(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0037192
(43) 공개일자 2024년03월21일

(51) 국제특허분류(Int. Cl.)
A61K 48/00 (2006.01) *A61K 38/00* (2006.01)
A61P 31/00 (2006.01) *A61P 35/00* (2006.01)
C07K 14/725 (2006.01) *C12N 15/85* (2006.01)
C12N 15/90 (2006.01)
(52) CPC특허분류
A61K 48/005 (2013.01)
A61K 38/00 (2013.01)
(21) 출원번호 10-2023-7042556
(22) 출원일자(국제) 2022년05월11일
심사청구일자 없음
(85) 번역문제출일자 2023년12월08일
(86) 국제출원번호 PCT/US2022/028831
(87) 국제공개번호 WO 2022/241029
국제공개일자 2022년11월17일
(30) 우선권주장
63/187,117 2021년05월11일 미국(US)
(뒷면에 계속)

(71) 출원인
마이얼로이드 테라퓨틱스, 인크.
미국 02139 매사추세츠주 캠브리지 테크놀로지 스퀘어 300 스위트 203
(72) 발명자
겟츠 대니얼
미국 02139 매사추세츠 캠브리지 테크놀로지 스퀘어 300 스위트 203 마이얼로이드 테라퓨틱스, 인크. 내
왕 유샤오
미국 02139 매사추세츠 캠브리지 테크놀로지 스퀘어 300 스위트 203 마이얼로이드 테라퓨틱스, 인크. 내
(뒷면에 계속)
(74) 대리인
김진희, 김태홍

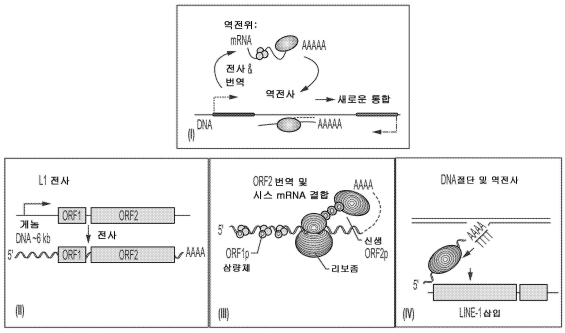
전체 청구항 수 : 총 48 항

(54) 발명의 명칭 **게놈 통합을 위한 방법 및 조성물**

(57) 요 약

표적 게놈을 조절하고 관심 있는 형질전이유전자를 세포의 게놈 내로 안정적으로 통합시키는 방법 및 조성물을 개시한다.

대 표 도



(52) CPC특허분류
A61P 31/00 (2018.01)
A61P 35/00 (2018.01)
C07K 14/7051 (2013.01)
C12N 15/85 (2013.01)
C12N 15/90 (2013.01)
C12N 2800/90 (2013.01)
C12N 2830/48 (2013.01)

(72) 발명자
비사리아 나미타
미국 02139 매사추세츠 캠브리지 테크놀로지 스퀘어 300 스위트 203 마이얼로이드 테라퓨틱스, 인크. 내
슈체르바코바 인나
미국 02139 매사추세츠 캠브리지 테크놀로지 스퀘어 300 스위트 203 마이얼로이드 테라퓨틱스, 인크. 내
리 소치타
미국 02139 매사추세츠 캠브리지 테크놀로지 스퀘어 300 스위트 203 마이얼로이드 테라퓨틱스, 인크. 내

(30) 우선권주장
63/254,791 2021년10월12일 미국(US)
63/274,907 2021년11월02일 미국(US)

명 세 서

청구범위

청구항 1

치료 유효량의 하나 이상의 폴리핵산, 또는 하나 이상의 폴리핵산을 코딩하는 적어도 하나의 벡터를 포함하는 약학 조성물로서, 하나 이상의 폴리핵산은

(a) 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소; 및

(b) 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함하는 삽입체 서열

을 포함하며, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 세포의 게놈 내로의 삽입체 서열의 통합을 촉진하고;

약학 조성물은 인간 대상체에 대해 실질적으로 면역원성을 나타내지 않는 것인 약학 조성물.

청구항 2

제1항에 있어서, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 하나 이상의 긴 산재된 핵 요소(long interspersed nuclear element)(LINE) 폴리펩티드를 포함하고, 하나 이상의 LINE 폴리펩티드는

(i) 인간 ORF1p 또는 이의 기능성 단편, 및

(ii) 인간 ORF2p 또는 이의 기능성 단편

을 포함하는 것인 약학 조성물.

청구항 3

제2항에 있어서, 삽입체 서열은 인간 세포의 게놈 내로 안정적으로 통합 및/또는 역전위되는 것인 약학 조성물.

청구항 4

제3항에 있어서, 인간 세포는 T 세포, B 세포, 골수 세포, 단핵구, 대식세포 및 수지상 세포로 구성된 군으로부터 선택된 면역 세포인 약학 조성물.

청구항 5

제3항에 있어서, 삽입체 서열은 (i) 하나 이상의 폴리핵산에 의해 코딩된 엔도뉴클레아제(endonuclease)에 의한 표적 부위의 DNA 가닥의 절단에 의해, (ii) 표적 프라이밍 역전사(TPRT)를 통해, 또는 (iii) 게놈의 DNA 표적 부위 내로의 삽입체 서열의 역스플라이싱을 통해 게놈 내로 통합되는 것인 약학 조성물.

청구항 6

제5항에 있어서, 삽입체 서열은 인간 ORF2p의 엔도뉴클레아제 도메인의 특이성을 이용함으로써 폴리 T 부위에서 게놈 내로 통합되는 것인 약학 조성물.

청구항 7

제6항에 있어서, 폴리 T 부위는 서열 TTTTTA를 포함하는 것인 약학 조성물.

청구항 8

제3항에 있어서, 하나 이상의 폴리핵산은 게놈의 표적 부위에 상보적인 상동성 아암(homology arm)을 포함하는 것인 약학 조성물.

청구항 9

제3항에 있어서, 삽입체 서열은 (a) 리보좀 RNA 유전자좌가 아닌 유전자좌에서 게놈 내로 통합되거나; (b) 게놈

의 유전자 또는 유전자의 조절 영역 내로 통합됨으로써, 유전자를 파괴하거나 유전자의 발현을 하향조절하거나; (c) 게놈의 유전자 또는 유전자의 조절 영역 내로 통합됨으로써, 유전자의 발현을 상향조절하거나; 또는 (d) 게놈 내로 통합되고 게놈의 유전자를 대체하는 것인 약학 조성물.

청구항 10

제3항에 있어서, 약학 조성물은 (i) 하나 이상의 siRNA 및/또는 (ii) RNA 가이드 서열, 또는 RNA 가이드 서열을 코딩하는 폴리핵산을 추가로 포함하고, RNA 가이드 서열은 게놈의 DNA 표적 부위를 표적화하고 삽입체 서열은 게놈의 DNA 표적 부위에서 게놈 내로 통합되는 것인 약학 조성물.

청구항 11

제1항에 있어서, 하나 이상의 폴리핵산은 3 kb 내지 20 kb의 총 길이를 가진 것인 약학 조성물.

청구항 12

제1항에 있어서, 하나 이상의 폴리핵산은 하나 이상의 폴리리보핵산, 하나 이상의 RNA 또는 하나 이상의 mRNA를 포함하는 것인 약학 조성물.

청구항 13

제1항에 있어서, 외생성 치료 폴리펩티드는 리간드, 항체, 수용체, 효소, 수송 단백질, 구조 단백질, 호르몬, 수축 단백질, 저장 단백질 및 전사 인자로 구성된 군으로부터 선택되는 것인 약학 조성물.

청구항 14

제13항에 있어서, 외생성 치료 폴리펩티드는 키메라 항원 수용체(CAR) 및 T 세포 수용체(TCR)로 구성된 군으로부터 선택된 수용체인 약학 조성물.

청구항 15

제1항에 있어서, 하나 이상의 폴리핵산은 프로모터 서열, 5' UTR 서열, 3' UTR 서열 및 폴리 A 서열을 포함하는 제1 발현 카세트를 포함하고,

(i) 프로모터 서열은 5' UTR 서열의 상류에 있고,

(ii) 5' UTR 서열은 폴리펩티드를 코딩하는 이동성 유전 요소의 서열의 상류에 있고,

(iii) 3' UTR 서열은 삽입체 서열의 하류에 있고;

(iv) 3' UTR은 폴리 A 서열의 상류에 있고;

5' UTR 서열, 3' UTR 서열 또는 폴리 A 서열은 인간 ORF2p 또는 이의 기능성 단편에 대한 결합 부위를 포함하는 것인 약학 조성물.

청구항 16

제15항에 있어서, 삽입체 서열은 프로모터 서열의 역상보체인 서열, 5' UTR 서열의 역상보체인 서열, 3' UTR 서열의 역상보체인 서열, 및 폴리 A 서열의 역상보체인 서열을 포함하는 제2 발현 카세트를 포함하고,

(i) 프로모터 서열의 역상보체인 서열은 5' UTR 서열의 역상보체인 서열의 하류에 있고,

(ii) 5' UTR 서열의 역상보체인 서열은 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 서열의 하류에 있고,

(iii) 3' UTR 서열의 역상보체인 서열은 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 서열의 상류에 있고,

(iv) 폴리 A 서열의 역상보체인 서열은 3' UTR 서열의 역상보체인 서열의 상류 및 폴리펩티드를 코딩하는 이동성 유전 서열의 하류에 있는 것인 약학 조성물.

청구항 17

제16항에 있어서, 제1 발현 카세트의 프로모터 서열은 제2 발현 카세트의 프로모터 서열과 상이한 것인 약학 조성물.

청구항 18

제2항에 있어서, 하나 이상의 LINE 폴리펩티드는 인간 ORF1p 또는 이의 기능성 단편을 포함하는 제1 LINE 폴리펩티드, 및 인간 ORF2p 또는 이의 기능성 단편을 포함하는 제2 LINE 폴리펩티드를 포함하고, 제1 LINE 폴리펩티드 및 제2 LINE 폴리펩티드는 상이한 오픈 리딩 프레임(open reading frame)(ORF)으로부터 번역되는 것인 약학 조성물.

청구항 19

제2항에 있어서, 하나 이상의 폴리핵산은 인간 ORF1p 또는 이의 기능성 단편을 코딩하는 제1 폴리핵산 분자, 및 인간 ORF2p 또는 이의 기능성 단편을 코딩하는 제2 폴리핵산 분자를 포함하는 것인 약학 조성물.

청구항 20

제1항에 있어서, 하나 이상의 폴리핵산은 5' UTR 서열 및 3' UTR 서열을 포함하고,

(a) 5' UTR은 LINE-1의 5' UTR, 또는 ACUCCUCCCCAUCCUCUCCCUCUGUCCCUCUGUCCCUCUGACCCUGCACUGUCCCAGCACC(서열번호 51)과 적어도 80% 서열 동일성을 가진 서열을 포함하고/하거나;

(b) 3' UTR은 LINE-1의 3' UTR, 또는 CAGGACACAGCCUUGGAUCAGGACAGAGACUUGGGGGCCAUCCUGCCCCUCCAACCCGACAUGUGUACCUCAGCUUUUUCCCUCACUUGCAUCAAUAAAGCUUCUGUGUUUGGAACAG(서열번호 52)와 적어도 80% 서열 동일성을 가진 서열을 포함하는 것인 약학 조성물.

청구항 21

제1항에 있어서, 외생성 치료 폴리펩티드를 코딩하는 서열은 인트론을 포함하지 않는 것인 약학 조성물.

청구항 22

제1항에 있어서, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 C-말단 핵 국소화 신호(NLS), N-말단 NLS, 또는 이들 둘 다를 포함하는 것인 약학 조성물.

청구항 23

제2항에 있어서, 외생성 폴리펩티드를 코딩하는 서열은 ORF1p 또는 이의 기능성 단편을 코딩하는 서열과 인 프레임(in frame)으로 있지 않고/않거나, ORF2p 또는 이의 기능성 단편을 코딩하는 서열과 인 프레임으로 있지 않는 것인 약학 조성물.

청구항 24

제1항에 있어서, 하나 이상의 폴리핵산은 뉴클레아제 도메인, ORF2p로부터 유래하지 않은 뉴클레아제 도메인, 메가TAL 뉴클레아제 도메인, TALEN 도메인, Cas9 도메인, Cas6 도메인, Cas7 도메인, Cas8 도메인, R2 역요소(retroelement)로부터의 징크 핑거(zinc finger) 결합 도메인, 또는 반복 서열에 결합하는 DNA 결합 도메인을 코딩하는 서열을 포함하는 것인 약학 조성물.

청구항 25

제24항에 있어서, 하나 이상의 폴리핵산은 뉴클레아제 도메인을 코딩하는 서열을 포함하고, 뉴클레아제 도메인은 뉴클레아제 활성을 갖지 않거나, 돌연변이가 없는 뉴클레아제 도메인에 비해 뉴클레아제 도메인의 활성을 감소시키는 돌연변이를 포함하는 것인 약학 조성물.

청구항 26

제2항에 있어서, ORF2p 또는 이의 기능성 단편은 엔도뉴클레아제 활성이 결여되어 있거나 S228P 및 Y1180A로 구성된 군으로부터 선택된 돌연변이를 포함하고/하거나, ORF1p 또는 기능성 단편은 K3R 돌연변이를 포함하는 것인 약학 조성물.

청구항 27

제1항에 있어서, 삽입체 서열은 2개 이상의 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함하는 것인 약학 조성물.

청구항 28

제2항에 있어서, 하나 이상의 폴리핵산은 하나 이상의 폴리리보핵산을 포함하고, 외생성 치료 폴리펩티드는 키메라 항원 수용체(CAR) 및 T 세포 수용체(TCR)로 구성된 군으로부터 선택된 수용체이며, 약학 조성물은 인간 대상체에게 전신 투여하기 위해 제제화된 것인 약학 조성물.

청구항 29

제1항에 있어서, 하나 이상의 폴리핵산은

(i) 지질 나노입자 및 중합체성 나노입자로 구성된 군으로부터 선택된 나노입자로 제제화되고/되거나;

(ii) 글리코실화된 RNA, 원형 RNA 및 자가복제 RNA로 구성된 군으로부터 선택된 하나 이상의 폴리핵산을 포함하는 것인 약학 조성물.

청구항 30

방법으로서,

(i) 질환 또는 병태의 치료를 필요로 하는 인간 대상체에서 질환 또는 병태를 치료하는 방법으로서, 제1항의 약학 조성물을 인간 대상체에게 투여하는 것을 포함하는 방법; 또는

(ii) 조성물을 생체외에서 인간 세포 집단과 접촉시킴으로써, 생체외 변형된 인간 세포 집단을 형성하는 것을 포함하는, 생체외에서 인간 세포 집단을 변형시키는 방법으로서, 조성물은 하나 이상의 폴리핵산, 또는 하나 이상의 폴리핵산을 코딩하는 적어도 하나의 벡터를 포함하고, 하나 이상의 폴리핵산은

(a) 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소; 및

(b) 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 삽입체 서열

을 포함하고, 생체외 변형된 인간 세포 집단은 인간 대상체에 대해 실질적으로 면역원성을 나타내지 않는 것인 방법인 방법.

청구항 31

제1항에 있어서, 하나 이상의 폴리핵산은 (i) 게놈 내로의 삽입체 서열의 부위 지정 통합을 위한 통합효소(integrase) 또는 이의 단편을 코딩하는 서열 및 (ii) 통합효소에 의해 작동될 수 있는 통합효소 게놈 랜딩(landing) 부위 서열을 추가로 포함하고, 게놈 랜딩 서열은 4개 초과의 연속 뉴클레오타이드 길이인 약학 조성물.

청구항 32

제31항에 있어서, ORF2와 통합효소는 별개의 폴리뉴클레오타이드에 존재하는 것인 약학 조성물.

청구항 33

제31항에 있어서, ORF2와 통합효소는 단일 폴리뉴클레오타이드에 존재하는 것인 약학 조성물.

청구항 34

제31항 또는 제32항에 있어서, 통합효소는 세포의 게놈 내로 통합되지 않는 것인 약학 조성물.

청구항 35

제31항 내지 제34항 중 어느 한 항에 있어서, 통합효소는 돌연변이되거나 절두된(truncated) 재조합 단백질인 약학 조성물.

