한 "조직-특이적"으로 불리우며, 본 발명과 함께 사용하도록 고려된다. 예를 들어, 다양한 중추 신경계(CNS) 뉴런에서 활성인 프로모터는 뇌의 여러 상이한 영역 중 임의의 영역에 영향을 줄 수 있는 뇌졸중으로 인한 손상으로부터 보호하는데 치료적으로 유용할 수 있다. 따라서, 본원에서 사용된 조직-특이적 조절 요소는 이중 단백질의 조절에 적용가능할 뿐만 아니라 본 레트로바이러스 벡터에서 표적화 폴리뉴클레오타이드 서열로서의 적용가능성을 갖는다.
- [0104] 또다른 구현예에서, 본원은 재조합 레트로바이러스 유래 구조체를 포함하는 플라스미드를 제공한다. 플라스미드는 HT1080, NIH 3T3 또는 다른 조직 배양물 세포와 같은 표적 세포 또는 세포 배양물에 직접 도입될 수 있다. 생성된 세포는 레트로바이러스 벡터를 배양 배지로 방출한다.
- [0105] 본원은 5' 내지 3'을 포함하는 폴리뉴클레오타이드 구조체를 제공한다: 전사를 개시하는데 유용한 프로모터 또는 조절 영역; *psi* 패키징 신호; *gag* 코딩 핵산 서열, *pol* 코딩 핵산 서열; *env* 코딩 핵산 서열; 2A 펩타이드 또는 2A 펩타이드-유사 코딩 서열; 마커, 치료용 또는 진단용 폴리펩타이드를 코딩하는 이중성 폴리뉴클레오타이드; 선택적인 IRES 또는 *poIII* 카세트; 및 LTR 핵산 서열을 포함한다. 상기 언급한 바와 같이, *gag*, *pol* 및 *env* 핵산 도메인은 ApoBec3에 의해 정지 코돈으로 변환된 트립토판 코돈을 제거하도록 변형될 수 있다. 특정 다른 구현예에서, 벡터는 이중성 폴리뉴클레오타이드의 하류 및 3' LTR의 상류에 *poIII* 카세트 또는 IRES 카세트를 추가로 포함할 수 있다. 본원의 다른 곳에서 기술된 바와 같이, 및 하기와 같이, 본원의 폴리뉴클레오타이드 구조체의 다양한 부분(예를 들어, 재조합 복제가능 레트로바이러스 폴리뉴클레오타이드)은 원하는 숙주 세포, 발현 타이밍 또는 양 및 이중성 폴리뉴클레오타이드에 부분적으로 의존하여 조작된다. 본 발명의 복제가능 레트로바이러스 구조체는 당업자에 의해 개별적으로 변형될 수 있는 다수의 도메인으로 나눌 수 있다.
- [0106] 예를 들어, 프로모터는 뉴클레오타이드 1 내지 약 뉴클레오타이드 582에서 서열식별번호: 2에 개시된 서열을 갖는 CMV 프로모터를 포함할 수 있고, 변형된 프로모터가 전사를 지향하고 개시할 수 있는한, 하나 이상(예를 들어, 2 내지 5, 5 내지 10, 10 내지 20, 20 내지 30, 30 내지 50, 50 내지 100 또는 그 이상의 핵산 염기)으로의 변형을 포함할 수 있다. 일 구현예에서, 프로모터 또는 조절 영역은 CMV-R-U5 도메인 폴리뉴클레오타이드를 포함한다. CMV-R-U5 도메인은 MLV R-U5 영역에 연결된 인간 사이토메갈로바이러스로부터의 즉시 초기 프로모터를 포함한다. 일 구현예에서, CMV-R-U5 도메인 폴리뉴클레오타이드는 약 뉴클레오타이드 1 내지 약 뉴클레오타이드 1202의, 서열식별번호: 2에 개시된 서열, 또는 서열식별번호: 2에 개시된 서열과 적어도 95% 동일한 서열을 포함하며, 상기 폴리뉴클레오타이드는 거기에 작동가능하게 연결된 핵산 분자의 전사를 촉진한다. 폴리뉴클레오타이드의 *gag* 도메인은 임의의 수의 레트로바이러스로부터 유도될 수 있지만, 전형적으로 온코레트로바이러스, 및 보다 특히 MLV와 같은 포유동물 온코레트로바이러스로부터 유래될 것이다. 일 구현예에서, *gag* 도메인은 약 뉴클레오타이드 번호 1203 내지 약 뉴클레오타이드 2819의 서열식별번호: 2의 서열 또는 그에 대하여 적어도 95%, 98%, 99% 또는 99.8%(가장 가까운 10번째에 어림됨) 동일성을 갖는 서열을 포함한다. 폴리뉴클레오타이드의 *pol* 도메인은 임의의 수의 레트로바이러스로부터 유래될 수 있지만, 전형적으로 온코레트로바이러스 및 보다 특히 MLV와 같은 포유동물 온코레트로바이러스로부터 유래될 것이다. 일 구현예에서, *pol* 도메인은 약 뉴클레오타이드 번호 2820 내지 약 뉴클레오타이드 6358의 서열식별번호: 2의 서열 또는 그에 대하여 적어도 95%, 98%, 99% 또는 99.9%(가장 가까운 10번째에 어림됨) 동일성을 갖는 서열을 포함한다. 폴리뉴클레오타이드의 *env* 도메인은 임의의 수의 레트로바이러스로부터 유래될 수 있지만, 전형적으로 온코레트로바이러스 또는 감마-레트로바이러스, 보다 특히 MLV와 같은, 포유동물 온코레트로바이러스 또는 감마-레트로바이러스로부터 유래될 것이다. 일부 구현예에서, *env* 코딩 도메인은 양친매성 *env* 도메인을 포함한다. 일 구현예에서, *env* 도메인은 약 뉴클레오타이드 번호 6359 내지 약 뉴클레오타이드 8323의 서열식별번호: 2의 서열 또는 그에 대하여 적어도 95%, 98%, 99% 또는 99.8%(가장 가까운 10번째에 어림됨) 동일성을 갖는 서열을 포함한다. 2A 펩타이드 또는 2A 펩타이드-유사 카세트는 *env* 도메인 다음에 삽입되고(예를 들어, 약 뉴클레오타이드 8324), 2A 또는 2A 유사-펩타이드의 C-말단에 연결된 이중성 폴리뉴클레오타이드의 말단까지 계속된다. 이중 도메인은 폴리퓨린 풍부한 도메인이 뒤따를 수 있거나, IRES 카세트 또는 *poIII* 카세트가 뒤따를 수 있다. 3' LTR은 임의의 수의 레트로바이러스, 전형적으로는 온코레트로바이러스 및 바람직하게는 MLV와 같은 포유동물 온코레트로바이러스로부터 유래될 수 있다. 일 구현예에서, 3' LTR은 U3-R-U5 도메인을 포함한다. 또다른 구현예에서, LTR은 약 뉴클레오타이드 9111 내지 약 11654의 서열식별번호: 2에 개시된 서열 또는 그에 대하여 적어도 95%, 98% 또는 99.5%(가장 가까운 10번째에 어림됨) 동일성을 갖는 서열을 포함한다.
- [0107] 본 발명은 또한 5'에서 3'까지 CMV-R-U5, 인간 사이토메갈로바이러스로부터 MLV R-U5 영역까지의 즉시 초기 프로

모터의 융합; PBS, 역전사효소용 프라이머 결합 부위; 5' 스플라이스 부위; ψ 패키징 신호; gag, MLV 그룹 특이 항원에 대한 ORF; pol, MLV 폴리메라제 다단백질에 대한 ORF; 3' 스플라이스 부위; 4070A env, MLV 균주 4070A의 외피 단백질에 대한 ORF; 2A 펩타이드 또는 2A 펩타이드-유사 서열; 트립토판 코돈에 대한 변형이 있거나 없는 변형된 시토신 데아미나제(열안정화 및 코돈 최적화); PPT, 폴리퓨린 관; 및 U3-R-U5, MLV 긴 말단 반복을 포함하는 재조합 레트로바이러스 벡터를 제공한다.

- [0108] 본원은 또한 하기 개시된 서열을 포함하는 레트로바이러스 벡터를 제공한다.
- [0109] 레트로바이러스 벡터는 다수의 세포 증식성 질병 및 장애를 포함하는 광범위한 질병 및 장애를 치료하는데 사용될 수 있다(예를 들어, 미국 특허 제4,405,712호 및 제4,650,764호; Friedmann, 1989, Science, 244:1275-1281; Mulligan, 1993, Science, 260:926-932, R. Crystal, 1995, Science 270:404-410 참조, 이들 각각은 본원에 참고문헌으로 전문 통합됨, 또한: The Development of HumanGene Therapy, Theodore Friedmann, Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1999. ISBN 0-87969-528-5; Concepts in Genetic Medicine, ed. Boro Dropulic and Barrie Carter, Wiley, 2008, Hoboken, NJ.; Gene & Cell Therapy-Therapeutic Mechanism and Strategies, 3rd edition ed. Nancy Smyth Templeton, CRC Press, Boca Raton FL 2008 참조, 이들 각각은 본원에 참고문헌으로 전문 통합됨).
- [0110] 본원은 또한 세포 증식성 질환의 치료를 위한 유전자 요법을 제공한다. 상기 치료는 증식성 장애를 갖는 개체의 세포 내로 적절한 치료용 폴리뉴클레오타이드(예를 들어, 안티센스, 리보자임, 자살 유전자, siRNA)를 도입함으로써 그의 치료 효과를 달성할 것이다. 폴리뉴클레오타이드 구조체의 전달은, 특히 분열 세포를 감염시킬 수 있는 MLV에 기초하는 경우, 본 발명의 재조합 레트로바이러스 벡터를 사용하여 달성될 수 있다.
- [0111] 또한, 본원에 기술된 치료 방법(예를 들어, 유전자 요법 또는 유전자 전달 방법)은 생체내 또는 생체외에서 수행될 수 있다. 유전자 치료 전에 종양의 대부분을 예를 들어 외과적으로 또는 방사선에 의해 제거하는 것이 바람직할 수 있다. 일부 양태에서, 레트로바이러스 치료법은 수술, 화학 요법 또는 방사선 요법이 선행되거나 이어서 수행될 수 있다.
- [0112] 따라서, 본원은 비분열 세포, 분열 세포 또는 종양 세포를 감염시킬 수 있는 재조합 레트로바이러스를 제공하며, 이 재조합 레트로바이러스는 바이러스성 GAG; 바이러스성 POL; 바이러스성 ENV; 2A 펩타이드 또는 펩타이드-유사 코딩 서열에 작동가능하게 연결된 이중 핵산; 및 패키징, 역전사 및 통합에 필요한 시스-작용 핵산 서열을 포함한다. 재조합 레트로바이러스는 HIV와 같은 렌티바이러스일 수 있거나, 오톰바이러스일 수 있다. 본 발명의 재조합 레트로바이러스는 재조합 레트로바이러스의 제조 방법에서 설명한 바와 같이, VPR, VIF, NEF, VPX, TAT, REV 및 VPU 단백질 중 적어도 하나를 추가로 포함할 수 있다. 특정 이론에 구속되지 않지만, 이들 유전자/단백질 생성물 중 하나 이상은 생산된 재조합 레트로바이러스(예를 들어, NEF)의 바이러스 역가를 높이는 데 중요하거나, 또는 비리온의 감염 및 패키징에 필요할 수 있다고 믿어진다.
- [0113] 본원은 또한 특정 핵산(예를 들어, 이중 서열)의 발현을 제공하기 위해 표적 세포로의 핵산 전달 방법을 제공한다. 따라서, 또다른 구현예에서, 본원은 표적 세포를 본원의 재조합 바이러스로 감염시키는 단계 및 표적 세포 내에서 이중 핵산을 발현시키는 단계를 포함하는, 표적 세포에서 이중 핵산의 도입 및 발현 방법을 제공하며, 상기 이중 핵산은 env 도메인 하류의 재조합 바이러스 벡터로 조작되고, 2A 또는 2A 유사-펩타이드와 작동가능하게 연결된다. 상기 언급된 바와 같이, 표적 세포는 레트로바이러스에 의한 감염이 가능한 분열, 비-분열, 종양 세포, 불멸화 세포, 변형된 세포 및 당업자에 의해 인식되는 기타 세포 유형을 포함하는 임의의 세포 유형일 수 있다.
- [0114] 생물학적 반응 변형제(예를 들어, 사이토카인)를 코딩하는 핵산을 세포 또는 개체로 전달하는 것이 바람직할 수 있다. 이 범주에는 "인터루킨"으로 분류되는 다수의 사이토카인을 코딩하는 핵산을 포함하는 면역증강제가 포함된다. 여기에는 예를 들어, 인터루킨 1 내지 38뿐만 아니라 본원의 다른 반응 변형제 및 인자도 포함된다. 또한, 이 범주에 포함되는 것은 동일한 메커니즘에 따라 작동하지는 않지만 인터페론, 특히 감마 인터페론, 종양 괴사 인자(TNF) 및 과립구-대식세포-콜로니 자극 인자(GM-CSF)이다. 다른 폴리펩타이드는 예를 들어, 혈관신생 인자 및 항-혈관신생 인자를 포함한다. 효소 결합 또는 면역 결합을 치료하기 위해 골수 세포 또는 대식세포에 상기 핵산을 전달하는 것이 바람직할 수 있다. 성장인자, 독성 펩타이드, 리간드, 수용체 또는 다른 생리적으로 중요한 단백질을 코딩하는 핵산 또한 특정 표적 세포에 도입될 수 있다. 상기 생물학적 반응 변형제 중 임의의 것은 하류에서 본원의 RRV로 조작되고 2A 또는 2A 유사-펩타이드에 작동가능하게 연결된다.
- [0115] 본원은 약물 특이적 표적화 및 효과를 촉진시키는 이중성 폴리뉴클레오타이드의 전달에 사용될 수 있다. 예를

들어, EGF 수용체 계열의 구성원인, HER2가 약물 트라스투주맙(trastuzumab)(Herceptin™, Genentech)의 결합 표적이다. 트라스투주맙은 항체-의존성 세포독성(ADCC)의 매개체이다. 활성은 1+ 및 비-발현 세포보다는 면역조직화학에 의한 2+ 및 3+ 수준의 과발현을 갖는 HER2-발현 세포를 우선적으로 표적으로 한다(Herceptin 처방 정보, Crommelin 2002). HER2 저중양에서 HER2 또는 분열된 HER2를 발현(세포의 및 막횡단 도메인만을 발현함)하는 백터의 도입에 의한 HER2 발현의 증진은 ADCC의 최적 유발을 촉진하고 임상 용도로 관찰되는 Herceptin에 대하여 신속하게 생기는 내성을 극복할 수 있다. 이 경우, 이종 유전자는 HER2를 코딩할 것이다.

[0116] 또다른 예에서, CD20은 약물 리투시맙(rituximab)(Rituxan™, Genentech)의 결합을 위한 표적이다. 리투시맙은 보체-의존성 세포독성(CDC) 및 ADCC의 매개체이다. 유동 세포계측법에 의한 평균 형광 강도가 높은 세포는 리투시맙에 대한 민감도가 증가한다(van Meerten 등, Clin Cancer Res 2006; 12(13): 4027-4035, 2006). CD20 저 B 세포에서 CD20을 발현하는 백터의 도입에 의한 CD20 발현의 증진은 ADCC의 최적의 유발을 촉진시킬 수 있다. 이 경우, 이종 유전자는 CD20을 코딩한다.

[0117] 본 발명은 본원의 RRV 백터를 투여한 후 화학요법제 또는 항암제로 치료하는 단계를 포함하는, 암 및 신생물과 같은 세포 증식성 질환을 치료하는 방법을 제공한다. 일 양태에서, RRV 백터는 RRV가 감염 및 복제하는 것을 허용하는 화학요법제 또는 항암제의 투여 전에 일정 기간 동안 개체에게 투여된다. 그 다음, 개체는 증식을 줄이거나 암세포를 죽이기 위해 일정 시간 및 투여량 동안 화학요법제 또는 항암제로 치료된다. 일 양태에서, 화학요법제 또는 항암제에 의한 치료가 암/중양을 감소시키지만 죽이지 않는 경우(예를 들어, 부분적인 차도 또는 일시적인 차도), 그 개체는 이어서 RRV로부터 세포독성 유전자(예를 들어, 시토신 데아미나제)를 세포발현하는데 있어서 독성 치료제로 전환되는 무독성 치료제(예를 들어, 5-FC)에 의해 치료될 수 있다.

[0118] 이러한 방법을 사용하여, 본 발명의 RRV 백터는 중양 세포의 복제 과정 중에 확산되며, 그런 다음 상기 세포는 항암제 또는 화학요법제로 치료함으로써 살해될 수 있고, 본원에 기술된 RRV 치료 방법을 사용하여 추가의 살해가 발생할 수 있다.

[0119] 본원의 또다른 구현예에서, 이종 유전자는 표적 항원(예를 들어, 암 항원)에 대한 코딩 서열을 포함할 수 있다. 이 구현예에서, 세포 증식성 장애를 포함하는 세포는 표적 항원의 발현(예를 들어, 암 항원의 과발현)을 제공하기 위해 표적 항원을 코딩하는 이종성 폴리뉴클레오타이드를 포함하는 RRV로 감염된다. 이어서, 표적 항원과 특이적으로 상호작용하는 표적 동족 부분(cognate moiety)을 포함하는 항암제가 대상에게 투여된다. 표적화 동족 부분은 세포독성 제제에 작동가능하게 연결될 수 있거나, 그 자체가 항암제일 수 있다. 따라서, 표적 항원 코딩 서열을 포함하는 RRV에 감염된 암 세포는 암세포상의 표적 발현을 증가시켜, 세포독성 표적화의 효율/효능을 증가시킨다.

[0120] 또다른 구현예에서, 본 발명의 RRV는 동족 항원 또는 리간드와 특이적으로 상호작용하는 결합 도메인(예를 들어, 항체, 항체 단편, 항체 도메인 또는 수용체 리간드)을 포함하는 코딩 서열을 포함할 수 있다. 결합 도메인에 대한 코딩 서열을 포함하는 RRV는 암세포 또는 중양 세포와 같은 세포 증식성 장애를 포함하는 개체에서 세포를 감염시키는데 사용될 수 있다. 그리고나서 감염된 세포는 결합 도메인 또는 항체를 발현할 것이다. 세포독성제에 작동가능하게 연결되거나 세포독성 그 자체가 항원 또는 동족체가 대상에게 투여될 수 있다. 세포독성 동족체는 결합 도메인을 발현하는 감염된 세포를 선택적으로 죽일 것이다. 대안적으로, 결합 도메인 자체가 항암제일 수 있다.

[0121] 본원은 세포 증식성 장애를 갖는 개체를 치료하는 방법을 제공한다. 개체는 임의의 포유동물일 수 있으며, 바람직하게는 인간이다. 개체는 본 발명의 재조합 복제가능 레트로바이러스 백터와 접촉한다. 접촉은 생체내 또는 생체외에서 이루어질 수 있다. 본 발명의 레트로바이러스 백터를 투여하는 방법은 당 분야에 공지되어 있으며, 예를 들어, 전신 투여, 국소 투여, 복강내 투여, 근육내 투여, 두개 내, 뇌척수뿐만 아니라 중양 또는 세포 증식성 장애 부위에 직접 투여하는 것을 포함한다. 당 분야에 공지된 다른 투여 경로도 포함한다.

[0122] 따라서, 본 발명은 세포 증식성 질환을 치료하는데 유용한 다양한 약학 조성물을 포함한다. 본 발명에 따른 약학 조성물은 담체, 부형제, 첨가제 또는 보조제를 사용하여 개체에 투여하기에 적합한 형태로 본 발명에 따른 세포 증식성 장애를 치료하거나 조절하는데 유용한 이종성 폴리뉴클레오타이드 서열을 함유하는 레트로바이러스 백터를 가져옴으로써 제조된다. 자주 사용되는 담체 또는 보조제는 탄산 마그네슘, 이산화 티탄, 락토스, 만니톨 및 다른 당, 활석, 우유 단백질, 젤라틴, 전분, 비타민, 셀룰로스 및 그의 유도체, 동물성 및 식물성 오일, 폴리에틸렌 글리콜 및 용매, 예컨대 멸균수, 알콜, 글리세롤 및 다가 알콜을 포함한다. 정맥내 비히클은 유체 및 영양 보충제를 포함한다. 방부제에는 항균제, 항산화제, 킬레이트제 및 불활성 가스가 포함된다. 다른 약학

적으로 허용가능한 담체는 예를 들어 Remington's Pharmaceutical Sciences, 15th ed. Easton: Mack Publishing Co., 1405-1412, 1461-1487 (1975) and The National Formulary XIV., 14th ed. Washington: American Pharmaceutical Association (1975)에 기술된 바와 같이, 수용액, 무독성 부형제, 예를 들어 염, 방부제, 완충제 등을 포함하며, 이의 내용은 본 명세서에 참고로 포함된다. 약학 조성물의 다양한 성분의 pH 및 정확한 농도는 당 분야의 통상적인 기술에 따라 조정된다. Goodman and Gilman's The Pharmacological Basis for Therapeutics (7th ed.)를 참조한다.

[0123] 다른 구현예에서, 본 발명의 복제가능 레트로바이러스 벡터로 형질감염된 숙주 세포가 제공된다. 숙주 세포는 효모 세포, 곤충 세포 또는 동물 세포와 같은 진핵 세포를 포함한다. 숙주 세포는 또한 박테리아 세포와 같은 원핵 세포를 포함한다.

[0124] 본원에서 제공된 벡터(예를 들어, 복제가능 레트로바이러스 벡터)로 형질도입(형질전환 또는 형질감염)되는 조작된 숙주 세포도 제공된다. 조작된 숙주 세포는 프로모터를 활성화시키고, 형질전환체를 선별하거나, 또는 코딩 폴리뉴클레오타이드를 증폭시키기 위해 적절하게 변형된 통상적인 영양 배지에서 배양될 수 있다. 온도, pH 등과 같은 배양 조건은 발현을 위해 선택된 숙주 세포와 함께 이전에 사용된 것들이며, 예를 들어, Sambrook, Ausubel and Berger, 뿐만 아니라 예를 들어, Freshney (1994) Culture of Animal Cells: A Manual of Basic Technique, 3rd ed. (Wiley-Liss, New York) 및 그 문헌에 인용된 참고 문헌을 포함한, 본원에 인용된 참고문헌들 및 당업자에게 명백할 것이다.

[0125] 적당한 발현 숙주의 예는: 박테리아 세포, 예컨대 대장균(*E. coli*), *B. 서브틸리스*(*B. subtilis*), 스트렙토마이세스(*Streptomyces*) 및 살모넬라 티피뮤리움(*Salmonella typhimurium*); 균류 세포, 예컨대 사카로마이세스 세레비시애(*Saccharomyces cerevisiae*), 피치아 파스토리스(*Pichia pastoris*) 및 뉴로스포라 크라사(*Neurospora crassa*); 곤충세포, 예컨대 드로소필라(*Drosophila*) 및 스포도테라 프루기페르다(*Spodoptera frugiperda*); 포유동물 세포, 예컨대, CHO, COS, BHK, HEK 293 br Bowes 흑색종; 또는 식물 세포 또는 외식편 등을 포함한다. 일반적으로, 인간 세포 또는 세포주가 사용될 것이지만; 시퀀싱, 증폭 및 클로닝의 목적으로 본 발명의 벡터 및 폴리뉴클레오타이드를 비-인간 숙주 세포 내로 클로닝하는 것이 바람직할 수 있다.

[0126] 하기 실시예는 본 발명을 예시하기 위한 것으로서, 본 발명을 제한하고자 하는 것은 아니다. 이러한 실시예는 사용될 수 있는 전형적인 것이지만, 당업자에게 알려진 다른 절차가 대안적으로 이용될 수도 있다.

[0127] **실시예**

[0128] 실시예 1: RRV-2A-GFPm, RRV-GSG-2A, RRV-2A-yCD2 및 RRV-GSG-2A-yCD2의 설계.

[0129] RRV-yCD2 및 RRV-GFP는 양친매성 외피 유전자 및 뇌심근염 바이러스 내부 리보솜 진입 부위(IRES) - env 유전자 하류의 이식유전자 카세트를 갖는 폴리 MLV-계 RRV이다(Perez et al, 2012). RRV-2A-GFP(일명 pAC3-2A-GFP) 및 RRV-2A-yCD2(pAC3-2A-yCD2) 벡터는 RRV-GFP 및 RRV-yCD2에 기초하지만, IRES 영역은 양친매성 외피 단백질 및 이식유전자(GFP 또는 yCD2)와 인-프레임(in-frame)으로 다양한 다른 2A 펩타이드로 대체되었다. RRV-2A-GFP 및 RRV-yCD2 벡터에 대한 클로닝 계획의 개요가 도 3에 도시되어있다. pAC3-T2A-GFP 구조체는 처음에 2개의 DNA 단편 및, BstB I 및 Not I 부위로 소화된 pAC3-emd 백본을 함유하는 김슨 어셈블리 클로닝 키트(Gibson Assembly Cloning Kit)(NEB)를 사용하여 생성되었다. 우선, 양친매성 env의 3' 말단, 토세아 아시그나 바이러스(T2A)의 2A 펩타이드, 및 5'에서 3'의 순서로 GFP의 5'의 서열을 포함하는 한 쌍의 센스 및 안티센스 올리고뉴클레오타이드를 합성하고(IDT), 하이브리드화하여, DNA 단편 2A-G를 생성시켰다. 김슨 어셈블리의 두번째 DNA 단편은 FP 단편이다(도 3). FP 단편은 하기 프라이머들을 사용하여 PCR에 의해 생성하였다: GFP-F-Gib (5'-GAAGTCG AGGGCGACAC-3') 및 GFP-R-Gib (5'-TAAATCTTTTATTTATCTGCGGCCGCAC-3').

[0130] 2A-G 단편에서, 5'는 pAC3 백본의 양친매성 env에서 BstBI 부위와 중첩되는 서열을 포함하며; 3'은 FP DNA 단편의 5'와 중첩되는 서열을 포함한다. 또한, AscI 제한 효소 부위는 제2 이식유전자인 GFP의 개시 코돈의 즉시 상류의 T2A의 3'-말단에 위치한다. AscI 부위의 포함은 T2A 펩타이드를 다른 2A 펩타이드로 후속 대체하기 위한 것이다. AscI 제한 부위가 뒤따르는, 추가적인 뉴클레오타이드 T를 갖는 AscI 제한 부위의 포함은 T2A 펩타이드의 마지막 프롤린 잔기에 추가로 3개의 아미노산(글리신-알라닌-프롤린) C-말단을 얻었다. 동시-번역 과정동안, T2A 펩타이드에 의해 매개되는 외피 단백질로부터 GFP 단백질을 분리하면, GFP의 N-말단에 4개의 아미노산 P, G, A 및 P가 추가로 생성된다. FP 단편에서, FP 단편의 5'-말단은 24개 뉴클레오타이드에 의해 2A-G 단편의 3'-말단에 중첩되는 서열을 함유하고, FP 단편의 3'-말단은 26개 뉴클레오타이드에 의해 Not I 부위에 걸친 pAC3-GFP 백본의 5'-말단에 중첩한다. 김슨 어셈블리 클로닝에서 생성된 플라스미드 DNA를 pAC3-T2A-GFP로 명명하였

다(도 3).

- [0131] 패지 테스코바이러스-1(P2A), 구제역 바이러스(F2A) 및 마비엄 A 바이러스(E2A)로부터 유래된 3개의 다른 통상적으로 사용되는 2A 펩타이드를 2개의 상이한 배열로 보유하는 추가의 RRV-2A-GFP 벡터를 이후에 합성하였다 (IDT). 각 DNA 단편은 BstBI 및 AscI 부위에 pAC3-T2A-GFP 백본의 T2A 대신에 지정된 2A 펩타이드 및 양친매성 env 유전자의 3'의 서열을 함유한다(도 3). 생성된 플라스미드 DNA를 pAC3-P2A-GFP, pAC3-F2A-GFP, pAC3-E2A-GFP, pAC3-GSG-T2A-GFP, pAC3-GSG-P2A-GFP, pAC3-GSG-F2A-GFP, 및 pAC3-GSG-E2A-GFP라고 명명한다.
- [0132] 기술된 RRV-2A-GFP 플라스미드 DNA(pAC3-E2A-GFP, pAC3-F2A-GFP, pAC3-P2A-GFP, pAC3-T2A-GFP, pAC3-GSG-E2A-GFP, pAC3-GSG-F2A-GFP, pAC3-GSG-P2A-GFP, 및 pAC3-GSG-T2A-GFP)는 모두 GFP의 3'-말단에 정지 코돈 돌연변이를 포함한다는 것이 나중에 측정되었다. FP PCR 단편을 생성할 때 GFP-R-Gib 프라이머(5'-TAAAATCTTTTATTTTATCTGCGGCCGCAC-3'(서열식별번호:4))에 돌연변이가 도입되었다. PCR로부터 유래된 GFP에서의 정지 코돈 돌연변이는 정지 코돈에 도달하기 전에 추가의 11개의 아미노산(C-A-A-A-D-K-I-K-D-F-I(서열식별번호: 5))에 대한 GFP ORF를 판독시켰다. 상기 플라스미드 DNA를 pAC3-E2A-GFPm, pAC3-F2A-GFPm, pAC3-P2A-GFPm, pAC3-T2A-GFPm, pAC3-GSG-E2A-GFPm, pAC3-GSG-F2A-GFPm, pAC3-GSG-P2A-GFPm, 및 pAC3-GSG-T2A-GFPm으로 다시 명명하였다. 이후, 2개의 명명법, pAC3-E2A-GFP/pAC3-E2A-GFPm, pAC3-F2A-GFP/pAC3-F2A-GFPm, pAC3-P2A-GFP/pAC3-P2A-GFPm, pAC3-T2A-GFP/pAC3-T2A-GFPm, pAC3-GSG-E2A-GFP/pAC3-GSG-E2A-GFPm, pAC3-GSG-F2A-GFP/pAC3-GSG-F2A-GFPm, pAC3-GSG-P2A-GFP/pAC3-GSG-P2A-GFPm, 및 pAC3-GSG-T2A-GFP/pAC3-GSG-T2A-GFPm이 상호교환적으로 사용된다.
- [0133] pAC3-P2A-GFPm, pAC3-GSG-P2A-GFPm, pAC3-T2A-GFPm 및 pAC3-GSG-T2A-GFPm 플라스미드 DNA의 각각의 2A 펩타이드 버전에서 GFPm 오픈 리딩 프레임을 yCD2 ORF로 대체함으로써 4개의 RRV-2A-yCD2 벡터의 동등한 세트를 생성시켰다(도 3). 프라이머:AscI-yCD2-F(5'-GATCGGCGCGCCTATGGTGACCGGCGGCATGGC-3'(서열식별번호:6) 및 3-37 (5'-CCCCTTTTCTGGAGACTAAATAA-3'(서열식별번호:7)을 사용하여 pAC3-yCD2 플라스미드 DNA로부터 AscI-yCD2-NotI PCR 단편을 생성하였다. PCR 산물과 4개의 pAC3-2A-GFPm 플라스미드 DNA 각각을 AscI 및 NotI로 제한 효소 소화시키고, AscI-yCD2-NotI 소화된 PCR 산물을 GFPm 대신에 서브클로닝하여, pAC3-P2A-yCD2, pAC3-GSG-P2A-yCD2, pAC3-T2A-yCD2, 및 pAC3-GSG-T2A-yCD2를 생성시켰다(표 3).

[표 3]

서열, 2A 펩타이드의 공급원 및 RRV 플라스미드-2A 펩타이드-이식유전자 명칭.

뉴클레오타이드서열 (GSG-링커서열밑줄침)	2A 의공급원 (감염된종)	RRV-2A-GFP 플라스미드
GAGGGCAGAGGAAGTCTTCTAACATGCGGTGACGTG GAGGAGAATCCCGGCCCT (SEQ ID NO:8)	토세아아시그나바이러스(곤충)	pAC3-T2A-GFP
GGAAGCGGAGAGGGCAGAGGAAGTCTTCTAACATGC GGTGACGTGGAGGAGAATCCCGGCCCT (SEQ ID NO:9)	토세아아시그나바이러스(곤충)	pAC3-GSG-T2A- GFP
GCTACTAATTTCAGCCTGCTGAAGCAGGCTGGAGAC GTGGAGGAGAACCCTGGACCT (SEQ ID NO:10)	돼지테스코바이러스-1 (포유동물)	pAC3-P2A-GFP
GGAAGCGGAGCTACTAATTTCAGCCTGCTGAAGCAG GCTGGAGACGTGGAGGAGAACCCTGGACCT (SEQ ID NO:11)	돼지테스코바이러스-1 (포유동물)	pAC3-GSG-P2A- GFP
GTGAAACAGACTTTGAATTTTGACCTTCTCAAGTTG GCGGGAGACGTGGAGTCCAACCCTGGACCT (SEQ ID NO:12)	구제역바이러스 (포유동물)	pAC3-F2A-GFP
GGAAGCGGAGTGAACAGACTTTGAATTTTGACCTT CTCAAGTTGGCGGAGACGTGGAGTCCAACCCTGGACCT (SEQ ID NO:13)	구제역바이러스 (포유동물)	pAC3-GSG-F2A- GFP
CAGTGTACTAATTATGCTCTCTTGAATTTGGCTGGA GATGTTGAGAGCAACCCTGGACCT (SEQ ID NO:14)	마비염 A 바이러스 (포유동물)	pAC3-E2A-GFP
GGAAGCGGACAGTGTACTAATTATGCTCTCTTGA TTGGCTGGAGATGTTGAGAGCAACCCTGGACCT (SEQ ID NO:15)	마비염 A 바이러스 (포유동물)	pAC3-GSG-E2A- GFP

뉴클레오타이드서열 (GSG-링커서열밑줄침)	2A 의공급원 (감염된종)	RRV-2A- yCD2 플라스미드
GAGGGCAGAGGAAGTCTTCTAACATGCGGTGACGTG GAGGAGAATCCCGGCCCT (SEQ ID NO:16)	토세아아시그나바이러스(곤충)	pAC3-T2A-yCD2
GGAAGCGGAGAGGGCAGAGGAAGTCTTCTAACATGC GGTGACGTGGAGGAGAATCCCGGCCCT (SEQ ID NO:17)	토세아아시그나바이러스(곤충)	pAC3-GSG-T2A- yCD2
GCTACTAATTTCAGCCTGCTGAAGCAGGCTGGAGAC GTGGAGGAGAACCCTGGACCT (SEQ ID NO:18)	돼지테스코바이러스-1 (포유동물)	pAC3-P2A-yCD2
GGAAGCGGAGCTACTAATTTCAGCCTGCTGAAGCAG GCTGGAGACGTGGAGGAGAACCCTGGACCT (SEQ ID NO:19)	돼지테스코바이러스-1 (포유동물)	pAC3-GSG-P2A- yCD2

[0134]

[0135]

[0136]

실시에 2: 293T 세포로부터 생산된 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 박터는 감염성이며, GFP 단백질을 발현한다.

HEK293T 세포를 형질감염(transfection)전 18 내지 20시간전에, 10cm 플레이트 당 2e6 세포로 시딩하였다. 다음날, 인산칼슘 방법을 사용하여 세포 시딩후 20시간 후에 20 μg의 플라스미드 DNA의 일시적인 형질감염을 위해 pAC3-2A-GFPm 및 pAC3-GSG-2A-GFPm 플라스미드를 사용하였다. 형질감염 18시간 후, 세포를 DMEM 완전 배지로 3회 세척하고, 신선한 완전한 배양 배지로 배양하였다. 형질감염 후 약 42시간 후에 바이러스 상등액을 수집하고, 0.45 μm 주사기 필터를 통해 여과하였다. HEK293T 세포의 일시적인 형질감염으로부터의 RRV-2A-GFPm, RRV-GSG-2A-GFPm 및 RRV-IRES-GFP의 바이러스 역가를 이전에 기술된 바와 같이 측정하였다(Perez 등, 2012). 간략하게, 박터의 단일주기 감염에 의해 PC3 세포상에서 박터 준비 역가를 측정하였다. 바이러스 박터 DNA에 특이적인 표적 세포 계능 DNA(MLV LTR 프라이머 세트; 5'-MLV-U3-R(5'-AGCCACAAACCCCTCACTC-3'(서열식별번호:20)), 3-MLV-Psi(5'-TCTCCGATCCCGACGA-3'(서열식별번호:21)), 및 프로브(5'-FAM-CCCCAAATGAAAGACCCCGCTGACG-BHQ1-3'(서열식별번호:22))를 감염 후 24시간후 아지도타미딘 치료하고, 그후 정량 PCR(qPCR)한후, 감염 후 48시간 후에 세포 계능 당 바이러스 DNA 복제본의 수를 정량화하여, 단일주기 감염을 보장하였다. 밀리리터당 형질도입 단위(TU)로 보고된 바이러스 역가(TU/ml)를 2×10⁷개의 복제본 내지 2×10¹개의 복제본의 플라스미드 DNA 및 공지량의 계능 DNA 입력, 세포 수 및 반응 혼합물 당 바이러스 스톱의 희석으로부터 구한 표준 곡선으로부터 유도된 임계주기(CT) 값의 계산에 의해 결정하였다. 표 4는 HEK293T 세포로부터 생산된 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm의 역가가 RRV-IRES-GFP의 역가와 유사하다는 것을 보여준다.

【표 4】

293T 세포로부터 생산된 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 벡터의 역가

	TU/mL	Stdv
pAC3-E2A-GFP	1.15E+06	2.55E+05
pAC3-F2A-GFP	1.63E+06	2.58E+05
pAC3-P2A-GFP	1.81E+06	3.11E+05
pAC3-T2A-GFP	3.31E+06	1.32E+05
pAC3-GSG-E2A-GFP	1.65E+06	2.76E+05
pAC3-GSG-F2A-GFP	1.32E+06	7.57E+04
pAC3-GSG-P2A-GFP	1.31E+06	1.22E+05
pAC3-GSG-T2A-GFP	2.66E+06	2.14E+05
pAC3emd	1.65E+06	2.12E+05

[0137]

[0138]

HEK293T 세포로부터 생산된 RRV-2A-GFPm 바이러스를 사용하여 0.01의 감염다중도(MOI)로 U87-MG를 감염시켰다. 초기 감염을 위해 6-웰 플레이트에 U87-MG 세포를 1×10^5 세포로 시딩하였다. 세포를 각 통로에서 1 내지 4 희석하여 6-웰 플레이트의 새로운 웰로 계대시키고, 각 샘플로부터의 나머지 세포를 수확하여, BD FACS Canto II(BD Biosciences)를 사용하여 GFPm 발현 세포의 비율 및 GFPm 평균 형광 세기를 측정함으로써 바이러스 확산을 평가하였다. 각 계대에서 GFP-양성 세포의 백분율을 도시하였다. 모든 RRV-2A-GFP 바이러스가 최대 감염도(~95% 이상의 GFP-양성 세포)에 도달할 때까지 분석의 길이를 측정했다. 도 4는 HEK293T 세포로부터 생산된 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm이 감염성임을 보여준다. RRV-2A-GFPm과 RRV-GSG-2A-GFPm 사이의 바이러스 확산 속도는 감염된 U87-MG 세포에서 지연(lag)을 나타내는 RRV-P2A-GFPm, RRV-T2A-GFPm 및 RRV-GSG-F2A-GFPm을 제외하고 RRV-IRES-GFP와 유사했다. 그럼에도 불구하고, 그들은 18일 이내에 최대한의 감염력에 도달했다. GFPm 발현 수준은 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 벡터에서도 다양했지만, 모두 RRV-IRES-GFP 감염된 U87-MG 세포에서 발현된 것의 약 20 내지 50%였다(도 5).

[0139]

실시예 3: RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 벡터는 U87-MG 세포에서 안정하다. RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 감염된 U87-MG 세포에서 감소된 GFP 발현이 바이러스 게놈내 GFP 유전자의 결실로 인한 것이 아님을 보장하기 위해, 프로바이러스 DNA의 3'env와 3'UTR 영역에 걸쳐있는 프라이머 세트를 사용한 종점 PCR에 의해, 2A-GFPm 부위의 완전성을 평가하였다. U87-MG 세포의 최대 감염력에서, T75 플라스크에서 컨플루언시(confluency)에 도달하기 위해 세포를 배양한 후, 배지를 새로운 배지로 교체하였으며, 배지 교체후 18 내지 24시간후 상등액 및 0.45 μ M의 여과물을 함유하는 바이러스를 수집하였다. 수집된 세포 상등액을 분주하고, 면역블로팅 및 재-감염 실험을 위해 사용될 때까지 -80°C 에서 보관하였다. 동시에, 세포를 2개의 분획으로 분할시켰다; 게놈 DNA 분리를 위한 1/10번째 및 전체 세포 용해물의 분리를 위한 9/10번째. 400 μ L 1X PBS에 재현탁하여 게놈 DNA를 세포 펠렛으로부터 추출하고, Promega 맥스웰 16 세포 DNA 정제 키트(Promega)를 사용하여 분리하였다. 프라이머 세트: IRES-F(5'-CTGATCTTACTCTTTGGACCTTG-3'(서열식별번호:23)) 및 IRES-R(5'-CCCCTTTTCTGGAGACTAAATAA-3'(서열식별번호:24))에 의한 PCR을 위한 주형으로서 100 ng의 게놈 DNA를 사용하였다. 생성된 PCR 생성물을 1% 아가로스 겔 상에서 분석하였다. 데이터는 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 벡터의 프로바이러스 DNA에서 2A-GFPm 및 GSG-2A-GFPm 영역이 바이러스 복제의 시간 경과 동안 U87-MG 세포에서 안정하다는 것을 보여준다(도 6).

[0140]

실시예 4: 최대 감염된 U87-MG 세포로부터 생산된 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm은 후속 감염 주기에서 감염성을 유지한다. 장기간 감염성이 RRV에 의해 전달되는 치료 효과를 유지하는 많은 중요한 기준 중 하나이기 때문에, 최대 감염된 U87-MG 세포에서 생산된 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm의 감염성은 나이브(naive) U87-MG 세포에서 추가적인 감염 주기를 수행함으로써 평가하였다. 최대 감염된 U87-MG 세포로부터 수집된 바이러스 상등액을 먼저 기술된 바와 같이 적정한 다음, 0.01의 MOI로 나이브 U87-MG 세포 상으로 재-감염시켰다. 도 7은 일시적으로 형질감염된 HEK293T 세포로부터 얻은 역가와 유사했던 최대 감염된 U87-MG 세포에서 생성된 역가가 RRV-IRES-GFP 벡터뿐만 아니라 RRV-2A-GFPm, RRV-GSG-2A-GFPm 벡터 중에서 비슷하다는 것을 보여준다.

[0141]

기술된 바와 같이 각 세포 계대에서 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm의 바이러스 확산을 모니터링하였다. 일시적으로 형질감염된 HEK293T 세포에서 생산된 바이러스 상등액을 사용하는 제1 감염주기에서 관찰된 바이러스 확산 속도와 대조적으로, 도 8은 모든 벡터가 RRV-IRES-GFP와 비슷한 속도로 퍼진다는 것을 보여준다. 그러나, 이 감염주기에서 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 감염된 U87-MG 세포로부터의 GFP 발현 수준은 이전에 관찰된 바

와 같이 RRV-IRES-GFP 세포에 의해 발현된 수준의 20 내지 50%였다(도 9).

[0142] 실시예 5: RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 벡터의 바이러스 외피 및 GFPm 단백질은 감염된 U87-MG 세포에서 상이한 효율로 처리된다.

[0143] GFPm 발현, 바이러스 외피 단백질로부터의 GFPm의 분리 효율 및 바이러스 외피 단백질의 적절한 처리를 평가하기 위해, 감염된 U87-MG 세포로부터 세포 용해물을 생성하였다. 세포 단일층이 최대 감염력으로 함유한 U87-MG 세포를 1X PBS에서 1회 세척하고, TrpZean(Sigma)에 의해 해리시키고, 완전한 DMEM에 재현탁시키고, 다시 1X PBS에서 세척한 후, 30분간 얼음위 200 μ l의 RIPA 용해 버퍼(Thermo Scientific)에서 세포용해시켰다. 용해물을 14,000 rpm에서 15분간 4°C에서 원심 분리하여 세포 파편을 제거하고, 상등액을 수집하여, 새로운 튜브로 옮겼다. BCA 침전 분석(Thermo Scientific)을 사용하여 세포 용해물을 단백질 농도에 대해 분석하고, 20 μ g 단백질을 SDS-PAGE에 적용하였다. 단백질을 4-12% XT-Tris SDS-PAGE 겔(BioRad)상에서 200볼트에서 45분 동안 용해시켰다. 이어서, 20볼트에서 7분동안 iBlot 건식 블롯팅 시스템을 사용하여 단백질을 PVDF 멤브레인(Life Technologies)으로 옮겼다. 항-gp70(래트 항-gp70, 클론 83A25; 1:500 희석) 및 항-GFP(토끼 항-GFP, 1:1000 희석)를 사용하여, 멤브레인을 외피 단백질의 gp70 서브유닛 및 GFPm의 발현에 대해 분석하였다. 홀스래디쉬 퍼옥시다아제에 접합된 대응하는 2차 항체를 사용하여, 단백질 발현을 검출하였다. 그 결과는 항-GFP 항체를 사용하여 ~ 120 kDa에서 env-2A-GFPm 융합 단백질의 고 분자량으로 표시된 바와 같이, RRV-F2A-GFPm, RRV-P2A-GFPm, 및 RRV-T2A-GFPm, RRV-GSG-F2A-GFPm 및 RRV-GSG-F2A-GFPm로부터의 GFPm 단백질이 바이러스 외피 단백질로부터 비효율적으로 분리되었음을 보여준다(도 10). 대조적으로, RRV-E2A-GFPm, RRV-GSG-P2A-GFPm 및 RRV-GSG-T2A-GFPm 벡터는 RRV-IRES-GFP와 비교하여 바이러스 외피 단백질로부터 GFPm의 분리가 상대적으로 효율적이었다(도 10). 이와 동시에, 항-gp70 항체를 사용하여, 감염된 U87-MG에서 바이러스 외피 단백질의 프로세싱을 조사하였다. 결과는 모든 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 벡터에서 전구체(Pr85) 또는 처리된 형태(gp70)로 둘러싸인 바이러스가 검출되었음을 보여주며(도 11), 이는 항-GFP 면역블롯에서 볼 수 있는 바와 같이, 바이러스 외피 단백질이 GFPm으로부터 분리되었음을 시사한다. 또한, 항-gp70 블롯에서 관찰된 분리 효율은 항-GFP 면역블롯에서 관찰된 것과 다소 일치한다. RRV-2A-GFPm 및 RRV-GSG-2A-GFPm 벡터 사이에서 융합 단백질인 Env-GFPm의 단백질 발현이 다양하지만, 항-GFP 및 항-gp70 면역블롯 모두에서 바이러스 외피-GFPm 융합 다단백질의 검출 부족으로 나타난 바와 같이, RRV-GSG-P2A-GFPm 및 RRV-T2A-GFPm은 가장 효율적인 분리를 가진 것으로 나타났다.

[0144] 실시예 6: 적당하게 가공된 바이러스 외피 단백질의 혼입 수준은 바이러스 외피와 GFPm 단백질 사이의 분리 효율과 상관 관계가 있다.

[0145] 최대 감염된 U87-MG 세포의 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm로부터의 바이러스 상등액을 14000 rpm에서 30분 동안 4°C에서 20% 수크로스 구배를 통해 펠렛화하고, 이어서 5% 2-메르캅토에탄올을 함유하는 20 μ l의 1X Laemmli 버퍼에 재현탁시키고, 4-20% 트리스 글리신 겔(BioRad)상에서 SDS-PAGE에 적용하였다. 전기 영동 및 단백질 전달을 기술된 바와 같이 수행하였다. 항-gp70(래트 사육항 항-gp70, 클론 83A25; 1:500 희석) 및 항-p15E(마우스 사육항 항-TM, 클론 372; 1:250 희석)를 사용하여, 적절하게 처리된 비리온(viron)-결합된 바이러스 외피 단백질 발현을 조사하였다. 단백질 발현은 홀스래디쉬 퍼옥시다아제에 접합된 대응하는 2차 항체를 사용하여 검출되었다. 데이터는 RRV-P2A-GFPm 및 RRV-T2A-GFPm 벡터를 제외한 RRV-2A-GFPm 및 RRV-GSG-2A-GFPm의 적당하게 처리된 외피 단백질, gp70 및 p12E/p15E가 비리온에서 RRV-IRES-GFP와 비슷한 수준으로 검출되었음을 보여준다(도 12). 예상했던대로 가장 낮은 수준의 비리온-결합된 외피 단백질을 나타낸 RRV-GSG-P2A-GFPm과 RRV-T2A-GFPm이 세포 용해물에서 가장 높은 수준의 융합 다단백질을 발현했다. 공개된 데이터와 일치하여, 이 데이터는 처리되지 않은 외피 단백질 전구체 단백질 Pr85 또는 이 경우 바이러스 외피-GFPm 융합 다단백질이 비리온에 통합되지 않는다는 표기법을 지원한다. 또한, "융해성(fusogenic)" p12E로 이어지는 2A 펩타이드를 지닌 R 펩타이드의 분열은 최대 감염된 U87-MG 세포에서 생산된 역가로 표시된 바와 같이, 비리온 성숙동안 감염성 바이러스 입자를 생성하기에 충분한 것으로 나타난다(도 7). 감염 동안, p15E/p12E 비율의 특성과 막 융합에서의 그의 역할은 불분명하다. 모두 함께, 데이터는 바이러스 외피 단백질 통합 수준이 표적 세포에서 측정된 역가 값과 상관 관계가 없다는 것을 시사한다. 벡터, 특히 RRV-GSG-P2A-GFPm 및 RRV-T2A-GFPm 벡터 사이의 역가 값의 차이가 예기치 않게 부족한 것은 RRV 입자 1에서 이들 세포의 역가에 영향을 주지 않으면서 외피 발현 수준의 범위가 용인될 수 있음을 시사한다.

[0146] 실시예 7: 293T 세포로부터 생산된 RRV-P2A-yCD2 및 RRV-T2A-yCD2, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 벡터는 감염성이며, yCD2 단백질을 발현한다.

[0147] HEK293T 세포를 10 cm 플레이트 당 2e6 세포로, 형질감염전 18 내지 20시간전에 시딩하였다. 다음날, 세포 시딩 후 20시간 후 인산칼슘법을 사용하여, 플라스미드 DNA 20 μg 을 일시적으로 형질감염시키기 위해 pAC3-P2A-yCD2, pAC3-T2A-yCD2, pAC3-GSG-P2A-yCD2 및 pAC3-GSG-T2A-yCD2 플라스미드를 사용하였다. 형질감염 후 18시간 후, 세포를 DMEM 완전 배지로 3회 세척하고, 신선한 완전한 배양 배지로 배양하였다. 형질감염 후 약 42시간 후에 바이러스 상등액을 수집하고, 0.45 μm 주사기 필터를 통해 여과하였다. HEK293T 세포의 일시적인 형질감염으로부터의 RRV-P2A-yCD2, RRV-T2A-yCD2, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2의 바이러스 역가를 이전에 기술된 바와 같이 측정하였다(Perez 등, 2012). 간략하게, 벡터의 단일주기 감염에 의해 PC3 세포상에서 벡터 제조 역가를 측정하였다. 바이러스 벡터 DNA에 특이적인 표적 세포 게놈 DNA(MLV LTR 프라이머 세트; 5'-MLV-U3-R(5'-AGCCCACAACCCCTCACTC-3'(서열식별번호:20)), 3'-MLV-Psi(5'-TCTCCCGATCCCGGACGA-3'(서열식별번호:21)) 및 프로브(5'-FAM-CCCAAATGAAAGACCCCGCTGACG-BHQ1-3'(서열식별번호:22))를 감염 후 24시간후 아지도타미딘 치료하고, 그후 정량 PCR(qPCR)한후, 감염 후 48시간 후에 세포 게놈 당 바이러스 DNA 복제본의 수를 정량화하여, 단일-주기 감염을 보장하였다. 밀리리터당 형질도입 단위(TU)로 보고된 바이러스 역가(TU/ml)를 2×10^7 개의 복제본 내지 2×10^1 개의 복제본 범위의 플라스미드 DNA 표준곡선 및 공지량의 게놈 DNA 입력, 세포 수 및 반응 혼합물 당 바이러스 스톱의 희석으로부터 유도된 임계주기(CT) 값의 계산에 의해 결정하였다. 표 5는 HEK293T 세포로부터 생산된 RRV-P2A-yCD2, RRV-T2A-yCD2, RRV-GSG-P2A-yCD2, 및 RRV-GSG-T2A-yCD2의 역가가 RRV-IRES-yCD2의 역가와 비슷하다는 것을 보여준다.

[표 5]

293T 세포로부터 생산된 RRV-P2A-yCD2, RRV-T2A-yCD2, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 벡터의 역가

	TU/mL	Stdv
pAC3P2AyCD2	3.06E+06	4.59E+05
pAC3GSGP2AyCD2	1.15E+06	2.45E+05
pAC3T2AyCD2	2.32E+06	3.78E+05
pAC3GSGT2AyCD2	1.88E+06	4.64E+05
pAC3-yCD2	1.76E+06	1.84E+05

[0148]

[0149] 또한, 최대 감염된 U87-MG 세포로부터 수집된 바이러스 상등액을 감염성을 유지할 수 있도록 기술된 바와 같이 적정하였다. 역가용으로 사용된 프라이머 세트는 5-MLV-U3-R, 3-MLV-Psi 프라이머 및 프로브를 함유하는 프라이머 세트와 유사한 프라이밍 효율을 갖는다. 감염된 U87-MG 세포로부터의 RRV-P2A-yCD2, RRV-T2A-yCD2, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 벡터를 적정하는데 사용되는 프라이머 세트는 다음과 같다: Env2 For: 5'-ACCCTCAACCTCCCTACAAGT-3'(서열식별번호:25), Env2 Rev: 5'-GTTAAGCGCTGATAGGCTC-3'(서열식별번호:26) 및 프로브 5'-FAM-CCCAAATGAAAGACCCCGCTGACG-BHQ1-3'(서열식별번호:27). 도 13은 최대 감염된 U87-MG 세포에서 생성된 역가가 일시적으로 형질감염된 HEK293T 세포로부터 수득된 역가와 유사하고, RRV-IRES-yCD2 벡터와 비슷하다는 것을 보여준다.