청구항 36

제31항 내지 제35항 중 어느 한 항에 있어서, 통합효소에 의해 작동될 수 있는 통합효소 게놈 랜딩 서열은 20개 초과의 뉴클레오타이드 길이 또는 30개 초과의 뉴클레오타이드 길이인 약학 조성물.

청구항 37

제31항 내지 제36항 중 어느 한 항에 있어서, 삽입체 서열은 통합효소에 의해 작동될 수 있는 부착 부위를 포함하는 것인 약학 조성물.

청구항 38

제31항 내지 제37항 중 어느 한 항에 있어서, 통합효소 게놈 랜딩 부위가 가이드 RNA 및 Cas 시스템을 이용하여 게놈 내로 삽입되는 것인 약학 조성물.

청구항 39

제31항 내지 제38항 중 어느 한 항에 있어서, 가이드 RNA, CAS 시스템 및 게놈 랜딩 서열은 LINE1-ORF를 코딩하는 서열 및 삽입체 서열을 포함하는 폴리뉴클레오타이드로부터 분리되어 있는 폴리뉴클레오타이드에 존재하는 것인 약학 조성물.

청구항 40

제31항 내지 제39항 중 어느 한 항에 있어서, 하나 이상의 ORF 폴리펩티드 서열은 돌연변이를 포함하는 것인 약학 조성물.

청구항 41

이종 게놈 삽입체 서열을 포유동물 세포의 게놈 내로 부위 특이적으로 통합시키는 방법으로서, (i) (a) 이종 삽입체 서열과 연결된 하나 이상의 인간 레트로트랜스포존(retrotransposon) 요소를 코딩하는 서열을 포함하는 폴리뉴클레오타이드, 및 (b) 가이드 RNA, RNA 가이딩(guided) 통합효소 또는 이의 단편, 및 통합효소에 의해 작동될 수 있는 랜딩 서열을 코딩하는 서열을 포함하는 폴리뉴클레오타이드를 세포 내로 도입하는 단계; 및 (ii) 이종 삽입체 서열이 게놈 부위 내로 통합되었는지를 확인하는 단계를 포함하는 방법.

청구항 42

LINE 레트로트랜스포존 시스템을 이용하여 이종 게놈 삽입체를 부위 특이적으로 통합시키는 방법으로서, LINE 레트로트랜스포존 시스템은 10개 초과의 연속 뉴클레오타이드 길이의 게놈 랜딩 서열을 인식할 수 있는 통합효소 단백질의 단편을 통합시키도록 변형되고, LINE 레트로트랜스포존 시스템은 이종 게놈 삽입체를 통합효소 단백질의 단편에 의해 인식되는 게놈 랜딩 서열 내로 통합시키는 것인 방법.

청구항 43

제41항 또는 제42항에 있어서, 4개 초과의 연속 뉴클레오타이드 길이의 게놈 랜딩 서열을 게놈 내로 통합시키는 단계를 추가로 포함하는 방법.

청구항 44

제43항에 있어서, 게놈 랜딩 서열을 게놈 내로 통합시키는 단계는 RNA 가이딩 CRISPR-Cas 시스템에 의해 수행되는 것인 방법.

청구항 45

제44항에 있어서, RNA 가이딩 CRISPR-Cas 시스템은 4개 초과의 연속 뉴클레오타이드 길이의 서열을 특정 게놈 부위 내로 통합시킬 수 있는 편집 기능을 가진 것인 방법.

청구항 46

제30항 또는 제44항에 있어서, RNA 가이딩 CRISPR-Cas 시스템은 ORF-mRNA 결합 서열을, 가이드 RNA의 서열에 대

해 서열 상동성을 가진 게놈 내의 특정 위치 내로 통합시키는 것인 방법.

청구항 47

제30항 및 제41항 내지 제46항 중 어느 한 항에 있어서, 삽입체는 약 10 킬로염기이거나 10 킬로염기보다 더 큰 것인 방법.

청구항 48

제30항 및 제41항 내지 제47항 중 어느 한 항에 있어서, 폴리뉴클레오타이드는 mRNA인 방법.

발명의 설명

기 술 분 야

[0001] **상호참조**

[0002] 본원은 2021년 5월 11일에 출원된 미국 가출원 제63/187,117호, 2021년 10월 12일에 출원된 미국 가출원 제63/254,791호, 및 2021년 11월 2일에 출원된 미국 가출원 제63/274,907호에 대한 우선권 이익을 주장하고, 이 출원들 각각은 전체적으로 본원에 참고로 포함된다.

[0003] **서열목록**

[0004] 본원은 ASCII 포맷으로 전자 제출되었고 전체적으로 본원에 참고로 포함된 서열목록을 함유한다. 2022년 6월 17일에 생성된 상기 ASCII 사본은 명칭이 56371-723_601_SL.txt이고 크기가 677,993 바이트이다.

배 경 기 술

[0005] 세포 요법은 치료하기 어려운 질환, 예컨대, 암, 지속적인 감염 및 다른 형태의 치료에 불응하는 특정 질환을 해결하기 위해 빠르게 발전하는 분야이다. 세포 요법은 종종 생체외에서 조작되고 체내의 결핍을 바로잡기 위해 유기체에게 투여되는 세포를 활용한다. 세포의 게놈을 조작하기 위한 효과적인 신뢰할 수 있는 시스템은 조작된 세포가 유기체에게 투여될 때 최적으로 장기간 효능을 가지면서 작용한다는 의미에서 매우 중요하다. 마찬가지로, 유전자 조작의 신뢰할 수 있는 기작은 유전자 요법의 성공에 있어서 초석을 형성한다. 그러나, 치료적으로 안전하고 효과적인 방식으로 핵산 카고(cargo)(예를 들어, 큰 카고)를 전달하는 방법에는 심각한 결함이 있다. 바이러스 전달 기작은 세포에서 큰 핵산 카고를 전달하는 데 자주 사용되나, 안전성 문제와 연관되어 있고 일부 세포 유형에서 카고를 발현하는 데 사용될 수 없다. 추가로, 세포를 반복적으로 유전자 조작하는 것은 세포 건강에 영향을 미칠 수 있고, 세포 주기의 변경을 유도할 수 있고, 세포를 치료 용도에 부적합하게 만들 수 있다. 치료 목적을 위해 외생적으로 도입된 유전 물질의 효과적인 전달 및 안정화 분야에서 진일보는 계속 추구된다.

발명의 내용

[0006] 본원은 치료 유효량의 하나 이상의 폴리핵산, 또는 이러한 하나 이상의 폴리핵산을 코딩하는 적어도 하나의 벡터를 포함하고 인간 대상체에서 실질적으로 면역원성을 나타내지 않는 약학 조성물을 제공하는 것으로, 상기 하나 이상의 폴리핵산은 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소; 및 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함하는 삽입체 서열을 포함하고, 이때 상기 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 세포의 게놈 내로의 삽입체 서열의 통합을 촉진한다.

[0007] 일부 실시양태에서, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 하나 이상의 긴 산재된 핵 요소(long interspersed nuclear element)(LINE) 폴리펩티드를 포함하며, 이때 하나 이상의 LINE 폴리펩티드는 인간 ORF1p 또는 이의 기능성 단편, 및 인간 ORF2p 또는 이의 기능성 단편을 포함한다.

[0008] 일부 실시양태에서, 삽입체 서열은 인간 세포의 게놈 내로 안정적으로 통합 및/또는 역전위된다.

[0009] 일부 실시양태에서, 인간 세포는 T 세포, B 세포, 골수 세포, 단핵구, 대식세포 및 수지상 세포로 구성된 군으로부터 선택된 면역 세포이다.

[0010] 일부 실시양태에서, 삽입체 서열은 (i) 하나 이상의 폴리핵산에 의해 코딩된 엔도뉴클레아제(endonuclease)에 의한 표적 부위의 DNA 가닥의 절단에 의해, (ii) 표적 프라이밍 역전사(TPRT)를 통해 또는 (iii) 게놈의 DNA 표적 부위 내로의 삽입체 서열의 역스플라이싱을 통해 게놈 내로 통합된다. 일부 실시양태에서, 삽입체 서열은 인

간 ORF2p의 엔도뉴클레아제 도메인의 특이성을 사용하여 폴리 T 부위에서 게놈 내로 통합된다. 일부 실시양태에서, 폴리 T 부위는 서열 TTTTTA를 포함한다. 일부 실시양태에서, 하나 이상의 폴리핵산은 게놈의 표적 부위에 상보적인 상동성 아암(homology arm)을 포함한다. 일부 실시양태에서, 삽입체 서열은 (a) 리보좀 RNA 유전자좌가 아닌 유전자좌에서 게놈 내로 통합되거나; (b) 게놈의 유전자 또는 유전자의 조절 영역 내로 통합됨으로써, 유전자를 파괴하거나 유전자의 발현을 하향조절하거나; (c) 게놈의 유전자 또는 유전자의 조절 영역 내로 통합됨으로써, 유전자의 발현을 상향조절하거나; 또는 (d) 게놈 내로 통합되어 게놈의 유전자를 대체한다. 일부 실시양태에서, 약학 조성물은 (i) 하나 이상의 siRNA 및/또는 (ii) RNA 가이드 서열, 또는 RNA 가이드 서열을 코딩하는 폴리핵산을 추가로 포함하고, 이때 RNA 가이드 서열은 게놈의 DNA 표적 부위를 표적화하고, 삽입체 서열은 게놈의 DNA 표적 부위에서 게놈 내로 통합된다. 일부 실시양태에서, 하나 이상의 폴리핵산은 3 kb 내지 20 kb의 총 길이를 가진다. 일부 실시양태에서, 하나 이상의 폴리핵산은 하나 이상의 폴리리보핵산, 하나 이상의 RNA 또는 하나 이상의 mRNA를 포함한다. 일부 실시양태에서, 외생성 치료 폴리펩티드는 리간드, 항체, 수용체, 효소, 수송 단백질, 구조 단백질, 호르몬, 수축 단백질, 저장 단백질 및 전사 인자로 구성된 군으로부터 선택된다. 일부 실시양태에서, 외생성 치료 폴리펩티드는 키메라 항원 수용체(CAR) 및 T 세포 수용체(TCR)로 구성된 군으로부터 선택된 수용체이다. 일부 실시양태에서, 하나 이상의 폴리핵산은 프로모터 서열, 5' UTR 서열, 3' UTR 서열 및 폴리 A 서열을 포함하는 제1 발현 카세트를 포함하고; 이때 상기 프로모터 서열은 5' UTR 서열의 상류에 있고, 5' UTR 서열은 폴리펩티드를 코딩하는 이동성 유전 요소의 서열의 상류에 있고, 3' UTR 서열은 삽입체 서열의 하류에 있고; 3' UTR은 폴리 A 서열의 상류에 있고; 5' UTR 서열, 3' UTR 서열 또는 폴리 A 서열은 인간 ORF2p 또는 이의 기능성 단편에 대한 결합 부위를 포함한다. 일부 실시양태에서, 삽입체 서열은 프로모터 서열의 역상보체인 서열, 5' UTR 서열의 역상보체인 서열, 3' UTR 서열의 역상보체인 서열, 및 폴리 A 서열의 역상보체인 서열을 포함하는 제2 발현 카세트를 포함하고; 이때 (i) 프로모터 서열의 역상보체인 서열은 5' UTR 서열의 역상보체인 서열의 하류에 있고, (ii) 5' UTR 서열의 역상보체인 서열은 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 서열의 하류에 있고, (iii) 3' UTR 서열의 역상보체인 서열은 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 서열의 상류에 있고, (iv) 폴리 A 서열의 역상보체인 서열은 3' UTR 서열의 역상보체인 서열의 상류에 있고 폴리펩티드를 코딩하는 이동성 유전 서열의 하류에 있다. 일부 실시양태에서, 제1 발현 카세트의 프로모터 서열은 제2 발현 카세트의 프로모터 서열과 상이하다. 일부 실시양태에서, 하나 이상의 LINE 폴리펩티드는 인간 ORF1p 또는 이의 기능성 단편을 포함하는 제1 LINE 폴리펩티드 및 인간 ORF2p 또는 이의 기능성 단편을 포함하는 제2 LINE 폴리펩티드를 포함하고, 이때 제1 LINE 폴리펩티드 및 제2 LINE 폴리펩티드는 상이한 오픈 리딩 프레임(ORF)으로부터 번역된다. 일부 실시양태에서, 하나 이상의 폴리핵산은 인간 ORF1p 또는 이의 기능성 단편을 코딩하는 제1 폴리핵산 분자 및 인간 ORF2p 또는 이의 기능성 단편을 코딩하는 제2 폴리핵산 분자를 포함한다. 일부 실시양태에서, 하나 이상의 폴리핵산은 5' UTR 서열 및 3' UTR 서열을 포함하고, 이때 5' UTR은 LINE-1의 5' UTR, 또는 ACUCCUCCCCAUCCUCUCCCUCUGUCCCUCUGUCCCUCUGACCCUGCACUGUCCCAGCACC(서열번호 51)과 적어도 80% 서열 동일성을 가진 서열을 포함하고/하거나; 3' UTR은 LINE-1의 3' UTR, 또는 CAGGACACAGCCUUGGAUCAGGACAGAGACUUGGGGGCCAUCCUGCCCCUCCAACCCGACAUGUGUACCUCAGCUUUUUCCCUCACUUGCAUCAAUAAAGCUUCUGUGUUUGGAACAG(서열번호 52)와 적어도 80% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, 외생성 치료 폴리펩티드를 코딩하는 서열은 인트론을 포함하지 않는다. 일부 실시양태에서, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 C-말단 핵 국소화 신호(NLS), N-말단 NLS 또는 이들 둘 다를 포함한다. 일부 실시양태에서, 외생성 폴리펩티드를 코딩하는 서열은 ORF1p 또는 이의 기능성 단편을 코딩하는 서열과 인 프레임(in frame)으로 있지 않고/않거나, ORF2p 또는 이의 기능성 단편을 코딩하는 서열과 인 프레임으로 있지 않다. 일부 실시양태에서, 하나 이상의 폴리핵산은 뉴클레아제(nuclease) 도메인, ORF2p로부터 유래하지 않은 뉴클레아제 도메인, 메가TAL 뉴클레아제 도메인, TALEN 도메인, Cas9 도메인, Cas6 도메인, Cas7 도메인, Cas8 도메인, R2 역요소(retroelement)의 징크 핑거(zinc finger) 결합 도메인, 또는 반복 서열에 결합하는 DNA 결합 도메인을 코딩하는 서열을 포함한다. 일부 실시양태에서, 하나 이상의 폴리핵산은 뉴클레아제 도메인을 코딩하는 서열을 포함하고, 이때 뉴클레아제 도메인은 뉴클레아제 활성을 갖지 않거나, 돌연변이가 없는 뉴클레아제 도메인에 비해 뉴클레아제 도메인의 활성을 감소시키는 돌연변이를 포함한다. 일부 실시양태에서, ORF2p 또는 이의 기능성 단편은 엔도뉴클레아제 활성이 결여되어 있거나 S228P 및 Y1180A로 구성된 군으로부터 선택된 돌연변이를 포함하고/하거나, ORF1p 또는 이의 기능성 단편은 K3R 돌연변이를 포함한다. 일부 실시양태에서, 삽입체 서열은 2개 이상의 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함한다. 일부 실시양태에서, 하나 이상의 폴리핵산은 하나 이상의 폴리리보핵산을 포함하고, 이때 외생성 치료 폴리펩티드는 키메라 항원 수용체(CAR) 및 T 세포 수용체(TCR)로 구성된 군으로부터 선택된 수용체이고, 약학 조성물은 인간 대상체에게 전신 투여하기 위해 제제화된다. 일부 실시양태에서, 하나 이상의 폴리핵산은 (i) 지질 나노입자 및 중합체성 나노입자로 구성된 군으로부터 선택된 나노입자로 제제화되고/되거나; (ii) 글리코실화된 RNA, 원형 RNA 및 자가복제 RNA로 구

성된 군으로부터 선택된 하나 이상의 폴리핵산을 포함한다.

[0011] 본원은 (i) 질환 또는 병태의 치료를 필요로 하는 인간 대상체에서 질환 또는 병태를 치료하는 방법으로서, 본원에 기재된 약학 조성물을 상기 인간 대상체에게 투여하는 단계를 포함하는 방법; 또는 (ii) 조성물을 생체외에서 인간 세포 집단과 접촉시킴으로써, 생체외 변형된 인간 세포 집단을 형성하는 단계를 포함하는, 생체외에서 인간 세포 집단을 변형시키는 방법으로서, 상기 조성물이 하나 이상의 폴리핵산, 또는 하나 이상의 폴리핵산을 코딩하는 적어도 하나의 벡터를 포함하고, 상기 하나 이상의 폴리핵산이 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소; 및 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 삽입체 서열을 포함하고, 생체외 변형된 인간 세포 집단이 인간 대상체에서 실질적으로 면역원성을 나타내지 않는 것인 방법을 제공한다. 일부 실시양태에서, 하나 이상의 폴리핵산은 (i) 게놈 내로의 삽입체 서열의 부위 지정 통합을 위한 통합효소(integrase) 또는 이의 단편을 코딩하는 서열 및 (ii) 통합효소에 의해 작동될 수 있는 통합효소 게놈 랜딩(landing) 부위 서열을 추가로 포함하고, 이때 게놈 랜딩 서열은 4개 초과의 연속 뉴클레오타이드 길이를 가진다. 일부 실시양태에서, ORF2 및 통합효소는 별개의 폴리뉴클레오타이드에 존재한다. 일부 실시양태에서, ORF2 및 통합효소는 단일 폴리뉴클레오타이드에 존재한다. 일부 실시양태에서, 통합효소는 세포의 게놈 내로 통합되지 않는다. 일부 실시양태에서, 통합효소는 돌연변이되거나 절두된(truncated) 재조합 단백질이다. 일부 실시양태에서, 통합효소에 의해 작동될 수 있는 통합효소 게놈 랜딩 서열은 20개 초과의 뉴클레오타이드 길이 또는 30개 초과의 뉴클레오타이드 길이를 가진다. 일부 실시양태에서, 삽입체 서열은 통합효소에 의해 작동될 수 있는 부착 부위를 포함한다. 일부 실시양태에서, 통합효소 게놈 랜딩 부위는 가이드 RNA 및 Cas 시스템을 이용하여 게놈 내로 삽입된다. 일부 실시양태에서, 가이드 RNA, CAS 시스템 및 게놈 랜딩 서열은 LINE1-ORF를 코딩하는 서열 및 삽입체 서열을 포함하는 폴리뉴클레오타이드로부터 분리되어 있는 폴리뉴클레오타이드에 존재한다. 일부 실시양태에서, 하나 이상의 ORF 폴리펩티드 서열은 돌연변이를 포함한다. 이종 게놈 삽입체 서열을 포유동물 세포의 게놈 내로 부위 특이적으로 통합시키는 방법으로서, (i) (a) 이종 삽입체 서열과 연결된 하나 이상의 인간 레트로트랜스포존(retrotransposon) 요소를 코딩하는 서열을 포함하는 폴리뉴클레오타이드, 및 (b) 가이드 RNA, RNA 가이딩(guided) 통합효소 또는 이의 단편, 및 통합효소에 의해 작동될 수 있는 랜딩 서열을 코딩하는 서열을 포함하는 폴리뉴클레오타이드를 세포 내로 도입하는 단계; (ii) 상기 이종 삽입체 서열이 게놈 부위 내로 통합되었는지를 확인하는 단계를 포함하는 방법을 제공한다.

[0012] 본원은 LINE 레트로트랜스포존 시스템을 이용하여 이종 게놈 삽입체를 부위 특이적으로 통합시키는 방법을 제공하는 것으로, 이때 LINE 레트로트랜스포존 시스템은 10개 초과의 연속 뉴클레오타이드 길이를 가진 게놈 랜딩 서열을 인식할 수 있는 통합효소 단백질의 단편을 통합시키도록 변형되고, LINE 레트로트랜스포존 시스템은 이종 게놈 삽입체를 통합효소 단백질의 단편에 의해 인식되는 게놈 랜딩 서열 내로 통합시킨다. 일부 실시양태에서, 이 방법은 4개 초과의 연속 뉴클레오타이드 길이를 가진 게놈 랜딩 서열을 게놈 내로 통합시키는 단계를 추가로 포함한다. 일부 실시양태에서, 게놈 랜딩 서열을 게놈 내로 통합시키는 단계는 RNA 가이딩 CRISPR-Cas 시스템에 의해 수행된다. 일부 실시양태에서, RNA 가이딩 CRISPR-Cas 시스템은 4개 초과의 연속 뉴클레오타이드 길이를 가진 서열을 특정 게놈 부위 내로 통합시킬 수 있는 편집 기능을 가진다. 일부 실시양태에서, RNA 가이딩 CRISPR-Cas 시스템은 ORF-mRNA 결합 서열을 가이드 RNA의 서열에 대해 서열 상동성을 가진 게놈 내의 특정 위치 내로 통합시킨다. 일부 실시양태에서, 삽입체는 약 10 킬로염기이거나 10 킬로염기보다 더 길다. 일부 실시양태에서, 폴리뉴클레오타이드는 mRNA이다.

[0013] 본원은 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 방법으로서, 폴리핵산을 포함하는 조성물을 표적 세포와 접촉시키는 단계; 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 단계; 및 표적 세포에서 외생성 폴리펩티드를 발현시키는 단계를 포함하는 방법을 제공하는 것으로, 이때 상기 폴리핵산은 외생성 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함하는 삽입체 서열, 및 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소를 포함하고, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 게놈 DNA 내로의 삽입체 서열의 통합을 촉진하고, 상기 표적 세포는 인간 간세포이다. 일부 실시양태에서, 인간 간세포는 1차 세포이다. 일부 실시양태에서, 인간 간세포는 배양된 간세포 세포주로부터 유래한다. 일부 실시양태에서, 통합은 인간 간세포에 최적인 조건 하에서 전기천공하는 단계를 포함한다. 일부 실시양태에서, 상기 방법은 약 2시간, 약 3시간, 약 4시간, 약 5시간, 약 6시간, 약 8시간, 약 10시간 또는 약 24시간 동안 통합시킨 후, 시험관내에서 인간 간세포를 배양하는 단계를 추가로 포함한다. 일부 실시양태에서, 상기 방법은 외생성 폴리펩티드를 발현하는 인간 간세포를, 이를 필요로 하는 인간 대상체 내로 도입하는 단계를 추가로 포함한다. 일부 실시양태에서, 인간 간세포의 적어도 2%는 통합 후 10일째 날에 외생성 폴리펩티드를 발현한다.

[0014] 본원은 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 방법으로서, 폴리핵산을 포함하는 조

성물을 표적 세포와 접촉시키는 단계; 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 단계; 및 표적 세포에서 외생성 폴리펩티드를 발현시키는 단계를 포함하는 방법을 제공하는 것으로, 이때 상기 폴리핵산은 외생성 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함하는 삽입체 서열, 및 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소를 포함하고, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 게놈 DNA 내로의 삽입체 서열의 통합을 촉진하고, 상기 표적 세포는 인간 심근세포이다. 일부 실시양태에서, 인간 심근세포는 1차 세포이다. 일부 실시양태에서, 인간 심근세포는 배양된 심근세포 세포주로부터 유래한다. 일부 실시양태에서, 통합은 인간 심근세포에 최적인 조건 하에서 전기천공하는 단계를 포함한다. 일부 실시양태에서, 상기 방법은 약 2시간, 약 3시간, 약 4시간, 약 5시간, 약 6시간, 약 8시간, 약 10시간 또는 최대 24시간 동안 통합시킨 후, 시험관내에서 심근세포를 배양하는 단계를 추가로 포함한다. 일부 실시양태에서, 상기 방법은 외생성 폴리펩티드를 발현하는 인간 심근세포를, 이를 필요로 하는 인간 대상체 내로 도입하는 단계를 추가로 포함한다. 일부 실시양태에서, 인간 심근세포의 적어도 2%는 통합 후 10일째 날에 외생성 폴리펩티드를 발현한다.

[0015] 본원은 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 방법으로서, 폴리핵산을 포함하는 조성물을 표적 세포와 접촉시키는 단계; 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 단계; 및 표적 세포에서 외생성 폴리펩티드를 발현시키는 단계를 포함하는 방법을 제공하는 것으로, 이때 상기 폴리핵산은 외생성 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함하는 삽입체 서열, 및 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소를 포함하고, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 게놈 DNA 내로의 삽입체 서열의 통합을 촉진하고, 상기 표적 세포는 인간 망막 색소 상피 세포이다. 일부 실시양태에서, 인간 망막 색소 상피 세포는 1차 세포이다. 일부 실시양태에서, 인간 망막 색소 상피 세포는 배양된 망막 색소 상피 세포주로부터 유래한다. 일부 실시양태에서, 통합은 인간 망막 색소 상피 세포에 최적인 조건 하에서 전기천공하는 단계를 포함한다. 일부 실시양태에서, 상기 방법은 약 2시간, 약 3시간, 약 4시간, 약 5시간, 약 6시간, 약 8시간, 약 10시간 또는 최대 24시간 동안 통합시킨 후, 시험관내에서 망막 색소 상피 세포를 배양하는 단계를 추가로 포함한다. 일부 실시양태에서, 상기 방법은 외생성 폴리펩티드를 발현하는 인간 망막 색소 상피 세포를, 이를 필요로 하는 인간 대상체 내로 도입하는 단계를 추가로 포함한다. 일부 실시양태에서, 인간 RPE의 적어도 2%는 통합 후 10일째 날에 외생성 폴리펩티드를 발현한다.

[0016] 본원은 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 방법으로서, 폴리핵산을 포함하는 조성물을 표적 세포와 접촉시키는 단계; 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 단계; 및 표적 세포에서 외생성 폴리펩티드를 발현시키는 단계를 포함하는 방법을 제공하는 것으로, 이때 상기 폴리핵산은 외생성 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함하는 삽입체 서열, 및 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소를 포함하고, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 게놈 DNA 내로의 삽입체 서열의 통합을 촉진하고, 상기 표적 세포는 인간 뉴런 세포이다. 일부 실시양태에서, 인간 뉴런 세포는 1차 세포이다. 일부 실시양태에서, 인간 뉴런 세포는 배양된 뉴런 세포주로부터 유래한다. 일부 실시양태에서, 통합은 인간 뉴런 세포에 최적인 조건 하에서 전기천공하는 단계를 포함한다. 일부 실시양태에서, 상기 방법은 약 2시간, 약 3시간, 약 4시간, 약 5시간, 약 6시간, 약 8시간, 약 10시간 또는 최대 24시간 동안 통합시킨 후, 시험관내에서 뉴런 세포를 배양하는 단계를 추가로 포함한다. 일부 실시양태에서, 상기 방법은 외생성 폴리펩티드를 발현하는 인간 뉴런 세포를 인간 내로 도입하는 단계를 추가로 포함한다. 일부 실시양태에서, 인간 뉴런 세포의 적어도 2%는 통합 후 10일째 날에 외생성 폴리펩티드를 발현한다. 일부 실시양태에서, 삽입체 서열은 인간 삽입체 서열이다. 일부 실시양태에서, 외생성 폴리펩티드는 외생성 치료 폴리펩티드이다. 일부 실시양태에서, 외생성 폴리펩티드는 외생성 인간 폴리펩티드이다. 일부 실시양태에서, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 표적 프라이밍 역전사(TPRT)를 통해 게놈 DNA 내로의 삽입체 서열의 통합을 촉진한다. 일부 실시양태에서, 폴리핵산은 mRNA 또는 mRNA 분자이다. 일부 실시양태에서, 이동성 유전 요소는 인간 LINE 1 레트로트랜스포존 요소를 포함한다. 일부 실시양태에서, ORF2p는 비인간 종으로부터 선택된다. 일부 실시양태에서, 비-인간 종으로부터 선택된 ORF2p는 역전위 효율 및/또는 번역 효율을 향상시키기 위해 더 변형된다. 일부 실시양태에서, 세포는 면역 세포, 간세포, 심근세포, 망막 색소 상피 세포 또는 뉴런이다. 일부 실시양태에서, ORF2p는 핵 국소화 서열(NLS)을 포함한다. 일부 실시양태에서, ORF2p는 동일하거나 상이한 적어도 2개의 NLS를 포함한다. 일부 실시양태에서, NLS는 ORF1p, ORF2p 또는 이들 둘 다를 코딩하는 서열의 N-말단에 있다. 일부 실시양태에서, NLS는 ORF1p, ORF2p 또는 이들 둘 다를 코딩하는 서열의 C-말단에 있다. 일부 실시양태에서, NLS는 SV40으로부터 유래한다. 일부 실시양태에서, NLS는 뉴클레오플라스민(nucleoplasmin)으로부터 유래한다. 일부 실시양태에서, 적어도 2개의 NLS 중 제1 NLS는 SV40으로부터 유래하고, 적어도 2개의 NLS 중 제2 NLS는 뉴클레오플라스민으로부터 유래한다. 일부 실시양태에서, 적어도 2개의 NLS 중 제1 NLS 및 제2 NLS는 SV40으로부터 유래한다. 일부 실시양태에서, 적어도 2개의 NLS 중 제1 NLS 및 제2 NLS는 뉴클레오플

라스민으로부터 유래한다. 일부 실시양태에서, 적어도 2개의 NLS 각각은 동일하다.

[0017] **참고에 의한 포함**

[0018] 본 명세서에서 언급된 모든 간행물들, 특허들 및 특허출원들은 각각의 개별 간행물, 특허 또는 특허출원이 본원에 참고로 포함되는 것으로 구체적 및 개별적으로 표시된 것처럼 동일한 정도로 본원에 참고로 포함된다. 참고로 포함된 간행물 및 특허 또는 특허출원이 본 명세서에 함유된 개시내용과 모순되는 경우, 본 명세서는 임의의 이러한 모순되는 자료를 대체하고/하거나 이러한 자료보다 우선한다.

도면의 간단한 설명

[0019] 본 발명의 신규 특징은 첨부된 청구범위에 구체적으로 기재되어 있다. 본 발명의 특징 및 장점은 본 발명의 원리가 이용된 예시적인 실시양태가 기재된 하기 상세한 설명과 첨부된 도면(본원에서 "도"로서도 지칭됨)을 참고함으로써 더 잘 이해될 것이다.

도 1a는 레트로트랜스포존의 일반적인 작용 기작을 예시한다. (I)는 자율 레트로트랜스포존의 전체 수명주기를 나타내는 개략도이다. (II) LINE-1 레트로트랜스포존은 mRNA로서 발현되는 2개의 단백질 ORF1p 및 ORF2p를 코딩하는 LINE-1 요소를 포함한다. 바이시스트론(bicistronic) mRNA는 2개의 단백질로 번역되고, ORF2p가 리보좀에 의한 리드-쓰로우(read-through) 이벤트에 의해 번역될 때, 폴리 A 꼬리를 통해 그 자신의 mRNA의 3' 말단에 결합한다(III). ORF2p는 컨센서스 서열 TAAAA에서 절단하고, 이때 mRNA의 3' 말단에서 폴리 A는 하이브리드화하고 ORF2 단백질의 역전사효소 활성을 프라이밍한다. 단백질은 mRNA를 DNA로 다시 역전사하여, LINE-1 서열을 게놈의 새로운 위치 내로 다시 삽입한다(IV).

도 1b는 게놈(오른쪽) 내로의 통합을 위해 디자인될 수 있는 유전 페이로드(왼쪽)를 포함하는 mRNA 구축물의 개략도의 예시이다.

도 1c는 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한 다양한 예시적인 디자인을 보여준다. 여기서 박스로 표시된 GFP는 예시적인 형질전이유전자이다.

도 1d는 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한 다양한 예시적인 디자인을 보여준다. 여기서 상자로 표시된 GFP는 예시적인 형질전이유전자이다.

도 1e는 LINE 트랜스포존의 작용 기작과 레트로트랜스포존 부위 내로의 형질전이유전자 카고의 도입을 보여주는 LINE-1 역전위 주기의 개략도의 예시이다. LINE-1 레트로트랜스포존은 ORF1 및 ORF2라는 2개의 단백질을 코딩하는 게놈 서열이다. 이 요소들은 LINE-1 mRNA, ORF1 삼량체, 및 역전사효소 엔도뉴클레아제인 ORF2와 함께 RNA-단백질 복합체를 형성하는 단백질로 전사되고 번역된다. 이 복합체는 뉴클레아제로 다시 이동하고, 이 뉴클레아제는 5'-TTTTN-3' 모티프에서 DNA를 절단하고, mRNA의 폴리 A 꼬리 및 절제된 절단 DNA와 함께 RNA-DNA 하이브리드를 만듦으로써 ORF2 단백질에 의한 LINE-1 RNA의 역전사를 위해 프라이밍된다. cDNA로의 LINE-1의 역전사는 새로운 LINE-1 통합 이벤트로 이어진다.