[0150] 실시예 8: 감염된 U87-MG 세포에서 RRV-P2A-yCD2 및 RRV-T2A-yCD2, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 벡터의 바이러스 외피 및 yCD2 단백질이 다른 효율로 처리된다.

[0151] yCD2 발현, 바이러스 외피 단백질로부터의 yCD2 단백질의 분리 효율 및 바이러스 외피 단백질의 적절한 처리를 평가하기 위해, 감염된 U87-MG 세포로부터 세포 용해물을 생성시켰다. 세포 단일층이 최대 감염력으로 함유한 U87-MG 세포를 1X PBS에서 1회 세척하고, TrpZean(Sigma)에 의해 해리시키고, 완전한 DMEM에 현탁시키고, 다시 1X PBS에서 세척한 후, 30분간 얼음위 200 μl 의 RIPA 용해 버퍼(Thermo Scientific)에서 세포용해시켰다. 용해물을 14,000 rpm에서 15분간 4°C에서 원심 분리하여 세포 파편을 제거하고, 상등액을 수집하여, 새로운 튜브로 옮겼다. BCA 침전 분석(Thermo Scientific)을 사용하여 세포 용해물을 단백질 농도에 대해 분석하고, 20 μg 단백질을 SDS-PAGE에 적용하였다. 단백질을 4-12% XT-Tris SDS-PAGE 겔(BioRad)상에서 200볼트에서 45분 동안 용해시켰다. 이어서, 20볼트에서 7분동안 iBlot 건식 블롯팅 시스템을 사용하여 단백질을 PVDF 멤브레인(Life Technologies)으로 옮겼다. 항-gp70(래트 항-gp70, 클론 83A25; 1:500 희석) 및 항-yCD2(마우스 항-yCD2, 1:1000 희석)를 사용하여, 멤브레인을 외피 단백질 및 yCD2의 gp70 서브유닛의 발현에 대해 분석하였다. 홀스래디쉬 퍼옥시다아제에 접합된 대응하는 2차 항체를 사용하여, 단백질 발현을 검출하였다. 그 결과는 항-yCD2 항

체를 사용하여 ~ 110 KDa에서 env-2A-yCD2 융합 단백질의 고 분자량으로 표시된 바와 같이, RRV-P2A-yCD2 및 RRV-T2A-yCD2로부터의 yCD2 단백질이 바이러스 외피 단백질로부터 비효율적으로 분리되었음을 보여준다(도 14). 대조적으로, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2의 경우 RRV-IRES-yCD2와 비교하여 바이러스 외피 단백질로부터 yCD2의 분리가 상대적으로 효율적이었다(도 14). 이와 동시에, 항-gp70 항체를 사용하여, 감염된 U87-MG에서 바이러스 외피 단백질의 프로세싱을 조사하였다. 결과는 모든 RRV-GSG-P2A-yCD2, RRV-GSG-T2A-yCD2 벡터에서 전구체(Pr85) 또는 처리된 형태(gp70)로 둘러싸인 바이러스가 쉽게 검출가능하지만 RRV-P2A-yCD2 및 RRV-T2A-yCD2 벡터에서는 훨씬 낮은 수준으로 검출가능함을 보여주었다(도 15). 또한, Pr85/gp70 바이러스 외피 단백질의 수준은 항-yCD2 면역블롯에서 관찰된 것과 다소 일치한다. 그러나, RRV-2A-GFPm 또는 RRV-GSG-2A-GFPm 벡터와는 달리, 항-gp70 항체 또는 항-A2 항체(Cat#ABS31, EMD Millipore)를 사용하여 바이러스 외피-yCD2 융합 다단백질을 검출할 수 없었다. 4개의 벡터 중, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 벡터는 항-yCD2 면역블롯에서 바이러스 외피-yCD2 융합 다단백질의 검출 부족으로 나타난 바와 같이, 융합 다단백질의 가장 효율적인 분리를 나타냈다. GSG-P2A 및 GSG-T2A 배열이 RRV 외피 단백질 오픈 리딩 프레임의 맥락에서 가장 효율적인 다단백질 분리를 야기한다는 것을 모든 데이터가 시사한다.

[0152] 실시예 9: RRV-G2G-P2A-YCD2 및 RRV-GSG-T2A-yCD2는 U87-MG 세포에서 장기간 안정성을 갖는다. U87-MG 세포에서 RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2의 장기간 벡터 안정성을 평가하기 위해 연속 감염을 수행하였다. 6-웰 플레이트에 시딩된 약 10^5 개의 나이브 U87-MG 세포를 0.1의 MOI에서 바이러스 벡터로 초기 감염시키고, 1주일 동안 배양하여, 감염의 단일주기를 완료하였다. 완전히 감염된 세포에서 얻은 2 ml의 바이러스 상등액 100 μ L를 사용하여 10^5 개의 나이브 세포를 감염시키고, 16개 이하의 사이클을 반복하였다. 게놈 DNA를 400 μ L 1X PBS에 재현탁하여 작은 펠렛에서 추출하고, Promega 맥스웰 16 세포 DNA 정제 키트(Promega)를 사용하여 분리했다. 이식유전자 카세트에 걸쳐있는 프라이머 쌍: IRES-F(5'-CTGATCTTACTCTTTGGACCTTG-3'(서열식별번호:23)) 및 IRES-R(5'-CCCCTTTTCTGGAGACTAAATAA-3'(서열식별번호:24))을 갖는 PCR의 주형으로 게놈 DNA의 100 나노그램을 사용하였다. 감염된 세포로부터의 통합된 프로바이러스의 PCR 증폭에 의해 2A-yCD2 영역의 벡터 안정성을 평가한다. 예상되는 PCR 생성물 크기는 약 0.73kb이다. 0.73kb보다 작은 밴드의 출현은 2A-yCD2 영역에서의 결실을 나타낸다. 도 16a는 RRV-yCD2의 IRES-yCD2(1.2 Kb) 영역이 이전에 보고된 바와 같이 감염 사이클 16까지 안정하다는 것을 보여준다(Perez 등, 2012). 유사하게, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2에서의, 2A-yCD2 영역은 감염 주기 16까지 안정을 유지한다. 그러나 RRV-GSG-T2A-yCD2의 2A-yCD2 영역은 감염 주기 13으로부터 나타난 결실(0.4 kb) 결실로서 RRV-GSG-P2A-yCD2보다 약간 덜 안정하지만, 주기 16 전체에 걸쳐 안정한 상태로 유지되었다(도 16b 및 16c).

[0153] 실시예 10: 적당하게 처리된 바이러스 외피 단백질의 혼입은 RRV-P2A-yCD2 및 RRV-T2A-yCD2, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 벡터에 의해 감염된 U87-MG 세포에서 바이러스 외피와 yCD2 단백질 사이의 분리 효율과 상관 관계가 있다.

[0154] RRV-2A-yCD2 및 RRV-GSG-2A-yCD2 최대 감염된 U87-MG 세포로부터 생성된 바이러스 상등액을 14000 rpm에서 30분 동안 4°C에서 20% 수크로스 구배를 통해 펠렛화하고, 이어서 5% 2-메르캅토에탄올을 함유하는 20 μ l의 1X Laemmli 버퍼에 재현탁시키고, 4-20% 트리스 글리신 겔(BioRad, Hercules CA)상에서 SDS-PAGE에 적용하였다. 전기 영동 및 단백질 전달을 기술된 바와 같이 수행하였다. 항-gp70(래트 사육항 항-gp70, 클론 83A25; 1:500 희석) 및 항-p15E(마우스 사육항 항-TM, 클론 372; 1:250 희석)를 사용하여, 적절하게 처리된 비리온-결합된 바이러스 외피 단백질 발현 및 성숙을 조사하였다. 단백질 발현은 홀스래디쉬 퍼옥시다아제에 접합된 대응하는 2차 항체를 사용하여 검출되었다. 데이터는 RRV-P2A-yCD2 및 RRV-T2A-yCD2를 제외한 RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2의 적당하게 처리된 외피 단백질인, gp70이 비리온내 RRV-IRES-yCD2와 비슷한 수준으로 검출되었음을 보여준다(도 17).

[0155] 중요하게는, 상기 데이터는 적절하게 처리된 바이러스 외피 단백질의 혼입 수준이 역가 값과 상관 관계가 없음을 시사한다.

[0156] 실시예 11: RRV-P2A-yCD2 및 RRV-T2A-yCD2, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 감염된 U87-MG 세포에서 yCD2 단백질 발현 수준이 다양하지만, RRV-IRES-yCD2 감염된 U87-MG 세포와 비슷한 5-FC 민감도를 나타냈다.

[0157] RRV-P2A-yCD2 및 RRV-T2A-yCD2의 면역블롯으로서, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2는 바이러스 외피로부터 분리된 단백질로서, 또는 감염된 U87-MG 세포에서 변화된 융합 다단백질로서 발현된 yCD2 단백질의 양이 LD₅₀ 실험을 수행함으로써 그들의 5-FC 민감도를 측정하였음을 보여주었다. RRV-P2A-yCD2 및 RRV-T2A-yCD2, RRV-

GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 벡터에 의해, 최대 감염된 U87-MG 세포는 MTS 분석법에 의해 5-FC LD₅₀을 측정하는데 사용하였다. 감염된 U87-MG 세포주 또는 감염되지 않은 U87-MG 세포주 각각에 대해, 1×10³ 세포/웰 /100 μL 배양 배지를 96-웰 플레이트에 3중으로 시딩했다. 세포를 0.00001 mM 내지 1 mM 범위의 1:10 희석액인 5-FC(cat#F7129, Sigma)로 처리하였다. 5-FC 치료는 대조군으로 포함되지 않았다. 플레이팅 후 1일후에 5-FC를 첨가한 다음, 2일마다 5-FC를 첨가한 완전한 배지로 보충하였다. 5-FC의 비-5-FU 매개된 세포독성 효과를 측정하기 위한 대조군으로서 나이브 U87-MG 세포가 포함되었다. 세포를 7일간의 배양 시간에 걸쳐 모니터링하고, CellTiter 96 수성 원 솔루션(Aqueous One Solution) 세포 증식 검정 시스템(Promega)을 사용하여 2일마다 세포 사멸을 측정하였다. MTS를 첨가한 후, MTS 배양 후 60분에 Infinite M200(Tecan) 플레이트 판독기를 사용하여 490 nm에서 OD 값을 얻었다. 각 샘플의 3회 복제물(triplicate)에서 얻은 평균 OD 값을 치료되지 않은, RRV-감염된 세포와 비교한 세포 생존율로 환산했다. 그 후, 백분율 값을 GraphPad Prim을 사용하여 로그 스케일로 5-FC 농도에 대해 플롯하여 LD₅₀ 그래프를 생성하였다. LD₅₀ 값은 획득한 데이터 포인트의 비선형 4-변수 적합을 사용한 소프트웨어에 의해 계산되었다. 데이터는 RRV-P2A-yCD2 및 RRV-T2A-yCD2 감염된 U87-MG 세포보다 RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 감염된 U87-MG 세포에서 "분리된" yCD2 단백질의 수준이 더 높았음에도 불구하고, RRV-P2A-yCD2 및 RRV-T2A-yCD2 감염된 U87-MG 세포에서 관찰되는 바이러스 외피-yCD2 융합 단백질은 5-FC를 5-FU로 전환시키는데 있어서 효소적으로 활성이어서, RRV-IRES-yCD2의 LC₅₀ 농도와 유사한 LC₅₀ 농도에서 세포독성 효과를 나타낸다(도 18).

[0158] 실시예 12: RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 감염된 Tu2449 세포는 RRV-IRES-yCD2와 비슷한 5-FC 민감도를 나타냈다

[0159] RRV-GSG-P2A-GMCSF-T2A-yCD2에 의해 최대 감염된 U87-MG 세포를 사용하여 기술된 바와 같이, 5-FC LD₅₀을 MTS 분석으로 측정하였다. RRV-IRES-yCD2를 대조군으로 포함시켰다. 0.00001 mM 내지 1 mM 범위의 1:10 희석액인 5-FC(cat#F7129, Sigma)로 처리하였다. 5-FC 치료는 대조군으로 포함되지 않았다. 플레이팅 후 1일후에 5-FC를 첨가한 다음, 2일마다 5-FC를 첨가한 완전한 배지로 보충하였다. 5-FC의 비-5-FU 매개된 세포독성 효과를 측정하기 위한 대조군으로서 나이브 U87-MG 세포가 포함되었다. 세포를 7일간의 배양 시간에 걸쳐 모니터링하고, CellTiter 96 수성 원 솔루션(Aqueous One Solution) 세포 증식 검정 시스템(Promega)을 사용하여 2일마다 세포 사멸을 측정하였다. MTS를 첨가한 후, MTS 배양 후 60분에 Infinite M200(Tecan) 플레이트 판독기를 사용하여 490 nm에서 OD 값을 얻었다. 각 샘플의 3회 복제물(triplicate)에서 얻은 평균 OD 값을 치료되지 않은, RRV-감염된 세포와 비교한 세포 생존율로 환산했다. 그 후, 백분율 값을 GraphPad Prim을 사용하여 로그 스케일로 5-FC 농도에 대해 플롯하여 LD₅₀ 그래프를 생성하였다. LD₅₀ 값은 획득한 데이터 포인트의 비선형 4-변수 적합을 사용한 소프트웨어에 의해 계산되었다. 데이터는 RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 감염된 Tu-2449세포에 의해 발현된 yCD2 단백질(도 20a)이 RRV-IRES-yCD2의 농도와 유사한 LC₅₀ 농도에서 세포독성 효과를 얻기 위해 5-FC를 5-FU로 전환시키는데 효소적으로 활성임을 나타낸다(도 20b).

[0160] 실시예 13: RRV-GSG-T2A-yCD2로 처리한 피하 동계의 신경교종 마우스는 RRV-IRES-yCD2와 유사한 지연된 종양 성장을 보였다.

[0161] 동계의 세포주 Tu-2449를 B6C3F1 마우스에서 동소 뇌종양 모델로 사용하였다(Ostertag 등, 2012). 피하 종양 모델링을 위해 Tu-2449 세포의 서브라인(Tu-2449SQ)을 확립했다. 98% 나이브 Tu-2449 SQ 세포와 2% RRV-GSG-T2A-yCD2 감염된 Tu-2449SQ 세포의 혼합물을 시험관내 제조하고, 피하 종양 이식을 위해 인산염 완충 식염수(PBS; Hyclone)에 재현탁했다. 98% 나이브 Tu-2449SQ 세포와 2% RRV-IRES-yCD2 감염된 Tu-2449SQ 세포의 혼합물을 양성 대조군 및 비교군으로 포함시켰다. 각 그룹(그룹당 n=10)의 B6C3F1 마우스는 0일에 1×10⁶ 종양 세포의 피하 이식을 받는다. 종양 이식 후 12일 후(약 75% 초과 종양이 RRV에 감염된 시점), 마우스에 45일 연속으로 PBS 또는 5-FC(500mg/kg 체중/용량, i.p., b.i.d.)를 투여한 후, 약물없이 2일간 투여하여 나머지 감염된 세포로부터의 벡터 확산을 허용하였다. 5일간 온(on), 2일간 오프(off) 주기의 약물 치료를 2회 더 반복했다. 종양 체적 측정은 매일 실시되었다. 결과는 5-FC 처리없이 RRV-IRES-yCD2 또는 RRV-GSG-T2A를 보유한 종양을 갖는 마우스가 계속 자라는 것을 나타낸다. 대조적으로, RRV-GSG-T2A 보유 후 5-FC 치료받은 종양보유 마우스를 미리 설정된 종양의 종양 성장을 지연시켰고, RRV-IRES-yCD2 + 5-FC로 치료한 종양과 비슷했다(도 21). 이 데이터는 피하 동계의 신경교종 마우스 모델에서 RRV-GSG-T2A-yCD2가 RRV-IRES-yCD2와 유사한 치료 효능을 가짐을 시사한다.

[0162] 실시예 14: HEK293T 세포로부터 생산된 RRV-GSG-T2A-GMCSF-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2-GSG-PS2-GMCSF

백터는 GMCSF 및 yCD2 단백질을 발현하고, 감염성이다.

- [0163] 각각 5' 및 3' 말단에 존재하는 AscI 및 NotI 제한 부위와 함께 화학 합성된 인간 GMCSF-GSG-P2A-yCD2 및 yCD2-GSG-P2A-GMCSF 카세트(Genewiz)를 AscI 및 NotI 제한 효소로 소화된 pAC3-GSG-T2A-yCD2 백본에 클로닝함으로써 pAC3-GSG-T2A-GMCSF-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2-GSG-P2A-GMCSF가 생성되었다. 생성된 GMCSF-GSG-P2A-yCD2 및 yCD2-GSG-P2A-GMCSF 카세트는 카세트의 N-말단(AscI 제한 부위의 5' 상류)에서 GSG-T2A와 인-프레임으로 있다.
- [0164] 형질감염 전 18 내지 20시간전에, HEK293T 세포를 10-cm 플레이트 당 2e6 세포로 시딩하였다. 다음날, 인산칼슘 방법을 사용하여 세포 시딩후 20시간후에 일시적인 형질감염을 위해 pAC3-GSG-T2A-GMCSF-GSG-P2A-yCD2 또는 pAC3-GSG-T2A-yCD2-GSG-P2A-GMCSF 플라스미드 20 μ g를 사용하였다. 형질감염후 18시간 후, 세포를 CMEM 배지로 3회 세척하고, 신선한 완전한 배지로 배양하였다. 형질감염 후 약 42시간 후에 바이러스 상등액을 수집하고, 0.45 μ m 주사기 필터를 통해 여과하였다. HEK293T 세포의 일시적인 형질감염에 의한 RRV-GSG-T2A-GMCSF-GSG-P2A-yCD2의 바이러스 역가를 기술된 바와 같이 측정하였다. 상기 데이터는 RRV-GSG-T2A-GMCSF-GSG-P2A-yCD2 및 pAC3-GSG-T2A-yCD2-GSG-P2A-GMCSF의 역가(\sim 2E6 TU/mL)가 RRV-IRES-yCD2의 역가와 비슷함을 보여준다.
- [0165] yCD2 단백질 발현을 평가하기 위해, pAC3-GSG-P2A-GMCSF-GSG-T2A-yCD2 또는 pAC3-GSG-T2A-yCD2-GSG-P2A-GMCSF 일시적으로 형질감염된 293T 세포로부터 세포 용해물을 생성하였다. 이 실험에서, pAC3-IRES-yCD2 및 pAC3-IRES-GMCSF 또한 대조군으로 포함되었다. GMCSF 발현을 위해, 일시적으로 형질감염된 293T 세포의 상등액을 ELISA(cat#DGM00, R & D Systems)로 측정하기 위해 수집하였다. 기술된 바와 같이, 전체 세포 용해물을 yCD2 단백질 발현에 대해 분석하였다. 항-yCD2 결과는 pAC3-GSG-P2A-GMCSF-GSG-T2A-yCD2 또는 pAC3-GSG-T2A-yCD2-GSG-P2A-GMCSF로부터의 yCD2 단백질이 \sim 15 kDa 밴드에 의해 나타난 바와 같이, GMCSF로부터 효율적으로 분리됨을 보여준다(도 22a). 그러나, 두 구성 모두 2A 펩타이드에 의해 매개된 GMCSF(pAC3-GSG-P2A-GMCSF-GSG-T2A-yCD2) 또는 바이러스 외피 단백질(pAC3-GSG-T2A-yCD2-GSG-P2A-GMCSF)로부터 yCD2의 분리는 RRV-IRES-yCD2로부터의 yCD2와 비교하여, yCD2의 크기에 의해 나타난 바와 같이, GMCSF로부터 yCD2 단백질의 적절한 분리, 현저히 상이하다(도 22a). 대조적으로, 바이러스 env로부터의 yCD2 단백질 분리는 약간 더 높은 분자량을 가지며(도 22a에서 2A-yCD2로 나타냄), 도 10 및 도 14에 도시된 RRV-GSG-P2A-GFP, RRV-GSG-T2A-GFP, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2 구조체와 일치한다. 데이터는 이론적으로 예상되는 아미노산 서열에서 Env로부터의 yCD2 분리가 정확하게 발생할 수 없다는 것을 시사한다. 그러나, yCD2가 다른 분비된 단백질(즉, GMCSF)의 하류에 놓일 때, yCD2 단백질의 적당한 분리가 관찰된다. 그러나, RRV-GSG-P2A-yCD2 및 RRV-GSG-T2A-yCD2로부터 발현된 2A-yCD2 단백질의 효소 활성은 시험관내 및 생체내 5-FC 민감성 및 세포독성 효과에 영향을 미치지 않는 것으로 보인다는 것이 중요하다(도 20).
- [0166] pAC3-GSG-P2A-GMCSF-GSG-T2A-yCD2 구조체내 바이러스 외피 단백질로부터 또는 pAC3-GSG-T2A-yCD2-GSG-P2A-GMCSF 구조체내 yCD2로부터의 GMCSF 단백질의 분리 효율은 미정이지만, GMCSF ELISA 결과는 분비된 GMCSF의 양이 RRV-GSG-P2A-GMCSF-GSG-T2A-yCD2의 경우 \sim 500 ng/mL이고, RRV-GSG-T2A-yCD2-GSG-P2A-GMCSF의 경우 \sim 760 ng/mL임을 보여준다(도 22b). 두 경우 모두에서 발현된 GMCSF의 양은 RRV-IRES-GMCSF(25 ng/mL)보다 약 20배 내지 30배 정도 많다. 이와 동시에, 감염된 U87-MG에서 바이러스 외피 단백질의 처리가 항-gp70 항체를 사용하여 검사된다. 그 결과는 전구체(Pr85) 또는 가공된 형태(gp70)의 바이러스 외피 단백질이 쉽게 검출가능하다는 것을 보여준다. 이와 함께 데이터는 Env-GSG-T2A-GMCSF-GSG-P2A-yCD2 및 Env-GSG-T2A-yCD2-GSG-P2A-GMCSF 다단백질 구성이 GMCSF 및 yCD2 단백질을 발현할 수 있음을 시사한다.
- [0167] 또한, 최대 감염된 U87-MG 세포로부터 수집된 바이러스 상등액을 바이러스가 감염성을 유지하도록 기술된 바와 같이 적정한다. 이 데이터는 최대 감염된 U87-MG 세포에서 생산된 역가(\sim 3E6 TU/mL)가 일시적으로 형질감염된 HEK293T 세포에서 얻은 역가와 유사하며, RRV-IRES-yCD2와 비슷함을 보여준다.
- [0168] 실시예 15: RRV-GSG-T2A-GMCSF-P2A-yCD2 및 RRV-GSG-T2A-yCD2-P2A-GMCSF 백터는 RRV-IRES-yCD2 감염된 U87-MG 세포의 5-FC 민감도와 유사한 5-FC 민감도를 나타낸다.
- [0169] RRV-GSG-T2A-GMCSF-GSG-P2A-yCD2 또는 RRV-GSG-T2A-yCD2-GSG-P2A-GMCS를 갖는 최대 감염된 U87-MG 세포를 기술된 바와 같은 MTS 분석법에 따라 그의 5-FC LD₅₀을 측정하는데 사용한다. RRV-IRES-yCD2는 대조군으로 포함된다. 이 데이터는 감염된 U87-MG 세포에서 검출된 "분리된" yCD2 단백질의 양이 0.008 mM의 LD₅₀ 농도에서 세포독성 효과를 나타낼 수 있고, RRV-IRES-yCD2와 유사함을 나타낸다.
- [0170] 실시예 16: RRV-GSG-T2A-GMCSF-RSV-yCD2 및 HEK293T 세포 및 최대 감염된 U87-MG 세포로부터 생산된 백터는 감

염성이며, GMCSF 및 yCD2 단백질을 발현한다.

- [0171] pAC3-GSG-T2A-GMCSF-RSV-yCD2는 각각 5' 및 3' 말단에 존재하는 AscI 및 NotI 제한 부위를 갖는 화학적으로 합성된 인간 GMCSF-RSV-yCD2 카세트(Genewiz)를 AscI 및 NotI 제한 효소로 소화된 pAC3-GSG-T2A-yCD2 백본으로 클로닝함으로써 생성되었다. 화학적으로 합성된 GMCSF-RSV-yCD2 카세트는 GMCSF ORF의 3' 말단에 정지 코돈을 함유한다.
- [0172] 형질감염전 18시간 내지 20시간 전에 HEK293T 세포를 10-cm 플레이트 당 2e6 세포로 시딩한다. 다음날, 20 μ g의 pAC3-GSG-T2A-GMCSF-RSV-yCD2 플라스미드를 인산칼슘 방법을 사용하여 세포 시딩후 20시간 후 일시적인 형질감염에 사용한다. 형질감염후 18시간 후, 세포를 DMEM 배지로 3회 세척하고, 신선한 완전한 배양 배지로 배양하였다. 형질감염 후 약 42시간 후에 바이러스 상등액을 수집하고, 0.45 μ m 주사기 필터를 통해 여과하였다. HEK293T 세포의 일시적인 형질감염으로부터 RRV-GSG-T2A-GMCSF-RSV-yCD2의 바이러스 역가를 기술된 바와 같이 측정하였다. 상기 데이터는 RRV-GSG-T2A-GMCSF-RSV-yCD2의 역가(~ 2E6 TU/mL)가 RRV-IRES-yCD2의 역가와 비슷함을 보여준다.
- [0173] 또한, 최대 감염된 U87-MG 세포로부터 수집된 바이러스 상등액을 바이러스가 감염성을 유지하도록 적정한다. 이 데이터는 최대 감염된 U87-MG 세포에서 생산된 역가(~ 2E6 TU/mL)가 일시적으로 형질감염된 HEK293T 세포에서 얻은 역가와 유사하며, RRV-IRES-yCD2와 비슷하다는 것을 보여준다.
- [0174] GMCSF 및 yCD2 단백질 발현을 평가하기 위해, RRV-GSG-T2A-GMCSF-RSV-yCD2 감염된 U87-MG 세포로부터 세포 용해물을 생성하였다. 이 실험에서, RRV-IRES-yCD2 및 RRV-IRES-GMCSF 또한 대조군으로 포함되었다. 최대 감염된 U87-MG 세포의 상등액을 ELISA(R & D Systems)에 의해 GMCSF의 단백질 발현 수준을 측정하기 위해 수집한다. 기술된 전체 세포 용해물을 yCD2 단백질 발현에 대해 분석한다. 항-yCD2 면역블롯 결과는 RRV-GSG-T2A-GMCSF-RSV-yCD2 감염된 U87-MG 세포로부터의 yCD2 단백질이 RRV-IRES-yCD2보다 2 내지 3배 정도 발현된다는 것을 보여준다. 이와 동시에, 감염된 U87-MG에서 바이러스 외피 단백질의 처리가 항-gp70 항체를 사용하여 검사된다. 그 결과는 전구체(Pr85) 또는 가공된 형태(gp70)의 바이러스 외피 단백질이 쉽게 검출가능하다는 것을 보여준다. 예상대로, 바이러스 외피-GMCSF 융합 다단백질은 또한 항-gp70 항체를 사용하여 세포 용해물에서 검출된다. 바이러스 외피 단백질로부터 GMCSF 단백질의 분리가 미정이지만, GMCSF ELISA 결과는 분비된 GMCSF의 양이 ~ 300 ng/mL이고 RRV-IRES-GMCSF(30 ng/mL)의 양보다 약 10배 더 많음을 나타낸다. 이와 함께 데이터는 바이러스 외피 단백질-GSG-T2A-GMCSF-RSV-yCD2 다단백질 구성이 RRV의 맥락에서 GMCSF 및 yCD2 단백질 뿐만 아니라 전염성 바이러스를 생성할 수 있음을 시사한다.
- [0175] 실시예 17: RRV-GSG-T2A-GMCSF-RSV-yCD2 벡터는 RRV-IRES-yCD2 감염된 U87-MG 세포의 5-FC 민감도와 동등한 5-FC 감도를 나타낸다.
- [0176] RRV-GSG-T2A-GMCSF-RSV-yCD2 벡터를 갖는 최대 감염된 U87-MG 세포를 사용하여, 기술된 바와 같이 5-FC LD50을 MTS 분석법으로 측정한다. 이 실험에서, RRV-IRES-yCD2가 대조군으로 포함된다. 데이터는 감염된 U87-MG 세포에서 발현된 yCD2 단백질의 양이 LD₅₀ 농도 0.010 mM에서 세포독성 효과를 달성할 수 있으며, RRV-IRES-yCD2와 유사하다는 것을 나타낸다.
- [0177] 실시예 18: 293T 세포 및 감염된 U87-MG 세포로부터 생성된 RRV-GSG-P2A-yCD2-RSV-PDL1miR30shRNA 벡터는 감염성이며, yCD2 단백질을 발현한다.
- [0178] pAC3-GSG-T2A-yCD2-RSV-miRPDL1은 각각 5' 및 3' 말단에 존재하는 AscI 및 NotI 제한 부위를 갖는 화학적으로 합성된 인간 yCD2-RSV-miRPDL1 카세트(Genewiz)를 AscI 및 NotI 제한 효소로 소화된 pAC3-GSG-T2A-yCD2 백본으로 클로닝함으로써 생성되었다. 화학적으로 합성된 yCD2-RSV-miRPDL1 카세트는 yCD2 ORF의 말단에 정지 코돈을 함유한다.
- [0179] 형질감염전 18시간 내지 20시간 전에 HEK293T 세포를 10-cm 플레이트 당 2e6 세포로 시딩한다. 다음날, 인산칼슘 방법을 사용하여 세포 시딩후 20시간후에 일시적인 형질감염을 위해 pAC3-GSG-T2A-yCD2-RSV-miRPDL1 플라스미드 20 μ g를 사용하였다. 형질감염후 18시간 후, 세포를 DMEM 배지로 3회 세척하고, 신선한 완전한 배양 배지로 배양하였다. 형질감염 후 약 42시간 후에 바이러스 상등액을 수집하고, 0.45 μ m 주사기 필터를 통해 여과하였다. HEK293T 세포의 일시적인 형질감염에 의한 RRV-GSG-T2A-yCD2-RSV-miRPDL1의 바이러스 역가를 기술된 바와 같이 측정하였다. 상기 데이터는 RRV-GSG-T2A-yCD2-RSV-miRPDL1의 역가(~ 2E6 TU/mL)가 RRV-IRES-yCD2의 역가와 유사함을 보여준다.

- [0180] 또한, 최대 감염된 U87-MG 세포로부터 수집된 바이러스 상등액을 바이러스가 감염성을 유지하도록 적정한다. 이 데이터는 최대 감염된 U87-MG 세포에서 생산된 역가(~ 2E6 TU/mL)가 일시적으로 형질감염된 HEK293T 세포에서 얻은 역가와 유사하며, RRV-IRES-yCD2와 유사하다는 것을 보여준다.
- [0181] yCD2 단백질의 발현 및 PDL1 세포 표면 발현의 발현을 측정하기 위해, 최대 감염된 U87-MG 세포를 수확하고, 전체 세포 용해물을 기술된 바와 같이 yCD2 단백질 발현에 대해 분석한다. 항-yCD2 면역블롯 결과는 RRV-GSG-T2A-yCD2-RSV-miRPDL1 감염된 U87-MG 세포로부터의 yCD2 단백질이 항-yCD2 항체를 사용하여 약 15 kDa 밴드로 표시되는 바와 같이, 바이러스 외피 단백질로부터 효율적으로 분리됨을 보여준다. 예상대로, 바이러스 외피-yCD2 융합 다단백질은 항-yCD2 및 항-gp70 항체 모두를 사용하여 세포 용해물에서 검출된다. 이와 동시에, 감염된 U87-MG에서 바이러스 외피 단백질의 처리가 항-gp70 항체를 사용하여 검사된다. 결과는 진구체(Pr85) 또는 가공된 형태(gp70)의 바이러스 외피 단백질이 쉽게 검출가능하다는 것을 보여준다. 또한, 융합 다단백질은 항-yCD2 면역블롯에서 보이는 바와 같이 검출된다.
- [0182] 실시예 19: RRV-GSG-T2A-yCD2-RSV-miRPDL1 감염된 U87-MG 세포는 RRV-IRES-yCD2 감염된 U87-MG 세포와 유사한 5-FC 민감도를 나타낸다.
- [0183] RRV-GSG-T2A-yCD2-RSV-miRPDL1 벡터를 갖는 최대 감염된 U87-MG 세포를 기술된 바와 같은 MTS 분석법에 따라 그의 5-FC LD₅₀을 측정하는데 사용한다. 이 실험에서 RRV-IRES-yCD2는 대조군으로 포함된다. 이 데이터는 감염된 U87-MG 세포에서 검출된 "분리된" yCD2 단백질의 양이 LD₅₀ 농도(0.008 mM)에서 RRV-IRES-yCD2와 유사한 세포독성 효과를 나타낼 수 있음을 나타낸다.
- [0184] 실시예 20: RRV-GSG-P2A-yCD2-RSV-miRPDL1 감염된 MDA-MB231 세포는 세포 표면에서 강력한 PD-L1 억압을 나타낸다.
- [0185] RRV-GSG-T2A-yCD2-RSV-miRPDL1의 PDL1 억압을 평가하기 위해, 0.1의 MOI를 사용하여 현저한 수준의 PDL1을 발현하는 것으로 나타난 MDA-MB231 세포를 감염시켰다. 이 실험에서 RRV-RSV-miRPDL1은 PDL1 억압을 평가하기 위한 양성 대조군으로 포함된다. 감염 후 대략 14일 후, 세포를 수확하고, 세포 표면 염색을 수행하여 FACS에 의한 PDL1 단백질의 수준을 측정한다. 데이터는 RRV-GSG-T2A-yCD2-RSV-miRPDL1에 의해 감염된 MDA-MB231 세포에서 PDL1의 세포 표면 발현이 약 75% 감소하였고, RRV-RSV-miRPDL1의 세포 표면 발현과 유사하다는 것을 보여준다. 이와 함께 데이터는 바이러스 외피 단백질-GSG-T2A-yCD2-RSV-miRPDL1 구성이 RRV의 맥락에서 감염성 바이러스, yCD2 단백질 및 miRPDL1을 생성할 수 있음을 시사한다.
- [0186] 실시예 21: HEK293T 세포 및 최대 감염된 U87-MG 세포로부터 생산된 RRV-P2A-TKO RRV-GSG-P2A-TKO, RRV-T2A-TKO 및 RRV-GSG-T2A-TKO 벡터는 감염성이며, TKO 단백질을 발현한다
- [0187] pAC3-P2A-TKO, pAC3-GSG-P2A-TKO, pAC3-T2A-TKO 및 pAC3-GSG-T2A-TKO는 인간 코돈 최적화(TKO)(국제 출원 공개 번호 WO2014/066700 참조, 본원에 참고로 포함됨)에 의해 Sr39-tk 카세트를 pAC3-2A 백본으로 클로닝함으로써 생성되었다(Black 등, *Cancer Res.*, 61:3022-3026, 2001; Kokoris 등, *Protein Science* 11:2267-2272, 2002). AscI 및 NotI 제한 효소로 분열된 pAC3-GSG-P2A-yCD2 또는 pAC3-GSG-T2A-yCD2 백본에 각각 5' 및 3' 말단에 존재하는 AscI 및 NotI 제한 효소 부위를 갖는 TKO의 서열을 화학적으로 합성하였다(Genewiz).
- [0188] 형질감염 전 18시간 내지 20시간 전에 HEK293T 세포를 10-cm 플레이트 당 2e6 세포로 시딩한다. 다음날, 인산칼슘 방법을 사용하여 세포 시딩 후 20시간 후에 일시적인 형질감염을 위해 pAC3-GSG-P2A-TKO 또는 pAC3-GSG-T2A-TKO 플라스미드 20 µg를 사용하였다. 형질감염 후 18시간 후, 세포를 DMEM 배지로 3회 세척하고, 신선한 완전한 배지로 배양하였다. 형질감염 후 약 42시간 후에 바이러스 상등액을 수집하고, 0.45 µm 주사기 필터를 통해 여과하였다. HEK293T 세포의 일시적인 형질감염에 의한 RRV-P2A-TKO, RRV-GSG-P2A-TKO, RRV-T2A-TKO 및 RRV-GSG-T2A-TKO의 바이러스 역가를 기술된 바와 같이 측정하였다. 상기 데이터는 상기 역가가 RRV-IRES-yCD2의 역가와 유사함을 보여준다(표 6).

【표 6】

HER293T 세포로부터 생성된 RRV-P2A-TKO RRV-GSG-P2A-TKO, RRV-T2A-TKO 및 RRV-GSG-T2A-TKO 벡터의 역가

적정된 샘플	회석	qPCR 복제물의 역가 (TU/ml)			회석 복제물의 평균		
		웰 1	웰 2	웰 3	트랜스 복제물	Std Dev	CV (%)
5 RRV-RSV-GFP	1	7.90E+05	6.97E+05	8.71E+05	8.05E+05	1.03E+05	12.80%
6 RRV-RSV-GFP	1	8.42E+05	6.81E+05	9.47E+05			
7 RRV-RSV-TKO	1	4.85E+05	5.63E+05	4.91E+05	4.97E+05	4.29E+04	8.63%
8 RRV-RSV-TKO	1	5.13E+05	4.31E+05	4.99E+05			
9 RRV-P2A-TKO	1	1.14E+06	1.26E+06	1.28E+06	1.12E+06	1.59E+05	14.21%
10 RRV-P2A-TKO	1	1.16E+06	8.69E+05	1.00E+06			
11 RRV-GSG-P2A-TKO	1	1.03E+06	9.75E+05	9.84E+05	1.07E+06	8.40E+04	7.85%
12 RRV-GSG-P2A-TKO	1	1.18E+06	1.14E+06	1.12E+06			
13 RRV-T2A-TKO	1	9.51E+05	1.09E+06	1.07E+06	1.15E+06	1.34E+05	11.66%
14 RRV-T2A-TKO	1	1.28E+06	1.21E+06	1.29E+06			
15 RRV-GSG-T2A-TKO	1	1.17E+06	1.55E+06	1.36E+06	1.53E+06	2.42E+05	15.78%
16 RRV-GSG-T2A-TKO	1	1.62E+06	1.88E+06	1.60E+06			
17 RRV-GSG-T2A-GFP	1	2.16E+06	1.80E+06	1.42E+06	1.65E+06	3.09E+05	18.70%
18 RRV-GSG-T2A-GFP	1	1.73E+06	1.38E+06	1.41E+06			
19 RRV-IRES-GFP	1	8.12E+05	9.68E+05	7.31E+05	7.73E+05	1.18E+05	15.25%
20 RRV-IRES-GFP	1	7.73E+05	7.45E+05	6.07E+05			
21 Mock 293T Sup	1	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
22 Mock 293T Sup	1	5.17E+06	#VALUE!	#VALUE!			
23 NBTGOT033pg01-R00137 (Exp 273)	200	2.38E+08	1.66E+08	1.64E+08	1.93E+08	4.11E+07	21.32%
24 NBTGOT033pg01-R00137 (Exp 273)	200	2.53E+08	1.70E+08	1.66E+08			

[0189]

[0190]

또한, 최대 감염된 U87-MG 세포로부터 수집된 바이러스 상등액을 바이러스가 감염성을 유지하도록 기술된 바와 같이 적정한다. 이 데이터는 최대 감염된 U87-MG 세포에서 생산된 역가가 일시적으로 형질감염된 HEK293T 세포로부터 얻은 역가와 유사함을 보여준다(도 23).

[0191]

TKO 단백질 발현을 평가하기 위해, RRV-P2A-TKO RRV-GSG-P2A-TKO, RRV-T2A-TKO 및 RRV-GSG-T2A-TKO 감염된 U87-MG 세포로부터 세포 용해물을 생성하였다. 전체 세포 용해물을 1:200에서 항-HSV-tk 항체(cat#sc28037, Santa Cruz Biotech Inc)를 사용하여 TKO 단백질 발현을 분석하였다. 그 결과, RRV-P2A-TKO 및 RRV-T2A-TKO 감염된 U87-MG 세포의 TKO 단백질은 GFP 및 yCD2 이식유전자에 의해 알 수 있는 바와 같이, RRV-GSG-P2A-TKO 및 RRV-GSG-T2A-TKO(도 24)보다 덜 효율적으로 분리된다는 것을 알 수 있다.

[0192]

실시예 22: RRV-P2A-TKO RRV-GSG-P2A-TKO, RRV-T2A-TKO 및 RRV-GSG-T2A-TKO 벡터는 U87-MG 세포에서 안정하다.

[0193]

최대 감염된 U87-MG 세포에서의 벡터 안정성을 평가하기 위해, 게놈 DNA를 Promega Maxwell 16 세포 DNA 정제 키트(Promega)를 사용하여 세포로부터 추출하였다. 이식유전자 카세트에 걸쳐있는 프라이머 쌍: IRES-F(5'-CTGATCTACTCTTTGGACCTTG-3'(서열식별번호:23)) 및 IRES-R(5'-CCCCTTTTCTGGAGACTAAATAA-3'(서열식별번호:24))을 갖는 PCR의 주형으로 게놈 DNA의 100 나노그램을 사용하였다. 모든 RRV-2A-TKO 구조체에 대한 예상 PCR 산물은 1.4kb이다. 데이터는 프로바이러스 DNA RRV-P2A-TKO RRV-GSG-P2A-TKO, RRV-T2A-TKO 및 RRV-GSG-T2A-TKO 벡터에서의 2A-TKO 및 GSG-2A-TKO 영역이 바이러스 복제 시간과정동안 U87-MG 세포에서 안정적임을 보여준다(도 25).

[0194]

실시예 23: RRV-P2A-TKO, RRV-GSG-P2A-TKO, RRV-T2A-TKO 및 RRV-GSG-T2A-TKO 감염된 U87-MG 세포는 RRV-S1-TKO에 대한 우수한 GCV 민감도를 나타냈다

[0195]

RRV-P2A-TKO, RRV-GSG-P2A-TKO, RRV-T2A-TKO 및 RRV-GSG-T2A-TKO를 갖는 최대 감염된 U87-MG 세포를 사용하여, MTS 분석에 의한 GCV LD₅₀를 측정하였다. 합성 최소 프로모터에 의해 TKO 발현이 구동된 RRV-S1-TKO(국제 공개 공보 WO2014/066700 참조, 본 명세서에 참고로 포함됨)가 대조군으로 포함되었다. GCV(cat#345700-50MG, EMD Millipore)를 사용한 처리는 0.0001 μM 내지 0.5 μM 범위의 일련의 1:2 희석으로 수행하였다. GCV 치료는 대조군으로 포함되지 않았다. 플레이트링 후 1일후에 GCV를 첨가한 다음, 2일마다 완전한 배지 + GCV를 보충하였다. 나이브 U87-MG 세포는 GCV의 세포독성 효과를 결정하는 대조군으로 포함되었다. 세포를 7일간의 배양 시간에 걸쳐 모니터링하고, CellTiter 96 Aqueous One Solution 세포 증식 검정 시스템(Promega)을 사용하여 2일마다 세포 사멸을 측정하였다. MTS를 첨가한 후, M90 배양 후 60분에 Infinite M200(Tecan) 플레이트 판독기를 사용하여 490 nm에서 OD 값을 얻었다. 각 샘플의 3개의 복제물에서 얻은 평균 OD 값을 치료되지 않은 RRV 감염 세포와

비교한 세포 생존율로 환산했다. GraphPad Prim을 사용하여 백분율 값을 로그 스케일에서 GCV 농도에 대해 플롯팅하여, LD₅₀ 그래프를 생성하였다. 획득한 데이터 포인트의 비선형 4-변수 핏(fit)을 사용한 소프트웨어에 의해 LD₅₀ 값을 계산하였다. 상기 데이터는 RRV-P2A-TKO, RRV-GSG-P2A-TKO, RRV-T2A-TKO 및 RRV-GSG-T2A-TKO에 의해 발현되는 TKO 단백질이 10분의 1 밀리몰 범위에서 GCV를 세포독성 GCV로 전환시키는데 효소적으로 활성이어서, 세포독성 효과를 달성하였음을 나타낸다(도 26). RRV-S1-TKO과 비교하여, RRV-P2A-TKO, RRV-GSG-P2A-TKO, RRV-T2A-TKO 및 RRV-GSG-T2A-TKO는 GCV 민감도가 12.5-20배 더 높음을 보여준다. 또한, Env-TKO 융합 다단백질로부터 TKO 분리의 차이가 있음에도 불구하고, RRV-P2A-TKO 대 RRV-GSG-P2A-TKO 또는 RRV-T2A-TKO 대 RRV-GSG-T2A-TKO 사이의 GCV LD₅₀에는 유의한 차이가 없었다. 2A-yCD2와 유사하게, 데이터는 세포에서 발현된 TKO 단백질의 양이 GCV를 세포독성 GCV로 전환시키기에 충분함을 시사한다.

[0196] 실시예 24: RRV-GSG-P2A-TKO 및 RRV-GSG-T2A-TKO로 처리한 피하, 동계의 신경교종 마우스는 RRV-IRES-yCD2의 것과 비슷한 종양 성장 지연을 나타낸다.

[0197] 동계의 세포주 Tu-2449를 B6C3F1 마우스에서 동소 뇌종양 모델로 사용하였다(Ostertag 등, 2012). 피하 종양 모델링을 위해 Tu-2449 세포의 서브라인(Tu-2449SQ)을 Tocagen에서 확립했다. 98% 나이브 Tu-2449 SQ 세포와 2% RRV-GSG-P2A-TKO, RRV-GSG-T2A-TKO 또는 RRV-S1-TKO 감염된 Tu-2449SQ 세포의 혼합물을 시험관내 제조하고, 피하 종양 이식을 위해 인산염 완충 식염수(PBS; Hyclone)에 재현탁했다. 98% 나이브 Tu-2449SQ 세포와 2% RRV-IRES-yCD2 감염된 Tu-2449SQ 세포의 혼합물을 양성 대조군 및 비교군으로 포함시켰다. 각 그룹(그룹당 n=10)의 B6C3F1 마우스는 0일에 1×10^6 종양 세포의 피하 이식을 받는다. 종양 이식 후 12일 후(약 75% 초과 종양이 RRV에 감염된 시점), 마우스에 5일 연속으로 PBS 또는 5-FC(500mg/kg 체중/용량, i.p., b.i.d.) 또는 GCV(50mg/kg 체중/용량, i.p., b.i.d.)를 투여한 후, 나머지 감염된 세포로부터의 백터 확산을 허용하는 약물없이 2일간 투여하였다. 5일간 온(on), 2일간 오프(off) 주기의 약물 치료를 2회 더 반복했다. 종양 체적 측정은 매일 실시되었다. 결과는 GCV 처리없이 RRV-GSG-P2A-TKO, RRV-GSG-T2A-TKO 또는 RRV-S1-TKO를 보유하거나 5-FC 처리없이 RRV-IRES-yCD2 또는 RRV-GSG-T2A를 보유한 종양을 갖는 마우스가 계속 자라는 것을 나타낸다. 대조적으로, 종양 치료된 RRV-GSG-P2A-TKO, RRV-GSG-T2A-TKO + GCV를 보유한 마우스는 사전 설정된 종양의 종양 성장을 지연시킨다. 게다가 RRV-S1-TKO + GCV로 치료한 종양 보유 마우스는 감소된 TKO 발현으로 인해, 종양 치료된 RRV-GSG-P2A-TKO, RRV-GSG-T2A-TKO + GCV보다 정도가 더 낮고 시간이 더 길더라도, 종양 성장의 지연을 나타낸다. 함께, 이 데이터는 RRV-GSG-P2A-TKO + GCV 및 RRV-GSG-T2A-TKO + GCV의 종양 성장 지연이 RRV-IRES-yCD2 + 5-FC로 치료한 것과 비슷하다는 것을 나타낸다. 이 데이터는 피하 동계의 신경교종 마우스 모델에서 RRV-GSG-P2A-TKO 및 RRV-GSG-T2A-TKO가 RRV-IRES-yCD2와 유사한 치료 효능을 가짐을 시사한다.

[0198] 실시예 25: HEK293T 세포 및 최대 감염된 U87-MG 세포로부터 생산된 RRV-GSG-T2A-PDL1scFv 및 RRV-GSG-T2A-PDL1scFvFc 벡터는 감염성이며, scFv 및 scFvFc 단백질을 발현한다.

[0199] 인간 및 마우스 PDL1에 대한 차단 단일 사슬 가변 단편(scFv)으로서 기능하도록 pAC3-T2A-PDL1scFv, pAC3-T2A-PDL1scFv-태그, pAC3-T2A-PDL1scFvFc 및 pAC3-T2A-PDL1scFvFc-태그가 생성되었다. PDL1scFv 카세트는 인간 IgG1의 단편 결정화가능한(Fc) 영역을 갖거나 갖지 않도록 설계된다. 또한, scFv 또는 scFvFc의 C-말단에 혼입된 HA 및 Flag 에피토프 태그를 갖는 매칭 카세트가 scFv 또는 scFvFc 단백질 발현의 검출을 위해 생성되었다. 각각의 카세트(PDL1scFv, PDL1scFv-태그, PDL1scFvFc 및 PDL1scFvFc-태그)의 서열을 각각 5' 및 3' 말단에 AscI 및 NotI 제한 효소 부위를 갖도록 화학적으로 합성하고(Genewiz), AscI 및 NotI 제한 효소로 소화된 pAC3-GSG-T2A-yCD2 백본으로 클로닝되었다.

[0200] 형질감염전 18시간 내지 20시간 전에 HEK293T 세포를 10-cm 플레이트 당 2e6 세포로 시딩한다. 다음날, 인산칼슘 방법을 사용하여 세포 시딩후 20시간후에 일시적인 형질감염을 위해 pAC3-T2A-PDL1scFv, pAC3-T2A-PDL1scFv-태그, pAC3-T2A-PDL1scFvFc 및 pAC3-T2A-PDL1scFvFc-태그 플라스미드 20 μ g를 사용하였다. 형질감염 18시간 후, 세포를 DMEM 배지로 3회 세척하고 신선한 완전한 배지로 배양하였다. 바이러스 상등액을 형질감염 후 약 42시간 후에 수집하고 0.45 μ m 주사기 필터를 통해 여과하였다. HEK293T 세포의 일시적인 형질감염으로부터의 RRV-GSG-T2A-GMCSF-GSG-P2A-yCD2의 바이러스 역가를 기술된 바와 같이 결정하였다. 데이터는 RRV-GSG-T2A-PDL1scFv, RRV-GSG-T2A-PDL1scFvFc, RRV-GSG-T2A-PDL1scFv-태그, RRV-GSG-T2A-PDL1scFvFc-태그의 역가가 RRV-IRES-yCD2의 역가와 비슷하다는 것을 보여준다(표 7).

【표 7】

일시적으로 형질감염된 HEK293T 세포로부터의 RRV-GSG-T2A-PDL1scFv, RRV-GSG-T2A-PDL1scFvFc, RRV-GSG-T2A-PDL1scFv-태그, RRV-GSG-T2A-PDL1scFvFc-태그의 역가

	TU/mL	Std Dev
RRV-PDL 1scFv	2.09E+06	4.80E+05
RRV-PDL 1scFv Fc	1.98E+06	4.38E+05
RRV-PDL 1scFv-태그	2.08E+06	6.73E+05
RRV-PDL 1scFv Fc-태그	1.29E+06	1.87E+05

[0201]

[0202]

scFv 단백질 발현을 평가하기 위해, RRV-GSG-T2A-PDL1scFv 및 RRV-GSG-T2A-PDL1scFvFc 형질감염된 HEK293T 세포로부터 세포 용해물을 생성하였다. 전체 세포 용해물을 항-Flag 및 항-HA 항체(cat#1804 및 cat#H3663, Sigma Aldrich)를 1:1,000으로 사용하여 scFv 단백질 발현에 대해 분석하였다. 그 결과, 이전에 GFP 및 yCD2 및 TKO 이식유전자로 관찰된 바와 같이, RRV-GSG-T2A-PDL1scFv-태그, RRV-GSG-T2A-PDL1scFvFc-태그 일시적으로 형질감염된 HEK293T 세포로부터의 PDL1scFv-태그 및 PDL1scFvFc-태그 단백질 발현이 Env-scFv 다단백질로부터 분리된다(도 27a).

[0203]

이와 동시에, 항-A2 항체를 사용하여, HEK293T 세포에서 바이러스 외피 단백질의 프로세싱을 조사하였다. 결과는 2A 펩타이드 서열을 함유하는 전구체(Pr85) 또는 처리된 형태(p15E)로 둘러싸인 바이러스가 모든 4개 벡터에서 검출되었음을 보여주며(도 27b), 이는 항-Flag 및 항-HA 면역블롯에서 볼 수 있는 바와 같이, 바이러스 외피 단백질이 scFv 및 scFvFc 단백질로부터 분리되었음을 시사한다. 융합 다단백질, Env-scFv 또는 Env-scFvFc 발현이 세포 용해물에서 검출되지만, 세포 용해물 및 상등액으로부터의 면역블롯에 의해 나타난 바와 같이, 상당량의 PDL1scFv 및 PDL1scFvFc 단백질이 융합 다단백질로부터 분리된다.

[0204]

유사하게, 풍부한 scFv-태그 및 scFvFc-태그 단백질 발현은 항-Flag 항체를 이용한 면역침전에 이어서, 항-HA에 의한 검출에 의해, 일시적으로 형질감염된 HEK293T 세포로부터의 상등액에서 검출되었으며, 역으로도 마찬가지이다. 또한, 상등액 뿐만 아니라 scFv-태그 및 scFvFc-태그 단백질 발현 세포 용해물 또한 최대 감염된 MDA-MB231(인간 유방암 세포주) 및 CT26(뮤린 대장 암 세포주) 세포에서 일시적으로 형질감염된 HEK293T 세포보다 약 2 내지 3배 이상 적은 수준으로 검출된다.