도 2a는 구축물을 사용하여 GFP를 코딩하는 서열을 안정적으로 통합시킴으로써 예시적인 형질전이유전자 GFP를 발현시키기 위한 세 가지 예시적인 디자인을 보여준다. 72시간에서 예상된 GFP 발현 수준은 오른쪽에 표시되어 있다.

도 2b는 구축물을 사용하여 RFP, RFP 및 GFP, 또는 ORF2p 및 GFP를 코딩하는 서열을 안정적으로 통합시킴으로써 예시적인 형질전이유전자 GFP를 발현시키기 위한 세 가지 예시적인 디자인을 보여준다. 72시간에서 예상된 GFP 및 RFP 발현 수준은 오른쪽에 표시되어 있다.

도 3a는 통상의 circRNA 구조 및 형성의 예시적인 도표를 보여준다.

도 3b는 예시적인 RL-GAAA tectoRNA 모티프 디자인의 두 도면을 보여준다. 도 3b는 기재된 순서대로 각각 서열번호 111 및 112를 개시한다.

도 3c는 잠재적인 tectoRNA를 시험하기 위한 플랫폼으로서 칩-유동 조각(chip-flow piece) RNA의 예시적인 구조를 보여준다.

도 4a는 ORF2 폴리 A 영역에 결합하는 ORF2p를 보여주는 예시적인 개략도를 보여준다.

도 4b는 ORF2p와 MS2 RNA 결합 도메인의 융합체가 어떻게 ORF2를 코딩하는 mRNA의 3' UTR 내의 MS2 결합 RNA

서열에 결합하여 특이성을 증가시키는지를 보여주는 예시적인 개략도를 보여준다.

도 4c는 핵산을 특정 부위에서 세포의 게놈 내로 안정적으로 통합시키는 레트로트랜스포존 시스템의 예시적인 디자인을 보여준다. 상부 패널은 ORFp2-메가TAL DNA 결합 도메인 융합체를 사용하는 다자인을 보여주고, 이때 ORF2p의 DNA 결합 및 엔도뉴클레아제 활성은 돌연변이되어, 불활성 상태가 된다. 중간 패널은 엔도뉴클레아제 도메인이 또 다른 단백질의 높은 특이성 및 높은 신뢰도 뉴클레아제 도메인으로 대체되어 있는 키메라 ORF2p를 보여준다. 하부 패널은 융합 단백질이 ORF2 결합 부위뿐만 아니라 ORF2 부위 근처의 추가 DNA 서열에도 결합하도록 이종 단백질의 DNA 결합 도메인과 ORF2p의 융합체를 보여준다.

도 5는 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한 예시적인 구축물 (I) 내지 (X)을 보여준다.

도 6a는 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한, ORF1p를 코딩하는 서열을 가진 예시적인 구축물을 보여준다.

도 6b는 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한, ORF1p를 코딩하는 서열을 갖지 않은 예시적인 구축물을 보여준다.

도 7a는 G-사중체(quadruplex) 또는 슈도노트(pseudoknot)(서열번호 113)에 상응하는 구조를 5' UTR 내로 도입하고/하거나, xrRNA, 삼중체 모티프(기재된 순서대로 각각 서열번호 74, 54 및 53) 및/또는 비-A 뉴클레오타이드 잔기를 3' UTR 내로 도입하여, 5'-3' 엑소뉴클레아제, 예컨대, XRN1에 의한 분해 또는 3'-5' 엑소좀 분해를 억제함으로써 mRNA 반감기를 개선하는 예시적인 방법을 보여준다.

도 7b는 암 세포에 결합하여 항암 활성을 유도하는 키메라 수용체를 코딩하는 형질전이유전자를 발현하는 골수 세포의 예시적인 개략도를 보여준다.

도 7c는 키메라 수용체의 증가 및 연장된 발현에 대한, 도 7b에 기재된 바와 같이 암 세포에 결합하는 키메라 수용체를 코딩하는 벌크 또는 정제된 RNA 도입의 예상된 결과를 보여준다.

도 8a는 예시적인 플라스미드 디자인, 및 카고 핵산 서열을 가진 예상된 LINE-1 mRNA 전사체를 보여준다. 플라스미드는 (ORF1 및 ORF2 단백질 코딩 서열을 포함하는) LINE-1 서열, 및 GFP를 코딩하는 핵산 서열인 카고 서열을 갖고, 이때 GFP의 코딩 서열은 인트론에 의해 불연속된다. GFP는 서열이 게놈 내로 통합되고 인트론이 스플라이싱될 때까지 발현되지 않는다.

도 8b는 도 8a에 나타낸 플라스미드에 의해 코딩된 mRNA 전사체의 성공적인 통합 및 모의 형질감염된 세포에 비해 GFP의 발현을 보여주는 예시적인 결과를 보여준다(GFP 양성 세포의 평균 형광 강도의 배수 증가가 표시되어 있다). 모의 형질감염된 세포는 GFP 카고 서열을 결여하는 벡터에 의해 형질감염되었다.

도 8c는 도 8b에 나타낸 결과로부터의 예시적인 유세포분석 결과를 보여준다.

도 9a는 예시적인 플라스미드 디자인, 및 카고 핵산 서열을 가진 예상된 LINE-1 mRNA 전사체를 보여준다. 플라스미드는 (ORF1 및 ORF2 단백질 코딩 서열을 포함하는) LINE-1 서열, 및 CD5에 결합할 수 있는 세포외 영역 및 FCR 세포내 도메인과 PI3 키나제 동원 도메인을 포함하는 세포내 영역을 가진 재조합 키메라 융합 수용체 단백질(ATAK 수용체)을 코딩하는 핵산 서열인 카고 서열을 가진다. ATAK 수용체의 코딩 서열은 인트론에 의해 불연속된다.

도 9b는 도 9a에 나타낸 플라스미드에 의해 코딩된 mRNA 전사체의 성공적인 통합 및 모의 형질감염된 세포에 비해 ATAK의 발현을 보여주는 예시적인 결과를 보여준다(ATAK 양성 세포의 평균 형광 강도의 배수 증가가 표시되어 있다). 모의 형질감염된 세포는 ATAK 카고 서열을 결여하는 벡터에 의해 형질감염되었다. ATAK 수용체 단백질의 발현은 표지부착된 CD5 항체와의 결합에 의해 검출되었다.

도 9c는 도 9b에 제시된 결과로부터의 예시적인 유세포분석 결과를 보여준다.

도 10a는 예시적인 플라스미드 디자인, 및 카고 핵산 서열을 가진 예상된 LINE-1 mRNA 전사체를 보여준다. 플라스미드는 (ORF1 및 ORF2 단백질 코딩 서열을 포함하는) LINE-1 서열, 및 재조합 키메라 융합 수용체 단백질(ATAK 수용체)을 코딩하는 핵산 서열인 카고 서열을 갖고, 그 뒤에 T2A 자가절단 서열에 이어 분할 GFP 서열을 가진다(모두 LINE-1 서열에 대해 역방향으로 존재함). GFP의 코딩 서열은 인트론에 의해 불연속된다. 카고의 역전사 및 통합 후 예상된 mRNA가 도시되어 있다.

도 10b는 도 10a에 나타낸 플라스미드에 의해 코딩된 mRNA 전구체의 성공적인 통합 및 모의 형질감염된 세포에 비해 ATAK-T2A-GFP의 발현을 보여주는 예시적인 결과를 보여준다(GFP 및 ATAK 이중 양성 세포의 배수 변화가 표시되어 있다). 모의 형질감염된 세포는 ATAK 카고 서열을 결여하는 벡터에 의해 형질감염되었다. ATAK 수용체 단백질의 발현은 표지부착된 CD5 항체와의 결합에 의해 검출되었다.

도 10c는 도 10a에 나타낸 실험 설정을 이용하여 GFP 및 CD5 결합제(ATAK) 둘 다를 발현시키기 위한 2회 별도의 실험 실행으로부터의 대표적인 유세포분석 데이터를 보여준다.

도 10d는 도 10a에 나타낸 실험 설정을 이용하여 GFP 및 CD5 결합제(ATAK) 둘 다를 발현시키기 위한 2회 별도의 실험 실행으로부터의 대표적인 유세포분석 데이터를 보여준다.

도 11a는 역전위 기반 유전자 전달을 위한 예시적인 mRNA 구축물을 보여준다. ORF1 및 ORF2 서열은 2개의 상이한 mRNA 분자에 존재한다. ORF2p(ORF2) 코딩 mRNA는 역위된 GFP 코딩 서열을 포함한다.

도 11b는 ORF2-FLAG-GFPai mRNA만의 전기천공으로 정규화된, ORF1-mRNA 및 ORF2-FLAG-GFPai mRNA 둘 다의 전기천공 시 GFP의 발현(GFP 양성 세포의 평균 형광 강도의 배수 증가가 표시되어 있음)을 보여주는 예시적인 데이터를 도시한다.

도 12a는 상이한 양으로 ORF1-mRNA 및 ORF2-FLAG-GFPai mRNA를 전기천공하였을 때 GFP의 발현(GFP 양성 세포의 평균 형광 강도의 배수 증가가 표시되어 있음)을 보여주는 예시적인 데이터를 도시한다. 배수 증가는 1x ORF2-GFPao 및 1x ORF1 mRNA를 기준으로 한 것이다.

도 12b는 도 11a에 도시된 mRNA의 전기천공 후 GFP+ 세포의 예시적인 형광 현미경관찰 영상을 보여준다.

도 13a는 유전자 전달을 위해 ORF1 및 ORF2 서열이 2개의 상이한 mRNA 분자에 있는 예시적인 mRNA 구축물(상부 패널) 및 단일 mRNA 분자 상에 ORF1 및 ORF2 단백질 코딩 서열을 포함하는 LINE-1 mRNA 전사체(하부 패널)를 보여준다. mRNA는 3'에서 5'으로 3' UTR 내에 CMV-GFP 서열과 함께 바이시스트론 ORF1 및 ORF2 서열을 함유한다. 전달된 ORF2-cmv-GFP 안티센스(LINE-1 mRNA)의 역전위 시, 세포는 GFP를 발현할 것으로 예상된다.

도 13b는 도 13a에 도시된 구축물을 전기천공하였을 때 GFP의 발현(GFP 양성 세포의 평균 형광 강도의 배수 증가가 표시되어 있음)을 보여주는 예시적인 데이터를 도시한다.

도 14a는 다회 전기천공이 역전위 효율을 증가시키는지를 시험하기 위한 예시적인 실험 디자인을 보여준다. HEK293T 세포를 맥스사이트(Maxcyte) 시스템으로 48시간마다 전기천공하고 24시간 내지 72시간 동안 배양한 후, 유동을 이용하여 GFP 양성 세포에 대해 평가하였다.

도 14b는 도 14a에 따라 1회 내지 5회 전기천공하였을 때 표시된 시간에서 GFP의 발현(GFP 양성 세포의 평균 형광 강도의 배수 증가가 표시되어 있음)을 보여주는 예시적인 데이터를 도시한다.

도 15a는 mRNA 전달을 통해 역전위를 향상시키는 예시적인 구축물을 도시한다. 한 구축물에서, 핵 국소화 신호(NLS) 서열은 ORF2 서열의 C-말단에 융합된다(ORF2-NLS 융합체). 한 구축물에서, 인간 ORF2 대신에 밍크 고래 ORF2 서열이 사용되었다. 한 구축물에서, Alu 요소의 최소 서열(AJL-H33델타)은 LINE-1 서열의 3' UTR에 삽입된다. 한 구축물에서, MS2 헤어핀은 LINE-1 서열의 3' UTR에 삽입되고, MS2 헤어핀 결합 단백질(MCP) 서열은 ORF2 서열에 융합된다.

도 15b는 도 15a에 도시된 구축물을 사용하여 GFP의 발현(GFP 양성 세포의 평균 형광 강도의 배수 증가가 표시되어 있음)을 보여주는 예시적인 데이터를 도시한다.

도 16a는 유전자 전달을 위해 ORF1 및 ORF2 서열이 2개의 상이한 플라스미드 분자에 있는 예시적인 플라스미드 구축물(상부 패널), 및 ORF1과 ORF2 사이의 ORF간 서열의 다양한 대체를 가진 단일 mRNA 분자 상에 ORF1 및 ORF2 단백질 코딩 서열을 포함하는 LINE-1 mRNA 전사체를 코딩하는 플라스미드(하부 패널)를 보여준다.

도 16b는 도 16a에 도시된 구축물을 사용하여 GFP의 발현(GFP 양성 세포의 평균 형광 강도의 배수 증가가 표시되어 있음)을 보여주는 예시적인 데이터를 도시한다.

도 17a는 GFP 서열을 가진 단일 mRNA 분자 상에 ORF1 및 ORF2 단백질 코딩 서열을 포함하는 LINE-1 mRNA 전사체를 코딩하는 예시적인 플라스미드 구축물(상부 패널), 및 GFP 서열을 가진 단일 mRNA 분자 상에 ORF1 및 ORF2 단백질 코딩 서열을 포함하는 예시적인 LINE-1 mRNA 전사체를 도시한다.

도 17b는 도 17a에 도시된 구축물을 사용하여 Jurkat 세포에서 GFP의 발현(GFP 양성 세포의 평균 형광 강도의

배수 증가가 표시되어 있음)을 보여주는 예시적인 데이터를 도시한다. 플라스미드 구축물을 형질감염시켰고, mRNA 구축물을 전기천공하였다.

도 18a는 예시적인 플라스미드 디자인, 및 카고 핵산 서열을 가진 예상된 LINE-1 mRNA 전사체를 보여준다. 플라스미드는 (ORF1 및 ORF2 단백질 코딩 서열을 포함하는) LINE-1 서열, 및 재조합 키메라 융합 수용체 단백질(ATAK 수용체)을 코딩하는 핵산 서열인 카고 서열을 갖고, 그 뒤에 T2A 자가절단 서열에 이어 분할 GFP 서열을 가진다(모두 LINE-1 서열에 대해 역방향으로 존재함). GFP의 코딩 서열은 인트론에 의해 불연속된다. 카고의 역전사 및 통합 후 예상된 mRNA가 도시되어 있다.

도 18b는 골수 세포주(THP-1)에서 도 10a에 나타낸 플라스미드에 의해 코딩된 mRNA 전사체의 성공적인 통합 및 모의 형질감염된 세포에 비해 ATAK-T2A-GFP의 발현을 보여주는 예시적인 결과를 보여준다(GFP 및 ATAK 이중 양성 세포의 배수 변화가 표시되어 있다). 데이터는 GFP 코딩 서열을 갖지 않은 모의 플라스미드에 의해 형질감염된 세포에 대해 정규화된, 형질감염 후 6일째 날 발현을 나타낸다.

도 19는 세포 동기화를 위한 예시적인 실험 설정을 보여준다. 외생성 핵산의 전달 전에 세포 주기 단계를 기반으로 이종 세포 집단을 분류한다. 세포 주기 동기화는 전달된 외생성 핵산의 더 높은 발현 및 안정화를 야기할 것으로 예상된다. 세포 분류 후, 세포가 균질하지 않은 경우, 단계에서 세포 주기를 정지시키는 적합한 작용제와 함께 세포를 더 인큐베이션할 수 있다.

도 20은 예컨대, DNA 리가제(ligase) 억제제 SCR7을 유도하거나, 예를 들어, HUSH 복합체 TASOR 단백질에 대한 miRNA를 사용하여 숙주 감시 단백질을 억제함으로써, DNA 복구 경로를 억제하거나 억제하지 않으면서 DNA 이중 가닥 절단을 유도하여 레트로트랜스포존 효율을 증가시키는 예시적인 방법을 보여준다.

도 21은 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한 예시적인 구축물을 보여준다.

도 22는 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한 예시적인 구축물을 보여준다.

도 23은 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한 예시적인 구축물을 보여준다.

도 24는 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한 예시적인 구축물을 보여준다.

도 25는 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한 예시적인 구축물을 보여준다.

도 26은 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한 예시적인 구축물을 보여준다.

도 27은 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한 예시적인 구축물을 보여준다.