[0205]

실시예 26: RRV-GSG-T2A-PDL1scFv 및 RRV-GSG-T2A-PDL1scFvFc는 PHA-자극된 T-세포 활성화를 회복시키고, 시험 관내에서 PDL1 차단 항체의 균등성을 보여준다.

[0206]

RRV-GSG-T2A-PDL1scFv 또는 RRV-GSA-T2A-PDL1scFvFc에 의한 종양 세포에서의 PDL1 차단이 PDL1-매개된 T-세포 억제제를 완화시킬 수 있는지를 결정하기 위해, 본 발명자들은 PDL1-매개된 트랜스-억제 공동-배양 실험을 수행한다. 본원에서, 다양한 종양 세포주에서 PDL1 발현의 조절이 IFN γ 의 세포내 발현 또는 상등액으로의 IFN γ 방출에 의해 측정된 바와 같이, 건강한 공여자 PBMC의 PHA-자극된 활성화를 변경할 수 있는지를 평가한다. 트랜스-억제 공동-배양분석에서 IFN γ 전처리의 잠재적인 다발성 영향을 제거하기 위해, 본 발명자들은 높은 PDL1 기저 세포 발현수준을 갖는, 인간 유방암 세포주인 MDA-MB-231을 사용하여 공동-배양 시스템을 구축했다. 이 분석에서 PDL1 결합의 필요성을 확인하기 위해, 항-PDL1 차단 항체가 포함되어 있다. 항 PDL1 차단 항체의 존재하에 PDL1⁺ 종양 세포 MDA-MB-231 세포는 IFN γ +/CD8⁺ T 세포의 빈도 증가에 의해 나타난 바와 같이, CD8⁺ T-세포 활성화를 억제할 수 없다. 유사하게, RRV-GSG-T2A-scFv 또는 RRV-GSG-T2A-scFvFc에 감염된 MDA-MB-231 세포는 CD8⁺ T-세포 활성화를 균등하게 회복시켰다. PDL1 차단 scFv에 의한 종양 세포 및 림프구의 PDL1: PD1 축의 파괴는 항-PDL1 차단 항체와 유사한 활성을 나타내며, RRV-GSG-T2A-PDL1scFv 및 RRV-GSG-T2A-PDL1scFvFc로부터의 실질적인 면역학적 이익에 대하여 입증을 제공한다.

[0207]

실시예 27: RRV, TOCA-511, 돌연변이 프로파일링.

[0208]

다양한 종양 유형은 가변적으로 신속한 RRV 복제를 지지할 수 있으며, 이 가변성은 RRV Toca 511(일명 T5.0002) 및 상급 신경교종을 위한 프로드럭 Toca FC 치료와 같은 RRV 기반 치료법에 대한 상이한 종양의 감수성을 변화시킬 수 있다(T.F. Cloughsey 등, Sci Transl Med., 8(341):341ra75, June1, 2016, doi: 10.1126/scitranslmed.aad9784.). 이 변화는 여러 가지 요인에 기인한 것 같지만 환자의 혈액 또는 종양으로부터 회수된 변형된 효모 시토신 데아미나제를 코딩하는 RRV의 시퀀싱 데이터로부터 APOBEC 기능, 특히 APOBEC3B

및 APOBEC3B에 의한 변형과 관련이 있는 것으로 보인다(B.P. Doehle 등, J.Virol.79: 8201-8207, 2005). 발현의 변형은 복제하는 레트로바이러스 벡터에서 종양 조직에서 점진적으로 복제됨에 따라 불활성화 또는 감쇠 돌연변이가 축적되는 빈도로부터 유추된다. 조사에 따르면, 가장 빈번한 반응 중 하나가 G에서 A로의 돌연변이이며, 이는 역전사 단계의 첫 번째 복제 단계에서 음성 가닥 단일 가닥 DNA에서 APOBEC 매개 돌연변이의 C에서 T로의 전이 특성에 해당한다. 이러한 돌연변이는 RRV 단백질의 아미노산 조성 변화, 예를 들어 TGG(트립토판)에서 정지 코돈(TAG, TGA 또는 TAA)으로의 엄청난 변화를 일으킬 수 있다. 일부 종양(특히 방광, 자궁 경부암, 폐암(선암종과 편평 세포 암종), 두경부암 및 유방암에서, APOBEC3B 활성이 상향 조절되며, 이 상향 조절은 APOBEC3B 활성과 일치하는 변화를 갖는 증가된 돌연변이 부하와 관련이 있는 것으로 나타났다(MB.Burns 등, Nature Genetics 45: 977-83, 2013; doi: 10.1038/ng.2701).이 상향 조절의 배경이 되는 추진 요인은 돌연변이율이 높을수록 종양의 진화와 종양 유익한 유전자형 및 표현형을 위한 선택을 지지하는 것으로 제안된다. 일 구 현예에서, APOBEC에 의해 전환되지 않을 폐닐알라닌 또는 티로신과 같은 유사한 화학적 또는 구조적 성질을 갖는 다른 아미노산에 대한 코돈의 치환에 의해 바이러스의 불활성화 변화가 회피된다. Toca 511은 프로드럭 5-FU의 세포독성 5-FU로의 전환을 촉매하는, IRES에 연결된 열안정성 코돈 최적화된 효모 시토신 데아미나제를 코딩하는 MLV 유래된 RRV이다. Toca 511 치료 과정동안, Toca 511은 APOBEC-매개된 시티딘 데아미나제와 같은 역전사 및 세포성 항-바이러스 방어 기전의 오류로 인해 돌연변이되기 쉽다. APOBEC 단백질은 주로 Toca 511 RNA 게놈의 역전사 동안 단일 가닥 DNA를 표적으로 하며, G에서 A 지점으로 나타난다.

[0209]

Toca 511 서열 돌연변이 스펙트럼은 종양 및 혈액으로부터 분리된 임상 샘플로부터의 Toca 511의 고효율 시퀀싱에 의해 프로파일링되었다. G에서 A로의 점 돌연변이는 Toca 511에서 가장 일반적인 돌연변이 유형으로, APOBEC 활성과 일치한다(도 28). 이것은 고효율 시퀀싱을 통한 인간 샘플의 감마-레트로바이러스 유전자 치료 돌연변이 스펙트럼의 첫 번째 특성 분석이다. G에서 A로의 돌연변이의 분석은 이들이 대개 코딩 서열의 비동적 변화를 유도한다는 것을 보여준다. 시토신 데아미나제 폴리펩타이드를 코딩하는 유전자 내에는 다수의 환자로부터의 샘플에서 재발성 G에서 A로의 돌연변이를 갖는 2개의 위치가 있었다(표 8). 이러한 돌연변이는 트립토판을 코딩하는 TGG를 TGA, TAG 또는 TAA 정지 코돈으로 변환하여, 9개의 아미노산만으로 CD 번역을 종결시킨다. 이러한 결과는 트립토판 코돈이 레트로바이러스 유전자 요법의 불활성화의 잠재적 원천임을 강조한다.

【표 8】

Toca 511의 재조합 시토신 데아미나제(서열식별번호: 28-29)에서의 점 돌연변이의 개요. 위치는 CD 단백질 내의 아미노산 위치이다. 샘플은 돌연변이를 나타낸 혈액 또는 종양의 임상 샘플 수를 나타낸다. 코돈과 변화는 원래의 코돈 순서와 후속 변화를 보여준다. AA는 원래의 코돈에 의해 코딩된 원래의 아미노산이며, 변화는 코돈 돌연변이 이후에 아미노산이 변화되는 것을 보여준다.

뉴클레오타이드	위치	샘플	코돈	변화	AA	변화
29	10	17	TGG	TAG	W	정지
30	10	5	TGG	TGA	W	정지
31	11	1	GAT	AAT	D	N
40	14	1	GGC	AGC	G	S
45	15	1	ATG	ATA	M	I
105	35	2	GGC	GAC	G	D
144	48	1	AGG	AAG	R	K
159	53	1	AGG	AAG	R	K
168	56	6	AAG	AAA	R	K
216	72	1	GGC	GAC	G	S
357	119	1	GAG	AAG	E	K
456	152	4	TGG	TAG	Q	정지

[0210]

- [0211] 따라서, 트립토판 코돈을 단백질 기능과 양립가능한 아미노산을 코딩하는 대체 코돈으로 변화시키는 것은 레트로바이러스 유전자 치료의 APOBEC 매개 불활성화를 완화시킬 수 있다.
- [0212] 안정성에 대한 돌연변이의 효과를 시험하기 위해, Toca 511 게놈 서열(예를 들어, 미국 특허 제8,722,867호, '867 특허의 서열식별번호: 19, 20 및 22 참조, 본원에 참고로 포함됨)은 안정성과 기능을 보존하는 대체 아미노산을 코딩하는 코돈으로 ApoBec 과돌연변이를 나타내는 코돈을 변경(예를 들어, 트립토판의 코돈을 일부 다른 허용가능한 아미노산으로 변경)하도록 설계되었다. 시토신 데아미나제 활성을 갖는 Toca 511 폴리펩타이드(서열 식별번호: 29 참조)는 자연 발생 균류 시토신 데아미나제 단백질과 밀접하게 관련되며, 이러한 시토신 데아미나제의 고해상도 구조가 이용가능하다. 따라서, 식물유전학적으로 다양한 균류 CD 단백질의 구조적 배열과 다중 서열 배열의 조합을 이용하여, ROSETTA, Provean, PSIPred 또는 이와 유사한 프로그램을 사용하여 생물학적 기능에 악영향을 미치지 않을 잠재적인 아미노산 치환을 식별할 수 있다. Toca 511 게놈을 변경하고, 효소 및 생물학적 활성, 용해도, 용액내 열 안정성뿐만 아니라 세포 배양 분석 및 5-FC에서 5-FU로의 전환과 같은 마우스 종양 모델에서 기능하는 능력을 측정하여 추정 아미노산 치환 세트를 시험하며, 세포 사멸을 개시하고, 종양에 대한 면역 반응을 활성화시켜 내구성있는 반응을 달성한다. 유사한 분석을 GAG, POL 및 ENV 서열에 사용하여 상기 서열을 변형시켜, ApoBec 과돌연변이되기 쉬운 코돈을 제거할 수 있다.
- [0213] 실시예 28: APOBEC 내성 yCD 바이러스 벡터는 누드 마우스에서 두개내 인간 이종 이식편(T98G)에서 치료제이다.
- [0214] APOBEC를 고도로 발현하는 T98G 인간 신경교종 세포주를 사용하는 두개내 이종 이식 모델은 높은 APOBEC 활성 조건 하에, 누드 마우스 숙주에서 APOBEC-내성 RCR-벡터 매개된 시토신 데아미나제 자살 유전자 치료 효능 뿐만 아니라 RRV 벡터 확산 및 생체 분포를 시험하기 위해 확립되었다 .
- [0215] 순응시킨 후, 마우스를 무작위로 9개의 치료 그룹 중 하나에 배정한다(하기 그룹 설명 참조). 8개 그룹은 0일째 마우스당 투여된 1×10^5 T98G 세포의 우측 선조체에 두개내 투여된다. 그룹 9 마우스에는 종양이 이식되지 않았다. 5일째에, 제형 버퍼만, 9×10^5 TU/5 μ l의 T5.0002(yCD를 나타내는 APOBEC-민감성 RRV; 그룹 3) 또는 9×10^5 TU/5 μ l, 9×10^4 TU/5 μ l 또는 9×10^3 TU/5 μ l의 APOBEC-내성 RCR 벡터(T5.002A)를 마우스에게 주사한다. 무작위 5-FC 투여를 500 mg/kg/day에서 시행하고, 19일부터는 단일 IP 주사로 투여하거나, 일부 그룹에게는 5-FC를 투여하지 않는다(그룹, 1, 4, 8). 중간 용량에서 벡터를 수신하는 마우스는 모두 5-FC를 투여받는다(즉, 이 용량에 대해 별도의 대조군이 없음). 5-FC 투여를 매일 연속 7일동안 계속하고, 이어서 15일간 치료받지 않는다. 약물과 휴식의 주기는 최대 4주기까지 반복된다. 그룹 8을 제외한 각 그룹으로부터 10마리의 마우스를 생존 분석 카테고리 무작위로 배정한다. 나머지 마우스는 예정된 스케줄에 따라 희생된다.

그룹 배정 및 투여량 수준

그룹	시험 물품	용량	약물 TX	N	분석 카테고리당 N	
					(A) 생존 분석	(B) 스케줄된 희생
1	형성 버퍼	5 µl	없음	4		제1 약물 주기 전에 4마리
2	형성 버퍼	5 µl	5-FC	10	10	
3	T5.0002	9e5/5µl	5FC	25	10	각 주기 개시전에 3마리, 총 15마리
4	T5.0002A	9e5/5µl	PBS	10	10	
5	T5.0002A	9e5/5µl	5FC	25	10	각 주기 개시전에 3마리, 총 15마리
6	T5.0002A	9e4/5µl	5FC	10	10	
7	T5.0002A	9e3/5µl	5FC	25	10	각 주기 개시전에 3마리, 총 15마리
8	T5.0002A	9e3/5µl	PBS	10	10	
9	종양 없음	없음	5FC	15		각 주기 개시전에 3마리, 총 15마리
총 동물 수				134	70	64

[0216]

[0217]

정맥내 투여는 꼬리 정맥 내로의 주입을 통해 수행된다. 복강내 투여는 방광을 피하기 위해 주의 깊게 복부에 주입하여 시행된다. 두개내 주입의 경우 마우스를 이소플루란으로 마취시키고, 블런트 이어 바아(blunt ear bar)를 갖춘 정위 장치(stereotaxic device)에 위치시킨다. 피부를 면도하고, 수술 부위를 준비하기 위해 두피를 치료하는데 베타디엔을 사용한다. 동물을 가열 패드 위에 놓고, 피부를 통해 정중선 절개를 하기 위해 살균 조건에서 메스를 사용한다. 피부의 후퇴와 절개 부위의 근막 반사는 두개골의 시각화를 가능하게 할 것이다. 3.5mm 돌출부가 있는 캡이 달린, 3mm 돌출부가 있는 가이드 캐놀라를 두개골의 작은 구멍에 삽입하고, 치과용 시멘트와 세개의 작은 나사를 두개골에 부착한다. 시멘트가 경화된 후, 피부를 봉합사로 봉합한다. 투시된 정위 좌표는 AP=0.5-1.0 mm, ML=1.8-2.0 mm, DV=3.0 mm이다. 동물 코호트에 대한 정확한 정위 좌표는 조종사 실험(2-3마리 동물)에서 염료를 주입하고, 그 위치를 결정함으로써 측정된다. 마취 회복동안 동물을 모니터링한다. 진통제인, 부프레노르핀(buprenorphine)을 시술이 끝나기 전에 피하(SC) 투여하고, 부프레노르핀을 약 12시간마다 최대 3일 동안 투여한다. 동물은 매일 모니터링한다. 세포 또는 벡터를 가이드 캐놀라를 통해 삽입된 3.5mm 돌출부가 있는 주입 캐놀라를 통해 두개내 주입된다. 속도는 헤밀턴 주사기 및 유연한 튜브가 장착된 주사기 펌프로 제어된다. 세포 주입의 경우, 1 마이크로리터의 세포가 분당 0.2 마이크로리터의 유속(총 5분)으로 전달된다. 벡터 주입의 경우, 5 마이크로리터의 벡터가 분당 0.33 마이크로리터의 유속(총 15분)으로 전달된다.

[0218]

APOBEC-내성 벡터는 마우스에게 뇌 중량 1g 당 형질전환 단위(TU)로서 전달되고 계산된다. 상기 계산을 사용하여, 사람을 포함한 다른 포유동물에 대해 용량의 변환을 계산할 수 있다. APOBEC-내성 벡터는 효과적인 용량 반응을 나타내지만, APOBEC 활성에 민감한 벡터는 유효 반응 감소를 나타낸다. 이중 이식 모델에 이식된 U87 자연 수준보다 적어도 3배 이상 상기 단백질 발현하는 인간 APOBEC3G 또는 APOBEC3B에 대한 발현 벡터로 형질감염된 U87 세포주에서 동일한 실험을 수행한다. 이 실험은 APOBEC 내성이 되도록 설계된 변형된 코돈 바이러스가

APOBEC 내성을 위한 코돈 변형이 없는 원래의 RRV보다 증가된 APOBEC 수준으로 U87 계통에서 복제 및/또는 치료 반응 이점을 가짐을 보여준다.

- [0219] 실시예 29: APOBEC 내성 yCD 바이러스 벡터는 뇌암의 동계의 마우스 모델에서 치료제이다.
- [0220] 동계의 동물 모델에서 본 발명의 방법 및 조성물을 입증하기 위한 추가의 실험이 수행된다.
- [0221] 동계의 BALB/c 마우스에서 유린 APOBEC3을 생산하도록 안정적으로 형질감염된 CT26 대장 암 세포주를 사용하는 두개내 주입 모델을 확립하여 APOBEC-내성 RRV 벡터 확산 및 생체 분포 뿐만 아니라 RRV-벡터 매개 시토신 데아미나제 자살 유전자 치료와 그 면역학적 영향의 치료 효능을 시험하였다.
- [0222] 이 연구에는 129마리의 동물, 0마리의 수컷, 119마리의 암컷 및 10마리의 우연성 동물(10마리의 암컷)이 포함된다. 순응시킨 후, 마우스를 무작위로 9개 치료 그룹 중 하나에 배정한다(아래 그룹 설명 참조). 8개의 그룹은 0 일째에 마우스에게 투여된 1×10^4 APOBEC-발현 CT26 세포의 우측 선조체 내로의 두개내 투여를 받는다. 그룹 9 마우스에는 종양 이식되지 않는다. 4일째에, 제형 버퍼만, 9×10^5 TU/5 μ l의 APOBEC에 대해 여전히 민감성인 대조군 벡터(T5.0002), 또는 9×10^5 TU/5 μ l, 9×10^4 TU/5 μ l 또는 9×10^3 TU/5 μ l의 APOBEC-내성 RCR 벡터(T5.0002A)를 마우스에게 주사한다. 벡터를 투여하지않은 마우스, 9×10^5 TU/5 μ l 또는 9×10^3 TU/5 μ l의 벡터를 투여한 마우스를 5-FC 무작위로 투여하고(500 mg/kg/BID), 13일째부터는 IP 주사로 투여하거나, 지시된 바와 같이 일부 그룹에게는 5-FC를 투여하지 않는다(PBS). 중간 용량에서 벡터를 수신하는 마우스는 모두 5-FC를 투여받는다(즉, 이 용량에 대해 별도의 대조군이 없음). 5-FC 투여를 매일 연속 7일동안 계속하고, 이어서 10일간 치료받지 않는다. 약물과 휴식의 주기는 최대 4주기까지 반복된다. 그룹 9를 제외한 각 그룹으로부터 10마리의 마우스를 생존 분석 카테고리에 무작위로 배정한다. 나머지 마우스는 예정된 스케줄에 따라 희생된다.
- [0223] 나이브 센티넬(sentinel) 마우스를 예정된 희생 동물과 공동 배양하고, 흘림(shedding)을 통한 벡터 전달을 평가하기 위해 동일한 시점에서 꺼냈다.

그룹 배정 및 투여량 수준

그룹	시험물품	용량	약물 TX	N	분석 카테고리 당 N		
					(A) 생존 분석	(B) 스케줄된 희생	(C) 센티넬
1	형성 버퍼	5 µl	PBS	4		제1 약물 주기전에 4마리	
2	형성 버퍼	5 µl	5FC	10	10		
3	T5.0002A	9E5/5µl	PBS	10	10		
4	T5.0002	9E5/5µl	5FC	10	10	각 주기 개시전에 3마리, 총 15마리	각 주기 개시전에 1마리, 총 5마리
5	T5.0002A	9E5/5µl	5FC	25	10	각 주기 개시전에 3마리, 총 15마리	각 주기 개시전에 1마리, 총 5마리
6	T5.0002A	9E4/5µl	5FC	10	10		
7	T5.0002A	9E3/5µl	5FC	25	10	각 주기 개시전에 3마리, 총 15마리	각 주기 개시전에 1마리, 총 5마리
8	T5.0002A	9E3/5µl	PBS	10	10		
9	중양 없음	없음	5FC	15		각 주기 개시전에 3마리, 총 15마리	
총 동물 수				119	70	64	15

[0224]

[0225]

정맥내 투여는 꼬리 정맥 내로의 주입을 통해 수행된다. 복강내 투여는 방광을 피하기 위해 주의 깊게 복부에 주입하여 시행된다. 두개내 투여를 위해, 3.2mm 돌출부가 오른쪽 선조체에 이식되고, 3.7mm 돌출부가 있는 캡이 장착된 있는 가이드 캐놀라가 있는 마우스를 사용한다. 투시된 정위 좌표는 AP=0.5-1.0 mm, ML=1.8-2.0 mm, DV=3.2 mm(브레그마(bregma)로부터)이다. 세포 또는 백터를 가이드 캐놀라를 통해 삽입된 3.7mm 돌출부가 있는 주입 캐놀라를 통해 두개내 주입된다. 속도는 해밀턴 주사기 및 유연한 튜브가 장착된 주사기 펌프로 제어된다.

[0226]

세포 주입의 경우, 1 마이크로리터의 세포가 분당 0.2 마이크로리터의 유속(총 5분)으로 전달된다. 백터 주입의 경우, 5 마이크로리터의 백터가 분당 0.33 마이크로리터의 유속(총 15분)으로 전달된다.

[0227]

백터는 마우스에게 뇌 중량 1g 당 형질전환 단위(TU)로서 전달되고 계산된다. 상기 계산을 사용하여, 사람을 포함한 다른 포유동물에 대해 용량의 변환을 계산할 수 있다. 이 연구의 결과는 APOBEC-내성 바이러스가 중양 전체에 퍼지고, yCD 무결성을 유지하며, APOBEC-민감성 RRV와 비교했을 때 5FC와 조합하여 중양 치료시에 더 효과적이라는 것을 보여줄 것이다. APOBEC-내성 RRV는 또한 나이브 케이지 메이트(cage mate)에게 수평적으로 퍼지지 않는다.

[0228]

상기 기술된 바와 같이, RRV는 "2A 카세트"를 함유한다. 예를 들어, 서열식별번호: AA는 2A 카세트를 포함하는 일반적인 구조를 제공한다. 카세트는 다수의 상이한 카세트로 교체할 수 있다. 예를 들어, 하기 카세트를 준비

하고 서열식별번호: AA의 카세트를 대체하는 서열식별번호: AA 벡터 백본으로 클로닝할 수 있다.

- [0229] 본원에서 제공된 방법 및 서열을 사용하여 다수의 벡터를 다음과 같이 설계하였다:
- [0230] pAC3-T2A-GFPm(서열식별번호:43)
- [0231] pAC3-GSG-T2A-GFPm(서열식별번호:44)
- [0232] pAC3-P2A-GFPm(서열식별번호:45)
- [0233] pAC3-GSG-P2A-GFPm(서열식별번호:46)
- [0234] pAC3-E2A-GFP(서열식별번호:47)
- [0235] pAC3-GSG-E2A-GFPm(서열식별번호:48)
- [0236] pAC3-F2A-GFPm(서열식별번호:49)
- [0237] pAC3-GSG-F2A-GFPm(서열식별번호:50)
- [0238] pAC3-T2A-yCD2(서열식별번호:51)
- [0239] pAC3-GSG-T2A-yCD2(서열식별번호:52)
- [0240] pAC3-P2A-yCD2(서열식별번호:53)
- [0241] pAC3-GSG-P2A-yCD2(서열식별번호:54)
- [0242] 본 발명의 많은 구현예가 설명되었다. 그럼에도 불구하고, 본 발명의 사상 및 범주를 벗어나지 않으면서 다양한 변형이 이루어질 수 있음을 이해할 것이다. 따라서, 다른 구현예들은 하기 청구범위의 범주 내에 있다.

도면

도면1

마비염 A 바이러스 (ERAV) 구제역 (FMDV) 돼지 테스코바이러스-1 (PTV1) 토세아아시그나 바이러스 (TaV)	E2A ---- QCTNYALLKLAGD VESNPG P- F2A - PVKQLLNFDLLKLAGD VESNPG P- P2A ----- ATNFsLLKQAGD VEENPG P- T2A ----- EGRGSLTCGD VEENPG P-
--	--

도면2

피코르나바이러스

EMC-B	GIFNAHYAGYFADLLIHNSIETNFG	(서열 번호 : 59)
EMC-D	GYFADLLIHNSIETNFG	(서열 번호 : 60)
EMC-PV21	RIFNAHYAGYFADLLIHNSIETNFG	(서열 번호 : 61)
MENGO	HVFETHYAGYFSDLLIHNSVETNFG	(서열 번호 : 62)
TME-GD7	KAVRGYHADYYKQRLIHNSVEMNFG	(서열 번호 : 63)
TME-DA	RAVRAYHADYYKQRLIHNSVEMNFG	(서열 번호 : 64)
TME-BEAN	KAVRGYHADYYRQRLIHNSVETNFG	(서열 번호 : 65)
테일러-유사 바이러스	KHVREYHAAYYKQRLMHNSVETNFG	(서열 번호 : 66)
르준간 바이러스 (174F)	MHSDEMDFAGGKFLNQCGNSVETNFG	(서열 번호 : 67)
르준간 바이러스 (145SL)	MHNDEMDSGGKFLNQCGNSVETNFG	(서열 번호 : 68)
르준간 바이러스 (87-012)	MHSDEMDFAGGKFLNQCGNSVETNFG	(서열 번호 : 69)
르준간 바이러스 (M1146)	YHDKDMDYAGGKFLNQCGNSVETNFG	(서열 번호 : 70)
FMD-A10	APAKQ LLNFDLLKLAGNSVSNFG	(서열 번호 : 71)
FMD-A12	APGKQ LLNFDLLKLAGNSVSNFG	(서열 번호 : 72)
FMD-C1	APAKQ LLNFDLLKLAGNSVSNFG	(서열 번호 : 73)
FMD-O1G	APVKQ LLNFDLLKLAGNSVSNFG	(서열 번호 : 74)
FMD-O1K	APVKQ LTNFDLLKLAGNSVSNFG	(서열 번호 : 75)
FMD-O (Taiwan)	APAKQ LLNFDLLKLAGNSVSNFG	(서열 번호 : 76)
FMD-O/SK	APVKQ LLSFDLLKLAGNSVSNFG	(서열 번호 : 77)
FMD-SAT3	KPDKQ MCNFDLLKLAGNSVSNFG	(서열 번호 : 78)
FMD-SAT2	GVAKQ LLNFDLLKLAGNSVSNFG	(서열 번호 : 79)
ERAV	NINKQ CTNYSLLKLAGNSVSNFG	(서열 번호 : 80)
ERBV	TILSE GATNFSLLKLAGNSVELNFG	(서열 번호 : 81)
ERV-3	NLLSQ GATNFDLLKLAGNSVSNFG	(서열 번호 : 82)
PTV-1	VMAFQ GPGATNFSLLKQAGNSVENFG	(서열 번호 : 83)
PTV-2	TMMLQ GPGATNFSLLKQAGDVEENFG	(서열 번호 : 84)
PTV-3	TMSFQ GPGASSFSLKQAGNSVENFG	(서열 번호 : 85)
PTV-4	TMMLQ GPGASNSFSLKQAGNSVENFG	(서열 번호 : 86)
PTV-5	TMLFQ GPGASNSFSLKQAGNSVENFG	(서열 번호 : 87)
PTV-6	TMSFQ GPGATNFSLLKQAGNSVENFG	(서열 번호 : 88)
PTV-7	VVSFQ GPGATNFSLLKQAGNSVENFG	(서열 번호 : 89)
PTV-8	TMSLQ GPGATNFSLLKQAGNSVENFG	(서열 번호 : 90)
PTV-9	TMAFQ GPGATNFSLLKQAGNSVENFG	(서열 번호 : 91)
PTV-10	TLSEFQ GPGATNFSLLKQAGNSVENFG	(서열 번호 : 92)
PTV-11	RMSFQ GPGATNFSLLKQAGNSVENFG	(서열 번호 : 93)

곤충 바이러스

CrPV	FLRKRTQLLMSGNSVSNFG	(서열 번호 : 94)
DCV	EAARQMLLLLSGNSVETNFG	(서열 번호 : 95)
ABPV	GSWTDILLLLSGNSVETNFG	(서열 번호 : 96)
ABPV는 폴란드 1을 분리한다	GSWTDILLLLSGNSVETNFG	(서열 번호 : 97)
ABPV는 헝가리 1을 분리한다	GSWTDILLLWSGNSVETNFG	(서열 번호 : 98)
IFV	TRAEIEDELIRAGNSVSNFG	(서열 번호 : 99)
TaV	RAEGRGSLTTCGNSVENFG	(서열 번호 : 100)
EEV	QGAGRGSVLTTCGNSVENFG	(서열 번호 : 101)
APV	NYPMPALQKIIISLNSNFG	(서열 번호 : 102)
KBV	GTWESVNLNLAGNSVSNFG	(서열 번호 : 103)
PnPV (a)	AQGWVPDLTVDGNSVSNFG	(서열 번호 : 104)
PnPV (b)	IGGGQKDLTQDGNISNFG	(서열 번호 : 105)
검은띠물결자나방	(a) AQGWAPDLTQDGNISNFG	(서열 번호 : 106)
피코르나-유사 바이러스	(b) IGGGQRDLTQDGNISNFG	(서열 번호 : 107)
프로비덴스 바이러스	(a) VGDRGSLTTCGNSVSNFG	(서열 번호 : 108)
	(c) SGGRGSLTAGNSVSNFG	(서열 번호 : 109)
	(b) GDPIEDLTDGNSVSNFG	(서열 번호 : 110)

도면2a

C형 로타바이러스

소 로타바이러스	SKFQIDRILISGDVIELNFG (서열 번호 : 111)
돼지 로타바이러스	AKFQIDKILISGDVIELNFG (서열 번호 : 112)
인간 로타바이러스	SKFQIDKILISGDVIELNFG (서열 번호 : 113)

레오바이러스 (사이포바이러스 1)

봄바익스 모리	FRSNYDLLKLCGDVIESNFG (서열 번호 : 114)
라이만트리아 디스파르	FRSNYDLLKLCGDVIESNFG (서열 번호 : 115)
팬드롤리무스 편크타투스	FRSNYDLLKLCGDVIESNFG (서열 번호 : 116)

트라이판소마 spp. 반복서열

T. 브루세이 TSR1 (CAB95325.1)	SSIIIRTKMLVSGDVSEENFG (서열 번호 : 117)
(CAB95559.1)	SSIIIRTKMLLSGDVSEENFG (서열 번호 : 118)
T. 크루지 AP 엔도뉴클레아제	SSIIIRTKILLSGDVSEENFG (서열 번호 : 119)
	CDAQRQKLLLSGDIQSNFG (서열 번호 : 120)

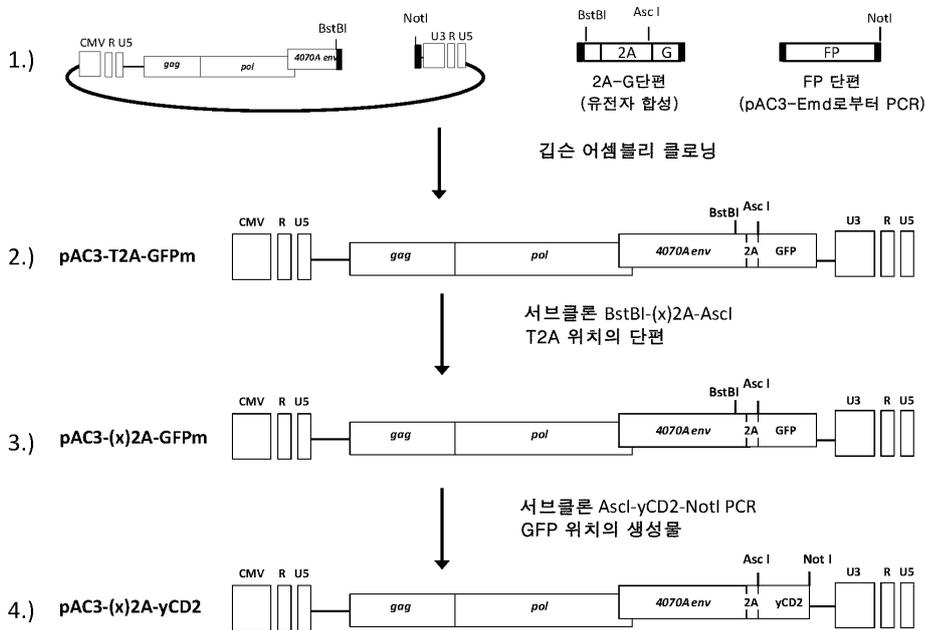
원핵 서열

T. 마리티마 aguA	YIPDFGGFLVKAQSEFNFG (서열 번호 : 121)
B. 브론치셀티카	VHCAGRGGPVRLLDKKGNGFG (서열 번호 : 122)

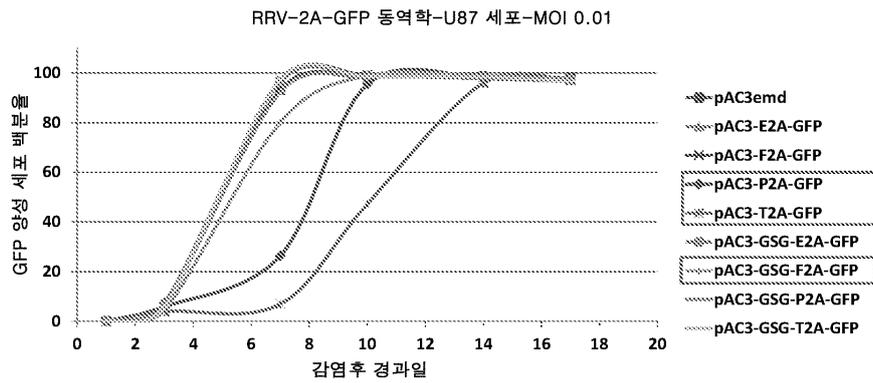
진핵 (세포) 서열

마우스 mor-1F	DLELETVGS HQAAANTNFG (서열 번호 : 123)
D. 멜라노가스테르 mod (mdg4)	TAADKIQGSWKMTTEGNFG (서열 번호 : 124)
A. 니돌란스 Ca 채널 MID1	PITNRPNSGLINTEINFG (서열 번호 : 125)

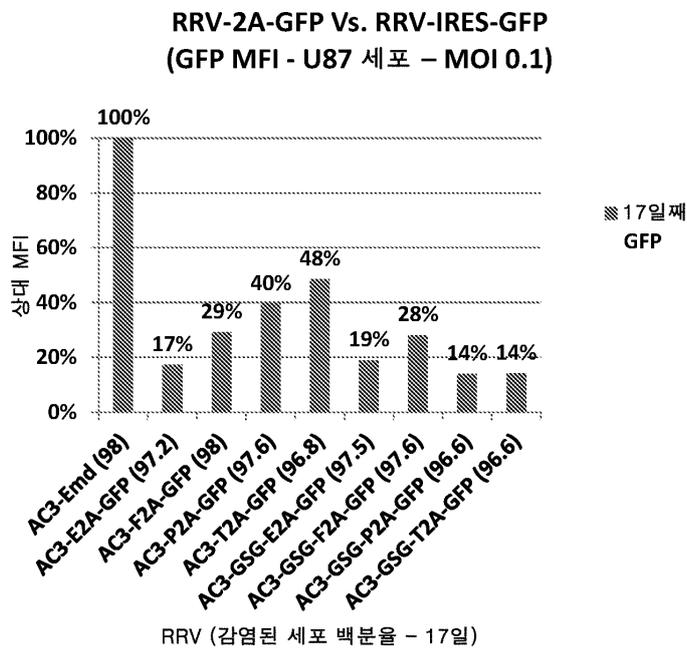
도면3



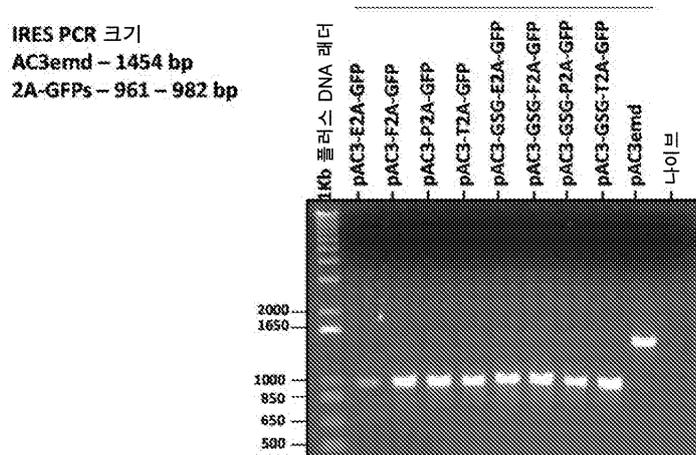
도면4



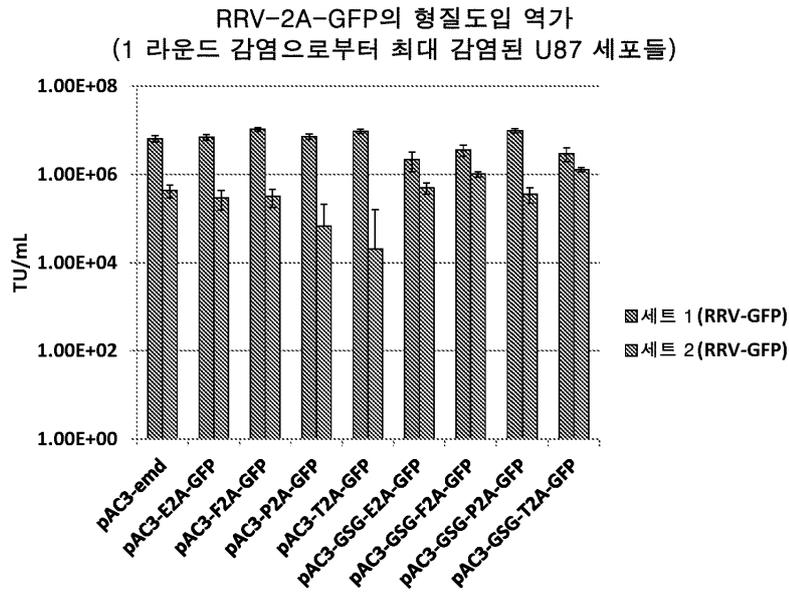
도면5



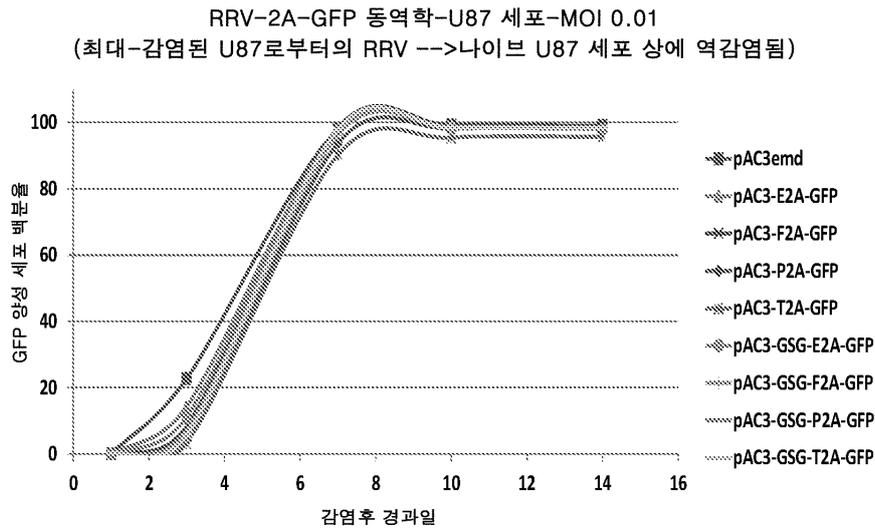
도면6



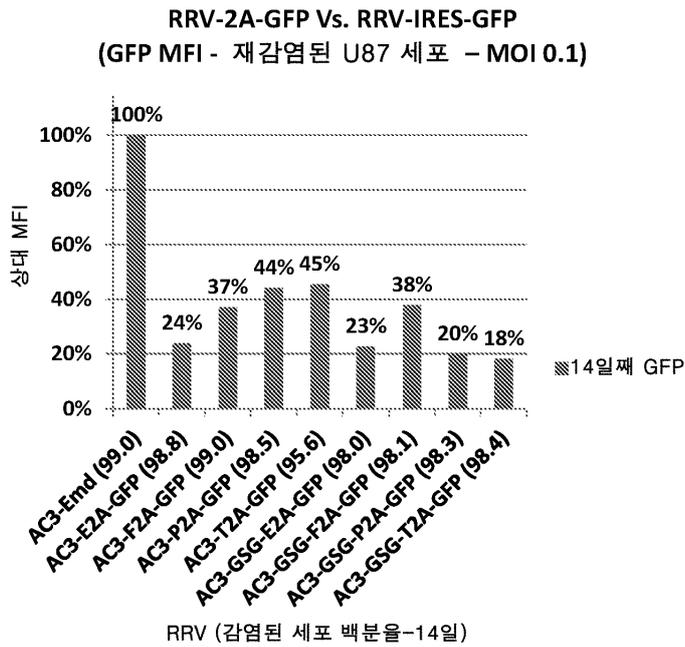
도면7



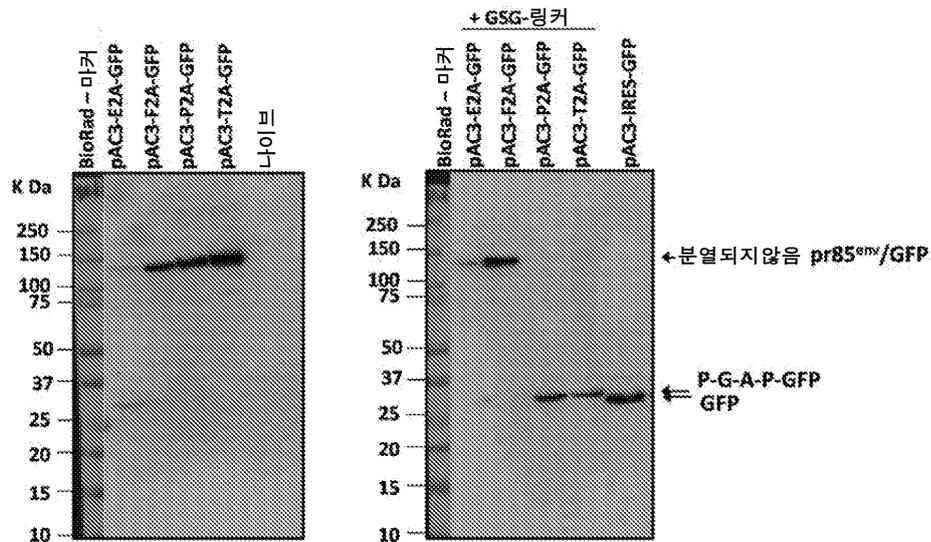
도면8



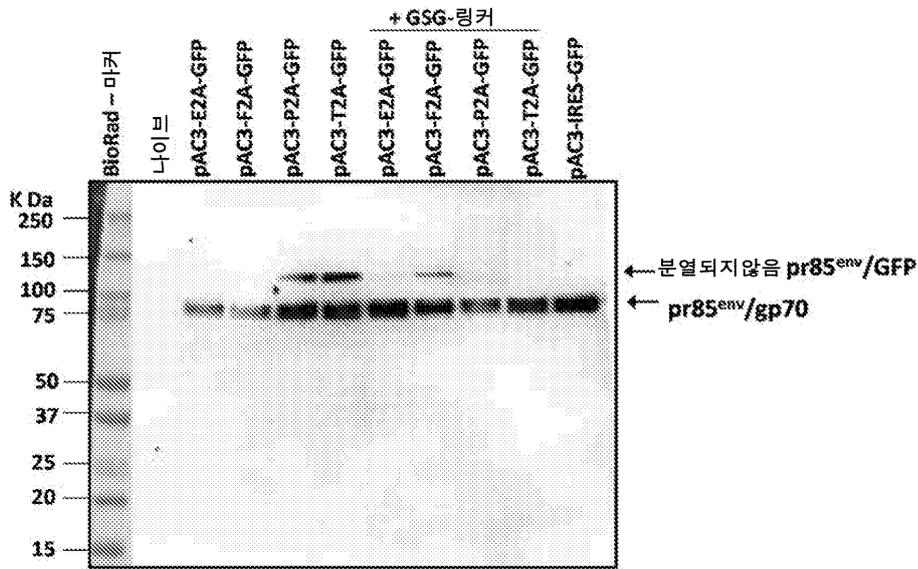
도면9



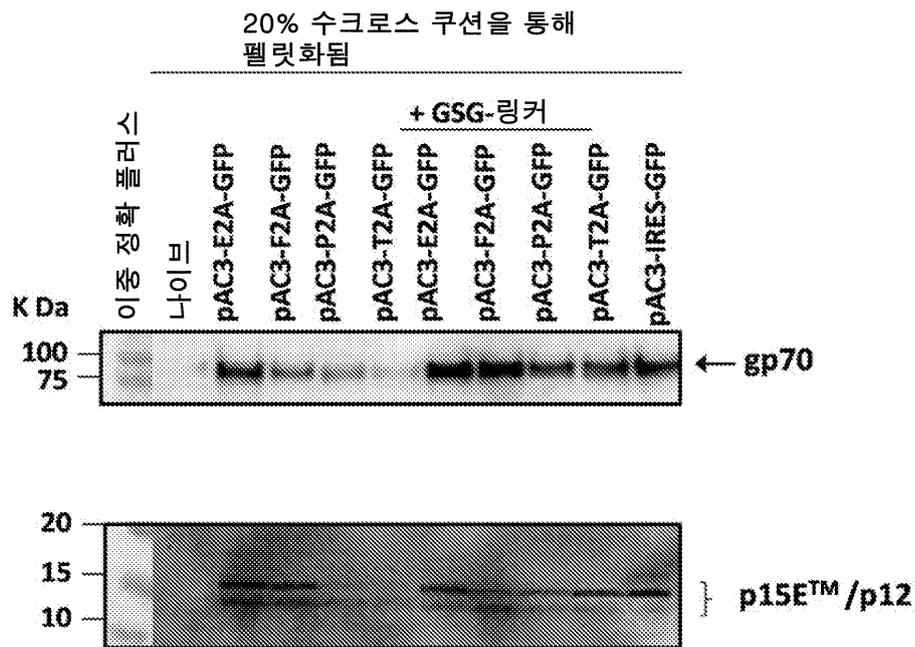
도면10



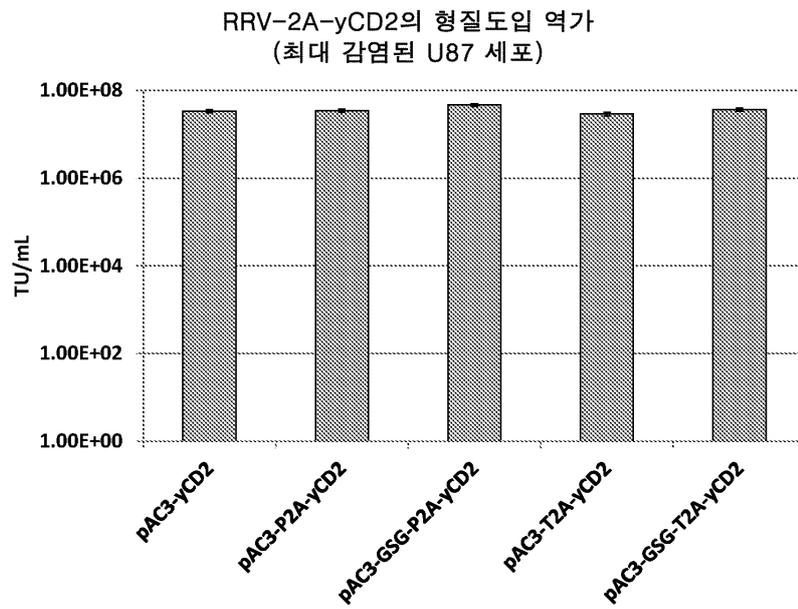
도면11



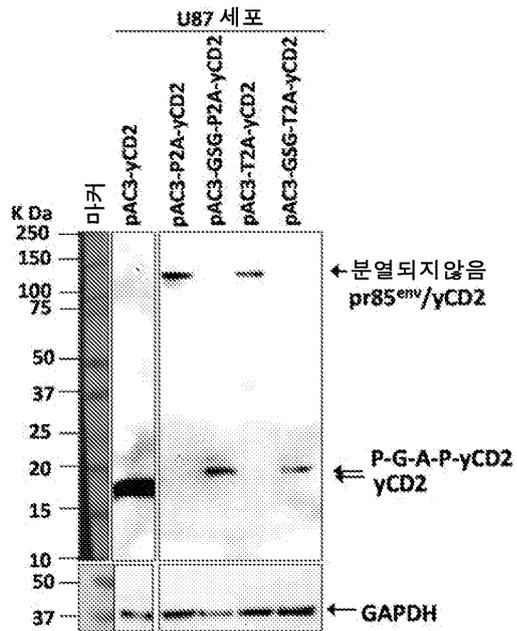
도면12



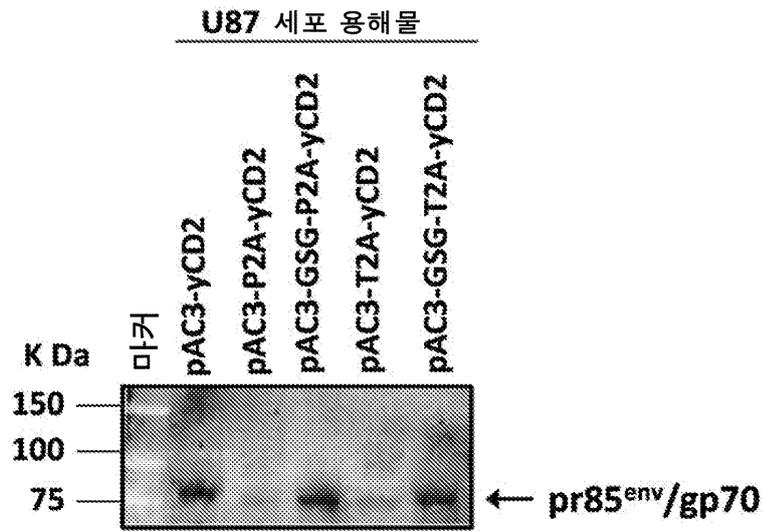
도면13



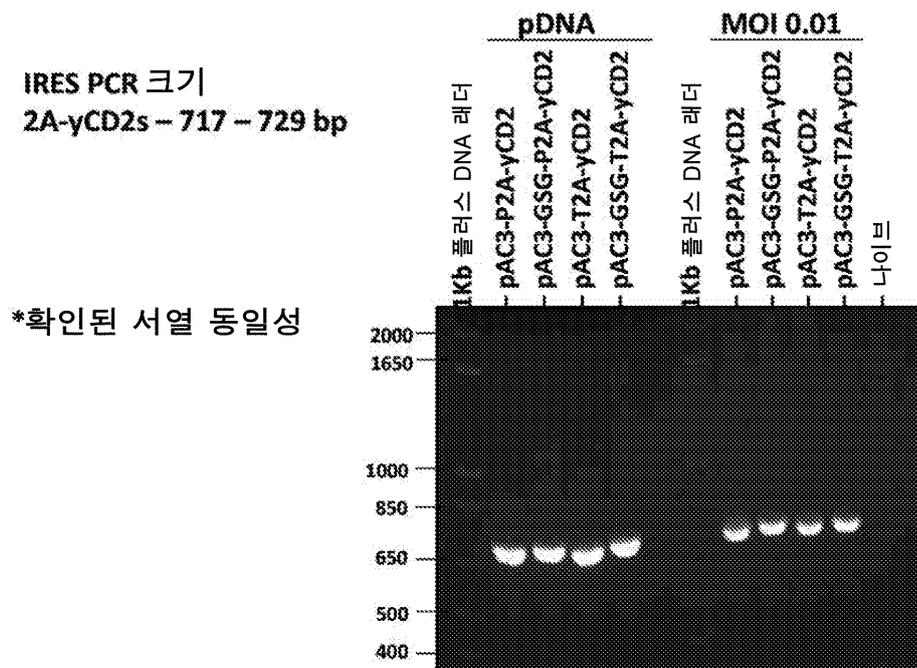
도면14



도면15

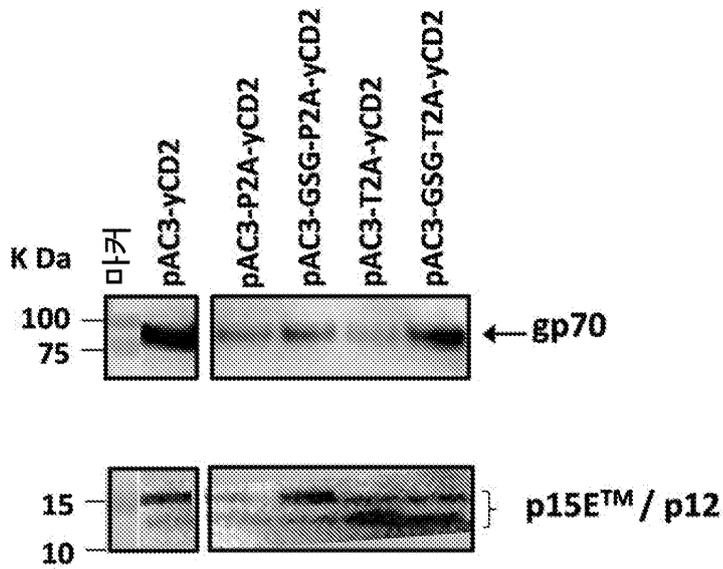


도면16

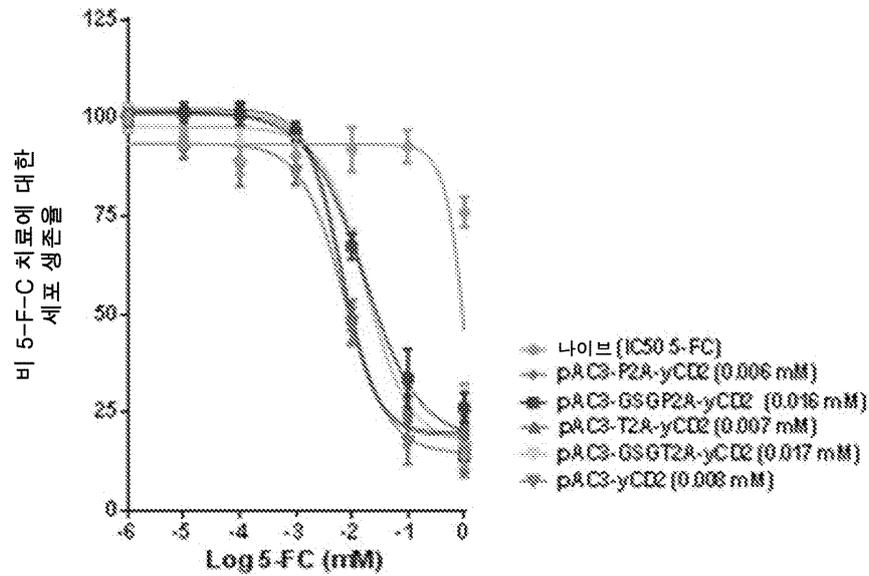


도면17

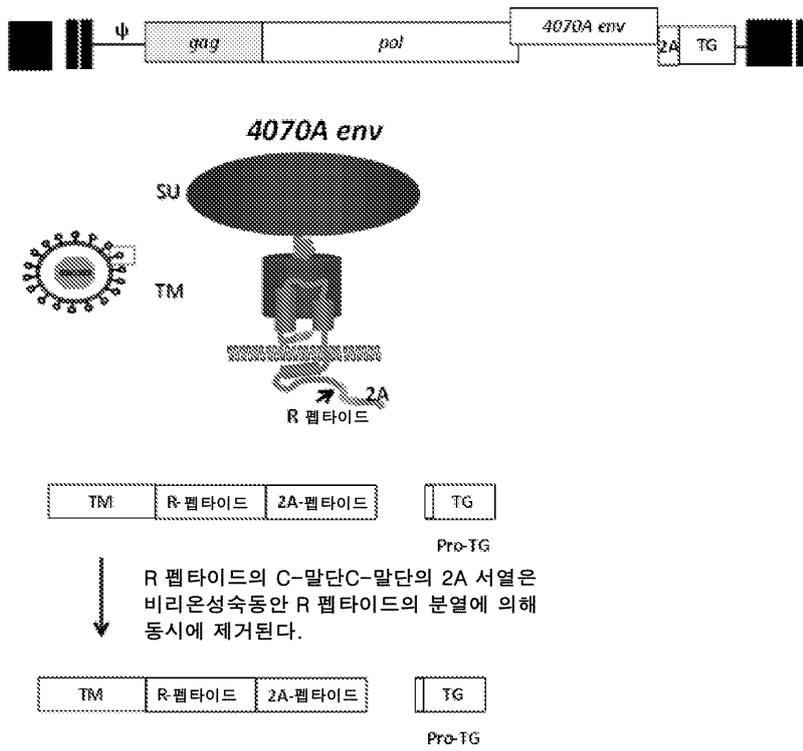
20% 수크로스 쿠션을 통해
펠릿화된 U87 sup(d13)



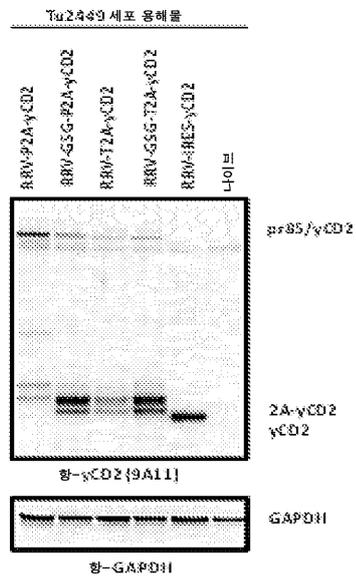
도면18



도면19

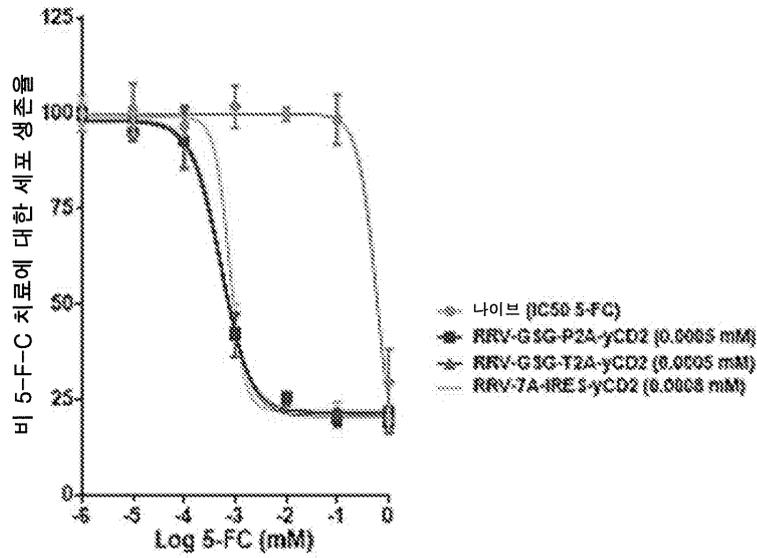


도면20a

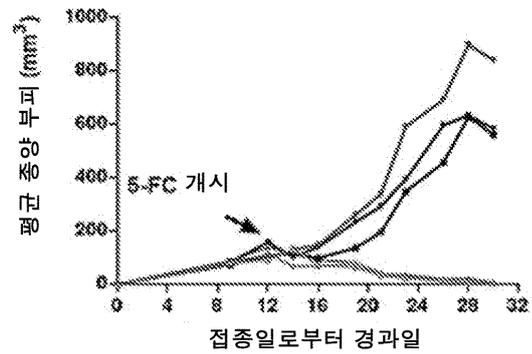


도면20b

Tu2449 세포 5-FC 민감도
(7일째)

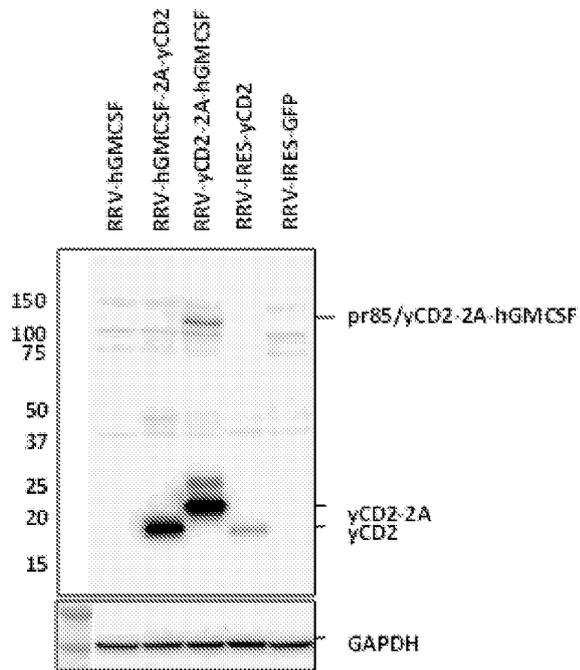


도면21

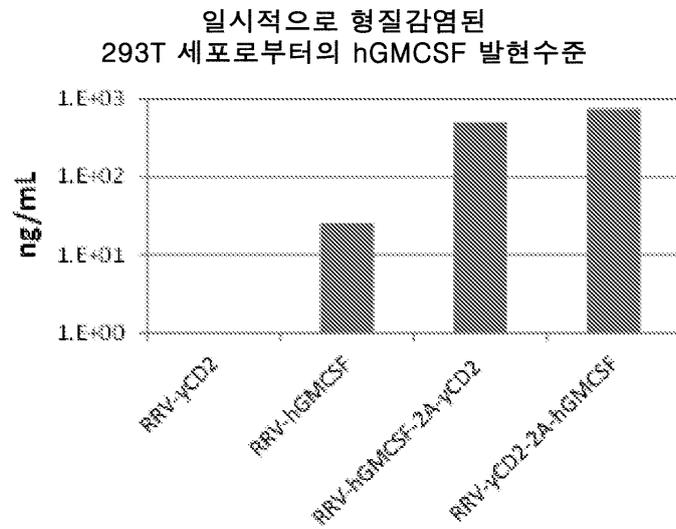


- RRV-yCD2+5-FC
- RRV-yCD2+ PBS
- RRV-RSV-yCD2+ 5-FC
- RRV-RSV-yCD2+ PBS
- RRV-GSG-T2A-yCD2+ 5-FC
- RRV-GSG-T2A-yCD2+ PBS

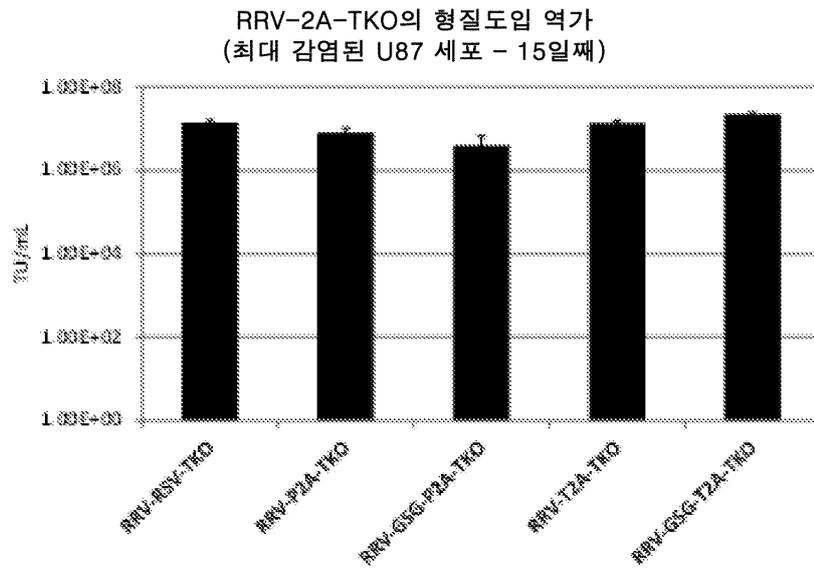
도면22a



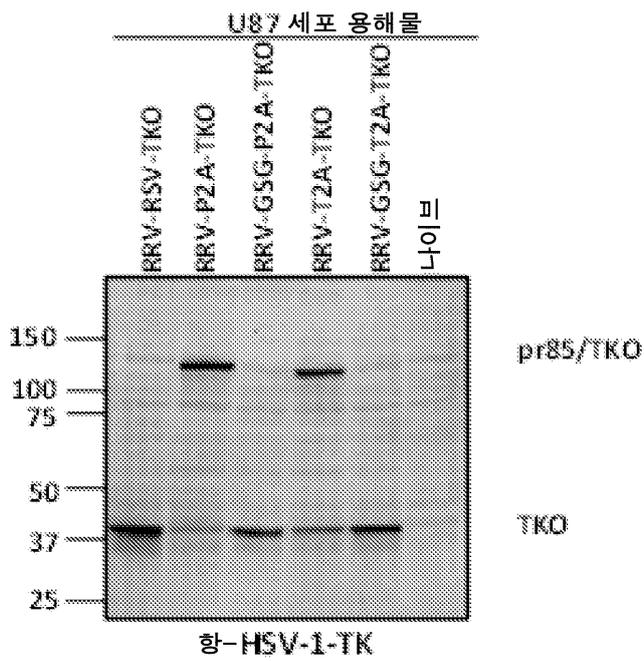
도면22b



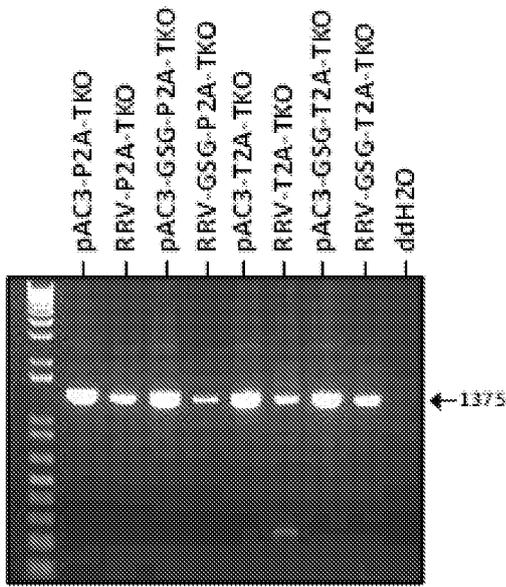
도면23



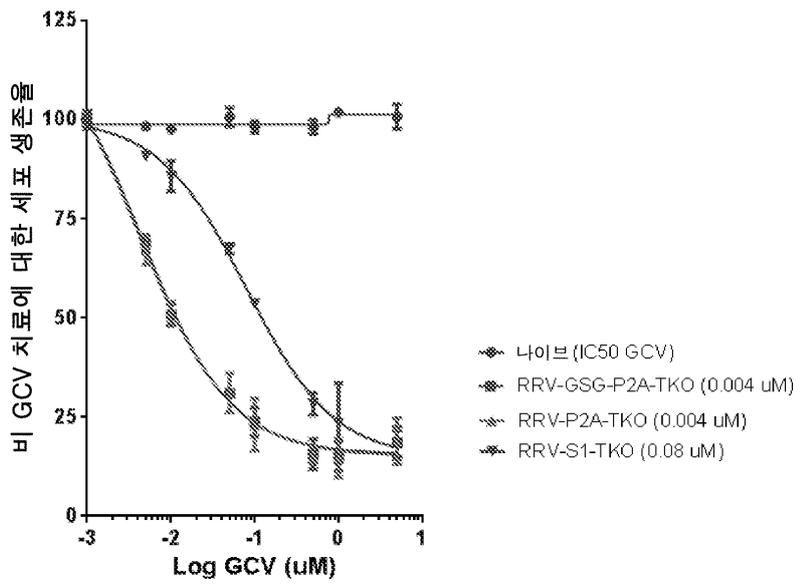
도면24



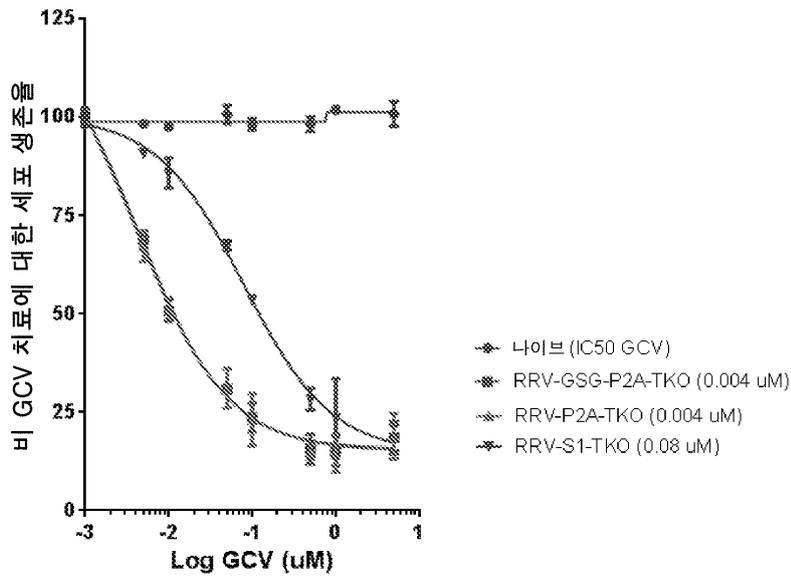
도면25



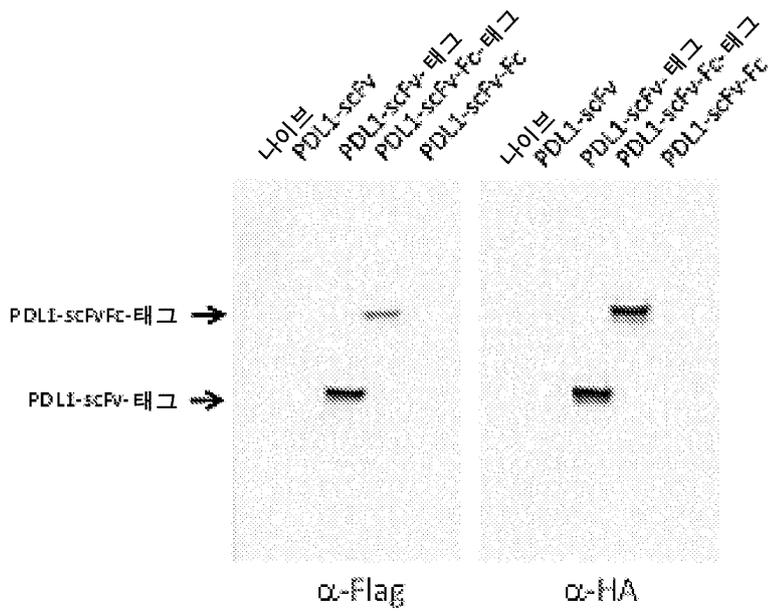
도면26a



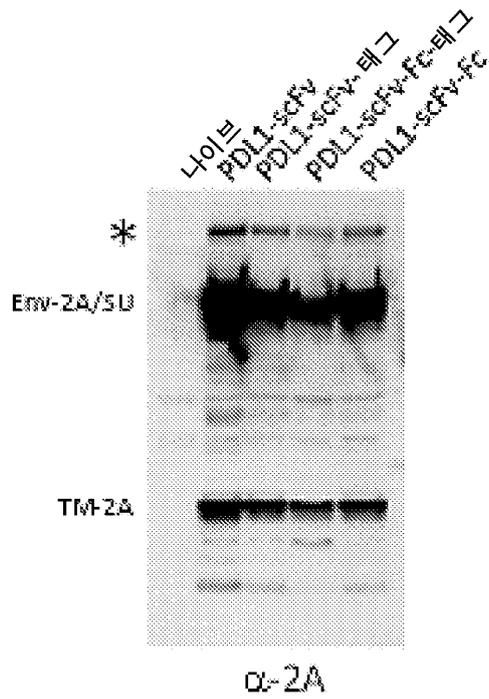
도면26b



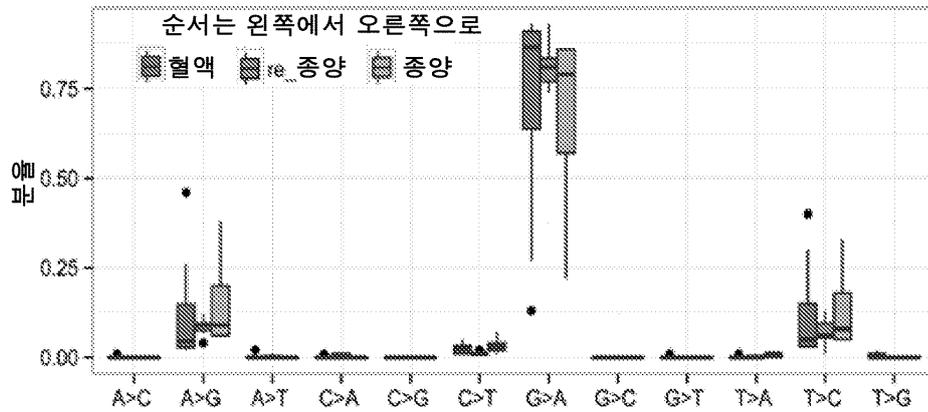
도면27a



도면27b



도면28



서열 목록

- <110> Tocagen Inc.
- <120> RECOMBINANT VECTORS COMPRISING 2A PEPTIDE
- <130> IPA180321-US
- <140> PCT/US2016/049947
- <141> 2016-09-01
- <150> US 62/214,884
- <151> 2015-09-04
- <160> 125
- <170> KoPatent In 3.0

<210> 1
 <211> 8
 <212> PRT
 <213> Artificial Sequence
 <220><223> 2A peptide consensus sequence
 <220><221> MISC_FEATURE
 <222> (2)
 <223> Xaa is V or I
 <220><221> MISC_FEATURE
 <222> (4)
 <223> Xaa is any amino acid
 <400> 1

Asp Xaa Glu Xaa Asn Pro Gly Pro

1 5
 <210> 2
 <211> 11654
 <212> DNA
 <213> Artificial Sequence
 <220><223> RRV vector containing 2A-cassette
 <400> 2

```

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg      60
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt      120
gacgtcaata atgacgatg ttccatagt aacccaata gggactttcc attgacgtca      180
atgggtggag tatttacggt aaactgcca cttggcagta catcaagtgt atcatatgcc      240

aagtacgcc cciattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta      300
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac      360
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg      420
atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg      480
ggactttcca aaatgtcgt acaactccgc cccattgacg caaatgggcg gtaggcgtgt      540
acggtgggag gtctataata gcagagctgg tttagtgaac cggcgccagt cctccgattg      600
actgagtcgc ccgggtacce gtgtatccaa taaacctct tgcagttgca tccgacttgt      660

ggctctcgtg ttcttggga gggctctctc tgagtattg actaccgctc agcggggggtc      720
    
```

tttcatttgg gggctcgtcc gggatcggga gacccctgcc cagggaccac cgaccacca 780
 ccgggaggtg agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840
 tgattttatg cgctgcgtc ggtactagt agctaactag ctctgtatct ggcggaccag 900
 tgggtggaact gacgagttcg gaacacccgg ccgcaaccct gggagacgtc ccagggactt 960
 cgggggccgt tttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg 1020
 gtgcaccccc cttagaggag ggatatgtgg tctctgtagg agacgagaac ctaaaacagt 1080

 tcccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgctcttg 1140
 tctgtctcag catcgttctg tgttctctct gtctgactgt gtttctgtat ttgtctgaga 1200
 atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg 1260
 agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct 1320
 ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380
 tcatcaccca ggtaagatc aaggcttttt cacctggccc gcatggacac ccagaccagg 1440
 tcccctacat cgtgacctgg gaagccttgg cttttgacct ccctccctgg gtcaagccct 1500

 ttgtacacce taagcctccg cctcctcttc ctccatecgc cccgtctctc ccccttgaac 1560
 ctctcgttc gaccccgct cgatcctccc tttatccagc cctcactcct tctctaggcg 1620
 ccaaacctaa acctcaagtt ctttctgaca gtggggggcc gctcatcgac ctacttacag 1680
 aagaccccc gccttatagg gaccecaagc caccccttc cgacagggac ggaaatggtg 1740
 gagaagcgac ccctgcggga gaggcaccgg acccctcccc aatggcatct cgctacgtg 1800
 ggagacggga gccccctgtg gccgactcca ctacctgca ggcatcccc ctccgcgcag 1860
 gaggaaacgg acagcttcaa tactggccgt tctcctcttc tgacctttac aactggaaaa 1920

 ataataacce ttcttttct gaagatccag gtaaactgac agctctgatc gagtctgttc 1980
 tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040
 gagaagaaaa acaacgggtg ctcttagagg ctgaaaaggc ggtgcggggc gatgatgggc 2100
 gccccactea actgccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg 2160
 attacaccac ccaggcaggf aggaaccacc tagtccacta tcgccagtgt ctcttagcgg 2220
 gtctcaaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280
 ggccaatga gtctccctcg gccttcttag agagacttaa ggaagcctat cgcaggtaca 2340

 ctcttatga ccctgaggac ccaggcaag aactaatgt gtctatgtct ttcatttggc 2400
 agtctgcccc agacattggg agaaagttag agaggttaga agatttaaaa aacaagacgc 2460
 ttggagattt ggtagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520
 gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580

agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640
 ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700
 atcgcgacca gtgtgcctac tgcaaagaaa aggggcactg ggctaagat tgtccaaga 2760

 aaccacgagg acctcgggga ccaagacccc agacctcct cctgacccta gatgactagg 2820
 gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac 2880
 ccgtcacctt cctgtagat actggggccc aacctccgt gctgaccaa aatcctggac 2940
 ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgctgga 3000
 ccacggatcg caaagtacat ctagctaccg gtaaggtcac ccactcttc ctccatgtac 3060
 cagactgtcc ctatcctcg ttaggaagag atttgctgac taaactaaaa gcccaaatcc 3120
 actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgtga 3180

 ccctaaatat agaagatgag catcggttac atgagacctc aaaagagcca gatgtttctc 3240
 taggggccac atggctgtct gattttctc aggcctgggc ggaaaccggg ggcattggac 3300
 tggcagttcg ccaagctcct ctgacatac ctctgaaagc aacctctacc cccgtgtcca 3360
 taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420
 tgttgacca ggaatactg gtacctgcc agtcccctg gaacacgccc ctgctaccg 3480
 ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc 3540
 ggggtggaaga catccacccc accgtgceca acccttaca cctcttgagc gggtccccac 3600

 cgtcccacca gtgttacct gtgcttgatt taaaggatgc cttttctgc ctgagactcc 3660
 accccaccag tcagcctctc ttgcctttg agtggagaga tccagagatg ggaatctcag 3720
 gacaattgac ctggaccaga ctcccacagg gtttcaaaa cagtcccacc ctgtttgatg 3780
 aggcactgca cagagacctc gcagacttcc ggatccagca cccagacttg atcctgtac 3840
 agtacgtgga tgacttactg ctggccgcca ctctgagct agactgcaa caaggtactc 3900
 gggccctgtt acaaaccta gggaacctcg ggtatcgggc ctcgccaag aaagccaaa 3960
 tttgccagaa acaggtcaag tatctgggt atcttctaaa agagggtcag agatggctga 4020

 ctgaggccag aaaagagact gtgatgggc agcctactcc gaagaccct cgacaactaa 4080
 gggagtctct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140
 cagccccctt gtacctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200
 aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag 4260
 atttgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgc aaaggtgtcc 4320
 taacgaaaa actgggacct tggcgtcggc cgttggecta cctgtccaaa aagctagacc 4380

cagtagcagc tgggtggccc ccttgccctac ggatggtagc agccattgcc gtactgacaa 4440

 aggatgcagg caagctaacc atgggacagc cactagtcac tctggccccc catgcagtag 4500
 aggcactagt caaacaaccc cccgaccgct ggttttccaa cgcccggatg actcactatc 4560
 aggccttgct ttiggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaaccggg 4620
 ctacgtctct cccactgctt gaggaagggc tgcaacacaa ctgccttgat atcctggccg 4680
 aagcccacgg aacccgaccc gacctaacgg accagccgct cccagacgcc gaccacacct 4740
 ggtacacgga tggaaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800
 ccaccgagac cgaggtaatc tgggctaaag ccctgccagc cgggacatcc gctcagcggg 4860

 ctgaactgat agcactcacc caggccctaa agatggcaga aggtagaag ctaaatgttt 4920
 atactgatag ccgttatgct ttgtctactg cccatatcca tggagaaata tacagaaggc 4980
 gtgggttgct cacatcagaa ggcaaaagaga tcaaaaataa agacgagatc ttggccctac 5040
 taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100
 gacacagcgc cgaggctaga ggcaaccgga tggtgacca agcggcccca aaggcagcca 5160
 tcacagagac tcagacaccc tctacccctc tcatagaaaa tteatcaccc tacacctcag 5220
 aacattttca ttacacagtg actgatataa aggacctaac caagtgggg gccatttatg 5280

 ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340
 ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc 5400
 tcctagagag aagccacagt ccctactaca tgtgaaccg ggatcgaaca ctcaaaaata 5460
 tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg 5520
 gaactagggc ccgcgggcat cggcccggca ctattggga gatcgattc accgagataa 5580
 agcccggatt gtatggctat aaatatcttc tagtttttat agataccttt tctggctgga 5640
 tagaacctt cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg 5700

 agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760
 tctccaaggt gagtcagaca gtggccgac tgttggggat tgattggaaa ttacattgtg 5820
 catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaacctc aaggagactt 5880
 taactaaatt aacgcttgca actggctcta gagactgggt gctcctact cccttagccc 5940
 tgtaccgagc ccgcaacacg ccgggcccc atggcctcac cccatagtag atcttatatg 6000
 gggcaccccc gcccttgta aacttccctg accctgacat gacaagagtt actaacagcc 6060
 cctctctcca agctcactta caggtctctt acttagtcca gcacgaagtc tggagacctc 6120

tggcggcagc ctaccaagaa caactggacc gaccggtggt acctcaccct taccgagtcg	6180
gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac	6240
cttacacagt cctgctgacc acccccaccg cctc aaaagt agacggcatc gcagcttggg	6300
tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat	6360
ggcgcgttca acgctctcaa aacccctca agataagatt aaccctgga agccctaat	6420
agtcatggga gtccctgttag gagtagggat ggacagagac ccccatcagg tctttaatgt	6480
aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct ccctcctggg	6540
aactgtacaa gatgccttcc caaaattata tttgatcta tgtgatctgg tggagagga	6600
gtgggacctt tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgacaggag	6660
acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcggggtg	6720
tgggggacca ggagagggt actgtggtaa atgggggtgt gaaaccaccg gacagcctta	6780
ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcgtaaca cccctggga	6840
cacgggatgc tctaaagttg cctgtggccc ctgctacgac ctctccaaag tatccaatc	6900
cttccaaggg gctactcgag ggggcagatg caacctcta gtctagaat tcaactgatc	6960
aggaaaaaag gctaaactggg acgggcccaa atcgtgggga ctgagactgt accggacagg	7020
aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccgagt	7080
ccccataggg cccaaccag tattaccega ccaagactc ccttctcac caatagagat	7140
tgtaccggtc ccacagccac ctagccccct caataccagt taccctctt ccaactaccg	7200
tacacctca acctcccta caagtccaag tgtcccacag ccacccccag gaaactggaga	7260
tagactacta gctctagta aaggagccta tcaggcgtt aacctacca atcccgaaa	7320
gaccaagaa tgttggctgt gcttagtgtc gggacctcct tattacgaag gagtagcgtt	7380
cgtgggcaact tataccaatc attccaccgc tccggccaac tgtacggcca cttccaaca	7440
taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac	7500
tcaccaggcc ttatgtaaca ccacccaaag cgcggctca ggatcctact acctgcagc	7560
accgccgga acaatgtggg cttgcagcac tggattgact cctgcttgt ccaccacgtt	7620
gctcaatcta accacagatt attgtatatt agttgaactc tggcccagag taatttacca	7680
ctccccgat tatatgtatg gtcagcttga acagcgtacc aatatataaa gagagccagt	7740
atcattgacc ctggcccttc tactaggagg attaaccatg ggagggatg cagctggaat	7800
agggacgggg accactgcct taattaaaac ccagcagttt gagcagcttc atgccctat	7860
ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc	7920
gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgctattcc taaaggaggg	7980

aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt 8040
 gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100
 aggccaagga tggttcgaag ggctgtttta tagatcccc tggtttacca ccttaatctc 8160
 caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct 8220

 caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac 8280
 tcagcaatat caccagctaa aacctataga gtacgagcca gtgaaacaga ctttgaattt 8340
 tgaccttctc aagtggcgg gagacgtgga gtccaacct ggacctggcg cgcctatggc 8400
 cagcaagggc gaggagctgt tcaccggggt ggtgccatc ctggtcgagc tggacggcga 8460
 cgtaaaccgc cacaagtca gcgtgtccgg cgaaggagag ggcatgcca cctacggcaa 8520
 gctgaccctg aagttcact gcaccaccgg caagctgcc gtgcctggc ccaccctct 8580
 gaccacctg acctacggcg tgcagtctt cggccctac cccgaccaca tgaagcagca 8640

 cgacttcttc aagtcgccca tgcccgaagg ctacgtccag gagcgacca tcttcttcaa 8700
 ggacgacggc aactacaaga cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa 8760
 ccgcatcgag ctgaaggga tcgaactcaa ggaggacggc aacatcctgg ggcacaagct 8820
 ggagtacaac tacaacagcc acaaggtcta taccaccgcc gacaagcaga agaacggcat 8880
 caaggtgaac ttcaagacc gccacaacat cgaggacggc agcgtgcagc tcgccacca 8940
 ctaccagcag aacaccccca tcggcgacgg ccccgctctg ctgcccgaca accactacct 9000
 gagcaccag tccgccctga gcaaagacc caacgagaag cgcgatcaca tggctctgct 9060

 ggagtctgtg accgcccgg ggatcactct cggcatggac gagctgtaca agtgtcggc 9120
 cgcagataaa ataaaagatt ttatttagtc tcagaaaaa ggggggaatg aaagaccca 9180
 cctgtaggtt tggcaagcta gcttaagtaa cgccatcttg caaggcatgg aaaaatacat 9240
 aactgagaat agagaagttc agatcaaggt caggaacaga tggaacagct gaatatgggc 9300
 caaacaggat atctgtggtg agcagttcct gccccggctc agggccaaga acagatggaa 9360
 cagctgaata tgggccaac aggatatctg tgtaagcag ttcttcccc ggctcagggc 9420
 caagaacaga tggccccag atgcggtcca gcctcagca gtttctagag aacctcaga 9480

 tgtttccagg gtgccccaa gacctgaaat gacctgtgc cttatttgaa ctaaccaatc 9540
 agttcgttc tcgttctgt tcgcgcgtt ctgctcccc agctcaataa aagagcccc 9600
 aaccctcac tcggggcgc agtctccga ttgactgagt cggccgggta cccgtgtatc 9660
 caataaacc tcttgagtt gcatccgact tgggtctcg ctgttcttg ggagggtctc 9720
 ctctgagtga ttgactacc gtcagcgggg gtctttcatt acatgtgagc aaaaggccag 9780
 caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc 9840

cctgacgagc atcacaaaaa tcgacgctca agtcagagggt ggcgaaaccc gacaggacta 9900

taaagatacc aggcgtttcc ccttgaagc tcctctgtgc gctctctgt tccgaccctg 9960

ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgt tctcaatgc 10020

tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac 10080

gaacccccgc ttcagccccg ccgctgcgcc ttatccgta actatcgtct tgagtccaac 10140

ccgtaagac acgacttacc gccactggca gcagccactg gtaacaggat tagcagagcg 10200

aggatgtag gcggtgctac agagtctctg aagtggggc ctaactacgg ctacactaga 10260

aggacagtat ttggtatctg cgtctctctg aagccagtta ccttcgaaa aagagtgggt 10320

agctcttgat ccggcaaaaa aaccaccgct ggtagcgggt gttttttgt ttgcaagcag 10380

cagattaccg gcagaaaaaa aggatctcaa gaagatcctt tgatctttc tacggggtct 10440

gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg 10500

atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat 10560

gagtaaacct ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgate 10620

tgtctatctt gttcatccat agttgcctga ctccccctcg ttagataac tacgatacgg 10680

gagggttac catctggccc cagtgtctga atgataccgc gagaccacg ctcaccggt 10740

ccagatttat cagcaataaa ccagccagcc ggaaggccg agcgcagaag tggctctgca 10800

actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg 10860

ccagtaata gtttgcgcaa cgtttgtgcc attgctgcag gcatcgtggt gtcacgctcg 10920

tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcagat tacatgatec 10980

cccatgttgt gcaaaaaagc ggttagctcc ttcggtctc cgatcgttgt cagaagtaag 11040

ttggcccgag tgttatact catggttatg gcagcactgc ataattctct tactgtcatg 11100

ccatccgtaa gatgcttttc tgtgactggt gactactcaa ccaagtcatt ctgagaatag 11160

tgtatcgggc gaccgagttg ctcttgcccg gcgtcaacac gggataatac cgcgccacat 11220

agcagaactt taaaagtgt catcattgga aaacgttctt cggggcgaaa actctcaagg 11280

atcttaccgc tgttgagatc cagttcgatg taaccactc gtgcacccaa ctgatcttca 11340

gcatctttta cttcaccag cgtttctggg tgagcaaaaa caggaaggca aatgccgca 11400

aaaaagggaa taagggcgac acggaatgt tgaatactca tactcttct tttcaatat 11460

tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag 11520

aaaaataaac aaatagggt tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa 11580

gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag gccccttcgt 11640
 cttcaagaat tcat 11654
 <210> 3
 <211> 9
 <212> DNA
 <213> Artificial Sequence
 <220><223> GSG linker sequence
 <400> 3
 ggaagcggga 9
 <210> 4
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> GFP-R-Gib primer
 <400> 4
 taaaatcttt tattttatct gcggccgcac 30

 <210> 5
 <211> 11
 <212> PRT
 <213> Artificial Sequence
 <220><223> Peptide readthrough sequence
 <400> 5
 Cys Ala Ala Ala Asp Lys Ile Lys Asp Phe Ile
 1 5 10
 <210> 6
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220><223> Ascl-yCD2 formard primer
 <400> 6
 gatcggcgcg cctatggtga ccggcggcat ggc 33
 <210> 7
 <211> 24
 <212> DNA

<213> Artificial Sequence
 <220><223> 3-37 primer

<400> 7
 cccctttttc tggagactaa ataa 24

<210> 8
 <211> 54
 <212> DNA
 <213> Artificial Sequence
 <220><223> 2A peptide sequence

<400> 8
 gagggcagag gaagtcttct aacatgcggt gacgtggagg agaatcccgg ccct 54

<210> 9
 <211> 63
 <212> DNA
 <213> Artificial Sequence
 <220><223> 2A peptide sequence

<400> 9
 ggaagcggag agggcagagg aagtcttcta acatgcggtg acgtggagga gaatcccggc 60
 cct 63

<210> 10
 <211> 57
 <212> DNA
 <213> Artificial Sequence
 <220><223> 2A peptide sequence

<400> 10
 gtactaact tcagcctgct gaagcaggct ggagacgtgg aggagaacct tggacct 57

<210> 11
 <211> 66
 <212> DNA
 <213> Artificial Sequence
 <220><223> 2A peptide sequence

<400> 11
 ggaagcggag ctactaactt cagcctgctg aagcaggctg gagacgtgga ggagaacct 60

ggacct 66

<210> 12

<211> 66

<212> DNA

<213> Artificial Sequence

<220><223> 2A peptide sequence

<400> 12

gtgaaacaga ctttgaattt tgaccttctc aagttggcgg gagacgtgga gtccaacct 60

ggacct 66

<210> 13

<211> 75

<212> DNA

<213> Artificial Sequence

<220><223> 2A peptide sequence

<400> 13

ggaagcggag tgaacagac tttgaattt gaccttctca agttggcggg agacgtggag 60

tccaacctg gacct 75

<210> 14

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> 2A peptide sequence

<400> 14

cagtgtacta attatgctct cttgaaattg gctggagatg ttgagagcaa cctggacct 60

60

<210> 15

<211> 69

<212> DNA

<213> Artificial Sequence

<220><223> 2A peptide sequence

<400> 15

ggaagcggac agtgtactaa ttatgctctc ttgaaattgg ctggagatgt tgagagcaac 60

cctggacct 69

<210> 16
 <211> 54
 <212> DNA
 <213> Artificial Sequence
 <220><223> 2A peptide sequence
 <400> 16
 gagggcagag gaagtcttct aacatgCGGT gacgtggagg agaatcccgg cct 54

<210> 17
 <211> 63
 <212> DNA
 <213> Artificial Sequence
 <220><223> 2A peptide sequence
 <400> 17
 ggaagcggag agggcagagg aagtcttcta acatgCGGTg acgtggagga gaatcccggc 60
 cct 63

<210> 18
 <211> 57
 <212> DNA
 <213> Artificial Sequence
 <220><223> 2A peptide sequence
 <400> 18
 gtactaact tcagcctgct gaagcaggct ggagacgtgg aggagaacc tggacct 57

<210> 19
 <211> 66
 <212> DNA
 <213> Artificial Sequence
 <220><223> 2A peptide sequence
 <400> 19
 ggaagcggag ctactaactt cagcctgctg aagcaggctg gagacgtgga ggagaaccct 60
 ggacct 66

<210> 20
 <211> 19

<212>	DNA	
<213>	Artificial Sequence	
<220><223>	5-MLV-U3-R primer	
<400>	20	
	agcccacaac ccctcactc	19
<210>	21	
<211>	18	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	3-MLV-Psi primer sequence	
<400>	21	
	tctcccgatc ccggacga	18
<210>	22	
<211>	26	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Probe sequence	
<400>	22	
	cccctaatga aagacccccg ctgacg	26
<210>	23	
<211>	23	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	IRES forward primer	
<400>	23	
	ctgatcttac tctttggacc ttg	23
<210>	24	
<211>	24	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	IRES reverse primer	
<400>	24	
	cccctttttc tggagactaa ataa	24

<210>	25	
<211>	22	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	ENV forward primer	
<400>	25	
accctcaacc tccctacaa gt		22
<210>	26	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	ENV reverse primer	
<400>	26	
gttaagcgcc tgataggctc		20
<210>	27	
<211>	26	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Env probe sequence	
<400>	27	
ccccaaatga aagacccccg ctgacg		26
<210>	28	
<211>		
>	477	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Human codon optimized heat stabilized CD coding sequence	
<400>	28	
atggtgaccg gcgcatggc ctccaagtgg gatcaaaagg gcatggatat cgcttacgag		60
gaggccctgc tgggtacaa ggaggcggc gtgcctatcg gcgctgtct gatcaacaac		120
aaggacggca gtgtgctggg caggggccac aacatgaggt tccagaaggg ctccgccacc		180
ctgcacggcg agatctccac cctggagaac tgtggcaggc tggagggcaa ggtgtacaag		240
gacaccacc tgiacaccac cctgtcccct tgtgacatgt gtaccggcgc tatcatcatg		300

tacggcatcc ctaggtgtgt gatcggcgag aacgtgaact tcaagtccaa gggcgagaag 360
 tacctgcaaa ccaggggcca cgaggtggtg gttgttgacg atgagaggtg taagaagctg 420
 atgaagcagt tcatcgacga gaggcctcag gactggttcg aggatcgcg cgagtaa 477

- <210> 29
- <211> 158
- <212> PRT
- <213> Artificial Sequence
- <220><223> Heat stabilized APOBEC modified CD polypeptide
- <220><221> MISC_FEATURE
- <222> (10)
- <223> Xaa is any amino acid
- <220><221> MISC_FEATURE
- <222> (152)
- <223> Xaa is any amino acid

<400> 29
 Met Val Thr Gly Gly Met Ala Ser Lys Xaa Asp Gln Lys Gly Met Asp
 1 5 10 15
 Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30
 Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45
 Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60
 Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80
 Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95
 Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
 100 105 110
 Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125
 Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe

130

135

140

Ile Asp Glu Arg Pro Gln Asp Xaa Phe Glu Asp Ile Gly Glu

145

150

155

<210> 30

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> 2A peptide coding sequence

<400> 30

gagggcagag gaagtcttct aacatgcggt gacgtggagg agaatcccgg ccct 54

<210> 31

<211> 1062

<212> DNA

<213> Artificial Sequence

<220><223> BstBI-env-T2A-GFPm of pAC3-T2A-GFPm

<400> 31

ttcgaagggc tgtttaatag atccccctgg ttaccacct taatctccac catcatggga 60

cctctaatag tactcttact gatcttactc tttggacctt gcattctcaa tcgattggtc 120

caatttgta aagacaggat ctcagtggtc caggctctgg ttttgactca gcaatatcac 180

cagctaaaac ccatagagta cgagccagag ggagaggaa gtcttctaac atgcggtgac 240

gtggaggaga atccccggcc tggcgcgcct atggccagca agggcgagga gctgttcacc 300

ggggtgggtc ccatcctggt cgagctggac ggcgacgtaa acggccacaa gttcagcgtg 360

tccggcgaag gagagggcga tgccacctac ggcaagctga ccctgaagtt catctgcacc 420

accggcaagc tgcccgtgcc ctggcccacc ctctgacca ccttgaccta cggcgtgcag 480

tgcttcgccc gctaccccga ccacatgaag cagcagcact tcttcaagtc cgccatgccc 540

gaaggctacg tcaggagcg caccatcttc ttcaaggacg acggcaacta caagaccgcg 600

gccgaggtga agttcgagg gcacacctg gtgaaccgca tcgagctgaa gggcatcgac 660

ttcaaggagg acggcaacat cctggggcac aagctggagt acaactaaa cagccacaag 720

gtctatatca ccgccgacaa gcagaagaac ggcatcaagg tgaacttcaa gaccgccac 780

aacatcgagg acggcagcgt cgagctcgcc gaccactacc agcagaacac ccccatcggc 840

gacggccccg tgctgctgcc cgacaaccac tacctgagca cccagtccgc cctgagcaaa 900

gacccaacg agaagcgca tcacatggc ctgctggagt tcgtgaccg cgcgggatc 960
actctcgca tggacgagct gtacaagtgt gcggccgag ataaaataa agattttatt 1020
tagtctccag aaaaagggg gaatgaaaga cccacctgt ag 1062

<210> 32
<211> 1026
<212> DNA
<213> Artificial Sequence
<220><223> BstBI-env-P2A-GFPm of pAC3-P2A-GFPm
<400> 32

ttcgaaggc tgttaatag atccccctg tttaccact taatctcac catcatggga 60
cctctaatag tactcttact gatcttactc tttggacctt gcattctcaa tcgattggtc 120

caatttgta aagacaggat ctcaaggctc caggctctgg tttgactca gcaatatcac 180
cagctaaaac ccatagagta cgagccagct actaactca gcctgctgaa gcaggctgga 240
gagctggagg agaaccctgg acctggcgcg cctatggcca gcaagggcga ggagctgttc 300
accgggtgg tgcccatcct ggtcgagctg gacggcgacg taaacggcca caagttcagc 360
gtgtccggc aaggagagg cgatgccacc tacggcaagc tgaccctgaa gttcatctgc 420
accaccgca agctgcccgt gccctggccc acctctgta ccacctgac ctaccgctg 480
cagtgtctc cccgctacc cgaccacatg aagcagcacg acttcttcaa gtccgcatg 540

cccgaaggc acgtccagga ggcaccatc ttcttcaagg acgacggcaa ctacaagacc 600
cgcgccgagg tgaagttcga gggcgacacc ctggtgaacc gcatcgagct gaaggcacc 660
gacttcaagg aggacggcaa catcctgggg cacaagctgg agtacaacta caacagccac 720
aaggtctata tcaccgcca caagcagaag aacggcatca aggtgaactt caagaccgc 780
cacaacatcg aggacggcag cgtgcagctc gccgaccact accagcagaa caccatc 840
ggcgacggc ccgtgctgct gcccgacaac cactacctga gcaccagtc cgcctgagc 900
aaagaccca acgagaagcg cgatcacatg gtctctgctg agttcgtgac cgcgccggg 960

atcactctc gcatggaca gctgtacaag tgtgcccgc cagataaaat aaaagatttt 1020
atttag 1026

<210> 33
<211> 1029
<212> DNA
<213> Artificial Sequence
<220><223> BstBI-env-E2A-GFPm of pAC3-E2A-GFPm

<400> 33

ttcgaagggc tgtttaatag atccccctgg tttaccacct taatctccac catcatggga 60
 cctctaatag tactcttact gatcttactc tttggacctt gcattctcaa tcgattggtc 120
 caatttgta aagacaggat ctcagtggtc caggctctgg ttttgactca gcaatatcac 180

cagctaaaac ccatagagta cgagccacag tgtactaatt atgctctctt gaaattggct 240
 ggagatgttg agagcaacce tggacctggc ggcctatgg ccagcaaggg cgaggagctg 300
 ttcaccgggg tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc 360
 agcgtgtccg gcgaaggaga gggcgatgcc acctacggca agctgacct gaagttcatc 420
 tgcaccaccg gcaagctgcc cgtgcctgg cccacctcg tgaccacctt gacctacggc 480
 gtgcagtgtc tcgcccgtc ccccgaccac atgaagcagc acgacttctt caagtccgcc 540
 atgccgaag gctactcca ggagcgcacc atcttcttca aggacgacgg caactacaag 600

acccgcgcc aggtgaagtt cgagggcgac acctggtga accgcatcga gctgaagggc 660
 atcgacttca aggaggacgg caacatcctg gggcacaagc tggagtaca ctacaacagc 720
 cacaaggtct atatcacgc cgacaagcag aagaacggca tcaagtgaa cttcaagacc 780
 cgccacaaca tcgaggacgg cagcgtgcag ctgcccggacc actaccagca gaacaccccc 840
 atcggcgacg gccccgtgt gctgcccgac aacctacc tgagcaccca gtccgcctg 900
 agcaaagacc ccaacgagaa gcgcgateac atggtctctg tggagtctgt gaccgcgcc 960
 gggatcactc tcggcatgga cgagctgtac aagtgtcgg cgcagataa aataaaagat 1020

tttatttag 1029

<210> 34

<211> 1035

<212> DNA

<213> Artificial Sequence

<220><223> BstBI-env-F2A-GFPm of pAC3-F2A-GFPm

<400> 34

ttcgaagggc tgtttaatag atccccctgg tttaccacct taatctccac catcatggga 60
 cctctaatag tactcttact gatcttactc tttggacctt gcattctcaa tcgattggtc 120
 caatttgta aagacaggat ctcagtggtc caggctctgg ttttgactca gcaatatcac 180
 cagctaaaac ccatagagta cgagccagtg aaacagactt tgaatttga ctttctcaag 240

ttggcgggag acgtggagtc caacctgga cctggcgcgc ctatggccag caaggcgag 300
 gagctgttca ccgggtgtt gccatcctg gtcgagctgg acggcgacgt aaacggccac 360

aagttcagcg tgcctggcga aggagagggc gatgccacct acggcaagct gaccctgaag 420
 ttcatctgca ccaccggcaa gctgcccgtg ccttggccca cctctgtgac caccttgacc 480
 tacggcgtgc agtgcttcgc ccgtacccc gaccacatga agcagcacga cttcttcaag 540
 tccgccatgc ccgaaggcta cgtccaggag cgcaccatct tcttcaagga cgacggcaac 600
 tacaagacc gcgccgaggt gaagttcgag ggcgacacc tggtaaccg catcgagctg 660

aaggcatcg acttcaagga ggacggcaac atcctggggc acaagctgga gtacaactac 720
 aacagccaca aggtctatat caccgccgac aagcagaaga acggcatcaa ggtgaacttc 780
 aagaccgcc acaacatcga ggacggcagc gtgcagctcg ccgaccacta ccagcagaac 840
 acccccatcg gcgacggccc cgtgctgctg cccgacaacc actacctgag cacccagtec 900
 gccctgagca aagaccccaa cgagaagcgc gatcacatgg tcctgctgga gttcgtgacc 960
 gccgccggga tcaactctcg catggacgag ctgtacaagt gtgcggccgc agataaata 1020
 aaagatttta tttag 1035

<210> 35

<211> 1032

<212> DNA

<213> Artificial Sequence

<220><223> BstBI-env-GSG-T2A-GFPm of pAC3-GSG-T2A-GFPm

<400> 35

ttcgaagggc tgtttaatag atccccctgg tttaccacct taatctccac catcatggga 60
 cctctaatag tactcttact gatcttactic ttggacctt gcattctcaa tcgattggtc 120
 caatttgta aagacaggat ctcaagtggc caggctctgg tttgactca gcaatatcac 180
 cagctaaaac ccatagagta cgagccagga agcggagagg gcagaggaag tcttctaaca 240
 tgcggtgacg tggaggagaa tcccggcctt ggcgcgccta tggccagcaa gggcgaggag 300

ctgttcaccg gggtggtgcc catcctggtc gagctggacg gcgacgtaa cggccacaag 360
 ttcagcgtgt ccggcgaagg agagggcgat gccacctacg gcaagctgac cctgaagttc 420
 atctgacca ccggcaagct gcccgtgcc tggcccacc tcgtgaccac cttgacctac 480
 ggcgtgcagt gcttcgccc ctaccgccac cacatgaagc agcacgactt cttcaagtec 540
 gccatgcccg aaggctactt ccaggagcgc accatcttct tcaaggacga cggcaactac 600
 aagaccgcg ccgaggtgaa gttcgagggc gacaccctgg tgaaccgat cgagctgaag 660
 ggcatcgact tcaaggagga cggcaacatc ctggggcaca agctggagta caactacaac 720

agccacaagg tctatcac cgccgacaag cagaagaacg gcatcaaggt gaacttcaag 780

acccgccaca acatcgagga cggcagcgtg cagctcgcg accactacca gcagaacacc 840
 cccatcggcg acggccccgt gctgctgccc gacaaccact acctgagcac ccagtccgcc 900
 ctgagcaaag accccaacga gaagcgcgat cacatggtcc tgctggagtt cgtgaccgcc 960
 gccgggatca ctctcgcat ggacgagctg tacaagtggt cggccgcaga taaaataaaa 1020
 gattttattt ag 1032
 <210> 36
 <211> 1035
 <212> DNA

<213> Artificial Sequence

<220><223> BstBI-env-GSG-P2A-GFPm of pAC3-GSG-P2A-GFPm

<400> 36
 ttcgaagggc tgtttaatag atccccctgg tttaccacct taatctccac catcatggga 60
 cctctaatag tactcttact gatcttactc tttggacctt gcattctcaa tcgattggtc 120
 caatttgta aagacaggat ctcaagtggc caggctctgg ttttgactca gcaatatcac 180
 cagctaaaac ccatagagta cgagccagga agcggagcta ctaacttcag cctgctgaag 240
 caggctggag acgtggagga gaacctgga cctggcgcgc ctatggccag caagggcgag 300
 gagctgttca ccgggggtgt gcccatctg gtcgagctgg acggcgacgt aaacggccac 360

 aagttcagcg tgtccggcga aggagagggc gatgccacct acggcaagct gacctgaag 420
 ttcatctgca ccaccggcaa gctgccctgt ccctggccca ccctcgtgac caccttgacc 480
 tacggcgtgc agtgcttcgc ccgctacccc gaccacatga agcagcacga cttcttcaag 540
 tccgcatgc ccgaaggcta cgtccaggag cgcaccatct tcttcaagga cgacggcaac 600
 tacaagacc gcgccagggt gaagttcgag ggcgacaccc tggtgaaccg catcgagctg 660
 aagggcatcg acttcaagga ggacggcaac atcctggggc acaagctgga gtacaactac 720
 aacagccaca aggtctatat caccgccgac aagcagaaga acggcatcaa ggtgaacttc 780

 aagaccgcc acaacatcga ggacggcagc gtgcagctcg ccgaccacta ccagcagaac 840
 acccccatcg gcgacggccc cgtgctgctg cccgacaacc actacctgag caccagttcc 900
 gccctgagca aagacccaa cgagaagcgc gatcacatgg tctgctgga gttcgtgacc 960
 gccgccggga tcaactcgg catggacgag ctgtacaagt gtgcggccgc agataaata 1020
 aaagatttta tttag 1035
 <210> 37
 <211> 1044
 <212> DNA