도 28은 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한 예시적인 구축물을 보여준다.

도 29는 레트로트랜스포존의 일반적인 작용 기작과 함께 2.4 kb 카고를 가진 예시적인 레트로트랜스포존 구축물(왼쪽), 및 HEK293 세포의 게놈 내로 통합된 핵산 서열로부터 카고에 의해 코딩된 형광 GFP 마커의 발현에 대한 대표적인 데이터(오른쪽)를 보여준다. 센스 방향으로 인트론에 의해 분할된 안티센스 GFP 유전자 및 프로모터 서열을 LINE-1의 3' UTR에 배치하면 GFP 카고의 재구성 및 역전위가 일어난다. 유세포분석에 의해 측정되고(오른쪽) 막대 그래프로 정량된(왼쪽 하부), 왼쪽에 표시된 구축물에 의해 형질감염된 293T 세포의 GFP 발현. ORF의 독시사이클린 유도 35일 후 수집된 데이터.

도 30은 막 단백질(CD5 결합제 키메라 항원 수용체, CD5-CAR)을 포함하는 3.0 kb 카고를 가진 예시적인 레트로트랜스포존 구축물(왼쪽), 및 HEK293 세포의 게놈 내로 통합된 핵산 서열로부터의 CD5 결합제의 발현에 대한 대표적인 유세포분석 데이터(오른쪽)를 보여준다. CD5 결합제 양성(+) 세포의 %가 삽입도에 표시되어 있다.

도 31은 막 단백질(CD5 결합제 키메라 항원 수용체, CD5-CAR, 및 자가절단 가능한 T2A 요소에 의해 분리된 GFP)을 포함하는 3.7 kb 카고를 가진 예시적인 레트로트랜스포존 구축물(상부), 및 CD5 결합제와 GFP의 발현을 입증하는 대표적인 유세포분석 데이터(하부)를 보여준다.

도 32는 막 단백질(HER2 결합제 키메라 항원 수용체, 및 자가절단 가능한 T2A 요소에 의해 분리된 GFP)을 포함하는 3.9 kb 카고를 가진 예시적인 레트로트랜스포존 구축물(상부), 및 HER2 결합제와 GFP의 발현을 입증하는 대표적인 유세포분석 데이터(하부)를 보여준다.

도 33a는 mRNA로서 전달된 레트로트랜스포존 요소의 전달에 대한 예시적인 데이터를 보여준다.

도 33b는 GFP 카고를 가진 LINE1 mRNA의 전달을 위한 트랜스(trans) 및 시스(cis) mRNA 디자인을 보여주는 개

략도를 보여준다(상부 패널). 별도의 ORF1 및 ORF2 mRNA와 함께 트랜스 mRNA를 사용한 293T 세포의 전기천공의 대표적인 결과. ORF2만을 가진 100 ㎍/㎖의 mRNA, 각각 100 ㎍/㎖의 ORF1 + ORF2 mRNA, 또는 ORF1 mRNA와 동일한 5' 및 3' UTR을 가진 GFP 코딩 mRNA로 293T 세포를 전기천공하였다(데이터 도표의 왼쪽 패널). 역전위 이벤트로 인해 GFP 양성 세포가 생성된다. 전기천공 4일 및 10일 후, 유세포분석으로 GFP 형광에 대해 세포를 어세이하였다. 모의 전기천공된 세포는 게이팅을 위한 음성 대조군 집단으로서 사용된다. 오른쪽 막대 그래프는 전기천공 동안 트랜스 mRNA, 및 시스 ORF1 및 ORF2 함유 mRNA 농도의 적정을 표시하는 대표적인 실험으로부터의 결과를 보여준다. 트랜스 mRNA는 실선 막대이고 시스 mRNA는 줄무늬이다. 20X는 전기천공 반응에서 2000 ㎍/㎖이다.

도 33c는 ORF1 및 ORF2-GFPai 트랜스 mRNA의 적정을 보여준다. 전기천공 동안 농도를 별도로 그리고 함께 200 ㎍/㎖까지 증가시키면 GFP 유전자 카고의 역전위가 증가된다.

도 33d는 도면에서 각각의 유세포분석 데이터 도표 위에 표시된 다양한 구축물들에 대한 예시적인 데이터를 보여주는 것으로, 상부 패널은 4일째 날 데이터이고, 하부 패널은 13일째 날 데이터이다. 오른쪽 도면은 배양물에서 GFP 발현 세포의 광학 현미경 영상과 형광 현미경 영상을 보여준다. 13일째 날 구축물당 통합된 카고의 카피는 오른쪽 하부에 표시되어 있다. 다양한 LINE-1 플라스미드에 의해 형질감염된, LINE-1 mRNA(레트로-mRNA), 및 ORF1 및 ORF2-GFP mRNA 전기천공된 세포로부터의 게놈 DNA 통합에 대한 qPCR 어세이가 표시되어 있다. 2개의 qPCR 프라이머-프로브 세트, 즉 하우스킵핑 유전자 RPS30을 위한 qPCR 프라이머-프로브 세트 및 GFP 유전자를 위한 qPCR 프라이머-프로브 세트가 사용되었다. 플라스미드에 의해 형질감염된 세포는 SV40 유지 서열을 함유하지 않는 플라스미드를 사용한다. 세포당 통합은 플라스미드 및 게놈 DNA의 표준 곡선의 보간을 통해 샘플당 카피 수를 측정하고 293T 세포당 2개의 RPS30 카피에 대해 정규화함으로써 계산된다. 오차 막대는 3회 기술적 재현실험 측정치의 표준 편차를 나타낸다.

도 34는 표시된 세포주에서의 예시적인 레트로트랜스포존 구축물(왼쪽) 및 발현 데이터(오른쪽)를 보여준다.

도 35는 K562, 293T 및 THP1 세포(상부 패널)에서 LINE 1 GFP 구축물의 발현을 보여주는 유세포분석 데이터; 및 K562 및 THP-1 세포주에서 세포당 LINE-2-GFP mRNA의 통합 수(하부 패널)를 보여준다.

도 36은 1차 T 세포에서 LINE 1 GFP 구축물의 발현을 보여주는 유세포분석 데이터를 보여준다(왼쪽). 세포당 통합은 오른쪽 그래프에 표시되어 있다. 데이터는 전기천공 후 6일째 날에 수집되었다.

도 37a는 단리된 1차 T 세포의 활성화, 배양 시간, 전기천공 및 GFP 발현 어세이의 개략도를 보여준다.

도 37b는 도면에 표시된 냉동-해동 전 및 후 표시된 농도에서 1차 T 세포에서의 LINE 1 GFP mRNA 구축물의 발현을 보여주는 유세포분석 데이터를 보여준다. 세포당 통합은 막대 도표로 제시되어 있다. GFP 카고를 사용한 레트로-mRNA 전기천공을 이용한 GFP 발현. 전기천공 후 4일째 날 및 전기천공 후 배양한 지 15일째 날에 GFP 발현을 어세이하였다. 이 시간 동안 1차 T 세포를 냉동보존하고 해동시켰다. GFP 통합에 대한 qPCR 통합 어세이. 20X 샘플로부터 게놈 DNA를 단리하고 GFP의 카피에 대해 어세이하였다.

도 38은 세포 유형에 걸쳐 레트로트랜스포존 통합 및 발현의 결과 요약을 보여준다.

도 39는 세포 요법을 위한 CART 세포, NK 세포, 뉴런 및 기타 세포의 용도, 및 유전자 요법, 유전자 편집, 전사 조절 및 게놈 조작을 포함하나 이들로 제한되지 않는 생체내 적용의 용도를 포함하나 이들로 제한되지 않는, 본원에 기재된 기술의 다양한 적용을 보여준다.

도 40은 2000 ng/㎕ LINE1-GFP mRNA로 전기천공한 GFP+ 293T 세포의 분류 및 농후화를 보여주는 예시적인 유세포분석 데이터를 도시한다. 첫 번째 패널은 LINE1-GFP mRNA의 부재 하에서 모의 전기천공된 세포에 대한 유세포분석 데이터를 보여준다. 두 번째 패널은 LINE1-GFP mRNA로 전기천공한 분류되지 않은 세포에 대해 전기천공 후 5일째 날에 수집된 유세포분석 데이터를 보여준다. 두 번째 패널의 GFP+ 세포를 분류하였고 유세포분석 데이터는 세 번째 패널에 제시되어 있다. 세 번째 패널의 GFP+ 세포를 분류 후 9일 동안 배양하고 10^3 또는 10^4 GFP 형광 강도 게이트를 사용하여 재분류하였다. 네 번째 패널은 재분류한 후 4일째 날에 수집된 10^3 GFP 게이트에서 GFP+를 사용하여 재분류한 세포에 대한 유세포분석 데이터를 보여준다. 다섯 번째 패널은 재분류한 후 4일째 날에 수집된 10^3 GFP 게이트에서 GFP+를 사용하여 재분류한 세포에 대한 유세포분석 데이터를 보여준다.

도 41a는 정량 PCR을 이용하여 세포당 GFP 코딩 핵산의 게놈 통합을 평가하기 위한 GFP(NB2 플라스미드) 및 하우스킵핑 유전자(FAU)에 대한 표준 곡선을 보여준다.

도 41b는 게놈 통합의 정량을 위한 도 41a의 표준 곡선의 보간을 도시하는 예시적인 그래프의 결과를 보여준다.

도 41c는 도 40에 나타낸 바와 같은 LINE1-GFP mRNA 전기천공 및 이중 분류 후 293T 세포의 게놈 내로 통합된 GFP 유전자의 수를 보여준다. qPCR에 따라 10^3 GFP+ 세포 및 10^4 GFP+ 세포에서 게이팅될 때 세포당 평균 GFP 통합 수가 표시되어 있다.

도 42는 전기천공 후 3일 동안 배양한 후 전기천공 용액 중의 LINE1-GFP mRNA를 표시된 적정량(ng/㎕)으로 사용하여 전기천공한 GFP+ 293T 세포를 보여주는 예시적인 유세포분석 데이터를 도시한다.

도 43은 전기천공 후 5일 동안 배양한 후 전기천공 용액 중의 LINE1-GFP mRNA를 표시된 적정량(ng/㎕)으로 사용하여 전기천공한 GFP+ 293T 세포를 보여주는 예시적인 유세포분석 데이터를 도시한다.

도 44는 전기천공 후 7일 동안 배양한 후 전기천공 용액 중의 LINE1-GFP mRNA를 표시된 적정량(ng/㎕)으로 사용하여 전기천공 GFP+ 293T 세포를 보여주는 예시적인 유세포분석 데이터를 도시한다.

도 45는 도 42 내지 44에 따라 전기천공 후 3일, 5일 또는 7일 동안 배양한 후, 전기천공 용액 중의 LINE1-GFP mRNA를 표시된 적정량(ng/㎕)으로 사용하여 전기천공한 293T 세포의 qPCR에 따른 게놈당 GFP 통합 수의 그래프(상부), 및 도 42 내지 44의 데이터에 따른 통합 동역학의 그래프(하부)를 보여준다.

도 46은 전기천공 후 6일 동안 배양한 후, 전기천공 용액 중의 LINE1-GFP mRNA를 표시된 적정량(ng/㎕)으로 사용하여 전기천공한 GFP+ K562 세포를 보여주는 예시적인 유세포분석 데이터(오른쪽), 및 qPCR에 따른 게놈당 GFP 통합 수의 그래프(왼쪽)를 도시한다.

도 47은 전기천공 후 3일 동안 배양한 후, 표시된 적정량의 LINE1-GFP mRNA로 전기천공한 GFP+ 인간 1차 단핵구를 보여주는 예시적인 유세포분석 데이터(상부), 및 qPCR에 따른 게놈당 GFP 통합 수의 그래프(하부)를 도시한다.

도 48은 전기천공 후 4일 동안 배양한 후, 2000 ng/㎕ LINE1-GFP mRNA 및 100 ng/㎕, 200 ng/㎕ 또는 300 ng/㎕의 BRCA1 표적화 siRNA(siBRCA1)로 전기천공한 GFP+ 293T 세포를 보여주는 예시적인 유세포분석 데이터(하부), 및 qPCR에 따른 게놈당 GFP 통합 수의 그래프(상부)를 도시한다.

도 49는 전기천공 후 6일 동안 배양한 후, 2000 ng/㎕ LINE1-GFP mRNA 및 100 ng/㎕의 RNASEL 표적화 siRNA(siRNASEL), ADAR1 표적화 siRNA(siADAR1) 또는 ADAR2 표적화 siRNA(siADAR2)로 전기천공한 GFP+ 293T 세포를 보여주는 예시적인 유세포분석 데이터(하부), 및 qPCR에 따른 게놈당 GFP 통합 수의 그래프(상부)를 도시한다.

50은 전기천공 후 6일 동안 배양한 후, 2000 ng/㎕ LINE1-GFP mRNA 및 100 ng/㎕의 APOBEC3C 표적화 siRNA(siAPOBEC3C) 또는 FAM208A 표적화 siRNA(siFAM208A)로 전기천공한 GFP+ 293T 세포를 보여주는 예시적인 유세포분석 데이터(하부), 및 qPCR에 따른 게놈당 GFP 통합 수의 그래프(상부)를 도시한다.

도 51은 전기천공 후 6일 동안 배양한 후, 1000 ng/㎕ 또는 1500 ng/㎕ LINE1-GFP mRNA, 및 25 ng/㎕, 50 ng/㎕ 또는 75 ng/㎕의 RNASEL 표적화 siRNA(siRNASEL), ADAR1 표적화 siRNA(siADAR1), ADAR2 표적화 siRNA(siADAR2) 및 BRCA1 표적화 siRNA(siBRCA1)를 함유하는 siRNA 칵테일로 전기천공한 GFP+ 293T 세포를 보여주는 예시적인 유세포분석 데이터(하부), 및 qPCR에 따른 게놈당 GFP 통합 수의 그래프(상부)를 도시한다.

도 52는 전기천공 후 5일 동안 배양한 후, 1000 ng/㎕ LINE1-GFP mRNA, 및 25 ng/㎕, 50 ng/㎕ 또는 75 ng/㎕의 RNASEL 표적화 siRNA(siRNASEL), ADAR1 표적화 siRNA(siADAR1), ADAR2 표적화 siRNA(siADAR2) 및 BRCA1 표적화 siRNA(siBRCA1)를 함유하는 siRNA 칵테일로 전기천공한 GFP+ K562 세포를 보여주는 예시적인 유세포분석 데이터(하부), 및 qPCR에 따른 세포당 GFP 통합 수의 그래프(상부)를 도시한다.

도 53은 예시적인 LINE1-GFP mRNA 구축물의 외래 핵 국소화 서열(NLS) 및 예시적인 ORF1p 및 ORF2p 돌연변이의 예시적인 위치를 보여주는 개략도를 도시한다.

도 54a는 NLS가 ORF1을 코딩하는 서열의 N-말단에서 삽입되어 있는 예시적인 LINE1-GFP 구축물을 보여주는 개략도를 도시한다.

도 54b는 표시된 구축물을 293T 세포 내로 전기천공한 후 4일째 날에 세포당 GFP 통합을 보여주는 막대 그래프를 도시한다.

도 54c는 표시된 구축물의 전기천공 후 4일째 날에 GFP+ 293T 세포를 보여주는 예시적인 유세포분석을

도시한다.

도 55a는 NLS가 ORF1을 코딩하는 서열의 C-말단에서 삽입되어 있는 예시적인 LINE1-GFP 구축물을 보여주는 개략도를 도시한다.

도 55b는 표시된 구축물을 293T 세포 내로 전기천공한 후 4일째 날에 세포당 GFP 통합을 보여주는 막대 그래프를 도시한다.

도 55c는 표시된 구축물의 전기천공 후 4일째 날에 GFP+ 293T 세포를 보여주는 예시적인 유세포분석을 도시한다.

도 56a는 NLS가 ORF2를 코딩하는 서열의 N-말단에서 삽입되어 있는 예시적인 LINE1-GFP 구축물을 보여주는 개략도를 도시한다.

도 56b는 표시된 구축물을 293T 세포 내로 전기천공한 후 4일째 날에 세포당 GFP 통합을 보여주는 막대 그래프를 도시한다.

도 56c는 표시된 구축물의 전기천공 후 4일째 날에 GFP+ 293T 세포를 보여주는 예시적인 유세포분석을 도시한다.