<213> Artificial Sequence

<220><223> BstBI-env-GSG-F2A-GFPm of pAC3-GSG-F2A-GFPm

<400> 37

ttcgaagggc tgtttaatag atccccctgg ttaccacct taatctccac catcatggga 60
 cctctaatag tactcttact gatcttactc ttggacctt gcattctcaa tcgattggtc 120
 caatttgta aagacaggat ctcaagtggc caggctctgg ttttgactca gcaatatcac 180
 cagctaaaac ccatagagta cgagccagga agcggagtga aacagacttt gaattttgac 240
 cttctcaagt tggcgggaga cgtggagtcc aacctggac ctggcgcgcc tatggccagc 300
 aaggcgagg agctgttcac cggggtggg cccatctgg tcgagctgga cggcgacgta 360
 aacggccaca agttcagcgt gtccggcga ggagaggcg atgccaccta cggcaagctg 420

acctgaagt teatctgac caccggcaag ctccccgtgc cctggccac cctcgtgacc 480
 accttgacct acggcgtgca gtgcttcgcc cgctaccccg accacatgaa gcagcacgac 540
 ttcttcaagt ccgcatgcc cgaaggctac gtccaggagc gcaccatctt cttcaaggac 600
 gagggcaact acaagaccgg cgccgaggtg aagtctgagg gcgacacct ggtgaaccgc 660
 atcgagctga agggcatcga cttcaaggag gacggcaaca tcctggggca caagctggag 720
 taaactaca acagccaaa ggtctatc accgccgaca agcagaagaa cggcatcaag 780
 gtgaacttca agaccggcca caacatcgag gacggcagcg tgcagctcgc cgaccactac 840

cagcagaaca ccccatcgg cgacggccc gtgctgctgc ccgacaacca ctacctgagc 900
 acccagtccg cctgagcaa agaccccaac gagaagcgg atcacatggt cctgctggag 960
 ttcgtgaccg ccgcccggat cactctcggc atggacgagc tgtacaagtg tgcggccgca 1020
 gataaataa aagattttat ttag 1044

<210> 38

<211> 1038

<212> DNA

<213> Artificial Sequence

<220><223> BstBI-env-GSG-E2A-GFPm of pAC3-GSG-E2A-GFPm

<400> 38

ttcgaagggc tgtttaatag atccccctgg ttaccacct taatctccac catcatggga 60
 cctctaatag tactcttact gatcttactc ttggacctt gcattctcaa tcgattggtc 120
 caatttgta aagacaggat ctcaagtggc caggctctgg ttttgactca gcaatatcac 180
 cagctaaaac ccatagagta cgagccagga agcggacagt gtactaatta tgctctcttg 240

aaattggctg gagatgttga gagcaaccct ggacctggcg cgcctatggc cagcaagggc 300
 gaggagctgt tcaccggggt ggtgcccatc ctggtcgagc tggacggcga cgtaacggc 360
 cacaagtcca gcgtgtccgg cgaaggagag ggcgatgcca cctacggcaa gctgaccctg 420
 aagttcatct gcaccaccgg caagctgccc gtgccctggc ccacctcgt gaccaccttg 480

 acctacggcg tgcagtgtt cccccgtac cccgaccaca tgaagcagca cgacttcttc 540
 aagtccgcca tgcccgaagg ctactccag gagcgcacca tcttcttcaa ggacgacggc 600
 aactacaaga cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa ccgcatcgag 660
 ctgaagggca tcgacttcaa ggaggacggc aacatcctgg ggcacaagct ggagtacaac 720
 tacaacagcc acaaggtcta tatcaccgcc gacaagcaga agaacggcat caaggtgaac 780
 ttcaagacc gccacaacat cgaggacggc agcgtgcagc tcgccgacca ctaccagcag 840
 aacaccccca tcggcgacgg ccccgctgtg ctgcccgaca accactacct gagcaccag 900

 tccgccctga gaaagacc caacgagaag cgcgatcaca tggctctgct ggagtctgtg 960
 accgccgccc ggatcactct cggcatggac gagctgtaca agtgtgcggc cgcagataaa 1020
 ataaaagatt ttatttag 1038

 <210> 39
 <211> 548
 <212> DNA
 <213> Artificial Sequence
 <220><223> T2A-AscI-yCD2 of pAC3-T2A-yCD2
 <400> 39

 gagggcagag gaagtcttct aacatcggt gacgtggagg agaatcccgg cctggcgcg 60
 cctatggtga ccggcggcat ggctccaag tgggatcaaa agggcatgga tatcgttac 120

 gaggaggccc tgctgggcta caaggaggc ggcgtgccta tcggcggctg tctgatcaac 180
 aacaaggacg gcagtgtgct gggcaggggc cacaacatga ggttcagaa gggctccgcc 240
 accctgcacg gcgagatctc caccctggag aactgtggca ggctggaggg caaggtgtac 300
 aaggacacca ccctgtacac caccctgtcc ctttgtgaca tgtgtaccgg cgctatcatc 360
 atgtacggca tccttaggtg tgtgatcggc gagaactga acttcaagtc caagggcgag 420
 aagtacctgc aaaccagggg ccacaggtg gtggttgttg acgatgagag gtgtaagaag 480
 ctgatgaagc agttcatcga cgagaggcct caggactggt tcgaggatat cggcgagtaa 540

 gcggccgc 548
 <210> 40

<211> 551
 <212> DNA
 <213> Artificial Sequence
 <220><223> P2A-AscI-yCD2 of pAC3-P2A-yCD2
 <400> 40
 gctactaact tcagcctgct gaagcagget ggagacgtgg aggagaacce tggacctggc 60
 gcgcctatgg tgaccggcgg catggcctcc aagtgggatc aaaaggcat ggatatcgct 120
 tacgaggagg ccctgctggg ctacaaggag ggcggcgtgc ctatcggcgg ctgtctgac 180
 aacaacaagg acggcagtgt gctgggcagg ggccacaaca tgaggttcca gaaggctcc 240
 gccaccctgc acggcgagat ctccaccctg gagaactgtg gcaggctgga gggcaaggtg 300
 tacaaggaca ccaccctgta caccaccctg tcccttgtg acatgtgtac cggcctatc 360
 atcatgtacg gcattccctag gtgtgtgac ggcgagaacg tgaacttcaa gtccaagggc 420
 gagaagtacc tgcaaaccag gggccacgag gtggtggtt ttgacgatga gaggtgtaag 480
 aagctgatga agcagttcat cgacgagagg cctcaggact ggttcgagga tatcggcgag 540
 taagcggccg c 551
 <210> 41
 <211> 557
 <212> DNA
 <213> Artificial Sequence
 <220><223> GSG-T2A-AscI-yCD2 of pAC3-GSG-T2A-yCD2
 <400> 41
 ggaagcggag agggcagagg aagtcttcta acatgcggtg acgtggagga gaatcccggc 60
 cctggcgcgc ctatggtgac cggcggcatg gcctccaagt gggatcaaaa gggcatggat 120
 atcgcttacg aggaggccct gctgggctac aaggaggcgc gcgtgcctat cggcggctgt 180
 ctgatcaaca acaaggacgg cagtgtgctg ggcaggggcc acaacatgag gttccagaag 240
 ggctccgcca ccctgcacgg cgagatctcc acctggaga actgtggcag gctggagggc 300
 aaggtgtaca aggacaccac cctgtacacc acctgtccc cttgtgacat gtgtaccgac 360
 gctatcatca tgiacggcat ccctaggtgt gtgatcggcg agaactgaa cttcaagtcc 420
 aaggcgcgaga agtacctgca aaccaggggc cagcaggtgg tggttgttga cgatgagagg 480
 tgtaagaagc tgatgaagca gttcatcgac gagaggcctc aggactggtt cgaggatc 540
 ggcgagtaag cggccgc 557
 <210> 42

<211> 560
 <212> DNA
 <213> Artificial Sequence
 <220><223> GSG-P2A-AscI-yCD2 of pAC3-GSG-P2A-yCD2
 <400> 42
 ggaagcggag ctactaactt cagcctgctg aagcaggctg gagacgtgga ggagaacct 60
 ggacctggcg cgcctatggt gaccggcggc atggcctcca agtgggatca aaaggcatg 120
 gatatcgctt acgaggagc cctgctgggc tacaaggagg gcggcgtgcc tatcggcggc 180
 tgtctgatca acaacaagga cggcagtgtg ctgggcaggg gccacaacat gaggttccag 240
 aagggtccg ccaccctgca cggcgagatc tccaccctgg agaactgtgg caggctggag 300
 ggcaaggtgt acaaggacac caccctgtac accaccctgt ccccttgtga catgtgtacc 360
 ggcgctatca tcatgtacgg catccctagg tgtgtgatcg gcgagaacgt gaacttcaag 420
 tccaagggcg agaagtacct gcaaaccagg ggccacgagg tgggtggtgt tgacgatgag 480
 aggtgtaaga agctgatgaa gcagttcatc gacgagagc ctcaggactg gttcgaggat 540
 atcggcagat aagcggccgc 560
 <210> 43
 <211> 11642
 <212> DNA
 <213> Artificial Sequence
 <220><223> pAC3-T2A-GFPm
 <400> 43
 tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagtccg 60
 cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc cccgcccatt 120
 gacgtcaata atgacgtatg ttccatagt aacccaata gggactttcc attgacgtca 180
 atgggtggag tatttacggt aaactgceca cttggcagta catcaagtgt atcatatgcc 240
 aagtagccc cctattgacg tcaatgacgg taatggccc gcctggcatt atgccagta 300
 catgacctta tgggacttc ctaactggca gtacatctac gtattagtca tcgtattac 360
 catggtgatg cggttttggc agtacatcaa tggcgtgga tagcggttg actcacgggg 420
 attccaagt ctccaccca ttgacgtcaa tgggagttg ttttggcacc aaaatcaacg 480
 ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg gtaggcgtgt 540
 acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgcagc cctccgattg 600

actgagtcgc ccgggtaccg gtgtatccaa taaacctct tgcagttgca tccgacttgt 660
 ggtctcgtg ttcttggga gggctctctc tgagtattg actaccctc agcgggggtc 720
 tttcatttgg ggctcgtcc gggatcggga gaccctgcc cagggaccac cgaccacca 780
 ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840
 tgattttatg cgctcgtc ggtactagt agctaactag ctctgtatct ggcggaccgc 900
 tggtggaact gacgagttcg gaacaccgg ccgcaacct gggagacgtc ccagggactt 960
 cgggggccgt tttgtggcc cgacctgagt ccaaaaatcc cgatcgttt ggactctttg 1020

gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt 1080
 tcccgcctcc gtctgaattt ttgcttccg tttgggaccg aagccgcgcc gcgcgtcttg 1140
 tctgtcgcag catcgttctg tgttctctc gtctgactgt gtttctgtat ttgtctgaga 1200
 atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg 1260
 agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct 1320
 ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380
 tcatcaccca ggtaagatc aaggctttt cacctggccc gcatggacac ccagaccagg 1440

tcccctacat cgtgacctgg gaagccttgg cttttgacc cctccctgg gtcaagcct 1500
 ttgtacacc taagcctccg cctctcttc ctccatcgc cccgtctctc cccctgaac 1560
 ctctcgttc gaccccgct cgatctccc ttatccage cctcactct tetctagggc 1620
 ccaaactaa acctcaagt ctttctgaca gtggggggcc gctcatcgac ctacttacag 1680
 aagaccccc gccttatagg gaccaagac cacccttc cgacaggac ggaaatggtg 1740
 gagaagcgac cctgcggga gaggcaccg accctcccc aatggcatct cgctacgtg 1800
 ggagacggga gcccctgtg gccactcca ctacctcga ggcatcccc ctccgcgac 1860

gaggaaacgg acagcttcaa tactggcctg tctctcttc tgaccttac aactggaaaa 1920
 ataataacc ttcttttct gaagatccag gtaaactgac agctctgac gactctgttc 1980
 tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040
 gagaagaaaa acaacgggtg ctcttagagg ctgaaaaggc ggtgcggggc gatgatgggc 2100
 gcccactca actgccaat gaagtcgat ccgctttcc cctcgagcgc ccagactggg 2160
 attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagtg ctcttagcgg 2220
 gtctcaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280

ggcccaatga gtctccctcg gccttctag agagacttaa ggaagcctat cgcaggtaca 2340
 ctcttatga cctgaggac ccagggaag aactaatgt gtctatgtct ttcatttggc 2400
 agtctgcccc agacattggg agaaagttag agaggtaga agatttaaa aacaagacgc 2460

ttggagattt ggtagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520
 gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccctagg acagaggatg 2580
 agcagaaaga gaaagaaaga gatcgttaga gacatagaga gatgagcaag ctattggcca 2640
 ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700

 atcgcgacca gtgtgcctac tgcaaagaaa aggggcactg ggctaaagat tgtccaaga 2760
 aaccacgagg acctcgggga ccaagacccc agacctcct cctgacccta gatgactagg 2820
 gaggtcaggg tcaggagccc ccccctgaac ccaggataac cctcaaagtc ggggggcaac 2880
 ccgtcacctt cctggtagat actggggccc aacctcctg gctgaccaa aatcctggac 2940
 ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgtgga 3000
 ccacggatcg caaagtacat ctactaccg gtaaggtcac ccactcttc ctccatgtac 3060
 cagactgtcc ctatcctctg ttaggaagag atttgctgac taaactaaaa gcccaaatcc 3120

 actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgtga 3180
 ccctaatat agaagatgag catcggctac atgagacctc aaaagagcca gatgtttctc 3240
 taggttccac atggctgtct gattttctc aggcctgggc ggaaaccggg ggcatgggac 3300
 tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360
 taaaacaata ccccatgtca caagaagcca gactggggat caagcccccac atacagagac 3420
 tgttgacca ggaatactg gtacctgcc agtccccctg gaacacgccc ctgctaccg 3480
 ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc 3540

 ggggtggaaga catccacccc accgtgccc acccttaca cctcttgagc gggctcccac 3600
 cgtcccacca gtggtacact gtgcttgatt taaaggatgc cttttctgc ctgagactcc 3660
 accccaccag tcagcctctc ttgcctttg agtggagaga tccagagatg ggaatctcag 3720
 gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780
 aggcactgca cagagacctc gcagacttcc ggatccagca cccagacttg atcctgctac 3840
 agtacgtgga tgacttactg ctggccgcca ctctgagct agactgcaa caaggtactc 3900
 gggcctgtt acaaaccta gggaacctcg ggtatcgggc ctggccaag aaagccaaa 3960

 tttgccagaa acaggtcaag tatctgggt atcttctaaa agagggtcag agatggctga 4020
 ctgagccag aaaagagact gtgatgggc agectactcc gaagaccct cgacaactaa 4080
 gggagtctct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140
 cagccccctt gtacctctc accaaaacgg ggactctgtt taattgggac ccagaccaac 4200
 aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgcag 4260
 atttgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc 4320

taacgcaaaa actgggacct tggcgtcggc cggtaggcta cctgtccaaa aagctagacc 4380

 cagtagcagc tgggtggccc ccttgectac ggatggtagc agccattgcc gtactgacaa 4440
 aggatgcagg caagctaacc atgggacagc cactagtcac tctggccccc catgcagtag 4500
 aggcactagt caaacaacct cccgaccgct ggctttccaa cgcccggatg actcactatc 4560
 aggccttgct tttggacacg gaccgggtcc agttcggacc ggtgtagacc ctgaaccgg 4620
 ctacgtgct ccactgctt gaggaagggc tgcaacacaa ctgccttgat atcctggcgg 4680
 aagcccacgg aaccggacct gacctaacgg accagccgct cccagacgcc gaccacacct 4740
 ggtacacgga tggaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800

 ccaccgagac cgaggtaatc tgggctaaag ccctgccagc cgggacatcc gctcagcggg 4860
 ctgaactgat agcactcacc caggccctaa agatggcaga aggtagaag ctaaatgttt 4920
 atactgatag ccgttatgct tttgctactg cccatatcca tggagaaata tacagaaggc 4980
 gtgggttgct cacatcagaa ggcaaaagaga tcaaaaataa agacgagatc ttggccctac 5040
 taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100
 gacacagcgc cgagctaga ggcaaccgga tgctgacca agcggcccga aaggcagcca 5160
 tcacagagac tccagacacc tctacctcc tcatagaaaa ttcacaccc tacacctcag 5220

 aacattttca ttacacagtg actgatataa aggacctaac caagtgggg gccatztatg 5280
 ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340
 ttgaattatt agactttctt catcagctga ctacctcag cttctcaaaa atgaaggctc 5400
 tcctagagag aagccacagt ccctactaca tctgaaccg ggatcgaaca ctcaaaaata 5460
 tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaacagg 5520
 gaactagggt ccgcccggat cggcccggca ctattggga gatcatttc accgagataa 5580
 agcccggatt gtatggetat aaatatcttc tagtttttat agatacttt tetggetgga 5640

 tagaacctt cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg 5700
 agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760
 tctcaaggt gactcagaca gtggccgac tgttgggat tgattgaaa ttacattgtg 5820
 catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaacctc aaggagactt 5880
 taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc 5940
 tgtaccgagc ccgcaacacg ccgggcccc atggcctcac cccatagag atcttatatg 6000
 gggcaccccc gcccttgta aacttccctg acctgacat gacaagagtt actaacagcc 6060

cctctctcca agctcactta caggctctct acttagtcca gcacgaagtc tggagacctc 6120
 tggcggcagc ctaccaagaa caactggacc gaccggtggt acctcaccct taccgagtcg 6180
 gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240
 cttacacagt cctgctgacc acccccaccg cctcaaaagt agacggcatic gcagcttggg 6300
 tacacgccgc ccacgtgaag gctgccgacc cgggggtgg accatcctct agactgacat 6360
 ggcgcttca acgctctcaa aaccacctca agataagatt aaccctgga agcccttaat 6420
 agtcatggga gtctgttag gtagtaggat ggcagagagc ccccatcagg tctttaatgt 6480

 aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct cctcctggg 6540
 aactgtacaa gatgccttcc caaaattata ttttgatcta tgtgatctgg tcggagagga 6600
 gtgggacctc tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcaggag 6660
 acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcggggtg 6720
 tgggggacca ggagagggtc actgtggtaa atgggggtgt gaaaccaccg gacagcctta 6780
 ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcgtaaca cccctggga 6840
 cacgggatgc tctaaagtig cctgtggccc ctgctacgac ctctccaaag tatccaattc 6900

 ctccaaggg gctactcag ggggcagatg caaccctcta gtcctagaat tcaactgatgc 6960
 aggaaaaag gctaactggg acgggcccga atcgtgggga ctgagactgt accggacagg 7020
 aacagatcct attaccatgt tctcctgac ccggcaggtc cttaatgtgg gaccccgagt 7080
 ccccataggg cccaaccag tattaccga ccaaagactc ccttcctcac caatagagat 7140
 tgtaccggtc ccacagccac ctageccct caataccagt taccacctt ccaactaccg 7200
 tacacctca acctcccta caagtccaag gtccccacag ccacccccag gaactggaga 7260
 tagactacta gctctagtca aaggagccta tcaggcgtt aacctacca atcccacaa 7320

 gaccaagaa tgttggctgt gcttagtgc gggacctcct tattacgaag gagtagcgg 7380
 cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca cttccaaca 7440
 taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac 7500
 tcaccaggcc ttatgtaaca ccacccaaag cgccggetca ggatcctact accttgcagc 7560
 acccggcga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccaggt 7620
 gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taatttacca 7680
 ctccccgat tatatgatg gtcagcttga acagcgtacc aaatataaaa gagagccagt 7740

 atcattgacc ctggcccttc tactaggagg attaacatg ggaggattg cagctggaat 7800
 agggacgggg accactgct taattaaaac ccagcagttt gagcagcttc atgccctat 7860
 ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920

gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgctattcc taaaggaggg 7980
 aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt 8040
 gagagacagc atggccaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100
 aggccaagga tggttcgaag ggctgtttaa tagatcccc tggtttacca ccttaatctc 8160

caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac ctgcattct 8220
 caatcgattg gtccaatttg ttaaagacag gatctcagtg gtcaggctc tggttttgac 8280
 tcagcaatat caccagctaa aacctataga gtacgagcca gagggcagag gaagtcttct 8340
 aacatgcggt gacgtggagg agaatcccgg cctggcgcg cctatggcca gcaagggcga 8400
 ggagctgttc accgggggtg tgcccatcct ggtcgagctg gacggcgacg taaacggcca 8460
 caagttcagc gtgtccggcg aaggagaggg cgatgccacc tacggcaagc tgaccctgaa 8520
 gttcatctgc accaccggca agctgcccgt gccctggccc accctcgtga ccacctgac 8580

ctacggcgtg cagtgcctcg cccgctacc cgaccacatg aagcagcacg acttcttcaa 8640
 gtccgcatg cccgaaggct acgtccagga gcgcaccatc ttcttcaagg acgacggcaa 8700
 ctacaagacc cgcgcgagg tgaagttega gggcgacacc ctggtgaacc gcatcgagct 8760
 gaagggcatc gacttcaagg aggacggcaa catcctgggg cacaagctgg agtacaacta 8820
 caacagccac aaggtctata tcaccgccga caagcagaag aacggcatca aggtgaactt 8880
 caagaccgc cacaacatcg aggacggcag cgtgcagctc gccgaccact accagcagaa 8940
 caccctcatc ggcgacggcc ccgtgctgct gcccgacaac cactacctga gcaccagtc 9000

cgccctgagc aaagacccca acgagaagcg cgatcacatg gtctctgctgg agttcgtgac 9060
 cgcccgccgg atcactctcg gcatggacga gctgtacaag tgtgcggccg cagataaaat 9120
 aaaagat ttttagtctc cagaaaaagg ggggaatgaa agaccccacc ttaggtttg 9180
 gcaagctagc ttaagtaacg ccattttgca aggcattgaa aaatacataa ctgagaatag 9240
 agaagttcag atcaaggtca ggaacagatg gaacagctga atatggcca aacaggatat 9300
 ctgtggtgaa cagttctctc cccgctcag ggccaagaac agatggaaca gctgaatatg 9360
 ggccaacag gatatctgtg gtaagcagtt cctgccccgg ctcagggcca agaacagatg 9420

gtccccagat gcggtccagc cctcagcagt ttctagagaa ccatcagatg tttccagggt 9480
 gcccgaagga cctgaaatga ccctgtgctt ttttgaact aaccaatcag ttcgcttctc 9540
 gcttctgttc gcgcgcttct gctccccgag ctcaataaaa gagcccacaa ccctcactc 9600
 ggggcgccag tcttccgatt gactgagtcg cccgggtacc cgtgtatcca ataaacctc 9660
 ttgcagttgc atccgacttg tggctctgct gttccttggg agggctctct ctgagtgatt 9720
 gactaccctg cagcgggggt ttttattac atgtgagcaa aaggccagca aaaggccagg 9780

aaccgtaaaa aggccgcggtt gctggcgcttt ttccataggc tccgcccccc tgacgagcat 9840

cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga caggactata aagataaccag 9900
gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gettaccgga 9960
tacctgtccg cttttctccc ttcgggaagc gtggcgcttt ctcaatgctc acgctgtagg 10020
tatctcagtt cgggtgtagt cgttcgctcc aagctgggct gtgtgcacga accccccgtt 10080
cagccccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaacce ggtaagacac 10140
gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc 10200
ggtgctacag agttcttgaa gtgggtggcct aactacggct aactagaag gacagtattt 10260

ggtatctcgc ctctcgtgaa gccagttacc ttcggaaaaa gaggttgtag ctcttgatcc 10320
ggcaaaaaaa ccaccgctgg tagcgggtgt tttttgttt gcaagcagca gattacgcgc 10380
agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga cgctcagtgg 10440
aacgaaaact cacgttaagg gatthttgtc atgagattat caaaaaggat cttcacctag 10500
atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga gtaaaacttg 10560
tctgacagtt accaatgctt aatcagttag gcacctatct cagcgatctg tetatctctg 10620
tcatccatag ttgcctgact ccccgctcgt tagataacta cgatacggga gggttacca 10680

tctggcecca gtgctgcaat gataaccgca gaccacgct caccggctcc agatttatca 10740
gcaataaacc agccagccgg aagggccgag cgcagaagtg gtcctgcaac tttatccgcc 10800
tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc agttaatagt 10860
ttgcgcaacg ttgttgccat tgctgcagge atcgtggtgt cacgctcgtc gtttggtatg 10920
gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc catgttgtgc 10980
aaaaaagcgg ttagctcctt cggctctcgg atcgttgtca gaagtaagtt ggccgcagtg 11040
ttatcactca tggttatggc agcactgeat aattctctta ctgtcatgcc atccgtaaga 11100

tgcttttctg tgactgggta gtactcaacc aagtcattct gagaatagtg tatgcggcga 11160
ccgagtgtct cttgcccgcc gtcaacacgg gataataacc cgccacatag cagaacttta 11220
aaagtgtca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat cttaccgctg 11280
ttgagatcca gttcagatga acccactcgt gcaccaact gatcttcagc atcttttact 11340
ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa aaaggaata 11400
agggcgacac ggaatgttg aatactcata ctcttctttt ttcaatatta ttgaagcatt 11460
tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacia 11520

ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt 11580
atcatgacat taacctataa aaatagcgct atcacgaggc cctttcgtct tcaagaattc 11640
at 11642
<210> 44
<211> 11651
<212> DNA
<213> Artificial Sequence
<220><223> pAC3-GSG-T2A-GFPm
<400> 44
tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagtccg 60
cgttacataa cttacggtaa atggcccggc tggtgaccg cccaacgacc ccgcccatt 120
gacgtcaata atgacgtatg ttcccatagt aacccaata gggactttcc attgacgtca 180
atgggtggag tatttacggt aaactgcca cttggcagta catcaagtgt atcatatgcc 240
aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 300
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 360
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420
atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480
ggactttcca aaatgtcgt acaactccgc cccattgacg caaatgggcg gtaggcgtgt 540
acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg 600
actgagtgcg ccgggtaccg gtgtatccaa taaacctct tgcagttgca tccgacttgt 660
ggtctcgtg ttcttggga gggctctctc tgagtattg actaccctc agcgggggctc 720
tttcatttgg gggctcgtcc gggatcgga gaccctgcc cagggaccac cgaccacca 780
ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840
tgattttatg cgctgcgtc ggtactagt agctaactag ctctgtatct ggcggaccg 900
tggtggaact gacgagttcg gaacaccgg cgcgaacct gggagacgtc ccagggactt 960
cgggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg 1020
gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt 1080
tcccgcctcc gtctgaatth ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg 1140
tctgctgcag catcgttctg tgttctctc gtctgactgt gtttctgtat ttgtctgaga 1200
atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg 1260
agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct 1320

ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380

tcatcaccca ggttaagatc aaggtctttt cacctggccc gcatggacac ccagaccagg 1440

tcccctacat cgtgacctgg gaagccttgg cttttgaccc cctccctgg gtcaagccct 1500

ttgtacacc taagcctccg cctcctcttc ctccatccgc cccgtctctc cccttgaac 1560

ctcctcgttc gacccccctc cgatcctccc tttatccagc cctcactcct tctctaggcg 1620

ccaaacctaa acctcaagtt ctttttgaca gtggggggcc gctcatcgac ctacttacag 1680

aagaccccc gccttatagg gacccaagac caccctctc cgacagggac ggaaatggtg 1740

gagaagcgac cctgcggga gaggcaccgg accctcccc aatggcatct cgcctacgtg 1800

ggagacggga gccccctgtg gccgactcca ctacctgca ggcatteccc ctccgcgacg 1860

gaggaaacgg acagcttcaa tactggccgt tctcctcttc tgacctttac aactggaaaa 1920

ataataacce ttctttttct gaagatccag gtaaaactgac agctctgac gagtctgttc 1980

tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040

gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100

gccccactca actgcccact gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg 2160

attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagtgt ctcttagcgg 2220

gtctcaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280

ggcccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaggtaca 2340

ctccttatga cctgaggac ccagggaag aaactaatgt gtctatgtct ttcatttggc 2400

agtctgcccc agacattggg agaaagttag agaggttaga agatttaaaa aacaagacgc 2460

ttggagattt ggtagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520

gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580

agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640

ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700

atcgcgacca gtgtgcctac tgcaaagaaa aggggcactg ggctaaagat tgtccaaga 2760

aaccacgagg acctcgggga ccaagacccc agacctcct cctgacccta gatgactagg 2820

gaggtcaggg tcaggagccc ccccctgaac ccaggataac cctcaaagtc ggggggcaac 2880

ccgtcacctt cctggtagat actggggccc aacctcctg gctgaccaa aatcctggac 2940

ccctaagtga taagtctgcc tgggtccaag gggtactgg aggaaagcgg tatcgctgga 3000

ccacgcatcg caaagtacat ctactaccg gtaaggtcac cactctttc ctccatgtac 3060

cagactgtcc ctatcctctg ttaggaagag atttgctgac taaactaaaa gcccaaatcc 3120
 actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga 3180
 ccctaaatat agaagatgag catcggttac atgagacctc aaaagagcca gatgtttctc 3240
 tagggctccac atggctgtct gattttcttc aggcctgggc ggaaaccggg ggcattgggac 3300
 tggcagttec ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360
 taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420
 tgttgaccac ggaataactg gtaccctgcc agtccccctg gaacacgccc ctgctacccg 3480

 ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc 3540
 ggggtgaaga catccacccc accgtgccca acccttacia cctcttgagc gggctcccac 3600
 cgtcccacca gtggtacct gtgcttgatt taaaggatgc cttttctgc ctgagactcc 3660
 accccaccag tcagcctctc ttgcctttg agtggagaga tccagagatg ggaatctcag 3720
 gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780
 aggcactgca cagagacctc gcagacttcc ggatccagca cccagacttg atcctgctac 3840
 agtacgtgga tgacttactg ctggccgcca ctctgagct agactgcca caaggtactc 3900

 gggccctgtt acaaaccta ggaacctcg ggtatcgggc ctcgccaag aaagcccaaa 3960
 tttgccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga 4020
 ctgagggcag aaaagagact gtgatggggc agcctactcc gaagaccctc cgacaactaa 4080
 gggagttcct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140
 cagccccctt gtaccctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200
 aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag 4260
 atttgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc 4320

 taacgcaaaa actgggacct tggcgtcggc cggtggccta cctgtccaaa aagctagacc 4380
 cagtagcagc tgggtggccc ccttgcttac ggatggtagc agccattgcc gtactgacaa 4440
 aggatgcagg caagctaacc atgggacagc cactagtcat tctggcccc catgcagtag 4500
 aggcactagt caaacaacc cccgaccgct ggctttccaa cgcccggatg actcactatc 4560
 aggccttgct ttggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaaccggg 4620
 ctacgtctct cccactgctt gaggaagggc tgcaacacaa ctgccttgat atcctggccc 4680
 aagcccacgg aaccgaccc gacctaacgg accagccgct cccagacgcc gaccacacct 4740

 ggtacacgga tggaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800
 ccaccgagac cgaggtaatc tgggctaaag cctgccagc cgggacatcc gctcagcggg 4860
 ctgaactgat agcactcacc caggcctaa agatggcaga aggtaagaag ctaaatgttt 4920

atactgatag ccgttatgct tttgctactg cccatatcca tggagaaata tacagaaggc 4980
 gtgggttgct cacatcagaa ggcaaaagaga tcaaaaataa agacgagatc ttggccctac 5040
 taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100
 gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggcccga aaggcagcca 5160

 tcacagagac tccagacacc tctaccctcc tcatagaaaa tteatcacc tacacctcag 5220
 aacattttca ttacacagtg actgatataa aggacctaac caagtgggg gccatttatg 5280
 ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340
 ttgaattatt agactttctt catcagctga ctacctcag cttctcaaaa atgaaggctc 5400
 tcctagagag aagccacagt ccctactaca tgetgaaccg ggatcgaaca ctcaaaaata 5460
 tcaactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaacagg 5520
 gaactaggtt ccgcgggcat cggcccggca ctattggga gatcgatttc accgagataa 5580

 agccccgatt gtatggctat aaatatcttc tagtttttat agataccttt tctggctgga 5640
 tagaagcctt cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg 5700
 agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgcctteg 5760
 tctccaaggt gactcagaca gtggccgac tgttggggat tgattggaaa ttacattgtg 5820
 catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaacctc aaggagactt 5880
 taactaaatt aacgcttgca actggctcta gagactgggt gctcctact cccttagccc 5940
 tgtaccgagc ccgcaacacg ccgggcccc atggcctcac cccatagag atcttatatg 6000

 gggcaccccc gcccttgta aacttccctg acctgacat gacaagagt actaacagcc 6060
 cctctctcca agctcactta caggctctct acttagtcca gcacgaagtc tggagacctc 6120
 tggcggcage ctaccaagaa caactggacc gaccgggtgt acctaccct taccgagtcg 6180
 gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240
 cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcatc gcagcttggg 6300
 tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatctct agactgacat 6360
 ggcgcttca acgctctcaa aaccctca agataagatt aaccctgga agcccttaat 6420

 agtcatggga gtctgttag gactaggat ggcagagagc cccatcagg tctttaatgt 6480
 aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct cctctctggg 6540
 aactgtacaa gatgccttcc caaaattata tttgatcta tgtgatctgg tcggagagga 6600
 gtgggacctc tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcaggag 6660
 acagcggacc cggacttttg acttttacgt gtgccctggg catacgtaa agtcggggtg 6720
 tgggggacca ggagaggct actgtggtaa atgggggtgt gaaaccaccg gacaggctta 6780

ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcggtaca cccctggga 6840

cacgggatgc tctaaagtg cctgtggccc ctgctacgac ctctccaaag tatccaattc 6900

cttccaaggg gctactcgag ggggcagatg caaccctcta gtcctagaat tcaactgatgc 6960

aggaaaaag gctaaactggg acgggcccaa atcgtgggga ctgagactgt accggacagg 7020

aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccgagt 7080

cccataggg cccaaccag tattaccega ccaagactc ccttctcac caatagagat 7140

tgtaccggct ccacagccac ctgccccct caataccagt taccctctt ccaactaccag 7200

tacacctca acctcccta caagtccaag tgtcccacag ccacccag gaactggaga 7260

tagactacta gctctagca aaggagccta tcaggcgtt aacctacca atcccacaa 7320

gaccaagaa tgttggctgt gcttagtgtc gggacctct tattacgaag gagtagcgtt 7380

cgtgggcact tataccaatc attccaccgc tccggccaac tgtaccgcca ctccaaca 7440

taagcttacc ctatctgaag tgacaggaca gggcctatgc atggggcag tacctaaaac 7500

tcaccaggcc ttatgtaaca ccaccaaag cgcggctca ggatcctact accttcagc 7560

accgcccga acaatgtggg cttgcagcac tggattgact cctgcttgt ccaccaggt 7620

gctcaatcta accacagatt atttgtatt agttgaactc tggcccagag taatttacca 7680

ctccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt 7740

atcattgacc ctggcccttc tactaggagg attaacatg ggaggattg cagctggaat 7800

agggacgggg accactgct taattaaaac ccagcagttt gagcagcttc atgccctat 7860

ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctgaaaagt cactgacctc 7920

gttgtctgaa gtagtctac agaaccgag aggcctagat ttgctattcc taaaggagg 7980

aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cgggctagt 8040

gagagacagc atggccaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100

aggccaagga tggttcgaag ggctgtttaa tagatcccc tggttacca ccttaatctc 8160

caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgacttct 8220

caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac 8280

tcagcaatat caccagctaa aaccataga gtacagcca ggaagcggag agggcagagg 8340

aagtcttcta acatgcggtg acgtggagga gaatcccggc cctggcgcgc ctatggccag 8400

caaggcgag gagctgttca ccgggtggt gccatcctg gtcgagctgg acggcgact 8460

aaacggccac aagttcagc tgtccggcga aggagagggc gatgccacct acggcaagct 8520

gaccctgaag ttcatctgca ccaccggcaa gctgcccgtg ccctggccca ccctcgtgac 8580
caccttgacc tacggcgtgc agtgcttcgc ccgtacccc gaccacatga agcagcacga 8640
cttcttcaag tccgcatgc ccgaaggcta cgtccaggag cgcacatct tcttcaagga 8700
cgacggcaac tacaagacc gcgccgaggt gaagttcgag ggcgacacce tggatgaaccg 8760
catcgagctg aaggcatcg acttcaagga ggacggcaac atcctggggc acaagctgga 8820
gtacaactac aacagccaca aggtctatat caccgcccac aagcagaaga acggcatcaa 8880
ggtgaacttc aagaccgcc acaacatcga ggacggcagc gtgcagctcg ccgaccacta 8940

ccagcagaac accccatcg gcgacggccc cgtgctgctg cccgacaacc actacctgag 9000
cacccagtcc gccctgagca aagaccctaa cgagaagcgc gatcacatgg tctgctgga 9060
gttcgtgacc gccgccggga tcaactcctg catggacgag ctgtacaagt gtgcggccgc 9120
agataaaata aaagatttta tttagtctcc agaaaaagg gggaatgaaa gacccacct 9180
gtaggtttgg caagctagct taagtaacgc cttttgcaa ggcatggaaa aatacataac 9240
tgagaataga gaagttcaga tcaaggtcag gaacagatgg aacagctgaa tatgggcca 9300
acaggatatac tgtgtaagc agttcctgcc ccggctcagg gccaagaaca gatggaacag 9360

ctgaatatgg gccaaacagg atatctgtgg taagcagttc ctgccccggc tcagggcca 9420
gaacagatgg tccccagatg cggctccagc ctcagcagtt tctagagaac catcagatgt 9480
ttccagggtg cccaaggac ctgaaatgac cctgtgcctt atttgaacta accaatcagt 9540
tcgcttctcg ctctgttcg cgcgcttctg ctccccgagc tcaataaaag agcccacaac 9600
ccctcactcg gggcgccagt cctccgattg actgagtcgc ccgggtaccg gtgtatccaa 9660
taaaccctct tgcagttgca tccgacttgt ggtctcctg ttccttggga gggctcctc 9720
tgagtgattg actaccctc agcgggggtc tttcattaca tgtgagcaa aggccagcaa 9780

aaggccagga accgtaaaaa ggccgcttg ctggcgtttt tccataggt cgcceccct 9840
gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaaccgac aggactataa 9900
agataccagg cgtttcccc tggaaactcc ctctgctgct ctctgttcc gaccctgccg 9960
cttaccggat acctgtccgc ctttctcct tcgggaagcg tggcgtttc tcaatgctca 10020
cgctgtaggt atctcagttc ggtgtaggtc gttcgtcca agctgggctg tgtgcacgaa 10080
cccccgctc agcccgacc ctgcgcctta tccgtaact atcgtcttga gtccaaccg 10140
gtaagacacg acttatgcc actggcagca gccactggtg acaggattag cagagcgagg 10200

tatgtaggcg gtgctacaga gttcttgaag tggtagccta actacggcta cactagaagg 10260
acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc 10320
tcttgatccg gaaacaac caccctggt agcgggtggt ttttgtttg caagcagcag 10380

attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac 10440
 gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc 10500
 ttcacctaga tctttttaa ttaaaaatga agttttaa atcaatctaaag tatatatgag 10560
 taaacttggc ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt 10620

 ctatttcggt catccatagt tgccctgactc cccgtcgtgt agataactac gatacgggag 10680
 ggcttaccat ctggccccag tgctgcaatg ataccgcgag acccacgctc accggetcca 10740
 gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg tctgcaact 10800
 ttatccgct ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca 10860
 gtaaatagtt tgcgcaactg tgttgcatt gctgcaggca tcgtggtgac acgctcgtcg 10920
 tttggtatgg ctccattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc 10980
 atgttgtgca aaaaagcggg tagctccttc ggtcctccga tcgttgtcag aagtaagttg 11040

 gccgcagtgt tatcactcat ggttatggca gcaactgcata attctcttac tgcacatcca 11100
 tccgtaagat gcttttctgt gactggtgag tactcaacca agtcattctg agaatagttg 11160
 atgcccggac cgagttgctc ttgcccggcg tcaacacggg ataataccgc gccacatagc 11220
 agaactttaa aagtgtcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc 11280
 ttaccgctgt tgagatccag ttcatgtaa cccactcgtg cacccaactg atcttcagca 11340
 tcttttactt tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa 11400
 aagggaataa gggcgacacg gaaatgttga atactcatal tcttctttt tcaatattat 11460

 tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 11520
 aataaacaaa taggggttcc ggcacattt ccccgaaaag tgccacctga cgtctaagaa 11580
 accattatta tcatgacatt aacctataa aataggcgta tcacgaggcc ctttcgtctt 11640
 caagaattca t 11651

 <210> 45
 <211> 11645
 <212> DNA
 <213> Artificial Sequence
 <220><223> pAC3-P2A-GFPm
 <400> 45
 tagttattaa tagtaatcaa ttacgggggc attagttcat agcccatata tggagttccg 60

 cgttacataa cttacggtaa atggccccgc tggctgaccg cccaacgacc cccgcccatt 120
 gacgtcaata atgacgtatg ttccatagc aacccaata gggactttcc attgacgtca 180

atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 240
aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta 300
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 360
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420
atccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480

ggactttcca aaatgctgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 540
acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg 600
actgagtgcg ccgggtaccc gtgtatccaa taaacctct tgcagttgca tccgacttgt 660
ggtctcgctg ttcttggga gggctctctc tgagtgattg actaccctgc agcgggggtc 720
tttcatttgg gggctctgcc gggatcggga gaccctgcc cagggaccac cgaccacca 780
ccgggaggtg agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840
tgattttatg cgctgcgctc ggtactagt agctaactag ctctgtatct ggcggacccg 900

tggtggaact gacgagttcg gaacaccgg cgcaaccct gggagacgtc ccagggactt 960
cgggggccgt ttttgtggc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg 1020
gtgcaccccc cttagaggag ggatatgtgg tttctgtagg agacgagaac ctaaacagtt 1080
tcccgcctcc gtctgaattt ttgcttccg tttgggaccg aagccgcgcc gcgctcttg 1140
tctgtctgag catcgttctg tgttctctct gtctgactgt gtttctgtat ttgtctgaga 1200
atatgggcca gactgttacc actccctaa gtttgactt aggtcactgg aaagatgtcg 1260
agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct 1320

ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380
tcatcaccca ggttaagatc aaggtctttt cacctgcccc gcattggacac ccagaccagg 1440
tcccctacat cgtgacctgg gaagccttgg cttttgacce cctccctgg gtcaagcct 1500
ttgtacacc taagcctccg cctcctcttc ctccatccgc cccgtctctc cccctgaac 1560
ctcctcttc gaccccgct cgatctctcc ttatccagc cctcactct tctctaggcg 1620
ccaaacctaa acctcaagtt cttttgaca gtggggggcc gctcatcgac ctacttacag 1680
aagaccccc gccttatagg gaccaagac cacccttc cgacaggac ggaaatggtg 1740

gagaagcgac cctgcgga gaggcaccgg acecctcccc aatggcatct cgctacgtg 1800
ggagacggga gccccctgtg gccgactcca ctacctgca ggcatcccc ctccgcgag 1860
gaggaacgg acagcttcaa tactggccgt tctcctctc tgacctttac aactgaaaa 1920
ataataacce ttcttttct gaagatccag gtaaaactgac agctctgac gagtctgttc 1980
tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040

gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100
 gccccactca actgcccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg 2160

 attacaccac ccaggcaggt aggaaccacc tagtccacta tgcaccagtg ctccatagcgg 2220
 gtctccaaaa cgcgggcaga agccccacca atttgccaa ggtaaaagga ataacacaag 2280
 ggcccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaggtaca 2340
 ctcttatga cectgaggac ccagggcaag aactaatgt gtctatgtct ttcatttggc 2400
 agtctgcccc agacattggg agaaagttag agaggtaga agatttaaaa aacaagacgc 2460
 ttggagattt ggttagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520
 gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580

 agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640
 ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700
 atcgcgacca gtgtgcctac tgcaagaaa aggggcactg ggctaaagat tgtccaaga 2760
 aaccacgagg acctcgggga ccaagacccc agacctcct cctgacccta gatgactagg 2820
 gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac 2880
 ccgtcacctt cctggtagat actggggccc aacactcct gctgacccaa aatcctggac 2940
 ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgtgga 3000

 ccacggatcg caaagtacat ctagctaccg gtaaggcac cactcttcc ctccatgtac 3060
 cagactgtcc ctatcctctg ttaggaagag atttgctgac taaactaaaa gcccaaatcc 3120
 actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga 3180
 ccctaataat agaagatgag catcggctac atgagacctc aaaagagcca gatgtttctc 3240
 taggtccac atggctgtct gattttctc aggctgggc ggaaccggg ggcattggac 3300
 tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360
 taaaacaata cccatgtca caagaagcca gactggggat caagccccac atacagagac 3420

 tgttgacca ggaataactg gtaccctgcc agtccccctg gaacacgccc ctgctaccg 3480
 ttaagaaacc agggactaat gattatagc ctgtccagga tctgagagaa gtcaacaagc 3540
 ggggtgaaga catccacccc accgtgcca acccttaca cctcttgagc gggctcccac 3600
 cgtcccacca gtgttacact gtgtttgatt taaaggatgc ctttttctgc ctgagactcc 3660
 accccaccag tcagcctctc ttcgcctttg agtggagaga tccagagatg ggaatctcag 3720
 gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780

aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atcctgctac 3840

agtacgtgga tgacttactg ctggccgcca cttctgagct agactgcca caaggtactc 3900

gggccctggt acaaaccta gggaacctcg ggtatcgggc ctggccaag aaagccaaa 3960

tttgccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga 4020

ctgagccag aaaagagact gtgatggggc agcctactcc gaagaccct cgacaactaa 4080

gggagttcct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140

cagccccctt gtaccctctc accaaaacgg ggactctggt taattggggc ccagaccaac 4200

aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag 4260

at ttgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaagggtcc 4320

taacgcaaaa actgggacct tggcgtcggc cggtaggcta cctgtccaaa aagctagacc 4380

cagtagcagc tgggtggccc ccttgccctac ggatggtagc agccattgcc gtactgacaa 4440

aggatgcagg caagctaacc atgggacagc cactagtcac tctggcccc catgcagtag 4500

aggcactagt caaacaacc cccgaccgt ggctttccaa cggccggatg actcactatc 4560

aggccttgct ttggacacg gaccgggtcc agttcggacc ggtgtagcc ctgaaccgg 4620

ctacgtgct cccactgct gaggaagggc tgcaacacaa ctgccttgat atcctggccg 4680

aagcccacgg aaccgaccc gacctaacgg accagccgt cccagacgcc gaccacacct 4740

ggtacacgga tgaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800

ccaccgagac cgaggaatc tgggctaaag ccttgccagc cgggacatcc gctcagcggg 4860

ctgaactgat agcactcacc cagccctaa agatggcaga agtaagaag ctaaatgttt 4920

atactgatag ccgttatgct ttgctactg cccatatcca tggagaaata tacagaaggc 4980

gtgggttgct cacatcagaa ggcaagaga tcaaaaataa agacgagatc ttggccctac 5040

taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100

gacacagcgc cgaggctaga ggcaaccgga tgctgacca agcggcccga aaggcagcca 5160

tcacagagac tcagacacc tctaccctcc tcatagaaaa ttcacaccc tacacctcag 5220

aacattttca ttacacagt actgatataa aggacctaac caagtgggg gccatttatg 5280

ataaaacaaa gaagtattgg gttaccaag gaaaacctgt gatgcctgac cagtttactt 5340

ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc 5400

tcctagagag aagccacagt ccctactaca tgctgaaccg ggatcgaaca ctcaaaaata 5460

tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaacagg 5520

gaactagggt ccgctggcat cggcccggca ctcatggga gatcgatttc accgagataa 5580
 agcccggatt gtatggctat aaatatcttc tagtttttat agataccttt tctggctgga 5640
 tagaacctt cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg 5700
 agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760
 tctccaaggt gagtcagaca gtggccgatc tgttggggat tgattggaaa ttacattgtg 5820
 catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaacctac aaggagactt 5880
 taactaaatt aacgcttgca actggctctc gagactgggt gctcctactc cccttagccc 5940

 tgtaccgagc ccgaacacg ccgggcccc atggcctcac cccatagag atcttatatg 6000
 gggcaccccc gcccttgta aacttccctg accctgacat gacaagagtt actaacagcc 6060
 cctctctcca agctcactta caggctctct acttagtcca gcacgaagtc tggagacctc 6120
 tggcggcagc ctaccaagaa caactggacc gaccgggtgt acctcaccct taccgagtcg 6180
 gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240
 cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcatc gcagcttggg 6300
 tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat 6360

 ggcgcttca acgctctcaa aaccctca agataagatt aaccctgga agccctaat 6420
 agtcatggga gtctgttag gagtagggat ggcagagagc ccccatcagg tctttaatgt 6480
 aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct ccctctggg 6540
 aactgtaca gatgccttcc caaaattata tttgatcta tgtgatctgg tggagagga 6600
 gtgggacctc tcagaccagg aacctatgt cgggtatggc tgcaagtacc ccgaggagg 6660
 acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcgggggtg 6720
 tgggggacca ggagagggt actgtggtaa atgggggtgt gaaaccaccg gacaggctta 6780

 ctggaagccc acatcatcgt gggaccta atctcccttaag cgcgtaaca cccctggga 6840
 cacgggatgc tctaaagttg cctgtggccc ctgctacgac ctctccaaag tatccaattc 6900
 cttccaagg gctactcgag ggggcagatg caacctcta gtcctagaat tactgatgc 6960
 aggaaaaag gctaactggg acgggccc aaatcgctggga ctgagactgt accggacagg 7020
 aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccgagt 7080
 ccccatagg cccaaccgatt tattaccga ccaaagactc ccttctcac caatagagat 7140
 tgtaccggt ccacagccac ctageccct caataccagt tacccectt cactaccag 7200

 tacacctca acctcccta caagtccaag tgtcccacag ccacccag gaactggaga 7260
 tagactacta gctctagtca aaggacctc tcaggcgtt aacctacca atcccagaa 7320
 gaccaagaa tgttggctgt gcttagtgtc gggacctcct tattacgaag gagtagcgtt 7380

cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca cttccaaca 7440
 taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaac 7500
 tcaccaggcc ttatgtaaca ccacccaaag cgcggctca ggatcctact accttcagc 7560
 acccgccgga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccacgt 7620

 gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taatttacca 7680
 ctccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt 7740
 atcattgacc ctggcccttc tactaggagg attaaccatg ggagggttg cagctggaat 7800
 agggacgggg accactgctc taattaaaac ccagcagttt gagcagcttc atgccctat 7860
 ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920
 gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgctattcc taaaggaggg 7980
 aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt 8040

 gagagacagc atggccaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100
 aggccaagga tggttcgaag ggctgtttaa tagatcccc tggtttacca ccttaatctc 8160
 caccatcatg ggacctetaa tagtactctt actgatetta ctctttggac cttgcattct 8220
 caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac 8280
 tcagcaatat caccagctaa aacctataga gtacgagcca gctactaact tcagcctgct 8340
 gaagcaggct ggagacgtgg aggagaacc tggacctggc gcgcctatgg ccagcaaggg 8400
 cgaggagctg ttcaccgggg tggtgcccat cctggctgag ctggacggcg acgtaaacgg 8460

 ccacaagttc agcgtgtccg gcgaaggaga gggcgatgcc acctacggca agctgacct 8520
 gaagttcatc tgcaccaccg gcaagctgcc cgtgccctgg cccacctcg tgaccacctt 8580
 gacctacgge gtgcagtctc tcgccccta cccgaccac atgaagcagc acgacttctt 8640
 caagtccgce atgcccgaag gctacgtcca ggagcgcacc atcttcttca aggacgacgg 8700
 caactacaag acccgcgccg aggtgaagtt cgagggcgac accctggtga accgcatcga 8760
 gctgaagggc atcgacttca aggaggacgg caacatctg gggcacaagc tggagtacia 8820
 ctacaacagc cacaaggtct atatcaccgc cgacaagcag aagaacggca tcaagtgaa 8880

 cttcaagacc cgccacaaca tcgaggacgg cagcgtgcag ctcgccgacc actaccagca 8940
 gaacaccccc atcggcgagc gccccgtgct gctgcccgac aaccactacc tgagcaccca 9000
 gtccgcctg agcaaagacc ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt 9060
 gaccgccgce gggatcactc tcggcatgga cgagctgtac aagtgtcgg ccgcagataa 9120
 aataaaaagt tttatttagt ctccagaaaa aggggggaat gaaagacccc acctgtaggt 9180
 ttggcaagct agcttaagta acgccatttt gcaaggcatg gaaaaataca taactgagaa 9240

tagagaagtt cagatcaagg tcaggaacag atggaacagc tgaatatggg ccaaacagga 9300

tatctgtggt aagcagttcc tgccccgct cagggccaag aacagatgga acagctgaat 9360

atgggccaaa caggatatct gtgtaagca gttcctgccc cggetcaggg ccaagaacag 9420

atggtcccca gatgcggtcc agccctcagc agtttctaga gaaccatcag atgtttccag 9480

ggtgccccaa ggacctgaaa tgacctgtg ccttatttga actaaccaat cagttcgctt 9540

ctcgtttctg ttcgcgct tctgetcccc gagctcaata aaagagcca caaccctca 9600

ctcggggcgc cagtcctccg attgactgag tcccccgggt acccgtgtat ccaataaacc 9660

ctcttgagc tgcateccag ttgtggtctc gctgttcctt gggaggtct cctctgagtg 9720

attgactacc cgtcagcggg ggtctttcat tacatgtgag caaaaggcca gcaaaaggcc 9780

aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgcc cctgacgag 9840

catcacaata atcgacgctc aagtcagagg tggcgaacc cgacaggact ataaagatac 9900

caggcgtttc cccctggaag ctccctcgtg cgctctctg ttccgaccct gccgcttacc 9960

ggatacctgt ccgctttct ccttcggga agcgtggcgc tttctcaatg ctcacgctgt 10020

aggtatctca gttcgggtga ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc 10080

gttcagcccc accgctgcgc cttatccggt aactatcgtc ttgagtcaa cccggttaaga 10140

cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 10200

ggcggtgcta cagagttctt gaagtgggtg ctaactacg gctacactag aaggacagta 10260

tttggtatct gcctctgct gaagccagtt accttcgaa aaagagttgg tagctcttga 10320

tccggcaaac aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg 10380

cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 10440

tggaacgaaa actcacgta agggattttg gtcagtagat tatcaaaaag gatcttcacc 10500

tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 10560

tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcagat ctgtctattt 10620

cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatac ggagggctta 10680

ccatctggcc ccagtgtgc aatgataccg cgagaccac gctcaccggc tccagattta 10740

tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc 10800

gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtgttc gccagttaat 10860

agtttgcgca acgttgttc cattgctgca ggcatcgtgg tgcacgctc gtcgtttggt 10920

atggcttcat tcagctccgg ttccaacga tcaaggcgag ttacatgatc ccccatgttg 10980

tgcaaaaaag cggttagctc cttcggctct ccgatcgttg tcagaagtaa gttggccgca 11040
 gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta 11100
 agatgctttt ctgtgactgg tgagtactca accaagtcac tctgagaata gtgtatgcgg 11160
 cgaccagatt gciccttgccc ggcgtcaaca cgggataata ccgcgccaca tagcagaact 11220
 ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg 11280
 ctgttgagat ccagttcgat gtaacccaact cgtgcacca actgatcttc agcatctttt 11340
 actttacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga 11400

ataagggcga cacgaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc 11460
 atttatcagg gttattgtct catgagcggga tacatatattg aatgtattta gaaaaataaa 11520
 caaatagggg ttccgcgcac atttccccga aaagtccac ctgacgtcta agaaaccatt 11580
 attatcatga cattaaccta taaaaatagg cgtatcacga ggccctttcg tcttcaagaa 11640
 ttcac 11645