도 57a는 NLS 및 링커가 ORF2를 코딩하는 서열의 N-말단에서 삽입되어 있는 예시적인 LINE1-GFP 구축물을 보여주는 개략도를 도시한다.

도 57b는 표시된 구축물을 293T 세포 내로 전기천공한 후 5일째 날에 세포당 GFP 통합을 보여주는 막대 그래프를 도시한다.

도 57c는 표시된 구축물의 전기천공 후 5일째 날에 GFP+ 293T 세포를 보여주는 예시적인 유세포분석을 도시한다.

도 58a는 NLS가 ORF2를 코딩하는 서열의 C-말단에서 삽입되어 있는 예시적인 LINE1-GFP 구축물을 보여주는 개략도를 도시한다.

도 58b는 표시된 구축물을 293T 세포 내로 전기천공한 후 5일째 날에 세포당 GFP 통합을 보여주는 막대 그래프를 도시한다.

도 58c는 표시된 구축물의 전기천공 후 5일째 날에 GFP+ 293T 세포를 보여주는 예시적인 유세포분석을 도시한다.

발명을 실시하기 위한 구체적인 내용

[0020] 본 발명은 부분적으로, 세포 게놈 내로의 유전 카고(큰 유전 카고)의 전달 및 통합을 달성하도록 폴리뉴클레오타이드를 디자인하고 개발할 수 있다는 흥미로운 발견으로부터 비롯된다. 일부 실시양태에서, 상기 폴리뉴클레오타이드는 (i) 안정한 발현을 위한 유전 물질, 및 (ii) 안전하고 효과적인 비-바이러스 수단으로 유전 물질을 세포 내로 안정적으로 통합시킬 수 있게 하는 자가통합 게놈 통합 기구를 포함한다. 더욱이, 유전 물질은 리보좀 유전자좌 이외의 유전자좌에서 통합될 수 있고/있거나; 유전 물질은 부위 특이적으로 통합될 수 있고/있거나; 통합된 유전 물질은 세포의 천연 침묵 기구를 유발하지 않으면서 발현되는 것으로 보인다.

[0021] 클러스터링된 규칙적으로 이격된 짧은 팔린드롬 반복부(CRISPR)는 분자생물학 분야에 혁명을 일으켰고 강력한 유전자 편집으로도 개발되었다. 이것은 상동성 유도 복구(HDR)를 이용하고 게놈 부위로 향할 수 있다. CRISPR/Cas9는 천연 생성 RNA 가이딩 엔도뉴클레아제이다. CRISPR/Cas9 시스템은 부위 특이적 유전자 편집 및 기타 적용에 대한 큰 가능성을 입증하였지만, 이것이 특히 생체내 인간 유전자 요법에 사용되는 경우 그의 효능에 영향을 미치는 몇 가지 요인이 해결되어야 한다. 이 요인은 표적 DNA 부위 선택, sgRNA 디자인, 오프-표적 절단, HDR 대 NHEJ의 발생률/효율, Cas9 활성 및 전달 방법을 포함한다. 전달은 생체내 적용을 위한 CRISPR의 사용에 대한 주요 장애물로 남아 있다. 징크 핑거 뉴클레아제 ZFN은 Cys2-His2 징크 핑거 단백질(ZFP)과, *Fok*I 엔도뉴클레아제로부터 유래한 비특이적 DNA 제한 효소의 융합 단백질이다. ZFP의 사용 시 과제는 원하는 서열의 고친화성 결합을 위한 ZFP의 디자인 및 조작을 포함하고, 이것은 사소하지 않다. 또한, 모든 서열들이 ZFP 결합에 사용될 수 있는 것은 아니므로, 부위 선택이 제한된다. 또 다른 유의미한 과제는 오프-표적 절단이다. 전사 활성화제 유사 이펙터 뉴클레아제(TALEN)는 TALE 및 *Fok*I 뉴클레아제로 구성된 융합 단백질이다. 오프-표적 절단이 우려사항으로 남아 있지만, TALEN은 한 병렬 비교 연구에서 ZFN보다 더 특이적이고 더 적은 세포독성을 나

타내는 것으로 밝혀졌다. 그러나, TALEN은 실질적으로 더 크고, TALEN만을 코딩하는 cDNA는 3 kb이다. 이것은 전달 비히클 카고 크기 제한으로 인해 한 쌍의 TALEN의 전달을 한 쌍의 ZFN보다 더 어렵게 만든다. 추가로, 일부 바이러스 벡터들에서 TALEN의 팩키징 및 전달은 TALEN 서열의 높은 반복 수준으로 인해 문제가 될 수 있다. 돌연변이체 Cas9 시스템인, 불활성 dCas9와 *Fok*I 뉴클레아제 이량체의 융합 단백질은 특이성을 증가시키고 오프-표적 절단을 감소시키며, 잠재적인 표적 부위의 수는 PAM 및 다른 sgRNA 디자인 제약으로 인해 더 낮다.

[0022] 본 발명은 유전자 요법을 비롯한 요법을 동물 및 인간에게 제공하기 위해 트랜스포존 기반 벡터를 포함하는, 효과적이고 효율적인 신규 조성물을 제공함으로써 전술된 문제점들을 해결한다. 본 발명은 요법을 동물 및 인간에게 제공하기 위해 이 조성물을 사용하는 방법을 제공한다. 이 트랜스포존 기반 벡터는 투여 후 원하는 효과를 수용자에게 제공하는 데 유용한 의약의 제조에 사용될 수 있다. 유전자 요법은 트랜스포존 기반 벡터를 사용하여 외생성 유전자와 같은 유전자를 동물 내로 도입하는 것을 포함하나, 이것으로 제한되지 않는다. 이 유전자는 수용자에서 다양한 기능들, 예컨대, 핵산, 예를 들어, RNA의 제조를 위한 코딩, 또는 단백질 및 펩티드의 제조를 위한 코딩에 기여할 수 있다. 본 발명은 관심 있는 유전자, 프로모터, 삽입체 서열, 폴리 A 및 임의의 조절 서열을 포함하는 폴리뉴클레오타이드 서열의 효율적인 통합을 용이하게 할 수 있다. 본 발명은 인간 LINE-1 요소가 인간 세포뿐만 아니라 다른 동물 종의 세포에서도 역전위를 할 수 있고 세포의 게놈 내로의 유전 카고의 효율적인 전달 및 통합을 달성하기 위해 다양한 방식으로 조작될 수 있다는 발견에 기반한다. 이러한 LINE-1 요소는 유전 장애 및 암의 진단 및 치료에서의 용도를 포함하나 이것으로 제한되지 않는, 인간 및 동물 유전학에서 다양한 용도를 가진다. 본 발명의 LINE-1 요소는 다양한 질환들의 다양한 표현형 효과의 치료에도 유용하다. 예를 들어, LINE-1 요소는 항-종양발생 유전자 생성물을 코딩하는 DNA를 암 세포 내로 전달하는 데 사용될 수 있다. 본 발명의 LINE-1 요소의 다른 용도는 본 명세서를 읽을 때 숙련된 당업자에게 명확해질 것이다.

[0023] 일반적으로, 인간 LINE-1 요소는 내부 프로모터를 가진 5' UTR, 2개의 비-중첩 리딩 프레임(ORF1 및 ORF2), 200 bp 3' UTR 및 3' 폴리 A 꼬리를 포함한다. LINE-1 레트로트랜스포존은 LINE-1 ORF2 N-말단에서 엔도뉴클레아제 도메인도 포함할 수 있다. LINE-1이 엔도뉴클레아제를 코딩한다는 발견은 상기 요소가 자율 역전위를 할 수 있음을 입증한다. LINE-1은 그의 역전사 및 통합을 매개하는 비-중첩 기능성 도메인을 함유하는 모듈식 단백질이다. 일부 실시양태에서, LINE-1 엔도뉴클레아제 그 자체의 서열 특이성은 변경될 수 있거나, LINE-1 엔도뉴클레아제는 또 다른 부위 특이적 엔도뉴클레아제로 대체될 수 있다.

[0024] LINE-1 레트로트랜스포존은 재조합 DNA 기술을 이용함으로써, 레트로트랜스포존이 상당한 길이(최대 1 kb, 또는 1 kb 초과, 예를 들어, 5, 6, 7, 8, 9 또는 10 kb 초과)의 이종 또는 동종 핵산 서열을 세포의 게놈 내로 삽입하는 데 적합해지게 만드는 다른 핵산 요소를 포함하고/하거나 이러한 핵산 요소와 인접하도록 조작될 수 있다. LINE-1 레트로트랜스포존은 동일한 유형의 기술을 이용함으로써, 세포의 게놈 내로의 이종 또는 동종 핵산의 핵산 서열의 삽입이 부위 지정되도록(이러한 DNA가 삽입되는 부위는 알려져 있음) 조작될 수도 있다. 대안적으로, LINE-1 레트로트랜스포존은 DNA의 삽입 부위가 무작위이도록 조작될 수 있다. 레트로트랜스포존은 정상적으로 전사 침묵 상태인 DNA 영역 내로의 원하는 DNA 서열의 삽입을 달성하도록 조작될 수도 있고, 이때 DNA 서열은 세포에서 유전자의 정상적인 발현을 파괴하지 않는 방식으로 발현된다. 일부 실시양태에서, 통합 또는 역전위는 트랜스 배향으로 일어난다. 일부 실시양태에서, 통합 또는 역전위는 시스 배향으로 일어난다.

[0025] LINE-1이 인간 세포에 천연적으로 존재하기 때문에, 구축물이 인간 세포 내에 배치될 때, 이 구축물은 면역 시스템에 의해 외부 물질로서 거절되지 않아야 한다. 추가로, LINE-1 역통합의 기작은 유전자의 한 카피만이 임의의 특정 염색체 위치에서 통합되게 한다. 따라서, 시스템으로 구축된 카피 수 조절이 있다. 대조적으로, 통상의 플라스미드를 사용하는 유전자 전달 절차는 카피 수에 대한 조절을 거의 또는 전혀 제공하지 않고 종종 동일한 게놈 위치 내로 직렬로 통합된 DNA 분자의 복잡한 어레이를 초래한다.

[0026] 모든 용어들은 당업자에 의해 이해될 바와 같이 이해되어야 한다. 달리 정의되어 있지 않은 한, 본원에서 사용된 모든 기술 용어들 및 과학 용어들은 본 개시내용이 속하는 분야에서 통상의 기술을 가진 자에 의해 통상적으로 이해되는 의미와 동일한 의미를 가진다.

[0027] 본원에서 사용된 단락 제목은 조직화 목적만을 위한 것이고 기재된 보호대상을 제한하는 것으로서 해석되어서는 안 된다.

[0028] 본원에서 사용된 바와 같이, 문맥이 달리 명시하지 않은 한, 단수형 용어는 복수형 용어도 포함하기 위한 것이다.

[0029] 본원에서, 달리 언급되어 있지 않은 한, "또는"의 사용은 "및/또는"을 의미한다. 본원에서 사용된 용어 "및/또

는" 및 "이들의 임의의 조합" 및 이들의 문법적으로 동등한 용어는 교환 가능하게 사용될 수 있다. 이 용어들은 임의의 조합이 구체적으로 예상됨을 전달할 수 있다. 오로지 예시 목적으로, 하기 어구 "A, B 및/또는 C" 또는 "A, B, C, 또는 이들의 임의의 조합"은 "개별적으로 A; 개별적으로 B; 개별적으로 C; A 및 B; B 및 C; A 및 C; 및 A, B 및 C"를 의미할 수 있다. 문맥이 구체적으로 이접적 사용을 지칭하지 않은 한, 용어 "또는"은 결합적 또는 분리적으로 사용될 수 있다.

[0030] 용어 "약" 또는 "대략"은 부분적으로 값이 어떻게 측정되거나 결정되는지, 즉 측정 시스템의 한계에 의해 좌우될, 당분야에서 통상의 기술을 가진 자에 의해 결정된 특정 값에 대한 허용 가능한 오차 범위 이내에 있음을 의미할 수 있다. 예를 들어, "약"은 당분야의 관행에 따라 1 또는 1 초과의 표준 편차 이내에 있음을 의미할 수 있다. 대안적으로, "약"은 주어진 값의 최대 20%, 최대 10%, 최대 5% 또는 최대 1%의 범위를 의미할 수 있다. 대안적으로, 특히 생물학적 시스템 또는 과정과 관련하여, 이 용어는 값의 한 자릿수 이내, 5배 이내, 보다 바람직하게는 2배 이내에 있음을 의미할 수 있다. 특정 값이 본원 및 청구범위에 기재되어 있는 경우, 달리 언급되어 있지 않은 한, 특정 값에 대한 허용 가능한 오차 범위 이내에 있음을 의미하는 용어 "약"이 가정되어야 한다.

[0031] 본 명세서 및 청구범위에서 사용된 바와 같이, 용어 "포함하는"(및 임의의 형태의 포함하는, 예컨대, "포함한다" 및 "포함하고"), "가진"(및 임의의 형태의 가진, 예컨대, "가진다" 및 "갖고"), "포괄하는"(및 임의의 형태의 포괄하는, 예컨대, "포괄한다" 및 "포괄하고") 또는 "함유하는"(및 임의의 형태의 함유하는, 예컨대, "함유한다" 및 "함유하고")은 포괄적이거나 제한이 없고 언급되지 않은 추가 요소 또는 방법 단계를 배제하지 않는다. 본 명세서에서 논의된 임의의 실시양태는 본 개시내용의 임의의 방법 또는 조성물과 관련하여 구현될 수 있고, 그 반대도 마찬가지일 것으로 예상된다. 더욱이, 본 개시내용의 조성물은 본 개시내용의 방법을 달성하는 데 사용될 수 있다.

[0032] 본 명세서에서 "일부 실시양태", "실시양태", "한 실시양태" 또는 "다른 실시양태"의 언급은 실시양태와 관련하여 설명된 특정 특징, 구조 또는 특성이 본 개시내용의 적어도 일부 실시양태에 포함되나, 반드시 모든 실시양태에 포함되는 것이 아님을 의미한다. 본 개시내용의 이해를 용이하게 하기 위해, 다수의 용어들 및 어구들이 이하에 정의된다.

[0033] 본 개시내용의 다양한 특징들이 단일 실시양태와 관련하여 설명될 수 있지만, 상기 특징들은 따로 또는 임의의 적합한 조합으로 제공될 수도 있다. 대조적으로, 본 개시내용이 명료성을 위해 별도의 실시양태와 관련하여 본원에서 설명될 수 있지만, 본 개시내용은 단일 실시양태로 구현될 수도 있다.

[0034] 본 개시내용의 적용은 세포에서의 외생성 핵산의 발현과 관련된 방법 및 조성물을 포함하나, 이들로 제한되지 않는다. 일부 실시양태에서, 외생성 핵산은 골수 세포와 같은 세포의 게놈 내로 안정적으로 통합되도록 구성된다. 일부 실시양태에서, 외생성 핵산의 안정한 통합은 게놈 내의 특정 표적에서 일어날 수 있다. 일부 실시양태에서, 외생성 핵산은 하나 이상의 코딩 서열을 포함한다. 일부 실시양태에서, 외생성 핵산은 면역 수용체를 코딩하는 핵산 서열을 포함하는 하나 이상의 코딩 서열을 포함할 수 있다. 일부 실시양태에서, 본 개시내용은 면역 반응 기능과 관련된 막횡단 수용체(예를 들어, 식세포 수용체 또는 합성 키메라 항원 수용체)를 코딩하는 핵산을 인간 대식세포 또는 수지상 세포 또는 적합한 골수 세포 또는 골수 전구체 세포 내로 안정적으로 통합시키는 방법 및 조성물을 제공한다. 외생성 핵산은 내생적으로 세포에 이미 존재할 수 있는 서열을 포함하는지와 관계없이 원래 세포에 존재하지 않았고 세포의 외부로부터 첨가된 핵산을 지칭할 수 있다. 외생성 핵산은 DNA 또는 RNA 분자일 수 있다. 외생성 핵산은 형질전이유전자를 코딩하는 서열을 포함할 수 있다. 외생성 핵산은 재조합 단백질, 예컨대, 재조합 수용체 또는 키메라 항원 수용체(CAR)를 코딩할 수 있다. 외생성 핵산은 외생성 핵산이 세포 내부로 전달된다는 점과 관련하여 "유전 카고"로서 지칭될 수 있다. 유전 카고는 DNA 또는 RNA일 수 있다. 유전 물질은 일반적으로 화학적($CaCl_2$ 매개 형질감염) 또는 물리적(전기천공) 또는 생물학적(예를 들어, 바이러스 감염 또는 형질도입) 수단을 이용함으로써 몇몇 상이한 공지된 기법들에 의해 생체외에서 세포 내부로 전달될 수 있다.

[0035] 본원은 세포 내로의 유전 물질의 안정한 비-바이러스 전달 및 통합을 위한 조성물 및 방법을 제공한다. 한 측면에서, 유전 물질은 자가통합 폴리뉴클레오타이드이다. 유전 물질은 세포의 게놈 내로 안정적으로 통합될 수 있다. 세포는 인간 세포일 수 있다. 유전 물질을 세포의 게놈 내로 안전하고 신뢰할 수 있게 통합시키는 방법이 디자인된다.

[0036] 본원은 치료 유효량의 하나 이상의 폴리핵산, 또는 하나 이상의 폴리핵산을 코딩하는 적어도 하나의 벡터를 포함하고 인간 대상체에서 실질적으로 면역원성을 나타내지 않는 약학 조성물을 제공하는 것으로, 이때 상기 하나

이상의 폴리핵산은 (a) 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소; 및 (b) 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함하는 삽입체 서열을 포함하고, 상기 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 세포의 게놈 내로의 삽입체 서열의 통합을 촉진한다.

[0037] 일부 실시양태에서, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 하나 이상의 긴 산재된 핵 요소(LINE) 폴리펩티드를 포함하고, 이때 하나 이상의 LINE 폴리펩티드는 (i) 인간 ORF1p 또는 이의 기능성 단편 및 (ii) 인간 ORF2p 또는 이의 기능성 단편을 포함한다.

[0038] 일부 실시양태에서, 삽입체 서열은 인간 세포의 게놈 내로 안정적으로 통합 및/또는 역전위된다.

[0039] 일부 실시양태에서, 인간 세포는 T 세포, B 세포, 골수 세포, 단핵구, 대식세포 및 수지상 세포로 구성된 군으로부터 선택된 면역 세포이다.

[0040] 일부 실시양태에서, 삽입체 서열은 (i) 하나 이상의 폴리핵산에 의해 코딩된 엔도뉴클레아제에 의한 표적 부위의 DNA 가닥의 절단에 의해, (ii) 표적 프라이밍 역전사(TPRT)를 통해, 또는 (iii) 게놈의 DNA 표적 부위 내로의 삽입체 서열의 역스플라이싱을 통해 게놈 내로 통합된다.

[0041] 일부 실시양태에서, 삽입체 서열은 인간 ORF2p의 엔도뉴클레아제 도메인의 특이성을 이용함으로써 폴리 T 부위에서 게놈 내로 통합된다.

[0042] 일부 실시양태에서, 폴리 T 부위는 서열 TTTTTA를 포함한다.

[0043] 일부 실시양태에서, 하나 이상의 폴리핵산은 게놈 내의 표적 부위에 상보적인 상동성 아암을 포함한다.

[0044] 일부 실시양태에서, 삽입체 서열은 (a) 리보좀 유전자좌가 아닌 유전자좌에서 게놈 내로 통합되거나; (b) 게놈의 유전자 또는 유전자의 조절 영역 내로 통합됨으로써, 유전자를 파괴하거나 유전자의 발현을 하향조절하거나; (c) 게놈의 유전자 또는 유전자의 조절 영역 내로 통합됨으로써, 유전자의 발현을 상향조절하거나; 또는 (d) 게놈 내로 통합되고 게놈의 유전자를 대체한다.

[0045] 일부 실시양태에서, 약학 조성물은 (i) 하나 이상의 siRNA 및/또는 (ii) RNA 가이드 서열, 또는 RNA 가이드 서열을 코딩하는 폴리핵산을 추가로 포함하고, 이때 RNA 가이드 서열은 게놈의 DNA 표적 부위를 표적화하고, 삽입체 서열은 게놈의 DNA 표적 부위에서 게놈 내로 통합된다.

[0046] 일부 실시양태에서, 하나 이상의 유전자는 본원에서 제공된 방법에 의해 녹다운된다. 일부 실시양태에서, 하나 이상의 siRNA는 본원에 기재된 조성물 또는 방법에 사용된다. 예를 들어, LINE-1을 억제할 수 있는 경로의 조절을 통해 통합을 향상시키도록 하나 이상의 유전자를 녹다운할 수 있다. 일부 실시양태에서, 녹다운되는 하나 이상의 유전자는 ADAR1, ADAR2(ADAR1B), APOBEC3C, BRCA1, let-7 miRNA, RNase L, TASHOR(HUSH 복합체) 및/또는 RAD51을 포함한다. 예를 들어, RNase L의 녹다운은 LINE-1로부터 전사된 mRNA와 같은 mRNA의 분해를 억제하거나 방해함으로써 통합을 향상시키는 데 사용될 수 있다. 예를 들어, ADAR1, ADAR2(ADAR1B) 및/또는 BRCA1의 녹다운은 ADAR1, ADAR2(ADAR1B) 및/또는 BRCA1이 L1 RNP 어셈블리를 위한 ORF2p와 폴리 A 꼬리의 시스 결합을 억제하는 것을 억제하거나 방해함으로써 통합을 향상시키는 데 사용될 수 있다. 예를 들어, let-7 miRNA의 녹다운은 let-7 miRNA가 번역, 예컨대, ORF2p의 번역을 억제하는 것을 억제하거나 방해함으로써 통합을 향상시키는 데 사용될 수 있다. let-7 miRNA. 예를 들어, RAD51 및/또는 BRCA1의 녹다운은 RAD51 및/또는 BRCA1에 의한 절단된 DNA의 복구를 억제하거나 방해함으로써 통합을 향상시키는 데 사용될 수 있다.

[0047] 일부 실시양태에서, 하나 이상의 폴리핵산은 3 kb 내지 20 kb의 총 길이를 가진다.

[0048] 일부 실시양태에서, 하나 이상의 폴리핵산은 하나 이상의 폴리리보핵산, 하나 이상의 RNA 또는 하나 이상의 mRNA를 포함한다.

[0049] 일부 실시양태에서, 외생성 치료 폴리펩티드는 리간드, 항체, 수용체, 효소, 수송 단백질, 구조 단백질, 호르몬, 수축 단백질, 저장 단백질 및 전사 인자로 구성된 군으로부터 선택된다.

[0050] 일부 실시양태에서, 외생성 치료 폴리펩티드는 키메라 항원 수용체(CAR) 및 T 세포 수용체(TCR)로 구성된 군으로부터 선택된 수용체이다.

[0051] 일부 실시양태에서, 하나 이상의 폴리핵산은 프로모터 서열, 5' UTR 서열, 3' UTR 서열 및 폴리 A 서열을 포함하는 제1 발현 카세트를 포함하고; 이때 (i) 프로모터 서열은 5' UTR 서열의 상류에 있고, (ii) 5' UTR 서열은 폴리펩티드를 코딩하는 이동성 유전 요소의 서열의 상류에 있고, (iii) 3' UTR 서열은 삽입체 서열의 하류에 있

고; (iv) 3' UTR은 폴리 A 서열의 상류에 있고; 5' UTR 서열, 3' UTR 서열 또는 폴리 A 서열은 인간 ORF2p 또는 이의 기능성 단편에 대한 결합 부위를 포함한다.

[0052] 일부 실시양태에서, 삽입체 서열은 프로모터 서열의 역상보체인 서열, 5' UTR 서열의 역상보체인 서열, 3' UTR 서열의 역상보체인 서열, 및 폴리 A 서열의 역상보체인 서열을 포함하는 제2 발현 카세트를 포함하고, 이때 (i) 프로모터 서열의 역상보체인 서열은 5' UTR 서열의 역상보체인 서열의 하류에 있고, (ii) 5' UTR 서열의 역상보체인 서열은 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 서열의 하류에 있고; (iii) 3' UTR 서열의 역상보체인 서열은 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 서열의 상류에 있고, (iv) 폴리 A 서열의 역상보체인 서열은 3' UTR 서열의 역상보체인 서열의 상류에 있고 폴리펩티드를 코딩하는 이동성 유전 서열의 하류에 있다.

[0053] 일부 실시양태에서, 제1 발현 카세트의 프로모터 서열은 제2 발현 카세트의 프로모터 서열과 상이하다.

[0054] 일부 실시양태에서, 하나 이상의 LINE 폴리펩티드는 인간 ORF1p 또는 이의 기능성 단편을 포함하는 제1 LINE 폴리펩티드 및 인간 ORF2p 또는 이의 기능성 단편을 포함하는 제2 LINE 폴리펩티드를 포함하고, 이때 제1 LINE 폴리펩티드 및 제2 LINE 폴리펩티드는 상이한 오픈 리딩 프레임(ORF)으로부터 번역된다.

[0055] 일부 실시양태에서, 하나 이상의 폴리핵산은 인간 ORF1p 또는 이의 기능성 단편을 코딩하는 제1 폴리핵산 분자 및 인간 ORF2p 또는 이의 기능성 단편을 코딩하는 제2 폴리핵산 분자를 포함한다.

[0056] 일부 실시양태에서, 하나 이상의 폴리핵산은 5' UTR 서열 및 3' UTR 서열을 포함하고, 이때 (a) 5' UTR은 LINE-1의 5' UTR, 또는 ACUCCUCCCCAUCCUCUCCCUCUGUCCCUCUGUCCCUCUGACCCUGCACUGUCCCAGCACC(서열번호 51)과 적어도 80% 서열 동일성을 가진 서열을 포함하고/하거나; (b) 3' UTR은 LINE-1의 3' UTR, 또는 CAGGACACAGCCUUGGAUCAGGACAGAGACUUGGGGGCCAUCCUGCCCCUCCAACCCGACAUGUGUACCUCAGCUUUUUCCCUCACUUGCAUCAAUAAAGCUUCUGUGUUUGGAACAG(서열번호 52)와 적어도 80% 서열 동일성을 가진 서열을 포함한다.

[0057] 일부 실시양태에서, 외생성 치료 폴리펩티드를 코딩하는 서열은 인트론을 포함하지 않는다.

[0058] 일부 실시양태에서, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 C-말단 핵 국소화 신호(NLS), N-말단 NLS 또는 이들 둘 다를 포함한다.

[0059] 일부 실시양태에서, 외생성 폴리펩티드를 코딩하는 서열은 ORF1p 또는 이의 기능성 단편을 코딩하는 서열과 인 프레임으로 있지 않고/않거나, ORF2p 또는 이의 기능성 단편을 코딩하는 서열과 인 프레임으로 있지 않다.

[0060] 일부 실시양태에서, 하나 이상의 폴리핵산은 뉴클레아제 도메인, ORF2p로부터 유래하지 않은 뉴클레아제 도메인, 메가TAL 뉴클레아제 도메인, TALEN 도메인, Cas9 도메인, Cas6 도메인, Cas7 도메인, Cas8 도메인, R2 역요소의 징크 핑거 결합 도메인, 또는 반복 서열에 결합하는 DNA 결합 도메인을 코딩하는 서열을 포함한다.

[0061] 일부 실시양태에서, 하나 이상의 폴리핵산은 뉴클레아제 도메인을 코딩하는 서열을 포함하고, 이때 뉴클레아제 도메인은 뉴클레아제 활성을 갖지 않거나, 돌연변이가 없는 뉴클레아제 도메인에 비해 뉴클레아제 도메인의 활성을 감소시키는 돌연변이를 포함한다.

[0062] 일부 실시양태에서, ORF2p 또는 이의 기능성 단편은 엔도뉴클레아제 활성이 결여되어 있거나 S228P 및 Y1180A로 구성된 군으로부터 선택된 돌연변이를 포함하고/하거나, ORF1p 또는 이의 기능성 단편은 K3R 돌연변이를 포함한다.

[0063] 일부 실시양태에서, 삽입체 서열은 2개 이상의 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함한다.

[0064] 일부 실시양태에서, 하나 이상의 폴리핵산은 하나 이상의 폴리리보핵산을 포함하고, 이때 외생성 치료 폴리펩티드는 키메라 항원 수용체(CAR) 및 T 세포 수용체(TCR)로 구성된 군으로부터 선택된 수용체이고, 약학 조성물은 인간 대상체에게 전신 투여하기 위해 제제화된다.

[0065] 일부 실시양태에서, 하나 이상의 폴리핵산은 (i) 지질 나노입자 및 중합체성 나노입자로 구성된 군으로부터 선택된 나노입자로 제제화되고/되거나; (ii) 글리코실화된 RNA, 원형 RNA 및 자가복제 RNA로 구성된 군으로부터 선택된 하나 이상의 폴리핵산을 포함한다.

[0066] 본원은 질환 또는 병태의 치료를 필요로 하는 인간 대상체에서 질환 또는 병태를 치료하는 방법으로서, 본원에 기재된 약학 조성물을 상기 인간 대상체에게 투여하는 단계를 포함하는 방법도 제공한다.

[0067] 또한, 본원은 조성물을 생체외에서 인간 세포 집단과 접촉시킴으로써, 생체외 변형된 인간 세포 집단을 형성하는 단계를 포함하는, 생체외에서 인간 세포 집단을 변형시키는 방법으로서, 이때 상기 조성물이 하나 이상의 폴리핵산, 또는 하나 이상의 폴리핵산을 코딩하는 적어도 하나의 벡터를 포함하고, 상기 하나 이상의 폴리핵산이 (a) 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소; 및 (b) 외생성 치료 폴리펩티드를 코딩하는 서열의 역상보체인 삽입체 서열을 포함하고, 생체외 변형된 인간 세포 집단이 인간 대상체에서 실질적으로 면역원성을 나타내지 않는 것인 방법을 제공한다.

[0068] 한 측면에서, 본원은 유전 물질을 세포의 게놈 내로 통합시키는 조성물 및 방법을 제공하는 것으로, 이때 통합될 수 있는 유전 물질은 크기에 의해 구체적으로 제한되지 않는다. 일부 측면에서, 본원에 기재된 방법은 세포의 게놈 내로의 유전 "카고"의 1-단계 단일 폴리뉴클레오타이드 매개 전달 및 통합을 제공한다. 유전 물질은 코딩 서열, 예를 들어, 형질전이유전자(transgene), 펩티드, 재조합 단백질, 또는 항체 또는 이의 단편을 코딩하는 서열을 포함할 수 있고, 이때 상기 방법 및 조성물은 코딩 서열에 의해 코딩된 전사된 생성물의 안정한 발현을 보장한다. 유전 물질은 비코딩 서열, 예를 들어, 조절 RNA 서열, 예를 들어, 조절 작은 억제 RNA(siRNA), 마이크로RNA(miRNA), 긴 비코딩 RNA(lncRNA), 또는 하나 이상의 전사 조절제, 예컨대, 프로모터 및/또는 인핸서를 포함할 수 있고, 구조 생체분자, 예컨대, 리보좀 RNA(rRNA), 전달 RNA(tRNA), 또는 이들의 단편 또는 이들의 조합도 포함할 수 있으나, 이들로 제한되지 않는다.

[0069] 또 다른 측면에서, 본원은 전달의 안전성 및 효능 둘 다를 보장하는 비-바이러스 전달을 통해 크기에 의해 구체적으로 제한되지 않을 수 있는 유전 물질을 세포의 게놈 내로 부위 특이적으로 통합시키는 방법 및 조성물을 제공한다. 제공된 방법 및 조성물은 유전 물질을 포함하는 폴리뉴클레오타이드를 포함하는 치료제와 같은 치료제, 및 상기 폴리뉴클레오타이드 또는 상기 폴리뉴클레오타이드를 코딩하는 mRNA가 세포 내로 전달될 수 있게 하고 전달받은 세포의 게놈 내로 안정적으로 통합될 수 있게 하는 기구를 개발하는 데 특히 유용할 수 있다. 일부 실시양태에서, 치료제는 본원에 기재된 방법 및 조성물을 사용함으로써 세포의 게놈 내로 안정적으로 통합된 폴리뉴클레오타이드를 포함하는 세포일 수 있다.

[0070] 한 측면에서, 본 개시내용은 유전자를 세포 내로 안정적으로 전달하는 조성물 및 방법을 제공한다. 일부 실시양태에서, 이 조성물 및 방법은 유전자를 면역 세포 내로 안정적으로 전달하기 위한 것이다. 일부 경우, 면역 세포는 골수 세포이다. 일부 경우, 본원에 기재된 방법은 면역요법을 위한 골수 세포의 개발에 관한 것이다.