<210> 46
 <211> 11654
 <212> DNA
 <213> Artificial Sequence
 <220><223> pAC3-GSG-P2A-GFPm
 <400> 46

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagtccg 60
 cgttacataa cttacggtaa atggcccgc tggtgaccg cccaacgacc cccgcccatt 120
 gacgtcaata atgacgtatg ttccatagc aacccaata gggactttcc attgacgtca 180
 atgggtggag tatttacggt aaactgcca cttggcagta catcaagtgt atcatatgcc 240
 aagtagccc cciattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 300
 catgacctta tgggactttc ctacttgca gtacatctac gtattagtca tcgtattac 360
 catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420

attccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480
 ggactttcca aaatgtcgta acaactccgc ccattgacg caaatggcg gtaggcgtgt 540
 acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgcagc cctccgattg 600
 actgagtcgc ccgggtaccg gtgtatcaa taaccctct tgcagttgca tccgacttgt 660
 ggtctcgctg ttcttggga gggctctctc tgagtgattg actaccctc agcgggggtc 720
 tttcatttgg gggctcgtcc gggatcgga gaccctgcc cagggaccac cgaccacca 780

ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840

 tgatTTTTatg cgctgcgtc ggtactagt agctaactag ctctgtatct ggcggaccg 900
 tggTggaact gacgagttcg gaacacccgg ccgcaaccct gggagacgtc ccagggactt 960
 cgggggccgt tttTgtggc cgacctgagt ccaaaaatcc cgatcgTTTT ggactctttg 1020
 gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt 1080
 tcccgcctcc gtctgaatTT ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg 1140
 tctgtctcag catcgTtctg tgtTgtctct gtctgactgt gTttctgtat ttgtctgaga 1200
 atatgggcca gactgttacc actcccttaa gTttgacctt aggtcactgg aaagatgtcg 1260

 agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgtTgggtt accttctgct 1320
 ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380
 tcatcaccca ggTtaagatc aaggTctttt cacctggccc gcatggacac ccagaccagg 1440
 tcccctacat cgtgacctgg gaagcctTgg cTtttgacc cctccctgg gtcaagccct 1500
 ttgtacacc taagcctccg cctcctcttc ctccatccgc cccgtctctc cccctgaac 1560
 ctctcgttc gacccegcct cgatctctcc tttatccage cctcactcct tetctaggcg 1620
 ccaaacctaa acctcaagtt cTttctgaca gtggggggcc gctcatcgac ctacttacag 1680

 aagaccccc gccttatagg gaccaagac caccccttc cgacagggac ggaaatggtg 1740
 gagaagcgac ccctgcggga gaggcaccgg acccctcccc aatggcatct cgcctacgtg 1800
 ggagacggga gccccctgtg gccactcca ctacctcga ggcatteccc ctccgcgag 1860
 gaggaaacgg acagcttcaa tactggcctg tctcctcttc tgacctttac aactggaaaa 1920
 ataataacc ttctTTTTct gaagatccag gTaaactgac agctctgac gagtctgttc 1980
 tcatcaccca tcagcccacc tgggacgact gtcagcagct gTtggggact ctgctgaccg 2040
 gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100

 gccccactca actgccaat gaagtcatg ccgcttttcc cctcgagcgc ccagactggg 2160
 attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagtTg ctcttagcgg 2220
 gtctcaaaaa cgcgggcaga agccccacca attTggccaa ggtaaaagga ataacacaag 2280
 ggccaatga gtctccctcg gccttcttag agagacttaa ggaagcctat cgcaggtaca 2340
 ctcttatga cctgaggac ccagggcaag aaactaatgt gtctatgtct ttcattTggc 2400
 agtctgcccc agacattggg agaaagttag agaggTtaga agatttaaaa aacaagacgc 2460
 ttggagattt ggttagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520

gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580
 agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640
 ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700
 atcgcgacca gtgtgcctac tgcaaagaaa aggggactcg ggctaaagat tgtccaaga 2760
 aaccacgagg acctcgggga ccaagacccc agacctcct cctgacctta gatgactagg 2820
 gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac 2880
 ccgtcacctt cctggtagat actggggccc aacctcctg gctgacccaa aatcctggac 2940

 ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgtgga 3000
 ccacggatcg caaagtacat ctagctaccg gtaaggcac ccactcttc ctccatgtac 3060
 cagactgtcc ctatcctctg ttaggaagag atttgctgac taaactaaaa gcccaatcc 3120
 actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgtga 3180
 ccctaaatat agaagatgag catcggttac atgagacctc aaaagagcca gatgtttctc 3240
 taggtccac atggctgtct gattttctc aggcctgggc ggaaacggg gcatgggac 3300
 tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360

 taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420
 tgttgacca ggaataactg gtaccctgcc agtccccctg gaacacgccc ctgctaccg 3480
 ttaagaaacc agggactaat gattatagc ctgtccagga tctgagagaa gtcaacaagc 3540
 ggggtggaaga catccacccc accgtgccc acccttaca cctcttgagc gggtcccac 3600
 cgtcccacca gtggtacact gtgcttgatt taaaggatgc cttttctgc ctgagactcc 3660
 accccaccag tcagcctctc ttcgcctttg agtggagaga tccagagatg ggaatctcag 3720
 gacaattgac ctggaccaga ctcccacagg gttcaaaaa cagtcccacc ctgtttgatg 3780

 aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atcctgctac 3840
 agtacgtgga tgacttactg ctggccgcca cttctgagct agactgcaa caaggactc 3900
 gggccctgtt acaaacctta gggaaacctg ggtatcgggc ctcggccaag aaagcccaaa 3960
 ttgcccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga 4020
 ctgaggccag aaaagagact gtgatggggc agcctactcc gaagaccct cgacaactaa 4080
 gggagtctct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140
 cageccccctt gtaccctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200

 aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag 4260
 atttgactaa gccctttgaa ctctttgtcg acgagaagca gggtacgcc aaaggtgtec 4320
 taacgcaaaa actgggacct tggcgtcggc cgggtggccta cctgtccaaa aagctagacc 4380

cagtagcagc tgggtggccc ccttgccctac ggatggtagc agccattgcc gtactgacaa 4440
 aggatgcagg caagctaacc atgggacagc cactagtcat tctggccccc catgcagtag 4500
 aggcactagt caaacaaccc cccgaccgct ggctttccaa cgcccggatg actcactatc 4560
 aggccttgct ttiggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaacccgg 4620

 ctacgtgct cccactgct gaggaaggc tgcaacacaa ctgccttgat atcctggccg 4680
 aagcccacgg aaccgaccc gacctaacgg accagccgct cccagacgcc gaccacacct 4740
 ggtacacgga tggaaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800
 ccaccgagac cgaggtaatc tgggctaaag ccttgccagc cgggacatcc gctcagcggg 4860
 ctgaactgat agcactcacc caggecctaa agatggcaga aggtagaag ctaaatgttt 4920
 atactgatag ccgttatgct ttgctactg ccatatcca tggagaaata tacagaaggc 4980
 gtgggttgct cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac 5040

 taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100
 gacacagcgc cgaggctaga ggcaaccgga tgctgacca agcggcccga aaggcagcca 5160
 tcacagagac tcagacacc tctacccctc tcatagaaaa tteatcaccc tacacctcag 5220
 aacattttca ttacacagtg actgatataa aggacctaac caagtgggg gccatttatg 5280
 ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340
 ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc 5400
 tcctagagag aagccacagt ccctactaca tgctgaaccg ggatcgaaca ctcaaaaata 5460

 tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg 5520
 gaactagggc ccgcccggcat cggcccggca ctattggga gatcgattc accgagataa 5580
 agcccggatt gtatggctat aaatatcttc tagtttttat agataccttt tctggctgga 5640
 tagaacctt cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg 5700
 agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760
 tctcaaggt gagtcagaca gtggccgac tgttggggat tgattggaaa ttacattgtg 5820
 catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaacctc aaggagactt 5880

 taactaaatt aacgcttgca actggctcta gagactgggt gctcctact cccttagccc 5940
 tgtaccgagc ccgcaacacg ccgggcccc atggcctcac cccatagag atcttatatg 6000
 gggcaccccc gcccttgta aacttccctg accctgacat gacaagagtt actaacagcc 6060
 cctctctcca agctcactta caggtctctt acttagtcca gcacgaagtc tggagacctc 6120
 tggcggcagc ctaccaagaa caactggacc gaccgggtgt acctcaccct taccgagteg 6180
 gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240

cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcatc gcagcttgga 6300

 tacacccgc ccacgtgaag gctgccgacc cgggggtgg accatcctct agactgacat 6360
 ggcgcttca acgctctcaa aaccacctca agataagatt aaccctgga agccctaat 6420
 agtcatggga gtctgttag gtagggat ggcagagagc cccatcagg tctttaatgt 6480
 aacctggaga gtcaccaacc tgatgactgg gctaccgcc aatgccacct cctcctggg 6540
 aactgtaca gatgcctcc caaaattata tttgatcta tgtgatctgg tcggagagga 6600
 gtgggacct tcagaccagg aaccgatgt cgggatggc tgcaagtacc ccgaggag 6660
 acagcggacc cggactttg acttttacgt gtccctggg cataccgtaa agtcggggtg 6720

 tggggacca ggagaggct acttggtaa atgggggtgt gaaaccaccg gacaggctta 6780
 ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcgtaaca cccctgga 6840
 cagggatgc tctaaagtg cctgtggccc ctgctacgac ctctcaaag tatccaatc 6900
 ctccaagg gctactcag gggcagatg caaccctca gtcctagaat tctactgatgc 6960
 aggaaaaag gtaactggg acgggccc aaatcgtggga ctgagactgt accggacagg 7020
 aacagatcct attaccatgt tctcctgac ccggcaggtc cttaatgtgg gaccccgagt 7080
 cccataggg ccaaccag tattaccga ccaaagactc cttcctcac caatagat 7140

 tgtaccggt ccacagccac ctageccct caataccagt taccacctt cactaccag 7200
 tacacctca acctcccta caagccaag tgcaccag ccacccag gaactggaga 7260
 tagactacta gctctagta aaggacctc tcaggcgtt aacctacca atcccacaa 7320
 gaccaagaa tgttggctgt gcttagtgc gggacctct tattacgaag gagtagcgt 7380
 cgtgggact tataccaatc attccaccg tccggccaac tgtaccgcca ctccaaca 7440
 taagcttacc ctatctgaag tgacaggaca gggctatgc atggggcag tacctaaac 7500
 tcaccaggc ttatgtaaca ccaccaag cgcggtca ggatcctact acctgcagc 7560

 acccgcgga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccaggt 7620
 gctcaatcta accacagatt attgtgatt agtgaactc tggcccagag taatttacc 7680
 ctccccgat tatatgatg gtcagcttga acagctacc aaatataaaa gagagccagt 7740
 atcattgacc ctggccttc tactaggagg attaacatg ggaggattg cagctggaat 7800
 agggacggg accactgct taattaaaac ccagcagttt gagcagctc atgccctat 7860
 ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920
 gttgtctgaa gtagtctac agaaccgag aggcctagat ttgctattcc taaaggagg 7980

aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt 8040
 gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100
 aggccaagga tggttcgaag ggctgtttta tagatcccc tggtttacca ccttaatctc 8160
 caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct 8220
 caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac 8280
 tcagcaatat caccagctaa aacctataga gtacgagcca ggaagcggag ctactaactt 8340
 cagcctgctg aagcaggctg gagacgtgga ggagaacct ggacctggcg cgcctatggc 8400

 cagcaagggc gaggagctgt tcaccggggt ggtgccatc ctggtcgagc tggacggcga 8460
 cgtaaaccgc cacaagttca gcgtgtccgg cgaaggagag ggcgatgcca cctacggcaa 8520
 gctgaccctg aagttcactc gcaccaccgg caagctgccc gtgcctggc ccaccctcgt 8580
 gaccacctg acctacggcg tgcagtctt cggccctac cccgaccaca tgaagcagca 8640
 cgacttcttc aagtccgcca tggcccgaag ctacgtccag gagcgacca tcttcttcaa 8700
 ggacgacggc aactacaaga cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa 8760
 ccgcatcgag ctgaagggca tcgacttcaa ggaggacggc aacatcctgg ggcacaagct 8820

 ggagtacaac tacaacagcc acaaggtcta taccaccgc gacaagcaga agaacggcat 8880
 caaggtgaac ttcaagacc gccacaacat cgaggacggc agcgtgcagc tcgccgacca 8940
 ctaccagcag aacaccccca tcggcgacgg ccccgctgct ctgcccgaca accactacct 9000
 gagcaccag tccgccctga gcaaagacc caacgagaag cgcgatcaca tggctctgct 9060
 ggagttcgtg accgccgccc ggatcactct cggcatggac gagctgtaca agtgtcggc 9120
 cgagataaa ataaaagatt ttatttagtc tccagaaaaa gggggggaatg aaagaccca 9180
 cctgtaggtt tggcaagcta gcttaagtaa cgccatcttg caaggcatgg aaaatacat 9240

 aactgagaat agagaagttc agatcaaggt caggaacaga tggaacagct gaatatgggc 9300
 caaacaggat atctgtggtg agcagttcct gccccggctc agggccaaga acagatggaa 9360
 cagctgaata tgggccaac aggatatctg tggtaagcag ttcttcccc ggctcagggc 9420
 caagaacaga tggccccag atgcggtcca gcctcagca gtttctagag aacctcaga 9480
 tgtttccagg gtgccccaa gacctgaaat gaccctgtgc cttatttgaa ctaaccaatc 9540
 agttcgcttc tcgcttctgt tcgctgctt ctgctcccc agctcaataa aagagccac 9600
 aaccctcac tcggggcgcc agtctctcga ttgactgagt cgccccggta cccgtgtatc 9660

 caataaacc tcttgcagtt gcatccgact tgtggtctcg ctgttcttg ggagggtctc 9720
 ctctgagtga ttgactacc gtcagcgggg gtctttcatt acatgtgagc aaaaggccag 9780
 caaaaggcca ggaaccgtaa aaagcccgcg ttgctggcgt ttttccatag gctccgcccc 9840

cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc gacaggacta 9900
 taaagatacc aggcgtttcc ccctggaagc tcctctgtgc gctctcctgt tccgaccctg 9960
 ccgcttaccg gatacctgtc cgcttttctc ccttcgggaa gcgtggcgct ttctcaatgc 10020
 tcacgctgta ggiatctcag ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac 10080

gaaccccccg ttcagcccga ccgctgcgcc ttatccgta actatcgtct tgagtccaac 10140
 ccggtaaagc acgacttacc gccactggca gcagccactg gtaacaggat tagcagagcg 10200
 aggtatgtag gcggtgctac agagtctctg aagtggggc ctaactacgg ctacactaga 10260
 aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagttggt 10320
 agctcttgat ccggcaaaaa aaccaccgct gtagcgggtg gttttttgt ttgcaagcag 10380
 cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct 10440
 gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg 10500

atcttcacct agatcctttt aaattaaana tgaagtttta aatcaatcta aagtatata 10560
 gagtaaacct ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc 10620
 tgtctatitc gttcatccat agttgcctga ctccccgtcg ttagataac tacgatacgg 10680
 gagggcttac catctggccc cagtgtcga atgataccgc gagaccacg ctaccggct 10740
 ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag tggctctgca 10800
 actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg 10860
 ccagttaata gtttgcgcaa cgtttgtgcc attgctgcag gcatcgtggt gtacgctcgc 10920

tcgtttggta tggttcatt cagctccggt tccaacgat caaggcagat tacatgatec 10980
 cccatgttgt gcaaaaaagc ggttagctcc ttcggtctc cgatcgttgt cagaagtaag 11040
 ttggcccgag tgtatcact catggttatg gcagcactgc ataattctct tactgtcatg 11100
 ccatccgtaa gatgcttttc tgtgactggt gactactcaa ccaagtcatt ctgagaatag 11160
 tgtatcggc gaccgagttg ctcttgcccg gcgtcaacac gggataatac cgcgccacat 11220
 agcagaactt taaaagtgt catcattgga aaacgttctt cggggcgaaa actctcaagg 11280
 atcttaccgc tgttgagatc cagttcgatg taaccacctc gtgcacccaa ctgatcttea 11340

gcatctttta cttcaccag cgtttctggg tgagcaaaaa caggaaggca aatgccgca 11400
 aaaaaggaa taagggcgac acggaatgt tgaatactca tactcttct tttcaatat 11460
 tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag 11520
 aaaaataaac aataggggt tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa 11580
 gaaaccatta ttatcatgac attaacctat aaaaatagge gtatcacgag gcctttcgt 11640
 cttcaagaat tcat 11654

<210> 47

<211> 11648

<212> DNA

<213> Artificial Sequence

<220><223> pAC3-E2A-GFP

<400> 47

```

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg      60
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt      120
gacgtcaata atgacgtatg ttccatagt aacccaata gggactttcc attgacgtca      180
atgggtggag tatttacggt aaactgccc cttggcagta catcaagtgt atcatatgcc      240
aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta      300
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac      360

catggtgatg cggttttggc agtacatcaa tggcgtgga tagcggtttg actcacgggg      420
atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg      480
ggactttcca aaatgtcgt acaactccgc ccattgacg caaatgggcg gtaggcgtgt      540
acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg      600
actgagtgcg cgggtaccg gtgtatccaa taaacctct tgcagttgca tccgacttgt      660
ggtctcgctg ttccttggga gggctcctc tgagtattg actaccctc agcggggggtc      720
tttcatttgg gggctcgtcc gggatcggga gaccctgcc cagggaccac cgaccacca      780

cgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac      840
tgattttatg cgctgcgct ggtactagt agctaactag ctctgtatct ggcggaccgg      900
tggtggaact gacgagttcg gaacacccgg ccgcaacct gggagacgtc ccagggactt      960
cgggggccgt ttttgtggcc cgacctgagt caaaaatcc cgatcgtttt ggactctttg      1020
gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt      1080
tcccgcctcc gtctgaattt ttgcttccgg tttgggaccg aagccgcgcc gcgcgtcttg      1140
tctgctgcag catcgttctg tgttctctct gtctgactgt gtttctgtat ttgtctgaga      1200

atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg      1260
agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct      1320
ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc      1380
tcatcaccca ggtaagatc aaggctcttt cacctggccc gcatggacac ccagaccagg      1440
tcccctacat cgtgacctgg gaagccttgg cttttgacce cctccctgg gtcaagccct      1500

```

ttgtacacc taagcctccg cctcctcttc ctccatccgc cccgtctctc cccttgaac 1560
 ctctcgttc gaccccgct cgatctccc tttatccagc cctcactcct tctctaggcg 1620

 ccaaacctaa acctcaagtt ctttctgaca gtggggggcc gctcatcgac ctacttacag 1680
 aagaccccc gccttatagg gaccaagac cacccttc cgacaggac ggaaatggtg 1740
 gagaagcgac cctgcgga gaggcaccg accctccc aatggcatct cgctacgtg 1800
 ggagacggga gccccctgtg gccgacteca ctacctgca ggcatcccc ctccgcgag 1860
 gaggaacgg acagcttcaa tactggccgt tctcctcttc tgaccttac aactggaaaa 1920
 ataataacce ttcttttct gaagatccag gtaaacctgac agctctgac gagtctgttc 1980
 tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040

 gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100
 gccccactca actgcccact gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg 2160
 attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagtgt ctcttagcgg 2220
 gtctcaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280
 ggccaatga gtctccctg gccttcctag agagacttaa ggaagcctat cgaggtaca 2340
 ctcttatga ccctgaggac ccaggcaag aaactaatgt gtctatgtct ttcatttggc 2400
 agtctgcccc agacattggg agaaagttag agaggtaga agatataaaa aacaagacgc 2460

 ttggagattt ggtagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520
 gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580
 agcagaaaga gaaagaaaga gatcgttaga gacatagaga gatgagcaag ctattggcca 2640
 ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700
 atcgcgacca gtgtgctac tgcaaagaaa aggggactg ggctaaagat tgtccaaga 2760
 aaccacgagg acctcgggga ccaagacccc agacctcct cctgacctc gatgactagg 2820
 gaggtcaggg tcaggagccc ccccctgaac ccaggataac cctcaaagtc ggggggcaac 2880

 ccgtcacctt cctgtagat actggggccc aacctcctg gctgaccaa aatcctggac 2940
 ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcg tatcgtgga 3000
 ccacggatcg caaagtacat ctactaccg gtaaggtcac ccactcttc ctccatgtac 3060
 cagactgtcc ctatctctg ttaggaagag atttctgac taaactaaaa gcccaaatcc 3120
 actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgtga 3180
 ccctaataat agaagatgag catcggtac atgagacctc aaaagagcca gatgtttctc 3240

tagggtccac atggctgtct gattttcctc aggcctgggc ggaaaccggg ggcattgggac 3300

tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360

taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420

tgttggacca gggaaactg gtaccctgcc agtccccctg gaacacgccc ctgctacccg 3480

ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc 3540

gggtggaaga catccacccc accgtgceca acccttacia cctcttgagc gggtccccac 3600

cgtcccacca gtggtacact gtgcttgatt taaaggatgc cttttctgc ctgagactcc 3660

accccaccag tcagcctctc ttgcctttg agtggagaga tccagagatg ggaatctcag 3720

gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780

aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atcctgctac 3840

agtacgtgga tgacttactg ctggccgcca ctctgagct agactgcaa caaggtactc 3900

gggccctgtt acaaaccta gggaaacctc ggtatcgggc ctcgccaag aaagccaaa 3960

tttgccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga 4020

ctgagggcag aaaagagact gtgatggggc agcctactcc gaagaccct cgacaactaa 4080

gggagttcct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140

cagccccctt gtaccctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200

aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag 4260

atgtgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc 4320

taacgcaaaa actgggacct tggcgtcggc cggtggccta cctgtccaaa aagctagacc 4380

cagtagcagc tgggtggccc ccttgccctac ggatggtagc agccattgcc gtactgacaa 4440

aggatgcagg caagctaacc atgggacagc cactagtcac tctggcccc catgcagtag 4500

aggcactagt caaacaacc cccgaccgt ggttttccaa cgcccggatg actcactatc 4560

aggccttgct ttggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaaccggg 4620

ctacgtctgt cccactgect gaggaagggc tgcaacacaa ctgccttgat atcctggccg 4680

aagcccacgg aaccgaccc gacctaacgg accagccgt cccagacgcc gaccacacct 4740

ggtacacgga tggaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800

ccaccgagac cgaggtaatc tgggctaaag cctgccagc cgggacatcc gctcagcggg 4860

ctgaactgat agcactcacc caggccctaa agatggcaga aggtagaag ctaaatgttt 4920

atactgatag ccgttatgct tttgctactg cccatatcca tggagaaata tacagaaggc 4980

gtgggttgcT cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac 5040
 taaaagccct ctttctgcc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100
 gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggcccga aaggcagcca 5160
 tcacagagac tccagacacc tctaccctcc tcatagaaaa ttcacacacc tacacctcag 5220
 aacattttca ttacacagtg actgatataa aggacctaac caagttaggg gccatttatg 5280
 ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340
 ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc 5400

tcctagagag aagccacagt ccctactaca tgctgaaccg ggatcgaaca ctcaaaaata 5460
 tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg 5520
 gaactagggt ccgctggcat cggcccggca ctattggga gatcgatttc accgagataa 5580
 agcccggatt gtatggctat aaatatcttc tagtttttat agataccttt tctggctgga 5640
 tagaacctt cccaaccaag aaagaaccg ccaaggtcgt aaccaagaag ctactagagg 5700
 agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760
 tctccaaggt gagtcagaca gtggccgatc tgtggggat tgattggaaa ttacattgtg 5820

catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaacctc aaggagactt 5880
 taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc 5940
 tgtaccgagc cgcaacacg ccgggcccc atggcctcac cccatagag atcttatatg 6000
 gggcaccccc gcccttgta aacttccctg accctgacat gacaagagtt actaacagcc 6060
 cctctctcca agctcactta caggtctctt acttagtcca gcacgaagtc tggagacctc 6120
 tggcggcagc ctaccaagaa caactggacc gaccgggtgt acctaccct taccgagtcg 6180
 gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240

cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcatc gcagcttggg 6300
 tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat 6360
 ggccgttca acgctctca aacccctca agataagatt aaccctgga agcccttaat 6420
 agtcatggga gtctgttag gagtagggat ggcagagagc ccccatcagg tetttaatgt 6480
 aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct cctcctggg 6540
 aactgtacaa gatgccttcc caaaattata tttgatcta tgtgatctgg tggagagga 6600
 gtgggacctt tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgagggag 6660

acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcgggggtg 6720
 tgggggacca ggagagggt actgtggtaa atgggggtgt gaaaccaccg gacaggctta 6780
 ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcggtaaca cccctggga 6840

cacgggatgc tctaaagtgg cctgtggccc ctgctacgac ctctccaaag tatccaattc 6900
cttccaaggg gctactcgag ggggcagatg caaccctcta gtcctagaat tctactgatgc 6960
aggaaaaaag gctaactggg acgggcccac atcgtgggga ctgagactgt accggacagg 7020
aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gacccccgagt 7080

ccccataggg cccaaccag tattaccga ccaagactc ccttcctcac caatagagat 7140
tgtaccgct ccacagccac ctageccct caataccagt taccctctt cactaccag 7200
tacacctca acctccccta caagtccaag tgtcccacag ccacccccag gaactggaga 7260
tagactacta gctctagtca aaggagccta tcaggcgctt aacctacca atcccacaa 7320
gaccaagaa tgttggctgt gcttagtgc gggacctctt tattacgaag gagtagcgg 7380
cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca ctccaaca 7440
taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac 7500

tcaccaggcc ttatgtaaca ccaccaaaag cgccggctca ggatcctact acctgcagc 7560
acccgccgga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccaggt 7620
gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taatttacca 7680
ctccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt 7740
atcattgacc ctggcccttc tactaggagg attaaccatg ggagggatg cagctggaat 7800
agggacgggg accactgct taattaaac ccagcagttt gagcagcttc atgcccetat 7860
ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920

gttgtctgaa gtagtctac agaaccgcag aggcctagat ttgctattcc taaaggaggg 7980
aggctctctg gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt 8040
gagagacagc atggccaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100
aggccaagga tggttcgaag ggctgtttaa tagatcccc tggtttacca cettaatctc 8160
caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct 8220
caatcgattg gtccaattg ttaaagacag gatctcagtg gtccaggctc tggttttgac 8280
tcagcaatat caccagctaa aacctataga gtacgagcca cagtgtacta attatgctct 8340

cttgaattg gctggagatg ttgagagcaa cctggacct ggcgcgcta tggccagcaa 8400
gggcgaggag ctgttcaccg ggggtgtgcc catcctggtc gagctggacg gcgacgtaaa 8460
cggccacaag ttcagcgtgt ccggcgaagg agagggcgat gccacctacg gcaagctgac 8520
cctgaagttc atctgacca ccggcaagct gccctgccc tggcccacc tcgtgaccac 8580
cttgacctac ggcgtgcagt gcttcgccc ctaccccac cacatgaagc agcagactt 8640
cttcaagtcc gccatgcccg aaggctacgt ccaggagcgc accatcttct tcaaggacga 8700

cggcaactac aagaccccgcg ccgaggtgaa gttcgagggc gacaccctgg tgaaccgcat 8760

 cgagctgaag ggcatcgact tcaaggagga cggcaacatc ctggggcaca agctggagta 8820
 caactacaac agccacaagg tctatatac cgccgacaag cagaagaacg gcatcaaggt 8880
 gaacttcaag acccgccaca acatcgagga cggcagcgtg cagctcgccg accactacca 8940
 gcagaacacc cccatcgcg acggccccgt gctgctgccc gacaaccact acctgagcac 9000
 ccagtccgcc ctgagcaaaag accccaacga gaagcgcgat cacatggtcc tgctggagtt 9060
 cgtgaccgcc gccgggatca ctctcggcat ggacgagctg tacaagtgtg cggcccaga 9120
 taaaataaaa gatTTTTttt agtctccaga aaaagggggg aatgaaagac cccacctgta 9180

 ggTTTTgcaa gctagcttaa gtaacgcat tttgcaaggc atggaaaaat acataactga 9240
 gaatagagaa gttcagatca aggtcaggaa cagatggaac agctgaatat gggccaaca 9300
 ggatactgtt ggtaagcagt tcttccccg gctcagggcc aagaacagat ggaacagctg 9360
 aatatgggcc aaacaggata tctgtggtaa gcagttctg cccccgctca gggccaagaa 9420
 cagatggtcc ccagatcgcg tccagccctc agcagtttct agagaacat cagatgttcc 9480
 caggtgcccc caagacctg aaatgacct gtgccttatt tgaactaacc aatcagttcg 9540
 cttctcgctt ctgttcgcg gcttctgctc cccgagctca ataaaagagc ccacaacccc 9600

 tcaactgggg cgccagtct ccgattgact gagtgcgccg ggtaccctg tatccaataa 9660
 accctcttgc agttgcatcc gacttgtggt ctgcgtgttc cttgggaggg tctcctctga 9720
 gtgattgact acccgtcagc ggggtcttt cattacatgt gagcaaaagg ccagcaaaag 9780
 gccaggaacc gtaaaaaggc cgcgttctg gcgtttttcc ataggctccg cccccctgac 9840
 gagcatcaca aaaatcgacg ctcaagttag aggtggcgaa acccgacagg actataaaga 9900
 taccaggcgt ttccccctg aagctccctc gtgcgctctc ctgttccgac cctgccgctt 9960
 accggatacc tgtccgctt tctcccttcg ggaagcgtgg cgctttctca atgctcacgc 10020

 tgtaggtatc tcagttcggg gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc 10080
 cccgttcagc ccgaccctg cgcttatcc ggtaactatc gtcttgagtc caaccggta 10140
 agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggtat 10200
 gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac tagaaggaca 10260
 gtatttggtt tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct 10320
 tgatccgca acaaaaccac cgctggtagc ggtggTTTTt ttgtttgcaa gcagcagatt 10380
 acgcgcagaa aaaaaggatc tcaagaagat cttttagctt tttctacggg gtctgacgct 10440

cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa aaggatcttc 10500
 acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa 10560
 acttggctcg acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta 10620
 tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc 10680
 ttaccatctg gccccagtcg tgcaatgata ccgcgagacc cacgctcacc ggctccagat 10740
 ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtggctc tgcaacttta 10800
 tccgcctcca tccagtctat taattgttgc cgggaagcta gagliaagtag ttcgccagtt 10860

aatagttgc gcaacgttgt tgccattgct gcaggcatcg tgggtgcacg ctgctgttt 10920
 ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg atccccatg 10980
 ttgtgcaaaa aagcggttag ctcttcggt cctccgatcg ttgtcagaag taagttggcc 11040
 gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc 11100
 gtaagatgct tttctgtgac tgggtgagtac tcaaccaagt cattctgaga atagtgtatg 11160
 cggcgaccga gttgctcttg cccggcgtca acacgggata ataccgcgc acatagcaga 11220
 actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta 11280

ccgctgttga gatccagttc gatgtaacc actcgtgcac ccaactgatc ttcagcatct 11340
 tttactttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaag 11400
 ggaataaggg cgacacggaa atgttgaata ctcatactct tcctttttca atattatga 11460
 agcatttate agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat 11520
 aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc 11580
 attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa 11640
 gaattcat 11648

<210> 48

<211> 11657

<212> DNA

<213> Artificial Sequence

<220><223> pAC3-GSG-E2A-GFPm

<400> 48

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg 60
 cgttacataa cttacggtaa atggcccgcc tggtgaccg cccaacgacc cccgccatt 120
 gacgtcaata atgacgtatg ttccatagc aacccaata gggactttcc attgacgtca 180
 atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 240

aagtacgccc cciattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 300

catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 360
catggtgatg cggttttggc agtacaatcaa tgggcgtgga tagcggtttg actcacgggg 420
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480
ggactttcca aaatgtcgtg acaactccgc ccatttgacg caaatgggcg gtaggcgtgt 540
acggtgggag gtctataata gcagagctgg tttagtgaac cggcgccagt cctccgattg 600
actgagtcgc ccgggtaccg gtgtatcaa taaacctct tgcagttgca tccgacttgt 660
ggtctcgtg ttccttggga gggctctctc tgagtattg actaccctc agcgggggtc 720

tttcatttgg gggctcgtcc gggatcggga gaccctgcc cagggaccac cgaccacca 780
ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840
tgattttatg cgctcgtc ggtaactagt agctaactag ctctgtatct ggcggacccg 900
tgggtgaact gacgagttcg gaacaccccg ccgaacctt gggagacgtc ccagggactt 960
cgggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg 1020
gtgcaccccc cttagaggag ggatatgtgg tttcgttagg agacgagaac ctaaaacagt 1080
tcccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg 1140

tctgctgcag catcgttctg tgttctctct gtctgactgt gtttctgtat ttgtctgaga 1200
atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg 1260
agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct 1320
ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380
tcatcaccca ggtaagatc aaggcttttt cacctggccc gcatggacac ccagaccagg 1440
tcccctacat cgtgacctgg gaagccttgg cttttgacce cctccctgg gtcaagcctt 1500
ttgtacacce taagcctccg cctcctcttc ctccatcgc cccgtctctc ccccttgaac 1560

ctcctcgttc gaccccgctt cgatcctccc tttatccagc cctcactcct tctctaggcg 1620
ccaaacctaa acctcaagtt ctttctgaca gtggggggcc gctcatcgac ctacttacag 1680
aagaccccc gccttatagg gaccaagac caccctctc cgacagggac ggaaatggtg 1740
gagaagcgac cctcggggga gaggcaccgg acccctccc aatggcatct cgctacgtg 1800
ggagacggga gcccctgtg gccactcca ctacctgca ggcatcccc ctccgcgcag 1860
gaggaaacgg acagcttcaa tactggccgt tctcctcttc tgacctttac aactggaaaa 1920
ataataacce ttcttttct gaagatccag gtaaactgac agctctgatc gagtctgttc 1980

tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040
 gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100
 gccccactca actgccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg 2160
 attacaccac ccaggcaggt aggaaccacc tagtccacta tgcccagtig ctcttagcgg 2220
 gtctcaaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280
 ggccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaggtaca 2340
 ctcttatga ccctgaggac ccagggcaag aaactaatgt gtctatgtct ttcatttggc 2400

 agtctgcccc agacattggg agaaagttag agaggtaga agatttaaaa aacaagacgc 2460
 ttggagattt ggttagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520
 gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580
 agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640
 ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700
 atcgcgacca gtgtgcctac tgcaaagaaa aggggactg ggctaaagat tgtccaaga 2760
 aaccacgagg acctcgggga ccaagacccc agacctcct cctgacccta gatgactagg 2820

 gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac 2880
 ccgtcacctt cctggtagat actggggccc aacctccgt gctgaccaa aatcctggac 2940
 ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcg tatcgtgga 3000
 ccacggatcg caaagtacat ctagctaccg gtaaggtcac ccactcttc ctccatgtac 3060
 cagactgtcc ctatcctctg ttaggaagag atttctgac taaactaaaa gcccaaatcc 3120
 actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga 3180
 ccctaatat agaagatgag catcggtac atgagacctc aaaagagcca gatgtttctc 3240

 taggttccac atggctgtct gattttctc aggcctgggc ggaaaccggg ggcatgggac 3300
 tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360
 taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420
 tgttggacca ggaatactg gtaccctgcc agtcccctg gaacacgccc ctgctacceg 3480
 ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc 3540
 ggggtggaaga catcccccc accgtgcca acccttaca cctcttgagc gggtccccac 3600
 cgteccacca gtggtacact gtgcttgatt taaaggatgc ctttttctgc ctgagactcc 3660

 accccaccag tcagcctctc ttcgcctttg agtggagaga tccagagatg ggaatctcag 3720
 gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780
 aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atcctgctac 3840

agtacgtgga tgacttactg ctggccgcca ctctctgagct agactgccaa caaggtactc 3900
 gggccctgtt acaaaccta gggaacctcg ggtatcgggc ctcgccaag aaagccaaa 3960
 ttgcccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga 4020
 ctgaggccag aaaagagact gtgatggggc agcctactcc gaagaccct cgacaactaa 4080

 gggagtctct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140
 cagccccctt gtacctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200
 aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag 4260
 atttgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc 4320
 taacgcaaaa actgggacct tggcgtcggc cgggtggccta cctgtccaaa aagctagacc 4380
 cagtagcagc tgggtggccc ccttgcttac ggatggtagc agccattgcc gtactgaaa 4440
 aggatgcagg caagctaacc atgggacagc cactagtcat tctggccccc catgcagtag 4500

 aggcactagt caacaaccc cccgaccgt ggctttccaa cgcccggatg actcactatc 4560
 aggccttgc tttggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaaccgg 4620
 ctacgtgct cccactgct gaggaagggc tgcaacacaa ctgccttgat atcctggcgc 4680
 aagcccacgg aaccgacc gacctaacgg accagccgt cccagacgcc gaccacact 4740
 ggtacacgga tgaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800
 ccaccgagac cgagtaatc tgggctaaag cctgccagc cgggacatcc gctcagcggg 4860
 ctgaactgat agcactacc caggcctaa agatggcaga aggtagaag ctaaatgttt 4920

 atactgatag ccgttatgt tttgctactg cccatatcca tggagaaata tacagaaggc 4980
 gtgggttgc cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac 5040
 taaaagccct ctttctgcc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100
 gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggcccga aaggcagcca 5160
 tcacagagac tcagacacc tctacctc tcataaaaa tcatcacc tacacctcag 5220
 aacatttca ttacacagt actgatataa aggacctaac caagttgggg gccatttatg 5280
 ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340

 ttgaattatt agactttct catcagctga ctacctcag cttctcaaaa atgaaggctc 5400
 tcctagagag aagccacagt ccctactaca tctgaaccg ggatcgaaca ctcaaaaata 5460
 tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaacagg 5520
 gaactagggt ccgcccgcac cggcccggca ctattggga gatcgattc accgagataa 5580
 agcccggatt gtatggetat aaatatctt tagttttat agatacttt tctggctgga 5640
 tagaagcctt cccaaccaag aaagaaccc ccaaggtcgt aaccaagaag ctactagagg 5700

agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760

 tctccaaggt gagtcagaca gtggccgata tgttgggat tgattggaaa ttacattgtg 5820
 catacagacc ccaaagetca ggccaggtag aaagaatgaa tagaacctac aaggagactt 5880
 taactaaatt aacgcttgca actggctcia gagactgggt gctcctactc cccttagccc 5940
 tgtaccgagc ccgcaacacg ccgggcccc atggcctcac cccatgatgag atcttatatg 6000
 gggcaccccc gcccttgta aacttccctg acctgacat gacaagagt actaacagcc 6060
 cctctctcca agctcactta caggctctct acttagtcca gcacgaagtc tggagacctc 6120
 tggcggcagc ctaccaagaa caactggacc gaccgggtgt acctcacct taccgagtcg 6180

 gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240
 cttacacagt cctgctgacc acccccaccg cctcaaagt agacggcatc gcagcttggg 6300
 tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat 6360
 ggcgcttca acgctctcaa aaccctca agataagatt aaccctgga agccctaat 6420
 agtcatggga gtctgttag gagtaggat ggcagagagc cccatcagg tctttaatgt 6480
 aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct cctcctggg 6540
 aactgtaca gatgccttcc caaaattata tttgatcta tgtgatctgg tggagagga 6600

 gtgggacct tcagaccagg aaccgatgt cgggtatggc tgcaagtacc ccgagggag 6660
 acagcggacc cggactttg acttttacgt gtgccctggg cataccgtaa agtcggggtg 6720
 tgggggacca ggagaggct acttgtgtaa atgggggtgt gaaaccaccg gacagctta 6780
 ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcgtaaca cccctggga 6840
 cacgggatgc tctaaagtg cctgtggccc ctgctacgac ctctccaaag tatccaattc 6900
 cttcaaggg gctactcgag ggggcagatg caacctcta gtctagaat tactgatgc 6960
 aggaaaaaag gctaactggg acgggcccc atcgtgggga ctgagactgt accggacagg 7020

 aacagatcct attaccatgt tctcctgac ccggcaggtc cttaatgtgg gaccccgagt 7080
 cccataggg cccaaccgag tattaccga ccaaagactc cttcctcac caatagagat 7140
 tgtaccgct ccacagccac ctageccct caataccagt taccctt cactaccag 7200
 tacacctca acctcccta caagtccaag tgtccacag ccacccag gaactggaga 7260
 tagactacta gctctagta aaggagcta tcaggcctt aacctacca atcccacaa 7320
 gaccaagaa tgttggctgt gcttagtgc gggacctct tattacgaag gagtagcgg 7380
 cgtgggcact tataccaatc attcaccgc tccggccaac tgtacggcca ctccaaca 7440

taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac 7500
 tcaccaggcc ttatgtaaca ccacccaaag cgccggctca ggatcctact accttgcagc 7560
 acccgccgga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccacggt 7620
 gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taatttacca 7680
 ctccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt 7740
 atcattgacc ctggcccttc tactaggagg attaacatg ggagggatg cagctggaat 7800
 agggacgggg accactgcct taattaaaac ccagcagttt gagcagcttc atgccgctat 7860

 ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920
 gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgctattcc taaaggaggg 7980
 aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt 8040
 gagagacagc atggccaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100
 aggccaagga tggttcgaag ggctgtttaa tagatcccc tggtttacca ccttaatctc 8160
 cacatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct 8220
 caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac 8280

 tcagcaatat caccagctaa aacccataga gtacgagcca ggaagcggac agtgtactaa 8340
 ttatgctctc ttgaaattgg ctggagatgt tgagagcaac cctggacctg gcgcgcctat 8400
 ggccagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg agctggacgg 8460
 cgacgtaaac ggccacaagt tcagcgtgtc cggcgaagga gagggcgatg ccacctacgg 8520
 caagctgacc ctgaagtcca tctgcaccac cggcaagctg cccgtgccct ggccccacct 8580
 cgtgaccacc ttgacctacg gcgtgcagtg cttcgcccgc taccgccacc acatgaagca 8640
 gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca ccatcttctt 8700

 caaggacgac ggcaactaca agaccgcgc cgaggtgaag ttcgagggcg acacctggt 8760
 gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc tggggcacia 8820
 gctggagtac aactacaaca gccacaaggt ctatataacc gccacaagc agaagaacgg 8880
 catcaaggtg aacttcaaga cccgccacia catcgaggac ggcagcgtgc agctcgccga 8940
 cactaccag cagaaccccc ccatcggcga cggccccgtg ctgctgcccg acaaccacta 9000
 cctgagcacc cagtccgcc tgagcaaaga cccaacgag aagcgcgatc acatggtcct 9060
 gctggagttc gtgaccgccg ccgggatcac tctcgcatg gacgagctgt acaagtgtgc 9120

 ggccgcagat aaaataaaag attttattta gtctccagaa aaagggggga atgaaagacc 9180
 ccacctgtag gtttggcaag ctagettaag taacgccatt ttgcaaggca tggaaaaata 9240
 cataactgag aatagagaag ttcagatcaa ggtcaggaac agatggaaca gctgaatatg 9300

ggccaaacag gatatctgtg gtaagcagtt cctgccccgg ctcagggcca agaacagatg 9360
 gaacagctga atatgggcca aacaggatat ctgtggaag cagttcctgc cccggtcag 9420
 ggccaagaac agatggtccc cagatgcggt ccagccctca gcagtttcta gagaaccatc 9480
 agatgtttcc aggggtgcccc aaggacctga aatgacctg tgccttattt gaactaacca 9540

 atcagttcgc ttctcgcttc tgttcgctgc cttctgctcc ccgagctcaa taaaagagcc 9600
 cacaaccctt cactcggggc gccagtcctc cgattgactg agtcgcccgg gtaccctgtt 9660
 atccaataaa ccctcttgca gttgcatccg acttgtggtc tcgctgttcc ttgggagggt 9720
 ctctctgag tgattgacta cccgtcagcg ggggtctttc attacatgtg agcaaaaggc 9780
 cagcaaaagg ccaggaaccg taaaagagcc gcgttgctgg cgtttttcca taggctccgc 9840
 ccccctgacg agcatcaca aaatcgacgc tcaagtcaga ggtggcgaac cccgacagga 9900
 ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttccgacc 9960

 ctgccctta ccgataact gtccgccttt ctccctcgg gaagcgtggc gctttctcaa 10020
 tgctcacgct gtaggtatct cagttcgggt taggtcgttc gtcceaagct gggctgtgtg 10080
 cacgaacccc cgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtec 10140
 aaccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga 10200
 gcgaggtatg taggcggtgc tacagagttc ttgaagtggg ggcctaacta cggtacact 10260
 agaaggacag tatttggat ctgcgctctg ctgaagccag ttacctcgg aaaaagagtt 10320
 ggtagctctt gatccggcaa acaaacacc gctggtagcg gtggtttttt tgtttgcaag 10380

 cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg 10440
 tctgacgctc agtggaaaca aaactcacgt taagggatct tggcatgag attatcaaaa 10500
 aggatcttca cctagatcct tttaaatata aaatgaagtt ttaatcaat ctaaagtata 10560
 tatgagtaaa cttagctga cagttaccaa tcttaataca gtgaggcacc tatctcagcg 10620
 atctgtctat ttcgttcac catagttgcc tgactccccg tcgtgtagat aactacgata 10680
 cgggagggtt taccatctgg ccccagtgct gcaatgatac cgcgagacc acgtcaccg 10740
 gctccagatt tctcagcaat aaaccageca gccggaaggg ccgagcgcag aagtgtcct 10800

 gcaactttat ccgctccat ccagctatc aattgttgc gggaagctag agtaagtagt 10860
 tcgccagtta atagtttgc caacgttgt gccattgtg caggcatcgt ggtgtcacgc 10920
 tcgtcgtttg gtatggcttc attcagctcc ggttccaac gatcaaggcg agttacatga 10980
 tccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt 11040
 aagttggcgg cagttttatc actcatggtt atggcagcac tgcaaatc tcttactgtc 11100
 atgccatccg taagatgctt ttctgtgact ggtgagtagt caaccaagtc attctgagaa 11160

tagtgtatgc ggcgaccgag ttgctcttgc ccggcgctcaa cacgggataa taccgcgcca 11220

catagcagaa ctttaaaagt gctcatcatt ggaaaacggt cttcggggcg aaaactctca 11280

aggatcttac cgtggttgag atccagttcg atgtaacca ctctgtcacc caactgatct 11340

tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc 11400

gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt cttttttcaa 11460

tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 11520

tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc 11580

taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt 11640

cgctctcaag aattcat 11657

<210> 49

<211> 11654

<212> DNA

<213> Artificial Sequence

<220><223> pAC3-F2A-GFPm

<400> 49

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagtccg 60

cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc cccgeccatt 120

gacgtcaata atgacgtatg ttccatagt aacgccaata gggactttcc attgacgtca 180

atgggtggag tatttacggt aaactgcca cttggcagta catcaagtgt atcatatgcc 240

aagtagccc cciattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 300

catgacctta tgggactttc ctacttgcca gtacatctac gtattagtca tcgctattac 360

catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420

atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480

ggactttcca aaatgtcgt acaactccgc ccattgacg caaatgggcg gtaggcgtgt 540

acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg 600

actgagtcgc ccgggtacc gtgtatcaa taaacctct tgcagttgca tccgacttgt 660

ggctctcgctg ttccctggga gggctctctc tgagtgattg actaccctc agcgggggtc 720

tttcatttgg gggctcgtcc gggatcggga gaccctgcc cagggaccac cgaccacca 780

ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840

tgattttatg cgctcgtc ggtactagtt agctaactag ctctgtatct ggcggacceg 900

tgggtgaact gacgagttcg gaacacccgg ccgcaacct gggagacgtc ccagggactt 960

cgggggccgt ttttgtggcc cgacctgagt ccaaaaaatcc cgatcgtttt ggactctttg 1020
 gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt 1080

 tcccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg 1140
 tctgctgcag catcgttctg tgttgtctct gtctgactgt gtttctgtat ttgtctgaga 1200
 atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg 1260
 agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttggggtt accttctgct 1320
 ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380
 tcatcaccca ggtaagatc aaggtctttt cacctggccc gcattggacac ccagaccagg 1440
 tcccctacat cgtgacctgg gaagccttgg cttttgacct cctccctgg gtcaagcctt 1500

 ttgtacacce taagcctccg cctcctcttc ctccatccgc cccgtctctc ccccttgaac 1560
 ctctcgttc gaccccgctt cgatctctcc ttatccagc cctcactctt tetctaggcg 1620
 ccaaacctaa acctcaagtt ctttctgaca gtggggggcc gctcatcgac ctacttacag 1680
 aagaccccc gccttatagg gacccaagac caccccttc cgacagggac ggaaatggtg 1740
 gagaagcgac cctgcggga gaggcaccgg acccctcccc aatggcatct cgcctacgtg 1800
 ggagacggga gccccctgtg gccgactcca ctacctgca ggcatctccc ctccgcgag 1860
 gaggaaacgg acagcttcaa tactggcctt tctcctcttc tgacctttac aactggaaaa 1920

 ataataacce ttctttttct gaagatccag gtaaactgac agctctgac gagtctgttc 1980
 tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040
 gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100
 gccccactca actgcccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg 2160
 attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagtgt ctcttagcgg 2220
 gtctcaaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280
 ggccaatga gtctccctcg gccttcttag agagacttaa ggaagcctat cgcaggtaca 2340

 ctcttatga cctgaggac ccagggaag aaactaatgt gtctatgtct ttcatttggc 2400
 agtctgcccc agacattggg agaaagttag agaggttaga agatttaaaa aacaagacgc 2460
 ttggagattt ggttagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520
 gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580
 agcagaaaga gaaagaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640
 ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccactcg 2700

atcgcgacca gtgtgcctac tgcaaagaaa agggggcactg ggctaaagat tgtccaaga 2760

aaccacgagg acctcgggga ccaagacccc agacctcctt cctgacccta gatgactagg 2820

gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac 2880

ccgtcacctt cciggtagat actggggccc aacctccgt gctgacccaa aatcctggac 2940

ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgtgga 3000

ccacggatcg caaagtacat ctagetaccg gtaaggtcac cactcttctt ctccatgtac 3060

cagactgtcc ctatcctctg ttaggaagag atttgctgac taaactaaaa gcccaaatcc 3120

actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga 3180

ccctaaatat agaagatgag catcggctac atgagacctc aaaagagcca gatgtttctc 3240

tagggtccac atggctgtct gattttcttc aggcctgggc ggaaacggg ggcattggac 3300

tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360

taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420

tgttgacca ggaataactg gtaccctgcc agtccccctg gaacacgccc ctgctaccg 3480

ttaagaaacc agggactaat gattatagc ctgtccagga tctgagagaa gtcaacaagc 3540

gggtggaaga catccacccc accgtgccca acccttaca cctcttgagc gggtccac 3600

cgteccacca gtggtacact gtgcttgatt taaaggatgc cttttctgc ctgagactcc 3660

accccaccag tcagcctctc ttcgcctttg agtggagaga tccagagatg ggaatctcag 3720

gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780

aggcactgca cagagacctc gcagacttcc ggatccagca cccagacttg atcctgctac 3840

agtacgtgga tgacttactg ctggccgcca cttctgagct agactgccaa caaggtactc 3900

gggccctgtt acaaacctc ggaacctcg ggtatcgggc ctcggccaag aaagcccaaa 3960

tttgccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga 4020

ctgaggccag aaaagagact gtgatggggc agcctactcc gaagaccctt cgacaactaa 4080

gggagtctct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140

cagccccctt gtaccctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200

aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag 4260

atltgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc 4320

taacgcaaaa actgggacct tggcgtcggc cgtggccta cctgtccaaa aagctagacc 4380

cagtagcagc tgggtggccc ccttgectac ggatggtagc agccattgcc gtactgacaa 4440

aggatgcagg caagctaacc atgggacagc cactagtcac tctggccccc catgcagtag 4500
 aggcactagt caaacaaccc cccgaccgct ggctttccaa cgcccggatg actcactatc 4560
 aggccttgct tttggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaaccggg 4620
 ctacgtgct cccactgctt gaggaagggc tgcaacacaa ctgccttgat atcctggccg 4680
 aagcccacgg aaccgaccc gacctaacgg accagccgct cccagacgcc gaccacacct 4740
 ggtacacgga tggaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800
 ccaccgagac cgaggtaatc tgggctaaag ccctgccagc cgggacatcc gctcagcggg 4860

 ctgaactgat agcactcacc caggccctaa agatggcaga aggtagaag ctaaatgttt 4920
 atactgatag ccgttatgct tttgctactg cccatatcca tggagaaata tacagaaggc 4980
 gtgggttgct cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac 5040
 taaaagccct ctttctgcc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100
 gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggcccga aaggcagcca 5160
 tcacagagac tccagacacc tctaccctcc tcatagaaaa ttcacacccc tacacctcag 5220
 aacatittca ttacacagtg actgatataa aggacctaac caagtggggg gccatttatg 5280

 ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340
 ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc 5400
 tcctagagag aagccacagt ccctactaca tgtgaaccg ggatcgaaca ctcaaaaata 5460
 tcaactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaacagg 5520
 gaactagggt ccgcccggcat cggcccggca ctattggga gatcgattc accgagataa 5580
 agcccggatt gtatggctat aaatatcttc tagtttttat agataccttt tctggctgga 5640
 tagaagcctt cccaaccaag aaagaaccg ccaaggtcgt aaccaagaag ctactagagg 5700

 agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cetgcctteg 5760
 tctccaaggt gactcagaca gtggccgac tgttggggat tgattgaaa ttacattgtg 5820
 catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaacctc aaggagactt 5880
 taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc 5940
 tgtaccgagc ccgcaacacg ccgggcccc atggcctcac cccatatgag atcttatatg 6000
 gggcaccccc gcccttgta aacttccctg accctgacat gacaagagtt actaacagcc 6060
 cctctctcca agctcactta caggetctct acttagtcca gcacgaagtc tggagacctc 6120

 tggcggcage ctaccaagaa caactggacc gaccgggtgt acctcacct taccgagtcg 6180
 gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240
 cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcatc gcagcttga 6300

tacacccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat	6360
ggcgcgttca acgctctcaa aacccccctca agataagatt aaccctgga agcccttaat	6420
agtcatggga gtccctgttag gagtagggat ggcagagagc ccccatcagg tctttaatgt	6480
aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct ccctcctggg	6540
aactgtacaa gatgccttcc caaaattata ttttgatcta tgtgatctgg tcggagagga	6600
gtgggacctt tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcaggag	6660
acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcggggtg	6720
tgggggacca ggagaggct actgtggtaa atgggggtgt gaaaccaccg gacaggctta	6780
ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcgtaaca cccctggga	6840
cacgggatgc tctaaagtg cctgtggccc ctgctacgac ctctccaaag tatccaattc	6900
cttccaaggg gctactcgag ggggcagatg caaccctcta gtccctagaat tcaactgatc	6960
aggaaaaag gctaacctggg acgggcccaa atcgtgggga ctgagactgt accggacagg	7020
aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gacccccgagt	7080
ccccataggg cccaaccagg tattaccega ccaagactc ccttctcac caatagagat	7140
tgtaccggt ccacagccac ctagccccct caataccagt taccctctt ccaactaccag	7200
tacacctca acctcccta caagtccaag tgtcccacag ccacccccag gaactggaga	7260
tagactacta gctctagta aaggagccta tcaggcgctt aacctacca atcccgacia	7320
gaccaagaa tgttgctgt gcttagtgtc gggacctcct tattacgaag gtagtagcgt	7380
cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca ctcccaaca	7440
taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac	7500
tcaccaggcc ttatgtaaca ccacccaaag cgcggctca ggatcctact accttgcage	7560
accgccgga acaatgtggg cttgcagcac tggattgact cctgcttgt ccaccaggt	7620
gctcaatcta accacagatt attgtgatt agttgaactc tggcccagag taatttacca	7680
ctccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt	7740
atcattgacc ctggcccttc tactaggagg attaaccatg ggagggatg cagctggaat	7800
agggacgggg accactgcct taattaaaac ccagcagttt gacgacctc atgccgctat	7860
ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc	7920
gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgctattcc taaaggagg	7980
aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cgggctagt	8040
gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac	8100
aggccaagga tggttcgaag ggctgtttaa tagatcccc tggtttacca ccttaatctc	8160

caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct 8220

 caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac 8280
 tcagcaatat caccagctaa aacctataga gtacgagcca gtgaaacaga ctttgaattt 8340
 tgaccttctc aagttagcgg gagacgtgga gtccaacct ggacctggcg cgcctatggc 8400
 cagcaagggc gaggagctgt tcaccggggt ggtgcccatc ctggtcgagc tggacggcga 8460
 cgtaaaccgc cacaagtca gcgtgtccgg cgaaggagag ggcgatgcca cctacggcaa 8520
 gctgaccctg aagttcactc gcaccaccgg caagctgccc gtgcctggc ccaccctcgt 8580
 gaccaccttg acctacggcg tgcagtctt cgcccgtac cccgaccaca tgaagcagca 8640

 cgacttcttc aagtccgcca tgcccgaagg ctacgtccag gagcgacca tcttcttcaa 8700
 ggacgacggc aactacaaga cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa 8760
 ccgcatcgag ctgaagggca tcgacttcaa ggaggacggc aacatcctgg ggcacaagct 8820
 ggagtacaac tacaacagcc acaaggtcta taccaccgcc gacaagcaga agaacggcat 8880
 caagtgaaac ttcaagacc gccacaacat cgaggacggc agcgtgcagc tcgccgacca 8940
 ctaccagcag aacaccccca tcggcgacgg ccccgctctg ctgcccgaca accactacct 9000
 gagcacccag tccgccctga gcaaagacc caacgagaag cgcgatcaca tggctctgct 9060

 ggagtctgtg accgccgccc ggatcactct cggcatggac gagctgtaca agtgtcggc 9120
 cgagataaa ataaaagatt ttatttagtc tcagaaaaa ggggggaatg aaagacccca 9180
 cctgtaggtt tggcaagcta gcttaagtaa cgccatcttg caaggcatgg aaaaatacat 9240
 aactgagaat agagaagttc agatcaaggt caggaacaga tggaacagct gaatatgggc 9300
 caaacaggat atctgtggta agcagttcct gccccggctc agggccaaga acagatggaa 9360
 cagctgaata tgggccaaac aggatatctg tggtaagcag ttcttggccc ggctcagggc 9420
 caagaacaga tggccccag atgcggtcca gcctcagca gtttctagag aaccatcaga 9480

 tgtttccagg gtgccccaa gacctgaaat gaccctgtgc cttatttgaa ctaaccaatc 9540
 agttcgttc tcgcttctgt tcgcgctt ctgctcccc agctcaataa aagagcccac 9600
 aaccctcac tcggggcgc agtctccga ttgactgagt cggccgggta cccgtgtatc 9660
 caataaacc tcttgcagtt gcatccgact tfggtctcg ctgttcttg ggagggtctc 9720
 ctctgagtga ttgactacc gtcagcgggg gtctttcatt acatgtgagc aaaaggccag 9780
 caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc 9840
 cctgacgagc atcacaaaa tcgacgctca agtcagaggt ggcgaaacc gacaggacta 9900

taaagataacc aggcgctttcc ccctggaagc tcctctgtgc gctctctgt tccgaccctg 9960
 ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct ttctcaatgc 10020
 tcacgctgta ggtatctcag ttccggtgtag gtcgcttcgct ccaagctggg ctgtgtgcac 10080
 gaaccccccg ttcagcccga ccgctgcgcc ttatccgta actatcgtct tgagtccaac 10140
 ccggtaaagc acgacttacc gccactggca gcagccactg gtaacaggat tagcagagcg 10200
 aggtatgtag gcggtgctac agagtctctg aagtggggc ctaactacgg ctacactaga 10260
 aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcgaaa aagagttggt 10320

agctcttgat ccggcaaaaa aaccaccgct gtagcggtg gttttttgt ttgcaagcag 10380
 cagattaccg gcagaaaaaa aggatctcaa gaagatcctt tgatctttc tacggggtct 10440
 gacgctcagt ggaacgaaaa ctacagttaa gggattttg tcatgagatt atcaaaaagg 10500
 atcttaccct agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat 10560
 gagtaaaact ggictgacag ttaccaatgc ttaatcagtg aggcacctat ctacgcatc 10620
 tgtctatttc gttcatccat agttgcctga ctccccgtcg ttagataac tacgatacgg 10680
 gagggcttac catctggccc cagtctgca atgataccgc gagaccacg ctaccggct 10740

ccagatttat cagcaataaa ccagccagcc ggaaggccg agcgcagaag tggctctgca 10800
 actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg 10860
 ccagtaata gtttgcgcaa cgttgttgc attgtgcag gcatcgtggt gtcacgctcg 10920
 tcgtttgta tggcttcatt cagctccggt tccaacgat caaggcagat tacatgatcc 10980
 cccatgtgt gcaaaaaagc ggttagctcc ttcggtctc cgatcgtgt cagaagtaag 11040
 ttggccgag tgitatcact catggttatg gcagcactgc ataattctct tactgtcatg 11100
 ccatccgtaa gatgctttc tgtgactggt gactactcaa ccaagtcatt ctgagaatag 11160

tgtatcggc gaccgagttg ctcttgcccg gcgtcaacac gggataaac cgcgccacat 11220
 agcagaactt taaaagtgt catcattgga aaacgttctt cggggcgaaa actctcaagg 11280
 atcttaccg tgttgagatc cagttcgatg taaccactc gtgcaccaa ctgatctca 11340
 gcatctttta cttcaccag cgtttctggg tgagcaaaaa caggaaggca aatgccgca 11400
 aaaaaggaa taaggcgac acggaatgt tgaatactca tactcttct tttcaatat 11460
 tattgaagca tttatcagg ttattgtctc atgagcggat acatattga atgtatttag 11520
 aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa 11580

gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag gcctttcgt 11640
 ctcaagaat teat 11654

<210> 50

<211> 11663

<212> DNA

<213> Artificial Sequence

<220><223> pAC3-GSG-F2A-GFPm

<400> 50

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg	60
cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc cccgcccatt	120
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca	180
atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc	240
aagtacccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta	300
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac	360
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	420
atltccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	480
ggactttcca aaatgtcgt acaactccgc ccattgacg caaatgggcg gtaggcgtgt	540
acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg	600
actgagtcgc cgggtaccc gtgtatccaa taaacctct tgcagttgca tccgacttgt	660
ggtctcgtg ttctttggga gggctcctc tgagtgatg actaccegtc agcggggggtc	720
tttcatttgg gggctcgtcc gggatcgga gaccctgcc cagggaccac cgaccacca	780
ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac	840
tgattttatg cgctgcgtc ggtactagt agtaactag ctctgtatct ggccggaccg	900
tgggtggaact gacgagttcg gaacaccgg cgcaaccct gggagacgtc ccagggactt	960
cgggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg	1020
gtgcaccccc cttagaggag ggatatgtgg tcttgtagg agacgagaac ctaaacagt	1080
tcccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg	1140
tctgtgcag catcgttctg tgttctctc gtctgactgt gtttctgtat ttgtctgaga	1200
atatgggcca gactgttacc actcccttaa gttgacctt aggtcactgg aaagatgtcg	1260
agcggatcgc tcacaaccag tcggtatag tcaagaagag acgttgggtt accttctgct	1320
ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc	1380
tcatcaccca ggtaagatc aaggtctttt cacctggccc gcatggacac ccagaccagg	1440
tcccctacat cgtgacctgg gaagccttgg cttttgacce ccctccctgg gtcaagccct	1500
ttgtacacce taagcctcgg cctcctcttc ctccatccgc cccgtctctc cccttgaac	1560

ctcctcgttc gacccccct cgatcctccc ttatccagc cctcactcct tctctaggcg 1620
 ccaaacctaa acctcaagtt ctttctgaca gtggggggcc gctcatcgac ctacttacag 1680
 aagaccccc gccttatagg gacccaagac ccccccttc cgacaggac ggaaatggtg 1740
 gagaagcgac ccctgcggga gaggcaccgg acccctcccc aatggcatct cgcttacgtg 1800
 ggagacggga gccccctgtg gccgactcca ctacctgca ggcatcccc ctccgcgcag 1860

gaggaaacgg acagcttcaa tactggcctg tctcctcttc tgaccttac aactggaaaa 1920
 ataataacce ttcttttct gaagatccag gtaaactgac agctctgac gagtctgttc 1980
 tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040
 gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100
 gccccactca actgccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg 2160
 attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagtgt ctcttagcgg 2220
 gtctcaaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280

ggccaatga gtctccctcg gccttcttag agagacttaa ggaagcctat cgcaggtaca 2340
 ctcttatga cctgaggac ccagggaag aaactaatgt gtctatgtct ttcatttggc 2400
 agtctgcccc agacattggg agaaagttag agaggtaga agatttaaaa aacaagacgc 2460
 ttggagattt ggttagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520
 gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580
 agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640
 ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700

atcgcgacca gtgtgcctac tgcaaagaaa aggggactg ggctaaagat tgtccaaga 2760
 aaccacgagg acctcggga ccaagacccc agacctcct cctgacctc gatgactagg 2820
 gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc gggggcaac 2880
 ccgtcacctt cctggtagat actggggccc aacctcctg gctgacccaa aatcctggac 2940
 ccctaagtga taagtctgcc tgggtccaag gggtactgg aggaagcgg tatcgtgga 3000
 ccacggatcg caaagtacat ctagetaccg gtaaggcac cactcttcc ctccatgtac 3060
 cagactgtcc ctatcctctg ttaggaagag atttgcctgac taaactaaa gcccaatcc 3120

actttgagg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgtga 3180
 ccctaatat agaagatgag catcggttac atgagacctc aaaagagcca gatgtttctc 3240
 taggtccac atggctgtct gattttctc aggcctgggc ggaaccggg ggcattggac 3300
 tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360
 taaaacaata ccccatgtca caagaagcca gactggggat caagccccc atacagagac 3420

tgttggacca gggaaactg gtaccctgcc agtccccctg gaacacgccc ctgctaccg	3480
ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc	3540
gggtggaaga catccacccc accgtgceca acccttacia cctcttgagc gggctcccac	3600
cgtcccacca gtggtacact gtgcttgatt taaaggatgc cttttctgc ctgagactcc	3660
accccaccag tcagcctctc ttgcctttg agtggagaga tccagagatg ggaatctcag	3720
gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg	3780
aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atcctgctac	3840
agtacgtgga tgacttactg ctggccgcca ctctgagct agactgcca caaggtactc	3900
gggcctgtt acaaacccta gggaaacctg ggtatcgggc ctggccaag aaagccaaa	3960
tttgccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga	4020
ctgagccag aaaagagact gtgatggggc agcctactcc gaagaccct cgacaactaa	4080
gggagttcct agggacgca ggcttctgtc gcctctggat ccctgggtt gcagaaatgg	4140
cagccccctt gtaccctctc accaaaacgg ggactctgtt taattggggc ccagaccaac	4200
aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag	4260
atttgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgcc	4320
taacgcaaaa actgggacct tggcgtcggc cggtggccta cctgtccaaa aagctagacc	4380
cagtagcagc tgggtggccc cttgcctac ggatggtagc agccattgcc gtactgacaa	4440
aggatgcagg caagctaacc atgggacagc cactagtcac tctggcccc catgcagtag	4500
aggcactagt caaacaacc cccgaccgt ggctttccaa cgcccggatg actcactatc	4560
aggccttgct ttiggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaaccgg	4620
ctacgtgct cccactgct gaggaaggc tgcaacacaa ctgccttgat atcctggccg	4680
aagcccagg aaccgacc gacctaacgg accagccgt cccagaccg gaccacacct	4740
gtacacgga tggaaagcgt ctcttacaag agggacagcg taaggcggga gctgcggtga	4800
ccaccgagac cgaggtaatc tgggctaaag cctgccagc cgggacatcc gctcagcggg	4860
ctgaactgat agcactcacc caggccctaa agatggcaga aggtagaag ctaaatgttt	4920
atactgatag ccgttatgct ttgctactg cccatatcca tggagaaata tacagaagc	4980
gtgggttgc cacatcagaa ggcaaaagaga tcaaaaataa agacgagatc ttggccctac	5040
taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg	5100
gacacagcgc cgaggctaga ggcaaccgga tgctgacca agcggcccga aaggcagcca	5160

tcacagagac tccagacacc tctacacctc tcatagaaaa ttcacacacc tacacctcag 5220

aacattttca ttacacagtg actgatataa aggacctaac caagttgggg gccatttatg 5280

ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340

ttgaattatt agactttctt catcagctga ctacacctcag cttctcaaaa atgaaggctc 5400

tcctagagag aagccacagt ccctactaca tgcctgaaccg ggatcgaaca ctcaaaaata 5460

tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg 5520

gaactagggt ccgcgggcat cggcccggca ctattggga gatcgatttc accgagataa 5580

agcccggatt gtatggctat aaatatcttc tagtttttat agataccttt tctggctgga 5640

tagaacctt cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg 5700

agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760

tctccaaggt gactcagaca gtggccgac tgttggggat tgattggaaa ttacattgtg 5820

catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaacctc aaggagactt 5880

taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc 5940

tgtaccgagc ccgcaacacg ccgggcccc atggcctcac cccatagag atcttatatg 6000

gggcaccccc gcccttgta aacttccctg acctgacat gacaagagtt actaacagcc 6060

cctctctcca agctcactta cagctctct acttagtcca gcacgaagtc tggagacctc 6120

tggcggcagc ctaccaagaa caactggacc gaccgggtgt acctaccct taccgagtcg 6180

gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240

cttacacagt cctgctgacc acccccaccg cctcaaagt agacggcatc gcagcttggg 6300

tacacgccg ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat 6360

ggcgcgttca acgctctcaa aaccccctca agataagatt aaccctgga agccctaat 6420

agtcatggga gtctgttag gagtagggat ggcagagagc ccccatcagg tctttaatgt 6480

aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct cctcctggg 6540

aactgtacaa gatgccttcc caaaattata tttgatcta tgtgatctgg tcggagagga 6600

gtgggacctc tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcagggag 6660

acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcggggtg 6720

tgggggacca ggagagggt actgtggtaa atgggggtgt gaaaccaccg gacaggctta 6780

ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcggtaaca ccccctggga 6840

cacgggatgc tctaaagttg cctgtggccc ctgctacgac ctctccaaag tatccaatc 6900

cttccaaggg gctactcgag ggggcagatg caaccctcta gtcctagaat tctactgatgc 6960
 aggaaaaaag gctaactggg acgggcccga atcgtgggga ctgagactgt accggacagg 7020
 aacagatcct attaccatgt tctccctgac ccggcaggtc cftaatgtgg gacccccagt 7080
 ccccataggg cccaaccag tattaccgca ccaaagactc ccttcctcac caatagagat 7140
 tgtaccggct ccacagccac ctagccccct caataccagt tacccccctt cactaccag 7200
 tacacctca acctccccta caagtccaag tgtcccacag ccacccccag gaactggaga 7260
 tagactacta gctctagtca aaggagccta tcaggcgctt aacctacca atcccgacaa 7320

 gaccaagaa tgttggctgt gcttagtgtc gggacctcct tattacgaag gtagtagcgt 7380
 cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca cttccaaca 7440
 taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac 7500
 tcaccaggcc ttatgtaaca ccacccaaag cgcggctca ggatcctact accttcagc 7560
 acccgccgga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccacgt 7620
 gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taatttacca 7680
 ctccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt 7740

 atcattgacc ctggcccttc tactaggagg attaaccatg ggagggattg cagctggaat 7800
 aggacgggg accactgcct taattaaaac ccagcagttt gagcagcttc atgccctat 7860
 ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920
 gttgtctgaa gtagtcctac agaaccgag aggcctagat ttgctattcc taaaggagg 7980
 aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cgggctagt 8040
 gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100
 aggccaagga tggttcgaag ggctgtttaa tagatcccc tggttacca ccttaatctc 8160

 caccatcatg ggacctetaa tagtactctt actgatetta ctctttggac cttgcattct 8220
 caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac 8280
 tcagcaatat caccagctaa aacctataga gtacgagcca ggaagcggag tgaacagac 8340
 ttgaatfff gaccttctca agttggcggg agacgtggag tccaacctg gacctggcgc 8400
 gcctatggcc agcaagggcg aggagctgtt caccggggtg gtgccatcc tggtcgagct 8460
 ggacggcgac gtaaaggcc acaagttcag cgtgtccggc gaaggagagg gcgatgccac 8520
 ctacggcaag ctgacctga agttcatctg caccaccggc aagctgccg tgcctggcc 8580

 caccctcgtg accaccttga cctacggcgt gcagtgttc gcccgtacc ccgaccat 8640
 gaagcagcac gacttctca agtccgceat gccgaaggc tacgtccagg agcgcacat 8700
 cttctcaag gacgacgca actacaagac ccgcgccgag gtgaagttcg agggcgacac 8760

cctggatgaac cgcatcgagc tgaagggcat cgacttcaag gaggacggca acatcctggg 8820
 gcacaagctg gactacaact acaacagcca caaggtctat atcaccgccg acaagcagaa 8880
 gaacggcatc aaggtgaact tcaagaccg ccacaacatc gaggacggca gcgtgcagct 8940
 cgccgaccac taccagcaga acaccccat cggcgacggc cccgtgctgc tgcccagaaa 9000

 ccaactactg agcaccagt cgcctctgag caaagacccc aacgagaagc gcgatcacat 9060
 ggtcctgctg gacttctgga ccgcccggg gatcactctc ggcatggacg agctgtacaa 9120
 gtgtcggcc gcagataaaa taaaagattt tatttagtct ccagaaaaag gggggaatga 9180
 aagacccac ctgtaggttt ggcaagctag cttaagtaac gccattttgc aaggcatgga 9240
 aaaatacata actgagaata gagaagtcca gatcaaggtc agaacagat ggaacagctg 9300
 aatatggcc aacaggata tctgtggtaa gcagttctg ccccggtca gggccaagaa 9360
 cagatggaac agctgaatat gggccaaaca ggatatctgt ggtaagcagt tctgccccg 9420

 gctcaggcc aagaacagat ggtccccaga tgcggtccag ccctcagcag tttctagaga 9480
 accatcagat gttccaggg tgcccgaag acctgaaatg acctgtgcc ttatttgaac 9540
 taaccaatca gttcgtctt cgcttctgtt cgcgctctc tgctccccga getcaataaa 9600
 agagcccaca acccctcact cggggcgcca gtctccgat tgactgagtc gcccggtac 9660
 ccgtgtatcc aataaacct cttgcagttg catccgactt gtggtctcgc tgttcttgg 9720
 gaggtctcc tctgagtgat tgactaccg tcagcggggg tctttcatta catgtgagca 9780
 aaaggccagc aaaaggccag gaaccgtaa aaggccgct tgctggcgtt tttcatagg 9840

 ctccgcccc ctgacgagca tcacaaaaat cgacgtcaa gtcagaggtg gcgaaacccg 9900
 acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt 9960
 ccgacctgc cgttaccgg atacctgtcc gcctttctcc cttcggaag cgtggcgctt 10020
 tctcaatgct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc 10080
 tgtgtgcacg aacccccgt tcagcccagc cgtcgcct tatccggtaa ctatcgtctt 10140
 gagtccaacc cgtaagaca cgacttatc cactggcag cagccactgg taacaggatt 10200
 agcagagcga ggtatgtagg cggctctaca gatttctga agtggtgcc taactacggc 10260

 tacactagaa ggacagtatt tggatctgc gctctgctga agccagttac cttcgaaaa 10320
 agagtggta gctcttgatc cggcaaaaa accaccctg gtagcggtag ttttttgg 10380
 tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatcctt gatctttct 10440
 acgggtctg acgtcagtg gaacgaaaac tcacgttaag ggattttgt catgagatta 10500
 tcaaaaagga tcttacctg gatctttta aattaaaaat gaagttttaa atcaatctaa 10560
 agtatatatg agttaaactg gtctgacagt taccaatgct taatcagtg ggcacctatc 10620

tcagcgatct gtctatttcg ttcacccata gttgcctgac tccccgtcgt gtagataact 10680

acgatacggg agggcttacc atctggcccc agtgcctgcaa tgataccgcg agaccacgc 10740

tcaccggctc cagatattac agcaataaac cagccagccg gaaggccga gcgcagaagt 10800

ggtcctgcaa cttatccgc ctccatccag tctattaatt gttgccggga agctagagta 10860

agtagttcgc cagttaatag tttgcgcaac gttgttgcca ttgctgcagg catcgtggtg 10920

tcacgctcgt cgtttggtat ggcttcattc agctccggtt cccaacgac aaggcgagtt 10980

acatgatccc ccatgttgct caaaaaagcg gttagctcct tcggtcctcc gatcgttgct 11040

agaagtaagt tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt 11100

actgtcatgc catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc 11160

tgagaatagt gtatgcccgc accgagttgc tcttggccgg cgtcaacacg ggataatacc 11220

gcgccacata gcagaacttt aaaagtgtc atcattggaa aacgttcttc ggggcgaaaa 11280

ctctcaagga tcttaccgct gttgagatcc agttcgatgt aaccactcg tgcaccaac 11340

tgatcttcag catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa 11400

aatgccgcaa aaaagggaat aaggcgaca cggaaatgtt gaatactcat actcttctt 11460

tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata catattgaa 11520

tgatattaga aaaataaaca aataggggtt ccgcgccat tccccgaaa agtgccacct 11580

gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 11640

cccttctgct tcaagaatt cat 11663

<210> 51

<211> 11399

<212> DNA

<213> Artificial Sequence

<220><223> pAC3-T2A-yCD2

<400> 51

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg 60

cgttacataa cttacggtaa atggcccgc tggtgaccg cccaacgacc cccgccatt 120

gacgtcaata atgacgtatg tcccatagt aacccaata gggactttcc attgacgtca 180

atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 240

aagtacccc cctattgacg tcaatgacgg taaatgccc gcctggcatt atgccagta 300

catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 360

catggtgatg cggttttggc agtacatcaa tggcgtgga tagcggttg actcacgggg 420

attccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480
 ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg gtaggcgtgt 540

 acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg 600
 actgagtcgc ccgggtaccc gtgtatcaa taaacctct tgcagttgca tccgacttgt 660
 ggtctcgtg ttcttggga gggctcctc tgagtattg actaccctc agcgggggtc 720
 tttcatttgg ggctcgtcc gggatcggga gaccctgcc cagggaccac cgaccacca 780
 ccgggaggta agctggccag caactatct gtgtctgtcc gattgtctag tgtctatgac 840
 tgattttatg cgctcgtc ggtactagt agctaactag ctctgtatct ggcggaccg 900
 tggtggaact gacgagttcg gaacacccgg ccgaaccct gggagacgtc ccagggactt 960

 cgggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg 1020
 gtgcacccc cttagaggag ggatatgtgg tttcgttagg agacgagaac ctaaacagt 1080
 tcccgcctcc gtctgaattt ttgcttccg tttgggaccg aagccgcgcc gcgcgtcttg 1140
 tctgtcag catcgttctg tgttctctc gtctgactgt gtttctgat ttgtctgaga 1200
 atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgctg 1260
 agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct 1320
 ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380

 tcatcaccca ggtaagatc aaggctttt cacctggccc gcatggacac ccagaccagg 1440
 tcccctacat cgtgacctgg gaagccttgg cttttgacct cctccctgg gtcaagcct 1500
 ttgtacacce taagcctccg cctcctctc ctccatccgc cccgtctctc ccccttgaac 1560
 ctctcgttc gaccccgct cgatcctccc ttatccagc cctcactct tctctaggcg 1620
 ccaaacctaa acctcaagtt ctttctgaca gtggggggcc gctcatcgac ctacttacag 1680
 aagaccccc gccttatagg gacccaagac caccctctc cgacagggac ggaaatggtg 1740
 gagaagcgac ccctgcggga gaggcaccgg acccctccc aatggcatct cgctacgtg 1800

 ggagacggga gcccctgtg gccgactcca ctacctgca ggcattcccc ctccgcgag 1860
 gaggaaacgg acagcttcaa tactggccgt tctcctctc tgaccttac aactggaaaa 1920
 ataataacce ttcttttct gaagatccag gtaaactgac agctctgac gactctgttc 1980
 tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040
 gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100
 gcccactca actgccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg 2160

attacaccac ccaggcaggt aggaaccacc tagtccacta tgcaccgttg ctcttagcgg 2220

 gtctcaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280
 ggcccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaggtaca 2340
 ctcttatga ccctgaggac ccagggaag aactaatgt gtctatgtct tcatttggc 2400
 agtctgcccc agacattggg agaaagttag agaggttaga agatttaaaa aacaagacgc 2460
 ttggagattt ggtagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520
 gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580
 agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640

 ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700
 atcgcgacca gtgtgcctac tgcaaagaaa aggggcactg ggctaaagat tgtccaaga 2760
 aaccacgagg acctcgggga ccaagacccc agacctcct cctgaccta gatgactagg 2820
 gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac 2880
 ccgtcacctt cctggtagat actggggccc aacctcctg gctgaccaa aatcctggac 2940
 ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgtgga 3000
 ccacggatcg caaagtacat ctagctaccg gtaaggtcac ccactcttc ctccatgtac 3060

 cagactgtcc ctatcctctg ttaggaagag atttgcctgac taaactaaaa gcccaaatcc 3120
 actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga 3180
 ccctaataat agaagatgag catcggttac atgagacctc aaaagagcca gatgtttctc 3240
 taggtccac atggctgtct gattttctc aggcctgggc ggaaaccggg ggcattggac 3300
 tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360
 taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420
 tgttggacca ggaatactg gtaccctgcc agtcccctg gaacacgccc ctgctacceg 3480

 ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc 3540
 ggggtggaaga catccacccc accgtgcca acccttaca cctcttgagc gggtccccac 3600
 cgtcccacca gtggtacact gtgcttgatt taaaggatgc cttttctgc ctgagactcc 3660
 accccaccag tcagcctctc ttcgcctttg agtggagaga tccagagatg ggaatctcag 3720
 gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780
 aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atcctgctac 3840
 agtacgtgga tgacttactg ctggccgcca ctctgagct agactgcaa caaggtactc 3900

gggccctgtt acaaaccta gggaacctcg ggtatcgggc ctcgccaag aaagccaaa 3960
 ttgcccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga 4020
 ctgaggccag aaaagagact gtgatggggc agcctactcc gaagaccct cgacaactaa 4080
 gggagtccct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140
 cagccccctt gtacctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200
 aaaagccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttccag 4260
 atttgactaa gcctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc 4320

 taacgcaaaa actgggacct tggcgtcggc cggtaggcta cctgtcaaaa aagctagacc 4380
 cagtagcagc tgggtggccc ccttgccctac ggatggtagc agccattgcc gtactgaaa 4440
 aggatgcagg caagctaacc atgggacagc cactagtcat tctggcccc catgcagtag 4500
 aggcactagt caaacaacc cccgaccgct ggctttccaa cgcccggatg actcactatc 4560
 aggccttgcct ttggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaacccgg 4620
 ctacgtgct cccactgct gaggaaggc tgcaacaaa ctgccttgat atcctggccg 4680
 aagcccacgg aaccgaccc gacctaacgg accagccgct cccagacgcc gaccacacct 4740

 ggtacacgga tggaaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800
 ccaccgagac cgaggaatc tgggctaaag cctgcccagc cgggacatcc gctcagcggg 4860
 ctgaactgat agcactcacc caggecctaa agatggcaga aggtagaag ctaaatgttt 4920
 atactgatag ccgttatgct ttgtctactg cccatatcca tggagaaata tacagaaggc 4980
 gtgggttgcct cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac 5040
 taaaagccct ctttctgcc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100
 gacacagcgc cgaggctaga ggcaaccgga tgctgacca agcggccccga aaggcagcca 5160

 tcacagagac tcagacacc tctacctcc tcatagaaaa tteatcacc tacacctcag 5220
 aacattttca ttacacagt actgatataa aggacctaac caagttgggg gccatttatg 5280
 ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340
 ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc 5400
 tcctagagag aagccacagt ccctactaca tgcgaaccg ggatcgaaca ctcaaaaata 5460
 tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg 5520
 gaactagggt ccgcccggcat cggcccggca ctattggga gatcgatttc accgagataa 5580

 agcccggatt gtatggctat aaatatctc tagttttat agatacttt tctggctgga 5640
 tagaacctt cccaaccaag aaagaaaccg ccaaggctgt aaccaagaag ctactagagg 5700
 agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760

tctccaaggt gaggcagaca gtggccgata tgttggggat tgattggaaa ttacattgtg 5820
catacagacc ccaagctca ggccaggtag aaagaatgaa tagaacctc aaggagactt 5880
taactaaatt aacgcttgca actggctcta gagactgggt gtcctactc cccttagccc 5940
tgtaccgagc ccgcaacacg ccgggcccc atggcctcac cccatatgag atcttatatg 6000

gggcaccccc gcccttgta aacttccctg accctgacat gacaagagtt actaacagcc 6060
cctctctcca agctcactta caggetctct acttagtcca gcacgaagtc tggagacctc 6120
tggcggcagc ctaccaagaa caactggacc gaccggtggg acctcacct taccgagtcg 6180
gcgacacagt gtgggtccg cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240
cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcctc gcagcttggg 6300
tacacgccg ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat 6360
ggcgcgttca acgctctcaa aacccctca agataagatt aaccctgga agccctaat 6420

agtcatggga gtctgttag gaggtaggat ggacagagc ccccatcagg tctttaatgt 6480
aacctggaga gtcaccaacc tgatgactgg gctaccgcc aatgccact cctcctggg 6540
aactgtacaa gatgccttc caaaattata tttgatcta tgtgatctgg tggagagga 6600
gtgggacct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgagggag 6660
acagcggacc cggacttttg acttttacgt gtcccctggg cataccgtaa agtcggggtg 6720
tgggggacca ggagagggt actgtggtaa atgggggtgt gaaaccaccg gacagctta 6780
ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcgtaaca cccctggga 6840

cacgggatgc tctaaagtg cctgtgccc ctgctacgac ctctcaaag tatccaattc 6900
cttccaaggg gctactcgag ggggcagatg caaccctcta gtcctagaat tactgatgc 6960
aggaaaaaag gctaaactgg acgggcccga atcgtgggga ctgagactgt accggacagg 7020
aacagatcct attaccatgt tctcctgac ccggcaggtc cttaatgtgg gaccccgagt 7080
ccccatagg ccaaccagc tattaccga ccaaagactc ccttctcac caatagagat 7140
tgtaccggt ccacagccac ctageccct caataccagt taccctt cactaccag 7200
tacacctca acctcccta caagtccaag gtccccacag ccacccag gaactggaga 7260

tagactacta gctctagca aaggagccta tcaggcgtt aacctacca atcccgaaa 7320
gaccaagaa tgttggctgt gcttagtgc gggacctct tattacgaag gaggtaggt 7380
cgtggcact tataccaatc attccaccg tccggccaac tgtacggca cttccaaca 7440
taagcttacc ctatctgaag tgacaggaca ggcctatgc atggggcag tacctaaaac 7500
tcaccagcc ttatgtaaca ccaccaaag cggcgtctca ggatcctact accttcagc 7560
acccgccga acaatgtggg cttgcagcac tggattgact cctgcttgt ccaccaggt 7620

gctcaatcta accacagatt atttgtatt agttgaactc tggcccagag taatttacca 7680

ctccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt 7740

atcattgacc ctggcccttc tactaggagg attaacatg ggagggatg cagctggaat 7800

agggacgggg accactgcct taattaaaac ccagcagttt gagcagcttc atgccgctat 7860

ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920

gttgtctgaa gtagtctac agaaccgcag aggcctagat ttgctattcc taaaggaggg 7980

aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt 8040

gagagacagc atggccaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100

aggccaagga tggttcgaag ggctgtttta tagatcccc tggtttacca ccttaatctc 8160

caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac ctgcatctt 8220

caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac 8280

tcagcaatat caccagctaa aacctataga gtacgagcca gagggcagag gaagtcttct 8340

aacatgcggt gacgtggagg agaatcccgg ccttggcgcg cctatggtga ccggcggcat 8400

ggcctccaag tgggatcaaa agggcatgga tatcgcttac gaggaggccc tgctgggcta 8460

caaggagggc ggcgtgccta tcggcggctg tctgatcaac aacaaggacg gcagtgtgct 8520

gggcaggggc cacaacatga ggttcagaa gggctccgcc accctgcacg gcgagatctc 8580

cacctggag aactgtgga ggctggaggg caaggtgtac aaggacacca ccctgiacac 8640

cacctgtcc cttgtgaca tgtgtaccgg cgctatcctc atgtacggca tccttaggtg 8700

tgtgatcggc gagaacgtga acttcaagtc caagggcgag aagtacctgc aaaccagggg 8760

ccacagggtg gtggttgttg acgatgagag gtgtaagaag ctgatgaagc agttcatcga 8820

cgagaggcct caggactggt tcgaggatat cggcgagtaa gcggccgcag ataaaataaa 8880

agatTTTTATT tagtctccag aaaaaggggg gaatgaaaga cccacactgt aggtttggca 8940

agctagctta agtaacgcca ttttgcaagg catggaaaaa tacataactg agaatagaga 9000

agttcagatc aaggtcagga acagatggaa cagctgaata tgggccaaac aggatatctg 9060

tggtaagcag ttctgcccc ggctcagggc caagaacaga tggaacagct gaatatgggc 9120

caaacaggat atctgtggtg agcagttcct gccccgctc agggccaaga acagatggtc 9180

cccagatcgc gtccagccct cagcagtttc tagagaacca tcagatgttt ccagggtgcc 9240

ccaaggacct gaaatgacce tgtgccttat ttgaactaac caatcagttc gcttctcgct 9300

tctgttcgcg cgcttctgct ccccagctc aataaaagag cccacaacce ctcaactcggg 9360

gcgccagtcc tccgattgac tgagtcgccc gggtagccgt gtatccaata aaccctcttg 9420
 cagttgcatc cgacttggg tctcgtgtt ccttgggagg gtctcctctg agtgattgac 9480
 taccgtcag cgggggtctt tcattacatg tgagcaaaag gccagcaaaa ggccaggaac 9540
 cgtaaaaagg ccgctgttgc ggcgtttttc cataggtcc gccccctga cgagcatcac 9600
 aaaaatcgac gctcaagtca gagtggcga aacccgacag gactataaag ataccaggcg 9660
 ttccccctg gaagctccct cgtgcgctct cctgttccga cctgccgct taccggatac 9720
 ctgtccgct ttctccctc gggaagcgtg gcgctttctc aatgctcacg ctgtaggtat 9780

 ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag 9840
 cccgaccgt gcgccttacc cggtaactat cgtcttgagt ccaaccggt aagacacgac 9900
 ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta ttaggcggt 9960
 gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatttgg 10020
 atctgcgctc tgctgaagcc agttaccttc gaaaaaagag ttgtagctc ttgatccggc 10080
 aaacaacca ccgctggtag cgggtggttt tttgtttgca agcagcagat tacgcgcaga 10140
 aaaaaaggat ctcaagaaga tcctttgac ttttctacgg ggtctgacgc tcagtggaac 10200

 gaaaactcac gtttaaggat tttggtcatg agattatcaa aaaggatctt cacctagatc 10260
 cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggctt 10320
 gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttegttca 10380
 tccatagtgg cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct 10440
 ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga ttatcagca 10500
 ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgctcc 10560
 atccagtcta ttaattgttg ccggaagct agagtaagta gttgccagt taatagttg 10620

 cgcaacgttg ttgccattgc tgcaggcatc gtggtgtcac gctcgtcgtt tggtaggct 10680
 tcattcagct ccggttccca acgatcaagg cgagttacat gatccccat gttgtgcaaa 10740
 aaagcgttga gtccttcgg tctccgacg gttgtcagaa gtaagttgce cgcagtgtta 10800
 tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 10860
 ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg 10920
 agttgctctt gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa 10980
 gtgctcatca ttgaaaaacg ttcttcgggg cgaactct caaggatctt accgctgttg 11040

 agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc 11100
 accagcgttt ctgggtgagc aaaaacagga aggcaaatg ccgcaaaaaa gggaataagg 11160
 gcgacacgga aatgttgaat actcactc ttccttttc aatattattg aagcatttat 11220

cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaata 11280
 ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 11340
 atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtcttca agaattcat 11399
 <210> 52
 <211> 11408
 <212> DNA

 <213> Artificial Sequence
 <220><223> pAC3-GSG-T2A-yCD2
 <400> 52

 tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagtccg 60
 cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 120
 gacgtcaata atgacgtatg ttccatagt aacccaata gggactttcc attgacgtca 180
 atgggtggag tatttacggt aaactgcca cttggcagta catcaagtgt atcatatgcc 240
 aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta 300
 catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 360

 catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420
 atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480
 ggactttcca aaatgtcgt acaactccgc ccattgacg caaatgggcg gtaggcgtgt 540
 acggtgggag gtctatataa gcagagctgg tttagtgaac cggcggcagt cctccgattg 600
 actgagtgcg ccgggtaccg gtgtatccaa taaacctct tgcagttgca tccgacttgt 660
 ggtctcgtg ttccttggga gggctctctc tgagtattg actaccctc agcggggggtc 720
 tttcatttgg gggctcgtcc gggatcggga gaccctgcc cagggaccac cgaccacca 780

 ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840
 tgattttatg cgctgcgtc ggtactagt agctaactag ctctgtatct ggcggaccgg 900
 tgggtggaact gacgagttcg gaacacccgg ccgcaacct gggagacgtc ccagggactt 960
 cgggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg 1020
 gtgcacccc cttagaggag ggatatgtgg tttcggtagg agacgagaac ctaaaacagt 1080
 tcccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg 1140
 tctgctgcag catcgttctg tgttctctct gtctgactgt gtttctgtat ttgtctgaga 1200

 atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg 1260
 agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct 1320

ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380
 tcatcaccca ggtaagatc aaggtctttt cacctggccc gcatggacac ccagaccagg 1440
 tcccctacat cgtgacctgg gaagccttgg cttttgacct ccctccctgg gtcaagccct 1500
 ttgtacacce taagcctccg cctcctcttc ctccatccgc cccgtctctc ccccttgaac 1560
 ctctcgttc gaccccgctt cgatctccc tttatccagc cctcactcct tctctaggcg 1620

 ccaaacctaa acctcaagtt ctttttgaca gtggggggcc gctcatcgac ctacttacag 1680
 aagaccccc gccttatagg gacccaagac caccctctc cgacagggac ggaaatggtg 1740
 gagaagcgac ccctgcggga gaggcaccgg acccctcccc aatggcatct cgcctacgtg 1800
 ggagacggga gccccctgtg gccgactcca ctacctcgca ggcatteccc ctccgcgacg 1860
 gaggaaacgg acagcttcaa tactggccgt tctcctcttc tgacctttac aactggaaaa 1920
 ataataacce ttctttttct gaagatccag gtaaaactgac agctctgatc gagtctgttc 1980
 tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040

 gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100
 gccccactca actgcccact gaagtcgatg ccgcttttcc cctcagcgc ccagactggg 2160
 attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagtgt ctcttagcgg 2220
 gtctcaaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280
 ggcccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaggtaca 2340
 ctcttatga ccctgaggac ccagggaag aactaatgt gtctatgtct ttcatttggc 2400
 agtctgcccc agacattggg agaaagttag agaggtaga agatttaaa aacaagacgc 2460

 ttggagattt ggtagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520
 gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580
 agcagaaaga gaaagaaaga gatcgttaga gacatagaga gatgagcaag ctattggcca 2640
 ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700
 atcgcgacca gtgtgcctac tgcaaagaaa aggggcactg ggctaaagat tgtccaaga 2760
 aaccacgagg acctcgggga ccaagacccc agacctcct cctgacccta gatgactagg 2820
 gaggtcaggg tcaggagccc ccccctgaac ccaggataac cctcaaagtc ggggggcaac 2880

 ccgtcacctt cctgtagat actggggccc aacctcctg gctgaccaa aatcctggac 2940
 ccctaagtga taagtctgcc tgggtccaag gggtactgg aggaagcgg tatcgctgga 3000
 ccacggatcg caaagtacat ctactaccg gtaaggtcac ccactcttc ctccatgtac 3060
 cagactgtcc ctatctctg ttaggaagag atttctgac taaactaaaa gcccaatec 3120
 actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgtga 3180

ccctaaatat agaagatgag catcggtac atgagacctc aaaagagcca gatgtttctc 3240
 taggtccac atggctgtct gattttctc aggcctgggc ggaaaccggg ggcattggac 3300

 tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360
 taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420
 tgttggacca gggaatactg gtacctgcc agtccccctg gaacacgccc ctgctaccg 3480
 ttaagaaacc agggactaat gattatagc ctgtccagga tctgagagaa gtcaacaagc 3540
 ggggtggaaga catccacccc accgtgccc acccttaca cctcttgagc gggtccac 3600
 cgtcccacca gtggtacct gtgcttgatt taaaggatgc cttttctgc ctgagactcc 3660
 accccaccag tcagcctctc ttgcctttg agtggagaga tccagagatg ggaatctcag 3720

 gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780
 aggcactgca cagagacctc gcgacttcc ggatccagca cccagacttg atcctgctac 3840
 agtacgtgga tgacttactg ctggccgcca cttctgagct agactgcaa caaggtactc 3900
 gggccctgtt acaaaccta gggaacctc ggtatcgggc ctcgccaag aaagcccaaa 3960
 tttgccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga 4020
 ctgagccag aaaagagact gtgatggggc agcctactcc gaagaccct cgacaactaa 4080
 gggagtctct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140

 cagccccctt gtacctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200
 aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgcag 4260
 atttgactaa gccctttgaa ctctttgtcg acgagaagca gggtacgcc aaaggtgtcc 4320
 taacgcaaaa actgggacct tggcgtcggc cggtagccta cctgtccaaa aagctagacc 4380
 cagtagcagc tgggtggccc ccttgectac ggatggtagc agccattgcc gtactgaaa 4440
 aggatgcagg caagctaacc atgggacagc cactagtcac tctggcccc catgcagtag 4500
 aggcactagt caacaaccc cccgaccgt ggctttccaa cgcccggatg actcactac 4560

 aggccttgtt ttggacagc gaccgggtcc agttcggacc ggtggtagcc ctgaaccgg 4620
 ctacgtgct cccactgct gaggaagggc tgcaacacaa ctgccttgat atcctggccg 4680
 aagcccacgg aaccgaccc gacctaacgg accagccgt cccagacgc gaccacact 4740
 ggtacacgga tggagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800
 ccaccgagac cgaggtaatc tgggctaaag ccctgccagc cgggacatcc gctcagcggg 4860
 ctgaactgat agcactcacc caggccctaa agatggcaga aggtaagaag ctaaatgttt 4920

atactgatag ccgttatgct tttgctactg cccatatcca tggagaaata tacagaaggc 4980

 gtgggttgct cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac 5040
 taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100
 gacacagcgc cgaggctaga ggcaaccgga tgctgacca agcggcccga aaggcagcca 5160
 tcacagagac tccagacacc tctaccctcc tcatagaaaa tteatcacc tacacctcag 5220
 aacattttca ttacacagtg actgatataa aggacctaac caagtgggg gccatttatg 5280
 ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340
 ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc 5400

 tcctagagag aagccacagt ccctactaca tgctgaaccg ggatcgaaca ctcaaaaata 5460
 tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaacagg 5520
 gaactaggtt ccgcccggat cggcccggca ctcatggga gatcgattc accgagataa 5580
 agcccggatt gtatggctat aaatatcttc tagtttttat agatacctt tctggctgga 5640
 tagaagcctt cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg 5700
 agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgcctteg 5760
 tctccaaggt gagtcagaca gtggccgatc tgttgggat tgattggaaa ttacattgtg 5820

 catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaacctc aaggagactt 5880
 taactaaatt aacgcttgca actggctctc gagactgggt gctcctact cccttagccc 5940
 tgtaccgagc ccgcaacacg ccgggcccc atggcctcac cccatagag atcttatatg 6000
 gggcaccccc gcccttgta aacttccctg acctgacat gacaagagt actaacagcc 6060
 cctctctcca agctcactta caggctctct acttagtcca gcacgaagtc tggagacctc 6120
 tggcggcagc ctaccaagaa caactggacc gaccgggtgt acctaccct taccgagtcg 6180
 gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240

 cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcatc gcagcttga 6300
 tacacgccg ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat 6360
 ggcgcttca acgctctcaa aacccctca agataagatt aaccctgga agccctaat 6420
 agtcatggga gtctgttag gtagtaggat ggcagagagc cccatcagg tctttaatgt 6480
 aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct cctcctggg 6540
 aactgtacaa gatgccttcc caaaattata tttgatcta tgtgatctgg tcggagagga 6600
 gtgggacctc tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcaggag 6660

acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcggggtg 6720
 tgggggacca ggagaggct actgtggtaa atgggggtgt gaaaccaccg gacaggctta 6780
 ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcggttaaca cccctggga 6840
 cacgggatgc tctaaagtgt cctgtggccc ctgctacgac ctctccaaag tatccaattc 6900
 cttccaaggg gctactcgag ggggcagatg caaccctcta gtccctagaat tctactgatgc 6960
 aggaaaaaag gctaactggg acgggcccga atcgtgggga ctgagactgt accggacagg 7020
 aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gacccccagt 7080

ccccataggg cccaaccag tattaccga ccaaagactc ccttcctcac caatagagat 7140
 tgtaccggt ccacagccac ctageccccct caataccagt tacccectt cactaccag 7200
 tacacctca acctccccta caagtccaag tgtccacag ccacccccag gaactggaga 7260
 tagactacta gctctagtca aaggagccta tcaggcgtt aacctacca atcccacaa 7320
 gaccaagaa tgttggctgt gcttagtgc gggacctct tattacgaag gagtagcgg 7380
 cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca cttccaaca 7440
 taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac 7500

tcaccaggcc ttatgtaaca ccaccaaag cgccggctca ggatcctact acctgcagc 7560
 acccgccgga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccaggt 7620
 gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taatttacca 7680
 ctccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt 7740
 atcattgacc ctggcccttc tactaggagg attaaccatg ggagggatg cagctggaat 7800
 agggacgggg accactgct taattaaaac ccagcagttt gagcagcttc atgccctat 7860
 ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920

gttgtctgaa gtagtctac agaaccgcag aggcctagat ttgctattcc taaaggagg 7980
 aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt 8040
 gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100
 aggccaagga tggttcgaag ggctgtttaa tagatcccc tggtttacca cettaatctc 8160
 caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct 8220
 caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac 8280
 tcagcaatat caccagctaa aacctataga gtacgagcca ggaagcggag agggcagagg 8340

aagtcttcta acatgcggtg acgtggagga gaatcccggc cctggcgcgc ctatggtgac 8400
 cggcggcatg gctccaagt gggatcaaaa gggcatggat atcgtttacg aggaggccct 8460
 gctgggctac aaggaggcgc gcgtgcctat cggcggctgt ctgatcaaca acaaggacgg 8520

cagtggtgctg ggcaggggcc acaacatgag gttccagaag ggctccgcca cctgacacgg 8580
 cgagatctcc accctggaga actgtggcag gctggagggc aaggtgtaca aggacaccac 8640
 cctgtacacc accctgtccc cttgtgacat gtgtaccggc gctatcatca tgtacggcat 8700
 ccctaggtgt gtgatcggcg agaacgtgaa cttcaagtcc aaggcgaga agtacctgca 8760

 aaccaggggc cagcaggtgg tggttgttga cgatgagagg tgtaagaagc tgatgaagca 8820
 gttcatcgac gagaggcctc aggactggtt cgaggatata ggcgagtaag cggccgcaga 8880
 taaaataaaa gatTTTTTTT agtctccaga aaaagggggg aatgaaagac cccacctgta 8940
 ggtttggcaa gctagcttaa gtaacccat tttgcaaggc atggaaaaat acataactga 9000
 gaatagagaa gttcagatca aggtcaggaa cagatggaac agctgaatat gggccaaaca 9060
 ggatatctgt ggtaagcagt tctgccccg gctcagggcc aagaacagat ggaacagctg 9120
 aatatgggcc aaacaggata tctgtggtaa gcagttctctg ccccgctca gggccaagaa 9180

 cagatggtcc ccagatcggg tccagccctc agcagtttct agagaacat cagatgttcc 9240
 cagggtgccc caaggacctg aaatgacct gtgccttatt tgaactaacc aatcagttcg 9300
 cttctcgctt ctgttcgcg gcttctgctc cccgagetca ataaaagagc ccacaacccc 9360
 tcactcgggg cgccagtctc ccgattgact gagtcgccc ggtaccctg tatccaataa 9420
 accctcttgc agttgcatcc gacttgtggt ctgcctgttc cttgggaggg tctcctctga 9480
 gtgattgact acccgtcagc gggggtcttt cattacatgt gagcaaaagg ccagcaaaag 9540
 gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg cccccctgac 9600

 gagcatcaca aaaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga 9660
 taccaggcgt ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt 9720
 accggatacc tgtccgctt tctcccttcg ggaagcgtgg cgctttctca atgctcacgc 9780
 tgtaggtate tcagttcggg ttaggtcgtt cgtccaagc tgggctgtgt gcacgaaccc 9840
 cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caaccggta 9900
 agacacgact tatgccact ggcagcagcc actggtaaca ggattagcag agcgaggat 9960
 gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac tagaaggaca 10020

 gtatttggta tctgcgctct gctgaagcca gttacctcg gaaaaagagt tggtagctct 10080
 tgatccggca acaaacacc cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt 10140
 acgcgagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct 10200
 cagtggaacg aaaactcacg ttaaggat tggatcatga gattatcaa aaggatcttc 10260
 acctagatcc ttttaatta aaaatgaagt tttaatcaa tctaaagtat atatgagtaa 10320
 acttggctg acagttacca atgcttaac agtgaggcac ctatctcagc gatctgtcta 10380

ttctgttcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc 10440

 ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat 10500
 ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtggctc tgcaacttta 10560
 tccgcctcca tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt 10620
 aatagtttgc gcaacgttgt tgccattgct gcaggcatcg tgggtgcacg ctctgctgtt 10680
 ggtatggctt cattcagctc cggttcccaa cgatcaagge gagttacatg atccccatg 10740
 ttgtgcaaaa aagcggttag ctcttcggf cctccgatcg ttgtcagaag taagtggcc 10800
 gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc 10860

 gtaagatgct ttictgtgac tggtagtac tcaaccaagt cattctgaga atagtgtatg 10920
 cggcgaccga gttgctcttg cccggcgtca acacgggata ataccgcgc acatagcaga 10980
 actttaaaag tgetcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta 11040
 ccgctgttga gatccagttc gatgtaacc actcgtgcac ccaactgatc ttcagcatct 11100
 ttactttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag 11160
 ggaataaggg cgacacggaa atgttgaata ctcatactct tcctttttca atattattga 11220
 agcatttacc agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat 11280

 aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc 11340
 attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa 11400
 gaattcat 11408

 <210> 53
 <211> 11402
 <212> DNA
 <213> Artificial Sequence
 <220><223> pAC3-P2A-yCD2
 <400> 53

 tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg 60
 cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc cccgccatt 120

 gacgtcaata atgacgtatg tteccatagt aacccaata gggactttcc attgacgtca 180
 atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 240
 aagtacccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta 300
 catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 360
 catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420

attccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480
 ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg gtaggcgtgt 540

 acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg 600
 actgagtcgc ccgggtaccc gtgtatcaa taaacctct tgcagttgca tccgacttgt 660
 ggtctcgctg ttcttggga gggctctctc tgagtattg actaccctc agcgggggtc 720
 tttcatttgg ggctctgccc gggatcggga gaccctgcc cagggaccac cgaccacca 780
 ccgggaggta agctggccag caactatct gtgtctgtcc gattgtctag tgtctatgac 840
 tgattttatg cgctgcgctc ggtactagt agctaactag ctctgtatct ggcggaccgc 900
 tggtggaact gacgagttcg gaacaccccg ccgaaccct gggagacgct ccagggactt 960

 cgggggccgt tttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg 1020
 gtgcaccccc cttagaggag ggatatgtgg tttcgttagg agacgagaac ctaaacagtt 1080
 tcccgcctcc gtctgaattt ttgcttccg tttgggaccg aagccgcgcc gcgcgtcttg 1140
 tctgtgcag catcgttctg tgttctctc gtctgactgt gtttctgtat ttgtctgaga 1200
 atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg 1260
 agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct 1320
 ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380

 tcatcaccca ggtaagatc aaggcttttt cacctggccc gcatggacac ccagaccagg 1440
 tcccctacat cgtgacctgg gaagccttgg cttttgacct cctccctgg gtcaagcctt 1500
 ttgtacacce taagcctccg cctctcttc ctccatccgc cccgtctctc ccccttgaac 1560
 ctctctgttc gaccccgctt cgatctctcc ttatccagc cctcactct tctctaggcg 1620
 ccaaacctaa acctcaagtt ctttctgaca gtggggggcc gctcatcgac ctacttacag 1680
 aagaccccc gccttatagg gacccaagac caccctctc cgacagggac ggaaatggtg 1740
 gagaagcgac ccctgcggga gaggcaccgg acccctccc aatggcatct cgctacgtg 1800

 ggagacggga gcccctgtg gccgactcca ctacctgca ggcattcccc ctccgcgacg 1860
 gaggaaacgg acagcttcaa tactggccgt tctctcttc tgacctttac aactggaaaa 1920
 ataataacce ttcttttct gaagatccag gtaaactgac agctctgatc gactctgttc 1980
 tcatcaccca tcagcccacc tgggacgact gtgagcagct gttggggact ctgctgaccg 2040
 gagaagaaaa acaacgggtg ctcttagagg ctgaaaggc ggtgcggggc gatgatgggc 2100
 gccccactca actgccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg 2160

attacaccac ccaggcaggt aggaaccacc tagtccacta tgc ccagttg ctcttagcgg 2220

 gtctcaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280
 ggcccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaggtaca 2340
 ctcttatga ccctgaggac ccagggaag aaactaatgt gtctatgtct tcatttggc 2400
 agtctgcccc agacattggg agaaagttag agaggttaga agatttaaaa aacaagacgc 2460
 ttggagattt ggtagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520
 gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580
 agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640

 ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg 2700
 atcgcgacca gtgtgcctac tgcaaagaaa aggggcactg ggctaaagat tgtccaaga 2760
 aaccacgagg acctcgggga ccaagacccc agacctcct cctgaccta gatgactagg 2820
 gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac 2880
 ccgtcacctt cctggtagat actggggccc aacctcctg gctgaccaa aatcctggac 2940
 ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgtgga 3000
 ccacggatcg caaagtacat ctagctaccg gtaaggtcac ccactcttc ctccatgtac 3060

 cagactgtcc ctatcctctg ttaggaagag atttgcctgac taaactaaaa gcccaaatcc 3120
 actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga 3180
 ccctaatat agaagatgag catcggttac atgagacctc aaaagagcca gatgtttctc 3240
 taggtccac atggctgtct gattttctc aggcctgggc ggaaaccggg ggcatgggac 3300
 tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360
 taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420
 tgttggacca gggaatactg gtaccctgcc agtcccctg gaacacgccc ctgctacceg 3480

 ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc 3540
 ggggtggaaga catccacccc accgtgcca acccttaca cctcttgagc gggtccccac 3600
 cgtcccacca gtggtacact gtgcttgatt taaaggatgc cttttctgc ctgagactcc 3660
 accccaccag tcagcctctc ttcgcctttg agtggagaga tccagagatg ggaatctcag 3720
 gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780
 aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atcctgctac 3840
 agtacgtgga tgacttactg ctggccgcca ctctgagct agactgcaa caaggtactc 3900

gggccctgtt acaaaccta gggaacctcg ggtatcgggc ctcgccaag aaagccaaa 3960
 ttgcccagaa acaggtcaag tatctgggt atcttctaaa agagggtcag agatggctga 4020
 ctgaggccag aaaagagact gtgatggggc agcctactcc gaagaccct cgacaactaa 4080
 gggagtccct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140
 cagccccctt gtacctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200
 aaaagccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttccag 4260
 atttgactaa gcctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc 4320

 taacgcaaaa actgggacct tggcgtcggc cggtaggcta cctgtcaaaa aagctagacc 4380
 cagtagcagc tgggtggccc ccttgccctac ggatggtagc agccattgcc gtactgaaa 4440
 aggatgcagg caagctaacc atgggacagc cactagtcat tctggcccc catgcagtag 4500
 aggcactagt caaacaacc cccgaccgt ggctttccaa cgcccggatg actcactatc 4560
 aggccttgc tttggacag gaccgggtcc agttcggacc ggtgtagacc ctgaaccgg 4620
 ctacgtgct cccactgct gaggaaggc tgcaacaaa ctgccttgat atcctggccg 4680
 aagcccagg aaccgacc gacctaacgg accagccgt cccagacc gaccacact 4740

 ggtacacgga tggaaagcgt ctcttacaag agggacagc taaggcggga gctgcggtga 4800
 ccaccgagac cgagtaatc tgggctaaag cctgccagc cgggacatcc gctcagcggg 4860
 ctgaactgat agcactcacc caggecctaa agatggcaga aggtagaag ctaaatgttt 4920
 atactgatag ccgttatgct tttgctactg cccatatcca tggagaaata tacagaaggc 4980
 gtgggttgc cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac 5040
 taaaagccct ctttctgcc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100
 gacacagcgc cgaggctaga ggcaaccgga tgctgacca agcggcccga aaggcagcca 5160

 tcacagagac tcagacacc tctacctcc tcatagaaaa tteatcacc tacacctcag 5220
 aacatttca ttacacagt actgatataa aggacctaac caagttgggg gccatttatg 5280
 ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340
 ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc 5400
 tcctagagag aagccacagt ccctactaca tctgaaccg ggatcgaaca ctcaaaaata 5460
 tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg 5520
 gaactagggt ccgcccgcgt cggcccggca ctattggga gatcgattc accgagataa 5580

 agcccggatt gtatggctat aaatatctc tagttttat agatacttt tctggctgga 5640
 tagaacctt cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg 5700
 agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760

tctccaaggt gaggcagaca gtggccgata tgttggggat tgattggaaa ttacattgtg 5820
catacagacc ccaagctca ggccaggttag aaagaatgaa tagaacctc aaggagactt 5880
taactaaatt aacgcttgca actggctcta gagactgggt gtcctactc cccttagccc 5940
tgtaccgagc ccgcaacacg ccgggcccc atggcctcac cccatatgag atcttatatg 6000

gggcaccccc gcccttgta aacttccctg accctgacat gacaagagt actaacagcc 6060
cctctctcca agctcactta caggetctct acttagtcca gcacgaagtc tggagacctc 6120
tggcggcagc ctaccaagaa caactggacc gaccggtggg acctcacct taccgagtcg 6180
gcgacacagt gtgggtccg cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240
cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcctc gcagcttggg 6300
tacacgccg ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat 6360
ggcgcgttca acgctctcaa aacccctca agataagatt aaccctgga agccctaat 6420

agtcatggga gtctgttag gaggtaggat ggacagagc ccccatcagg tctttaatgt 6480
aacctggaga gtcaccaacc tgatgactgg gctaccgcc aatgccact cctcctggg 6540
aactgtacaa gatgccttc caaaattata tttgatcta tgtgatctgg tggagagga 6600
gtgggacct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgagggag 6660
acagcggacc cggacttttg acttttacgt gtcccctggg cataccgtaa agtcggggtg 6720
tgggggacca ggagagggt actgtggtaa atgggggtgt gaaaccaccg gacagctta 6780
ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcgtaaca cccctggga 6840

cacgggatgc tctaaagtg cctgtgccc ctgctacgac ctctcaaag tatccaattc 6900
cttccaaggg gctactcgag ggggcagatg caaccctcta gtcctagaat tactgatgc 6960
aggaaaaaag gtaactggg acgggcccga atcgtgggga ctgagactgt accggacagg 7020
aacagatcct attaccatgt tctcctgac ccggcaggtc cttaatgtgg gaccccgagt 7080
ccccatagg ccaaccagc tattaccga ccaaagactc ccttctcac caatagagat 7140
tgtaccggt ccacagccac ctageccct caataccagt taccctt cactaccag 7200
tacacctca acctcccta caagtccaag gtccccacag ccacccag gaactggaga 7260

tagactacta gctctagca aaggagccta tcaggcgtt aacctacca atcccacaa 7320
gaccaagaa tgttggctgt gcttagtgc gggacctct tattacgaag gaggtaggt 7380
cgtggcact tataccaatc attcaccgc tccggccaac tgtacggca cttccaaca 7440
taagcttacc ctatctgaag tgacaggaca ggcctatgc atggggcag tacctaaac 7500
tcaccagcc ttatgtaaca ccaccaaag cgcggctca ggatcctact acctgcagc 7560
acccgccga acaatgtggg cttgcagcac tggattgact cctgcttgt ccaccaggt 7620

gctcaatcta accacagatt atttgtatt agttgaactc tggcccagag taatttacca 7680

ctccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt 7740

atcattgacc ctggcccttc tactaggagg attaaccatg ggagggatg cagctggaat 7800

agggacgggg accactgcct taattaaaac ccagcagttt gagcagcttc atgccgctat 7860

ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920

gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgctattcc taaaggaggg 7980

aggctctctg gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt 8040

gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100

aggccaagga tggttcgaag ggctgtttta tagatcccc tggtttacca ccttaatctc 8160

caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac ctgcatctt 8220

caatcgattg gtccaatttg ttaaagacag gatctcagtg gtcacagctc tggttttgac 8280

tcagcaatat caccagctaa aacctataga gtacgagcca gctactaact tcagcctgct 8340

gaagcaggct ggagacgtgg aggagaacc tggacctggc gcgcctatgg tgaccggcgg 8400

catggcctcc aagtgggatc aaaaggcat ggatatcgt tacgaggagg cctgctggg 8460

ctacaaggag ggcggcgtgc ctatcggcgg ctgtctgac aacaacaagg acggcagtgt 8520

gctgggcagg ggccacaaca tgaggttcca gaaggctcc gccacctgc acggcgagat 8580

ctccacctg gagaactgtg gcaggctgga gggcaagggtg tacaaggaca ccacctgta 8640

caccacctg tccccttgat acatgtgtac cggcgtatc atcatgtac gcatccctag 8700

gtgtgtgatc ggcgagaacg tgaacttcaa gtccaagggc gagaagtacc tgcaaaccag 8760

gggccacgag gtggtggttg ttgacgatga gagggtgaag aagctgatga agcagttcat 8820

cgacgagagg cctcaggact ggttcgagga tatcggcgag taagcggccg cagataaaat 8880

aaaagatfff atttagtctc cagaaaaagg ggggaatgaa agaccccacc ttaggtttg 8940

gcaagctagc ttaagtaacg ccattttgca aggcatggaa aaatacataa ctgagaatag 9000

agaagttcag atcaaggta ggaacagatg gaacagctga atatgggcca aacaggatat 9060

ctgtgtaag cagttcctgc cccggctcag ggccaagaac agatggaaca gctgaatatg 9120

ggccaacag gatatctgtg gtaagcagtt cctgccccg ctgaggcca agaacagatg 9180

gtccccagat gcggtccagc cctcagcagt ttctagagaa ccatcagatg tttccagggt 9240

gccccaaagga cctgaaatga ccctgtgcct tatttgaact aaccaatcag ttcgcttctc 9300

gcttctgttc gcgcgttct gctccccgag ctcaataaaa gagcccacaa cccctcactc 9360

ggggcgccag tcctccgatt gactgagtcg cccgggtacc cgtgtatcca ataaacctc 9420
 ttgcagttgc atccgacttg tggctcgcct gttccttggg agggctctct ctgagtgatt 9480
 gactaccctg cagcgggggt ctttcattac atgtgagcaa aaggccagca aaaggccagg 9540
 aaccgtaaaa aggccgcggt gctggcgctt ttccataggc tccgcccc tgacgagcat 9600
 cacaaaaatc gagctcaag tcagaggtgg cgaaacccga caggactata aagataccag 9660
 gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga 9720
 tacctgtccg cttttctccc ttcgggaagc gtggcgcttt ctcaatgctc acgctgtagg 9780

 tatctcagtt cgggtgtagt cgttcgctcc aagctgggct gtgtgcacga accccccgtt 9840
 cagcccagacc gctcgcctt atccggtaac tatcgtcttg agtccaacc ggtaagacac 9900
 gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc 9960
 ggtgctacag agttcttgaa gtgggtggcct aactacggct aactagaag gacagtattt 10020
 ggtatctcgc ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc 10080
 ggcaaaaaa ccaccgctgg tagcgggtgt tttttgttt gcaagcagca gattacgcgc 10140
 agaaaaaag gatctcaaga agatcctttg atcttttcta cggggctctga cgctcagtgg 10200

 aacgaaaact cacgttaagg gatthtggtc atgagattat caaaaaggat cttcacctag 10260
 atccttttaa attaaaaatg aagttttaa tcaatctaaa gtatatatga gtaaaacttg 10320
 tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg tctatttctg 10380
 tcatccatag ttgcctgact ccccgctcgt tagataacta cgatacggga gggttacca 10440
 tctggcccca gtgctgcaat gataccgcga gaccacgct caccggctcc agatttatca 10500
 gcaataaacc agccagccgg aagggccgag cgcagaagtg gtctctgcaac tttatccgcc 10560
 tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc agttaatagt 10620

 ttgcgcaacg ttgttgccat tgctgcaggc atcgtggtgt cacgctcgtc gtttggtatg 10680
 gcttcattca gctccggttc ccaacgatca aggcgagtta catgatccc catgttgtgc 10740
 aaaaaagcgg ttagctcctt cggctctcgc atcgttgtca gaagtaagt ggccgcagtg 10800
 ttatcactca tggttatggc agcactgcat aattctetta ctgtcatgcc atccgtaaga 10860
 tgcttttctg tgactgggta gtactcaacc aagtcattct gagaatagt tatgcggcga 10920
 ccgagttgct cttgcccggt gtaacacgg gataataccg cgccacatag cagaacttta 10980
 aaagtgctca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat cttaccgctg 11040

 ttgagatcca gttcgatgta acccactcgt gcaccaact gatcttcagc atcttttact 11100
 ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata 11160
 agggcgacac ggaatgttg aatactcata ctcttccttt ttcaatatta ttgaagcatt 11220

tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacia 11280
 ataggggttc cgcgcacatt tccccgaaaa gtgccactg acgtctaaga aaccattatt 11340
 atcatgacat taacctataa aaataggcgt atcacgagge cctttcgtct tcaagaattc 11400
 at 11402

<210> 54

<211> 11411

<212> DNA

<213> Artificial Sequence

<220><223> pAC3-GSG-P2A-yCD2

<400> 54

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagtccg 60
 cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 120
 gacgtcaata atgacgtatg ttcccatagt aacccaata gggactttcc attgacgtca 180
 atgggtggag tatttacggt aaactgccca ctggcgagta catcaagtgt atcatatgcc 240
 aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta 300

 catgacctta tgggacttcc ctacttgcca gtacatctac gtattagtca tcgctattac 360
 catggtgatg cggttttggc agtatacaaa tgggcgtgga tagcggtttg actcacgggg 420
 attccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480
 ggactttcca aaatgctgta acaactccgc ccattgacg caaatgggcg gtaggcgtgt 540
 acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg 600
 actgagtgcg ccgggtaccg gtgtatcaa taaacctct tgcagttgca tccgacttgt 660
 ggtctcgtg ttcttggga gggctcctc tgagtgattg actaccctc agcgggggtc 720

 tttcatttgg gggctcgtcc gggatcggga gaccctgcc cagggaccac cgaccacca 780
 ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840
 tgattttatg cgctgcgtc ggtactagt agctaaactag ctctgtatct ggcggaccag 900
 tgggtggaact gacgagttcg gaacacccgg ccgcaaccct gggagacgtc ccagggactt 960
 cgggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg 1020
 gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt 1080
 tcccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgctcttg 1140

 tctgtgcag catcgttctg tgttctctct gtctgactgt gtttctgtat ttgtctgaga 1200
 atatgggcca gactgttacc actcccttaa gtttgactt aggtcactgg aaagatgtcg 1260

agcggatcgc tcacaaccag tCGgtagatg tcaagaagag acgttgggtt accttctgct 1320
 ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc 1380
 tcatcaccca ggftaagatc aaggctcttt cacctggccc gcatggacac ccagaccagg 1440
 tcccctacat cgtgacctgg gaagccttgg cttttgacct ccctccctgg gtcaagccct 1500
 ttgtacacce taagcctccg cctcctcttc ctccatccgc cccgtctctc ccccttgaac 1560

 ctctctgctt gaccccgctt cgatectecc ttatccage cctcactcct tetctaggcg 1620
 ccaaacttaa acctcaagtt ctttctgaca gtggggggcc gctcatcgac ctacttacag 1680
 aagaccccc gccttatagg gacccaagac caccctctc cgacagggac ggaatggtg 1740
 gagaagcgac ccctgcggga gaggcaccgg accctcccc aatggcatct cgcctacgtg 1800
 ggagacggga gccccctgtg gccgactcca ctacctgca ggcatcccc ctccgcgacg 1860
 gaggaacagg acagcttcaa tactggcctt tctcctcttc tgacctttac aactggaaaa 1920
 ataataacce ttcttttct gaagatccag gtaaactgac agctctgac gagtctgttc 1980

 tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg 2040
 gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc 2100
 gccccactca actgcccact gaagtcgatg ccgcttttcc cctcagcgc ccagactggg 2160
 attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagtgt ctcttagcgg 2220
 gtctccaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag 2280
 ggcccactga gtctccctcg gccttcttag agagacttaa ggaagcctat cgcaggtaca 2340
 ctcttatga ccctgaggac ccagggaag aaactaatgt gtctatgtct ttcatttggc 2400

 agtctgcccc agacattggg agaaagttag agaggtaga agatttaaaa aacaagacgc 2460
 ttggagattt ggtagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa 2520
 gagaggaacg tctcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg 2580
 agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca 2640
 ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccactcg 2700
 atcgcgacca gtgtgcctac tgcaaagaaa aggggcactg ggctaaagat tgtccaaga 2760
 aaccacgagg acctcgggga ccaagacccc agacctcct cctgacccta gatgactagg 2820

 gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac 2880
 ccgtcacctt cctggtagat actggggccc aacctcctg gctgacccaa aatcctggac 2940
 ccctaagtga taagtctgcc tgggtccaag gggtactgg aggaaagcgg tctcgtgga 3000
 ccacggatcg caaagtacat ctagnetacc gtaaggtcac cactcttct ctccatgtac 3060
 cagactgtcc ctatcctctg ttaggaagag atttgctgac taaactaaaa gcccaatcc 3120

actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga 3180
ccctaataat agaagatgag catcggttac atgagacctc aaaagagcca gatgtttctc 3240

tagggtccac atggctgtct gattttcttc aggcctgggc ggaaaccggg ggcattgggac 3300
tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360
taaaacaata ccccatgtca caagaagcca gactggggat caagcccccac atacagagac 3420
tgttggacca gggaatactg gtaccctgcc agtccccctg gaacacgccc ctgctacccg 3480
ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc 3540
gggtggaaga catcccccc accgtgcca acccttacia cctcttgagc gggtccccac 3600
cgtcccacca gtggtacact gtgcttgatt taaaggatgc ctttttctgc ctgagactcc 3660

accccaccag tcagcctctc ttgcctttg agtggagaga tccagagatg ggaatctcag 3720
gacaattgac ctggaccaga ctccccaggg gtttcaaaaa cagtcccacc ctgtttgatg 3780
aggcactgca cagagacctc gcagacttcc ggatccagca cccagacttg atcctgctac 3840
agtacgtgga tgacttactg ctggccgcca ctctgagct agactgcca caaggtactc 3900
gggcctgtt acaaacctc gggaacctcg ggtatcgggc ctggccaag aaagcccaaa 3960
tttgccagaa acaggtcaag tatctggggg atcttctaaa agagggtcag agatggctga 4020
ctgaggccag aaaagagact gtgatggggc agcctactcc gaagaccctc cgacaactaa 4080

gggagttcct agggacggca ggcttctgtc gcctctggat ccctgggttt gcagaaatgg 4140
cagccccctt gtacctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200
aaaagccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttccag 4260
atgtgactaa gcctttgaa ctctttgtcg acgagaagca gggctacgcc aaagggttcc 4320
taacgcaaaa actgggacct tggcgtcggc cgggtggccta cctgtccaaa aagctagacc 4380
cagtagcagc tgggtggccc ccttgccctac ggatggtagc agccattgcc gtactgacaa 4440
aggatgcagg caagctaacc atgggacagc cactagtcat tctggcccc catgcagtag 4500

aggcactagt caaacaacc cccgaccgt ggctttccaa cgcccggatg actcactatc 4560
aggccttgct ttggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaaccgg 4620
ctacgtgct cccactgct gaggaaggc tgcaacacaa ctgccttgat atcctggccc 4680
aagcccacgg aaccgacc gacctaacgg accagccgt cccagacgcc gaccacacct 4740
ggtacacgga tggaagcagt ctcttacaag agggacagcg taaggcggga gctgcggtga 4800
ccaccgagac cgaggaatc tgggctaaag cctgccagc cgggacatcc gctcagcggg 4860

ctgaactgat agcactcacc caggccctaa agatggcaga aggtaagaag ctaaagtgtt 4920

atactgatag ccgttatgct tttgctactg cccatatcca tggagaaata tacagaaggc 4980

gtgggttgc cacaacagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac 5040

taaaagccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg 5100

gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggcccga aaggcagcca 5160

tcacagagac tcagacacc tctacacctc tcatagaaaa tteatcaccc tacacctcag 5220

aacattttca ttacacagtg actgatataa aggacctaac caagtgggg gccatttatg 5280

ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt 5340

ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc 5400

tcctagagag aagccacagt ccctactaca tgcgaaccg ggatcgaaca ctcaaaaata 5460

tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg 5520

gaactagggt ccgctggcat cggcccggca ctattggga gatcgattc accgagataa 5580

agccccgatt gtatggctat aaatatctt tagttttat agatacctt tctggctgga 5640

tagaacctt cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg 5700

agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg 5760

tctccaaggt gactcagaca gtggccgatc tgttgggat tgattgaaa ttacattgtg 5820

catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaacctc aaggagactt 5880

taactaaatt aacgcttga actgctctc gagactgggt gctcctact cccttagccc 5940

tgtaccgagc ccgaacacg ccgggcccc atggcctcac cccatagag atcttatatg 6000

gggcaccccc gcccttcta aacttccctg acctgacat gacaagagt actaacagcc 6060

cctctctcca agctcactta cagctctct acttagtcca gcacgaagtc tggagacctc 6120

tggcggcagc ctaccaagaa caactggacc gaccggtggt acctcaccct taccgagteg 6180

gacacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac 6240

cttacacagt cctgctgacc acccccaccg cctcaaaagt agacggcatc gcagcttga 6300

tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat 6360

ggcgcgttca acgctctcaa aaccctca agataagatt aaccctgga agccctaat 6420

agtcatggga gtctgttag gactagggat ggcagagagc cccatcagg tctttaatgt 6480

aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct cctcctggg 6540

aactgtacaa gatgccttcc caaaattata tttgatcta tgtgatctgg tcggagagga 6600

gtgggaccct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc cgcagggag 6660
acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcggggtg 6720
tgggggacca ggagagggt actgtggtaa atgggggtgt gaaaccaccg gacaggctta 6780
ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcggtaaca cccctggga 6840
cacgggatgc tctaaagtgt cctgtggccc ctgctacgac ctctcaaag tatccaattc 6900
cttccaaggg gctactcgag ggggcagatg caacctcta gtctagaat tctactgatgc 6960
aggaaaaaag gctaaactggg acgggcccaa atcgtgggga ctgagactgt accggacagg 7020

aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccaggt 7080
ccccataggg cccaaccag tattaccga ccaaagactc cttcctcac caatagagat 7140
tgtaccggt ccacagccac ctagccccct caataccagt taccctctt cactaccag 7200
tacacctca acctcccta caagtccaag tgtccacag ccacccag gaactggaga 7260
tagactacta gcctagica aaggagccta tcaggcgctt aacctacca atccgacaa 7320
gaccaagaa tgttggctgt gcttagtgc gggacctct tattacgaag gtagcgggt 7380
cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca ctccaaca 7440

taagcttacc ctatctgaag tgacaggaca gggcctatgc atggggcag tacctaaac 7500
tcaccaggcc ttatgtaaca ccaccaaag cgcggctca ggatcctact accttcagc 7560
accgccgga acaatgtggg cttgcagcac tggattgact cctgcttgt ccaccaggt 7620
gctcaatcta accacagatt attgtatatt agttgaactc tggcccagag taatttacc 7680
ctccccgat tatatgtatg gtcagcttga acagcgtacc aatataaaa gagagccagt 7740
atcattgacc ctggcccttc tactaggagg attaaccatg ggagggattg cagctggaat 7800
agggacgggg accactgct taattaaaac ccagcagttt gacgacttc atgccgctat 7860

ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc 7920
gttgtctgaa gtagtctac agaaccgag aggcctagat ttgctattcc taaaggagg 7980
aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cgggctagt 8040
gagagacagc atggccaat taagagaaag gcttaatcag agacaaaaac tatttgagac 8100
aggccaagga tggttcgaag ggctgtttaa tagatcccc tggttacca ccttaatctc 8160
caccatcatg ggacctctaa tagtactctt actgatctta ctcttggac cttgactct 8220
caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggtttgac 8280

tcagcaatat caccagctaa aaccataga gtacgagcca ggaagcggag ctactaactt 8340
cagcctgctg aagcagctg gagactgga ggagaacct ggacctggcg gcctatggt 8400
gaccggcggc atggcctcca agtgggatca aaaggcatg gatatcgctt acgaggaggc 8460

cctgctgggc tacaaggagg gcggcgtgcc tatcggcggc tgtctgatca acaacaagga 8520
 cggcagtgtg ctgggcaggg gccacaacat gaggttccag aagggtccg ccaccctgca 8580
 cggcgagatc tccaccctgg agaactgtgg caggctggag ggcaaggtgt acaaggacac 8640
 caccctgtac accaccctgt ccccttgtga catgtgtacc ggcgctatca tcatgtacgg 8700

 catccctagg tgtgtgatcg gcgagaacgt gaacttcaag tccaagggcg agaagtacct 8760
 gcaaaccagg ggccacgagg tgggtggtgt tgacgatgag aggtgtaaga agctgatgaa 8820
 gcagttcatc gacgagaggc ctcaggactg gttcgaggat atcggcgagt aagcggccgc 8880
 agataaaata aaagatttta tttagtctcc agaaaaaggg gggaatgaaa gaccccacct 8940
 gtaggtttgg caagctagct taagtaacgc cttttgcaa ggcatggaaa aatacataac 9000
 tgagaataga gaagttcaga tcaaggtcag gaacagatgg aacagctgaa tatgggccaa 9060
 acaggatatac tgtgtaagc agttctctgcc cggctcagg gccaagaaca gatggaacag 9120

 ctgaatatgg gccaaacagg atatctgtgg taagcagttc ctgccccggc tcagggccaa 9180
 gaacagatgg tccccagatg cggctccagcc ctcagcagtt tctagagaac catcagatgt 9240
 ttccagggtg cccaaggac ctgaaatgac cctgtgectt atttgaacta accaatcagt 9300
 tcgcttctcg ctctgttcg cgcgcttctg ctccccgagc tcaataaaag agcccacaac 9360
 ccctcactcg gggcgccagt cctccgattg actgagtcgc cgggtaccg gtgtatccaa 9420
 taaacctct tgcagttgca tccgacttgt ggtctcctg ttccttggga gggtctctc 9480
 tgagtgattg actaccctgc agcgggggtc tttcattaca tgtgagcaaa aggccagcaa 9540

 aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggtc cgcceccct 9600
 gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaaccgac aggactataa 9660
 agataccagg cgtttcccc tggaaactcc ctctgtcgtc ctctgttcc gaccctgccg 9720
 cttaccggat acctgtcgc ctttctccct tcgggaagcg tggcgcttcc tcaatgctca 9780
 cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa 9840
 cccccgttc agcccgaccg ctgcgcctta tccgtaact atcgtcttga gtccaaccg 9900
 gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg 9960

 tatgtaggcg gtgctacaga gttcttgaag tggtagccta actacggcta cactagaagg 10020
 acagtatttg gtatctgcgc tctgtgaag ccagttacct tcggaaaaag agttggtagc 10080
 tcttgatccg gaaacaaac caccctggt agcgggtggt tttttgttg caagcagcag 10140
 attacgcga gaaaaaagg atctcaagaa gatcctttga tctttctac ggggtctgac 10200
 gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc 10260
 ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag 10320

taaacttggg ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt 10380

 ctatttcggt catccatagt tgcctgactc cccgtcgtgt agataactac gatacgggag 10440
 ggcttaccat ctggccccag tgctgcaatg ataccgcgag acccacgctc accggetcca 10500
 gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg tcttgcaact 10560
 ttatccgctt ccatccagtc tattaattgt tgcgggaag ctagagtaag tagttcgcca 10620
 gttaatagtt tgcgcaacgt tgttgccatt gctgcaggca tcgtgggtgc acgctcgtcg 10680
 tttggatagg ctccattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc 10740
 atgttgtgca aaaaagcggg tagctccttc ggtcctccga tcgttgtcag aagtaagttg 10800

 gccgcagtgt taccactcat ggttatggca gcaactgcata attctcttac tgtcatgcca 10860
 tccgtaagat gcttttctgt gactgggtgag tactcaacca agtcattctg agaatagttg 10920
 atgcccggac cgagttgctc ttgcccggcg tcaaacgggg ataataccgc gccacatagc 10980
 agaactttaa aagtgctcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc 11040
 ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca 11100
 tcttttactt tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa 11160
 aagggaataa gggcgacacg gaaatgttga atactcatal tcttcctttt tcaatattat 11220

 tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 11280
 aataaaciaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 11340
 accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtctt 11400
 caagaattca t 11411

 <210> 55
 <211> 20
 <212> PRT
 <213> Artificial Sequence
 <220><223> Equine rhinitis A virus 2A peptide
 <400> 55
 Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly Asp Val Glu Ser

1 5 10 15
 Asn Pro Gly Pro
 20
 <210> 56
 <211> 23

<212> PRT
 <213> Artificial Sequence
 <220><223> Foot-and-mouth disease 2A peptide
 <400> 56
 Pro Val Lys Gln Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp
 1 5 10 15
 Val Glu Ser Asn Pro Gly Pro
 20

<210> 57
 <211> 19
 <212> PRT
 <213> Artificial Sequence
 <220><223> Porcine teschovirus-1 2A peptide

<400> 57
 Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn
 1 5 10 15
 Pro Gly Pro

<210> 58
 <211> 18
 <212> PRT
 <213> Artificial Sequence
 <220><223> Thosea asigna virus 2A peptide

<400> 58
 Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro
 1 5 10 15
 Gly Pro

<210> 59
 <211> 25
 <212> PRT
 <213> Artificial Sequence
 <220><223> Encephalomyocarditis virus-B 2A peptide

<400> 59
 Gly Ile Phe Asn Ala His Tyr Ala Gly Tyr Phe Ala Asp Leu Leu Ile
 1 5 10 15

His Asp Ile Glu Thr Asn Pro Gly Pro
 20 25

<210> 60

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> Encephalomyocarditis virus-D 2A peptide

<400> 60

Gly Tyr Phe Ala Asp Leu Leu Ile His Asp Ile Glu Thr Asn Pro Gly
 1 5 10 15

Pro

<210> 61

<211> 25

<212> PRT

<213> Artificial Sequence

<220><223> Encephalomyocarditis virus-PV21 2A peptide

<400> 61

Arg Ile Phe Asn Ala His Tyr Ala Gly Tyr Phe Ala Asp Leu Leu Ile
 1 5 10 15

His Asp Ile Glu Thr Asn Pro Gly Pro
 20 25

<210> 62

<211> 25

<212> PRT

<213> Artificial Sequence

<220><223> Mengovirus 2A peptide

<400> 62

His Val Phe Glu Thr His Tyr Ala Gly Tyr Phe Ser Lys Leu Leu Ile
 1 5 10 15

His Asp Val Glu Thr Asn Pro Gly Pro

20 25

<210> 63

<211> 25

<212> PRT

<213> Artificial Sequence

<220><223> Theiler's encephalomyelitis virus-GD7 2A peptide

<400> 63

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

1 5 10 15

His Asp Val Glu Met Asn Pro Gly Pro

20 25

<210> 64

<211

> 25

<212> PRT

<213> Artificial Sequence

<220><223> Thieler's encephalomyelitis virus-DA 2A peptide

<400> 64

Arg Ala Val Arg Ala Tyr His Ala Asp Tyr Tyr Lys Gln Arg Leu Ile

1 5 10 15

His Asp Val Glu Met Asn Pro Gly Pro

20 25

<210> 65

<211> 25

<212> PRT

<213> Artificial Sequence

<220><223> Thieler's encephalomyelitis virus-BEAN 2A peptide

<400> 65

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

1 5 10 15

His Asp Val Glu Thr Asn Pro Gly Pro

20 25

<210> 66

<211> 25
 <212> PRT
 <213> Artificial Sequence
 <220><223> Theiler's-Like Virus 2A peptide
 <400> 66
 Lys His Val Arg Glu Tyr His Ala Ala Tyr Tyr Lys Gln Arg Leu Met
 1 5 10 15
 His Asp Val Glu Thr Asn Pro Gly Pro
 20 25

<210> 67
 <211> 26
 <212> PRT

<213> Artificial Sequence
 <220><223> Ljungan virus-174F 2A peptide
 <400> 67
 Met His Ser Asp Glu Met Asp Phe Ala Gly Gly Lys Phe Leu Asn Gln
 1 5 10 15
 Cys Gly Asp Val Glu Thr Asn Pro Gly Pro
 20 25

<210> 68
 <211> 26
 <212> PRT

<213> Artificial Sequence
 <220><223> Ljungan virus-145SL 2A peptide
 <400> 68
 Met His Asn Asp Glu Met Asp Tyr Ser Gly Gly Lys Phe Leu Asn Gln
 1 5 10 15

Cys Gly Asp Val Glu Ser Asn Pro Gly Pro
 20 25

<210> 69
 <211> 26
 <212> PRT

<213> Artificial Sequence

<220><223> Ljungan virus-(87-012) 2A peptide

<400> 69

Met His Ser Asp Glu Met Asp Phe Ala Gly Gly Lys Phe Leu Asn Gln

1 5 10 15

Cys Gly Asp Val Glu Thr Asn Pro Gly Pro

20 25

<210> 70

<211> 26

<212> PRT

<213> Artificial Sequence

<220>

<223> Ljungan Virus - (M1146) 2A peptide

<400> 70

Tyr His Asp Lys Asp Met Asp Tyr Ala Gly Gly Lys Phe Leu Asn Gln

1 5 10 15

Cys Gly Asp Val Glu Thr Asn Pro Gly Pro

20 25

<210> 71

<211> 24

<212> PRT

<213> Artificial Sequence

<220><223> Foot and Mouth Disease Virus 2A Peptide

<400> 71

Ala Pro Ala Lys Gln Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly

1 5 10 15

Asp Val Glu Ser Asn Pro Gly Pro

20

<210> 72

<211> 24

<212> PRT

<213> Artificial Sequence

<220><223> Foot and Mouth Disease Virus-A12 2A Peptide

<400> 72

Ala Pro Gly Lys Gln Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly

<210> 76
 <211> 24
 <212> PRT
 <213> Artificial Sequence
 <220><223> Foot and Mouth Disease Virus - O (Taiwan) 2A Peptide

<400> 76
 Ala Pro Ala Lys Gln Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly
 1 5 10 15
 Asp Val Glu Ser Asn Pro Gly Pro
 20

<210> 77
 <211> 24
 <212> PRT
 <213> Artificial Sequence
 <220><223> Foot and Mouth Disease Virus - O/SK 2A Peptide

<400> 77
 Ala Pro Val Lys Gln Leu Leu Ser Phe Asp Leu Leu Lys Leu Ala Gly
 1 5 10 15
 Asp Val Glu Ser Asn Pro Gly Pro

20
 <210> 78
 <211> 24
 <212> PRT
 <213> Artificial Sequence
 <220><223> Foot and Mouth Disease Virus - SAT3 2A Peptide

<400> 78
 Lys Pro Asp Lys Gln Met Cys Asn Phe Asp Leu Leu Lys Leu Ala Gly
 1 5 10 15
 Asp Val Glu Ser Asn Pro Gly Pro

20
 <210> 79
 <211> 24
 <212> PRT

<213> Artificial Sequence

<220><223> Foot and Mouth Disease Virus - SAT2 2A Peptide

<400> 79

Gly Val Ala Lys Gln Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly

1 5 10 15

Asp Val Glu Ser Asn Pro Gly Pro

20

<210> 80

<211> 24

<212> PRT

<213> Artificial Sequence

<220><223> Equine Rhinitis A Virus 2A Peptide

<400> 80

Asn Ile Asn Lys Gln Cys Thr Asn Tyr Ser Leu Leu Lys Leu Ala Gly

1 5 10 15

Asp Val Glu Ser Asn Pro Gly Pro

20

<210> 81

<211> 25

<212> PRT

<213> Artificial Sequence

<220><223> Equine Rhinitis B Virus 2A Peptide

<400> 81

Thr Ile Leu Ser Glu Gly Ala Thr Asn Phe Ser Leu Leu Lys Leu Ala

1 5 10 15

Gly Asp Val Glu Leu Asn Pro Gly Pro

20

25

<210> 82

<211> 25

<212> PRT

<213> Artificial Sequence

<220><223> Endogenous Retrovirus-3 2A Peptide

<400> 82

Asn Leu Leu Ser Gln Gly Ala Thr Asn Phe Asp Leu Leu Lys Leu Ala

1 5 10 15

Gly Asp Val Glu Ser Asn Pro Gly Pro

20 25

<210> 83

<211> 27

<212> PRT

<213> Artificial Sequence

<220><223> Punta Toro Virus-1 2A Peptide

<400> 83

Val Met Ala Phe Gln Gly Pro Gly Ala Thr Asn Phe Ser Leu Leu Lys

1 5 10 15

Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro

20 25

<210> 84

<211> 27

<212> PRT

<213> Artificial Sequence

<220><223> Punta Toro Virus-2 2A Peptide

<400> 84

Thr Met Met Leu Gln Gly Pro Gly Ala Thr Asn Phe Ser Leu Leu Lys

1 5 10 15

Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro

20 25

<210> 85

<211> 27

<212> PRT

<213> Artificial Sequence

<220><223> Punta Toro Virus-3 2A Peptide

<400> 85

Thr Met Ser Phe Gln Gly Pro Gly Ala Ser Ser Phe Ser Leu Leu Lys

1 5 10 15

Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro

20 25

<210> 86

<211> 27

<212> PRT

<213> Artificial Sequence

<220><223> Punta Toro Virus-4 2A Peptide

<400> 86

Thr Met Met Leu Gln Gly Pro Gly Ala Ser Asn Phe Ser Leu Leu Lys

1 5 10 15

Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro

20 25

<210>

87

<211> 27

<212> PRT

<213> Artificial Sequence

<220><223> Punta Toro Virus-5 2A Peptide

<400> 87

Thr Met Leu Phe Gln Gly Pro Gly Ala Ala Asn Phe Ser Leu Leu Arg

1 5 10 15

Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro

20 25

<210> 88

<211> 27

<212> PRT

<213> Artificial Sequence

<220><223> Punta Toro Virus-6 2A Peptide

<400> 88

Thr Met Ser Phe Gln Gly Pro Gly Ala Thr Asn Phe Ser Leu Leu Lys

1 5 10 15

Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro

20 25

<210> 89

<211> 27
 <212> PRT
 <213> Artificial Sequence
 <220><223> Punta Toro Virus-7 2A Peptide
 <400> 89
 Val Val Ser Phe Gln Gly Pro Gly Ala Thr Asn Phe Ser Leu Leu Lys
 1 5 10 15
 Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro
 20 25

<210>
 90
 <211> 27
 <212> PRT
 <213> Artificial Sequence
 <220><223> Punta Toro Virus-8 2A Peptide
 <400> 90
 Thr Met Ser Leu Gln Gly Pro Gly Ala Thr Asn Phe Ser Leu Leu Lys
 1 5 10 15
 Gln Ala Gly Asp Ile Glu Glu Asn Pro Gly Pro
 20 25

<210> 91
 <211> 27
 <212> PRT
 <213> Artificial Sequence
 <220><223> Punta Toro Virus-9 2A Peptide
 <400> 91
 Thr Met Ala Phe Gln Gly Pro Gly Ala Thr Asn Phe Ser Leu Leu Lys
 1 5 10 15
 Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro
 20 25

<210> 92
 <211> 27
 <212> PRT
 <213> Artificial Sequence

<220><223> Punta Toro Virus-10 2A Peptide

<400> 92

Thr Leu Ser Phe Gln Gly Pro Gly Ala Thr Asn Phe Ser Leu Leu Lys

1 5 10 15

Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro

20 25

<210

> 93

<211> 27

<212> PRT

<213> Artificial Sequence

<220><223> Punta Toro Virus-11 2A Peptide

<400> 93

Arg Met Ser Phe Gln Gly Pro Gly Ala Thr Asn Phe Ser Leu Leu Lys

1 5 10 15

Arg Ala Gly Asp Val Glu Glu Asn Pro Gly Pro

20 25

<210> 94

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Ciricket Paralysis Virus 2A Peptide

<400> 94

Phe Leu Arg Lys Arg Thr Gln Leu Leu Met Ser Gly Asp Val Glu Ser

1 5 10 15

Asn Pro Gly Pro

20

<210> 95

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Drosophila C Virus 2A Peptide

<400> 95

Glu Ala Ala Arg Gln Met Leu Leu Leu Leu Ser Gly Asp Val Glu Thr

1 5 10 15

Asn Pro Gly Pro

 20

<210> 96

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Acute Bee Paralysis Virus 2A Peptide

<400> 96

Gly Ser Trp Thr Asp Ile Leu Leu Leu Leu Ser Gly Asp Val Glu Thr

1 5 10 15

Asn Pro Gly Pro

 20

<210> 97

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Acute Bee Paralysis Virus Poland 1 isolate 2A Peptide

<400> 97

Gly Ser Trp Thr Asp Ile Leu Leu Leu Leu Ser Gly Asp Val Glu Thr

1 5 10 15

Asn Pro Gly Pro

 20

<210> 98

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Acute Bee Paralysis Virus Hungary 1 isolate 2A Peptide

<400> 98

Gly Ser Trp Thr Asp Ile Leu Leu Leu Trp Ser Gly Asp Val Glu Thr

1 5 10 15

Asn Pro Gly Pro

 20

<210> 99
 <211> 20
 <212> PRT
 <213> Artificial Sequence
 <220><223> Infectious Flacherie Virus 2A Peptide
 <400> 99
 Thr Arg Ala Glu Ile Glu Asp Glu Leu Ile Arg Ala Gly Ile Glu Ser

1 5 10 15
 Asn Pro Gly Pro
 20

<210> 100
 <211> 20
 <212> PRT
 <213> Artificial Sequence
 <220><223> Tomato Aspermy Virus 2A Peptide
 <400> 100
 Arg Ala Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu

1 5 10 15
 Asn Pro Gly Pro
 20

<210> 101
 <211> 20
 <212> PRT
 <213> Artificial Sequence
 <220><223> Equine Encephalosis Virus 2A Peptide
 <400> 101

Gln Gly Ala Gly Arg Gly Ser Leu Val Thr Cys Gly Asp Val Glu Glu
 1 5 10 15
 Asn Pro Gly Pro
 20
 <210> 102
 <211> 20
 <212> PRT

<213> Artificial Sequence
 <220><223> Avian Polyoma Virus 2A Peptide
 <400> 102
 Asn Tyr Pro Met Pro Glu Ala Leu Gln Lys Ile Ile Asp Leu Glu Ser
 1 5 10 15
 Asn Pro Pro Pro
 20

<210> 103

<211> 20

<212

> PRT

<213> Artificial Sequence
 <220><223> Kashmir bee virus 2A Peptide
 <400> 103
 Gly Thr Trp Glu Ser Val Leu Asn Leu Leu Ala Gly Asp Ile Glu Leu
 1 5 10 15
 Asn Pro Gly Pro
 20

<210> 104

<211> 20

<212> PRT

<213> Artificial Sequence
 <220><223> Perina Nuda Picorna-like Virus (a) 2A Peptide
 <400> 104
 Ala Gln Gly Trp Val Pro Asp Leu Thr Val Asp Gly Asp Val Glu Ser
 1 5 10 15
 Asn Pro Gly Pro
 20

<210> 105

<211> 20

<212> PRT

<213> Artificial Sequence
 <220><223> Perina Nuda Picorna-like Virus (b) 2A Peptide
 <400> 105

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

1 5 10 15

Asn Pro Gly Pro

20

<210> 106

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Ectropis Obliqua Picorna-like Virus (a) 2A Peptide

<400> 106

Ala Gln Gly Trp Ala Pro Asp Leu Thr Gln Asp Gly Asp Val Glu Ser

1 5 10 15

Asn Pro Gly Pro

20

<210> 107

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Ectropis Obliqua Picorna-like Virus (b) 2A Peptide

<400> 107

Ile Gly Gly Gly Gln Arg Asp Leu Thr Gln Asp Gly Asp Ile Glu Ser

1 5 10 15

Asn Pro Gly Pro

20

<210> 108

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Providence Virus (a) 2A Peptide

<400> 108

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

1 5 10 15

Pro Gly Pro

<210> 109
 <211> 19
 <212> PRT
 <213> Artificial Sequence
 <220><223> Providence Virus (b) 2A Peptide
 <400> 109
 Gly Asp Pro Ile Glu Asp Leu Thr Asp Asp Gly Asp Ile Glu Lys Asn
 1 5 10 15
 Pro Gly Pro

<210> 110
 <211> 19
 <212> PRT
 <213>
 > Artificial Sequence
 <220><223> Providence Virus (c) 2A Peptide
 <400> 110
 Ser Gly Gly Arg Gly Ser Leu Leu Thr Ala Gly Asp Val Glu Lys Asn
 1 5 10 15
 Pro Gly Pro

<210> 111
 <211> 20
 <212> PRT
 <213> Artificial Sequence
 <220><223> Bovine Rotavirus 2A Peptide
 <400> 111
 Ser Lys Phe Gln Ile Asp Arg Ile Leu Ile Ser Gly Asp Ile Glu Leu
 1 5 10 15
 Asn Pro Gly Pro

20

<210> 112
 <211> 20

<212> PRT
 <213> Artificial Sequence
 <220><223> Porcine Rotavirus 2A Peptide
 <400> 112
 Ala Lys Phe Gln Ile Asp Lys Ile Leu Ile Ser Gly Asp Val Glu Leu
 1 5 10 15
 Asn Pro Gly Pro
 20

<210> 113
 <211> 20
 <212> PRT
 <213> Artificial Sequence
 <220><223> Human Rotavirus 2A Peptide
 <400> 113
 Ser Lys Phe Gln Ile Asp Lys Ile Leu Ile Ser Gly Asp Ile Glu Leu
 1 5 10 15

Asn Pro Gly Pro
 20
 <210> 114
 <211> 20
 <212> PRT
 <213> Artificial Sequence
 <220><223> Bombyx Mori Reovirus 2A Peptide
 <400> 114
 Phe Arg Ser Asn Tyr Asp Leu Leu Lys Leu Cys Gly Asp Ile Glu Ser
 1 5 10 15

Asn Pro Gly Pro
 20
 <210> 115
 <211> 20
 <212> PRT
 <213> Artificial Sequence
 <220><223> Lymantria Dispar Reovirus 2A Peptide

<400> 115
Phe Arg Ser Asn Tyr Asp Leu Leu Lys Leu Cys Gly Asp Val Glu Ser
1 5 10 15
Asn Pro Gly Pro
20

<210> 116

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Dendrolimus Punctatus Reovirus 2A Peptide

<400> 116

Phe Arg Ser Asn Tyr Asp Leu Leu Lys Leu Cys Gly Asp Val Glu Ser
1 5 10 15
Asn Pro Gly Pro
20

<210> 117

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Trypanosoma Brucei TSR1 2A Peptide

<400> 117

Ser Ser Ile Ile Arg Thr Lys Met Leu Val Ser Gly Asp Val Glu Glu
1 5 10 15
Asn Pro Gly Pro
20

<210> 118

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Trypanosoma Spp. CAB95325.1 2A Peptide

<400> 118

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

Asn Pro Gly Pro

20

<210> 119

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Trypanosoma Spp. CAB95559.1 2A Peptide

<400> 119

Ser Ser Ile Ile Arg Thr Lys Ile Leu Leu Ser Gly Asp Val Glu Glu

1 5 10 15

Asn Pro Gly Pro

20

<210> 120

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Trypanosoma Cruzi 2A Peptide

<400> 120

Cys Asp Ala Gln Arg Gln Lys Leu Leu Leu Ser Gly Asp Ile Glu Gln

1 5 10 15

Asn Pro Gly Pro

20

<210> 121

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> T. maritima aguA 2A Peptide

<400> 121

Tyr Ile Pro Asp Phe Gly Gly Phe Leu Val Lys Ala Asp Ser Glu Phe

1 5 10 15

Asn Pro Gly Pro

20

<210> 122

<211> 21

<212> PRT
 <213> Artificial Sequence
 <220><223> B. bronchiseptica 2A Peptide

<400> 122
 Val His Cys Ala Gly Arg Gly Gly Pro Val Arg Leu Leu Asp Lys Glu
 1 5 10 15
 Gly Asn Pro Gly Pro
 20

<210> 123
 <211> 20
 <212> PRT
 <213> Artificial Sequence
 <220><223> Murine mor-1F 2A Peptide

<400> 123
 Asp Leu Glu Leu Glu Thr Val Gly Ser His Gln Ala Asp Ala Glu Thr
 1 5 10 15
 Asn Pro Gly Pro
 20

<210> 124
 <211> 20
 <212>
 > PRT
 <213> Artificial Sequence
 <220><223> D. melanogaster mod(mdg4) 2A Peptide

<400> 124
 Thr Ala Ala Asp Lys Ile Gln Gly Ser Trp Lys Met Asp Thr Glu Gly
 1 5 10 15
 Asn Pro Gly Pro
 20

<210> 125
 <211> 20
 <212> PRT
 <213> Artificial Sequence
 <220><223> A. nidulans Ca Channel MID1 2A Peptide

<400> 125

Pro Ile Thr Asn Arg Pro Arg Asn Ser Gly Leu Ile Asp Thr Glu Ile

1 5 10 15

Asn Pro Gly Pro

20