[0071] 본원은 질환의 치료를 필요로 하는 대상체에서 질환을 치료하는 방법으로서, L1 레트로트랜스포존을 코딩하는 서열에 작동 가능하게 연결된, 유전자 또는 이의 단편을 코딩하는 폴리시스트론 mRNA 서열을 포함하는 약학 조성물을 상기 대상체에게 투여하는 단계를 포함하는 방법을 제공하는 것으로; 이때 상기 유전자 또는 이의 단편은 길이가 적어도 10.1 kb이다.

[0072] 본원은 핵산 서열을 세포의 게놈 내로 통합시키는 방법으로서, L1 레트로트랜스포존을 코딩하는 서열에 작동 가능하게 연결된, 유전자 또는 이의 단편을 코딩하는 폴리시스트론 mRNA 서열을 포함하는 조성물을 세포와 접촉시키는 단계를 포함하는 방법을 제공하는 것으로; 이때 상기 유전자 또는 이의 단편은 길이가 적어도 10.1 kb이다. 일부 실시양태에서, 상기 유전자 또는 이의 단편(예를 들어, 페이로드)은 길이가 적어도 약 10.2 kb, 10.3 kb, 10.4 kb, 10.5 kb, 10.6 kb, 10.7 kb, 10.8 kb, 10.9 kb, 11 kb, 12 kb, 13 kb, 14 kb, 15 kb, 16 kb, 17 kb, 18 kb, 19 kb, 20 kb 이상이다.

[0073] 본원은 핵산 서열을 세포의 게놈 내로 통합시키는 방법으로서, L1 레트로트랜스포존을 코딩하는 서열에 작동 가능하게 연결된, 유전자 또는 이의 단편을 코딩하는 폴리시스트론 mRNA 서열을 포함하는 조성물을 세포와 접촉시키는 단계를 포함하는 방법을 제공하는 것으로; 이때 상기 유전자 또는 이의 단편은 ABCA4, MY07A, CEP290, CDH23, EYS, USH2a, GPR98, ALMS1, GDE, OTOF 및 F8로 구성된 군으로부터 선택된다.

[0074] 본원은 세포에서 재조합 핵산에 의해 코딩된 단백질을 발현시키는 방법으로서, L1 레트로트랜스포존을 코딩하는 서열에 작동 가능하게 연결된, 유전자 또는 이의 단편을 코딩하는 폴리시스트론 mRNA 서열을 포함하는 조성물을 세포와 접촉시킴으로써, 핵산 서열을 세포의 게놈 내로 통합시키는 단계; 및 상기 유전자 또는 이의 단편에 의해 코딩된 단백질을 발현시키는 단계를 포함하는 방법을 제공하는 것으로, 이때 상기 단백질의 발현은 (a)로부터 30일 이상 후에 검출될 수 있다.

[0075] 본원에 기재된 방법의 한 실시양태에서, 질환은 유전 질환이다.

[0076] 본원은 대상체에서 스타가르트병(Stargardt disease), LCA10, USH1D, DFNB12, 색소성 망막염(RP) USH2A, USH2C, 알스트롬(Alstrom) 증후군, 글리코겐 저장 질환 III, 비-증후군성 난청, A형 혈우병 또는 레버(Leber)

선천성 흑암시를 치료하는 방법으로서, (i) 인간 L1 트랜스포존에 작동 가능하게 연결된, 적합한 유전자 또는 이의 단편을 코딩하는 mRNA를 상기 대상체 내로 도입하는 단계, 또는 (ii) 인간 L1 트랜스포존에 작동 가능하게 연결된, 적합한 유전자 또는 이의 단편을 코딩하는 mRNA를 포함하는 세포 집단을 상기 대상체 내로 도입하는 단계를 포함하는 방법을 제공한다.

[0077] 본원에 기재된 방법의 한 실시양태에서, 방법은 스타가르트병의 치료를 필요로 하는 대상체에서 스타가르트병을 치료하는 단계를 포함하고, 이때 mRNA는 ABCA4 유전자 또는 이의 단편을 코딩한다.

[0078] 본원에 기재된 방법의 한 실시양태에서, 방법은 어셔(Usher) 증후군 1b형(어셔 1b) 질환의 치료를 필요로 하는 대상체에서 이러한 질환을 치료하는 단계를 포함하고, 이때 mRNA는 MY07A 유전자 또는 이의 단편을 코딩한다.

[0079] 본원에 기재된 방법의 한 실시양태에서, 방법은 레버 선천성 흑암시(LCA)10 질환의 치료를 필요로 하는 대상체에서 이러한 질환을 치료하는 단계를 포함하고, 이때 mRNA는 CEP290 유전자 또는 이의 단편을 코딩한다.

[0080] 본원에 기재된 방법의 한 실시양태에서, 방법은 어셔 증후군 1D형(USH1D) 비-증후군성 난청 또는 청력 상실 USH1D, DFN12 질환의 치료를 필요로 하는 대상체에서 이러한 질환을 치료하는 단계를 포함하고, 이때 mRNA는 CDH23 유전자 또는 이의 단편을 코딩한다.

[0081] 본원에 기재된 방법의 한 실시양태에서, 방법은 색소성 망막염(RP) 질환의 치료를 필요로 하는 대상체에서 이러한 질환을 치료하는 단계를 포함하고, 이때 mRNA는 EYS 유전자 또는 이의 단편을 코딩한다.

[0082] 본원에 기재된 방법의 한 실시양태에서, 방법은 어셔 증후군 2A형(USH2A)을 치료하는 단계를 포함하고, 이때 mRNA는 USH2a 유전자 또는 이의 단편을 코딩한다.

[0083] 본원에 기재된 방법의 한 실시양태에서, 방법은 어셔 증후군 2C형(USH2C)을 치료하는 단계를 포함하고, 이때 mRNA는 GPR98 유전자 또는 이의 단편을 코딩한다.

[0084] 본원에 기재된 방법의 한 실시양태에서, 방법은 알스트롬 증후군을 치료하는 단계를 포함하고, 이때 mRNA는 ALMS1 유전자 또는 이의 단편을 코딩한다.

[0085] 본원에 기재된 방법의 한 실시양태에서, 방법은 글리코겐 저장 질환 III을 치료하는 단계를 포함하고, 이때 mRNA는 GDE 유전자 또는 이의 단편을 코딩한다.

[0086] 본원에 기재된 방법의 한 실시양태에서, 방법은 비-증후군성 난청 또는 청력 상실을 치료하는 단계를 포함하고, 이때 mRNA는 OTOF 유전자 또는 이의 단편을 코딩한다.

[0087] 본원에 기재된 방법의 한 실시양태에서, 방법은 A형 혈우병을 치료하는 단계를 포함하고, mRNA는 인자 VIII(F8) 유전자 또는 이의 단편을 코딩한다.

[0088] 본원은 (A) 하나 이상의 돌연변이를 포함하는 핵산 서열을 세포의 게놈으로부터 절제하기 위한 표적화 절제 기구를 포함하는 제1 단백질 복합체를 코딩하는 폴리뉴클레오타이드 서열, 및 (B) 제2 단백질 복합체를 코딩하는 재조합 mRNA를 세포 내로 도입하는 단계를 포함하는, 세포의 게놈 핵산 서열의 표적화 대체 방법을 제공하는 것으로, 이때 상기 재조합 mRNA는 (i) 하나 이상의 돌연변이를 함유하지 않는, (A)의 절제된 핵산 서열을 포함하는 핵산 서열, 및 (ii) 독립적인 프로모터의 영향 하에서 L1 레트로트랜스포존 ORF2 단백질을 코딩하는 서열을 포함한다.

[0089] 본원에 기재된 방법의 한 실시양태에서, 하나 이상의 돌연변이를 포함하는 핵산 서열은 세포 유전자의 병원성 변이체를 포함한다.

[0090] 본원에 기재된 방법의 한 실시양태에서, 하나 이상의 돌연변이를 함유하지 않는 핵산 서열을 포함하는 (B)의 핵산 서열은 ORF2 서열에 작동 가능하게 연결된다.

[0091] 본원에 기재된 방법의 한 실시양태에서, 방법은 절제 부위에서 복수의 타이미딘 잔기를 포함하는 서열을 도입하는 단계를 추가로 포함한다.

[0092] 일부 실시양태에서, 상기 서열을 도입하는 단계는 적어도 4개의 타이미딘 잔기를 도입하는 단계를 포함한다.

[0093] 본원에 기재된 방법의 한 실시양태에서, 표적화 절제 기구는 서열 가이딩 부위 특이적 절제 엔도뉴클레아제를 포함한다.

[0094] 본원에 기재된 방법의 한 실시양태에서, 표적화 절제 기구는 CRISPR-CAS 시스템을 포함한다.

[0095] 일부 실시양태에서, 표적화 절제 기구는 변형된 재조합 LINE 1(L1) 엔도뉴클레아제이다.

[0096] 일부 실시양태에서, 복수의 타이미딘 잔기를 포함하는 서열을 도입하는 단계는 절제 부위에서의 프라임 편집에 의한 염기 연장을 포함한다.

[0097] 일부 실시양태에서, L1 레트로트랜스포존 ORF2 단백질을 코딩하는 mRNA 서열은 L1 레트로트랜스포존 ORF1 단백질을 코딩하는 서열을 추가로 포함한다.

[0098] 일부 실시양태에서, mRNA는 유도성 프로모터에 대한 서열을 포함한다.

[0099] 본원에 기재된 방법의 한 실시양태에서, 절제된 서열은 1000개의 염기보다 더 크다.

[0100] 본원에 기재된 방법의 한 실시양태에서, 절제된 서열은 6 kb보다 더 크다.

[0101] 본원에 기재된 방법의 한 실시양태에서, 절제된 서열은 약 10 kb이다.

[0102] 일부 실시양태에서, 세포는 림프구이다. 일부 실시양태에서, 세포는 골수 세포이다. 일부 실시양태에서, 세포는 상피 세포이다. 일부 실시양태에서, 세포는 암 세포이다.

[0103] 일부 실시양태에서, 핵산 서열은 ATP 결합 카세트(ABC) 수송제 유전자인 (ABCA4) 유전자, 또는 이의 단편을 코딩한다.

[0104] 일부 실시양태에서, 핵산 서열은 MY07A, CEP290, CDH23, EYS, USH2a, GPR98, ALMS1, GDE, OTOF 또는 F8 유전자, 또는 이의 단편을 코딩한다.

[0105] 일부 실시양태에서, 도입은 생체외에서 세포 내로 도입하는 것을 포함한다. 일부 실시양태에서, 도입은 전기천공을 포함한다. 일부 실시양태에서, 도입은 생체내에서 세포 내로 도입하는 것을 포함한다. 일부 실시양태에서, 하나 이상의 돌연변이를 함유하지 않는 서열을 포함하는 핵산 서열의 발현은 세포 내로 도입된 지 적어도 35일 후에 검출될 수 있다. 일부 실시양태에서, 대상체 내로의 도입은 mRNA의 전신 직접 투여를 포함한다.

[0106] 일부 실시양태에서, 대상체 내로의 도입은 mRNA의 국소 투여를 포함한다.

[0107] 일부 실시양태에서, mRNA 서열은 세포 표적화 모이어티를 포함한다.

[0108] 일부 실시양태에서, 세포 표적화 모이어티는 앱타머이다.

[0109] 일부 실시양태에서, 대상체 내로의 도입은 mRNA를 대상체의 망막 내로 도입하는 것을 포함한다.

[0110] 본원은 핵산 서열을 세포의 게놈 내로 통합시키는 방법으로서, 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 세포 내로 도입하는 단계를 포함하는 방법을 제공하는 것으로, 이때 mRNA는 (a) (i) 외생성 서열, 또는 (ii) 외생성 서열의 역상보체인 서열을 포함하는 삽입체 서열; (b) 5' UTR 서열, 및 5' UTR 서열의 하류에 있는 3' UTR 서열을 포함하고; 이때 5' UTR 서열 또는 3' UTR 서열은 인간 ORF 단백질에 대한 결합 부위를 포함하고, 삽입체 서열은 세포의 게놈 내로 통합되고, 삽입체 서열은 ABCA4, MY07A, CEP290, CDH23, EYS, USH2a, GPR98, ALMS1, GDE, OTOF 및 F8로 구성된 군으로부터 선택된 유전자이다.

[0111] 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 인간 ORF2p에 대한 결합 부위를 포함한다.

[0112] 본원은 핵산 서열을 면역 세포의 게놈 내로 통합시키는 방법으로서, 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 도입하는 단계를 포함하는 방법을 제공하는 것으로, 이때 mRNA는 (a) (i) 외생성 서열 또는 (ii) 외생성 서열의 역상보체인 서열을 포함하는 삽입체 서열; (b) 5' UTR 서열, 및 5' UTR 서열의 하류에 있는 3' UTR 서열을 포함하고, 이때 5' UTR 서열 또는 3' UTR 서열은 엔도뉴클레아제 결합 부위 및/또는 역전사효소 결합 부위를 포함하고, 삽입체 서열은 면역 세포의 게놈 내로 통합되고, 삽입체 서열은 ABCA4, MY07A, CEP290, CDH23, EYS, USH2a, GPR98, ALMS1, GDE, OTOF 및 F8로 구성된 군으로부터 선택된 유전자이다.

[0113] 본원은 핵산 서열을 세포의 게놈 내로 통합시키는 방법으로서, 재조합 mRNA, 또는 mRNA를 코딩하는 벡터를 도입하는 단계를 포함하는 방법을 제공하는 것으로, 이때 mRNA는 (a) (i) 외생성 서열 또는 (ii) 외생성 서열의 역상보체인 서열을 포함하는 삽입체 서열; (b) 5' UTR 서열, 5' UTR 서열의 하류에 있는 인간 레트로트랜스포존의 서열, 및 인간 레트로트랜스포존의 서열의 하류에 있는 3' UTR 서열을 포함하고; 이때 5' UTR 서열 또는 3' UTR 서열은 엔도뉴클레아제 결합 부위 및/또는 역전사효소 결합 부위를 포함하고, 인간 레트로트랜스포존의 서열은 2개의 ORF를 함유하는 단일 RNA로부터 번역되는 2개의 단백질을 코딩하고, 삽입체 서열은 세포의 게놈 내로 통합되고, 삽입체 서열은 ABCA4, MY07A, CEP290, CDH23, EYS, USH2a, GPR98, ALMS1, GDE, OTOF 및 F8로 구성된

군으로부터 선택된 유전자이다.

[0114] 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 ORF2p 결합 부위를 포함한다. 일부 실시양태에서, ORF2p 결합 부위는 3' UTR 서열의 폴리 A 서열이다.

[0115] 일부 실시양태에서, mRNA는 인간 레트로트랜스포존의 서열을 포함한다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 5' UTR 서열의 하류에 있다.

[0116] 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 3' UTR 서열의 상류에 있다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 2개의 ORF를 함유하는 단일 RNA로부터 번역되는 2개의 단백질을 코딩한다. 일부 실시양태에서, 2개의 ORF는 비-중첩 ORF이다.

[0117] 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 비-LTR 레트로트랜스포존의 서열을 포함한다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 LINE-1 레트로트랜스포존을 코딩한다. 일부 실시양태에서, LINE-1 레트로트랜스포존은 인간 LINE-1 레트로트랜스포존이다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 서열을 포함한다.

[0118] 일부 실시양태에서, 엔도뉴클레아제 및/또는 역전사효소는 ORF2p이다.

[0119] 일부 실시양태에서, 역전사효소는 II 군 인트론 역전사효소 도메인이다.

[0120] 일부 실시양태에서, 엔도뉴클레아제 및/또는 역전사효소는 밍크 고래 엔도뉴클레아제 및/또는 역전사효소이다.

[0121] 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 ORF2p를 코딩하는 서열을 포함한다. 일부 실시양태에서, 삽입체 서열은 ORF2p의 엔도뉴클레아제 도메인의 특이성을 이용함으로써 폴리 T 부위에서 게놈 내로 통합된다. 일부 실시양태에서, 폴리 T 부위는 서열 TTTTTA를 포함한다. 일부 실시양태에서, 레트로트랜스포존은 핵 체류 서열에 융합된 ORF1p 및/또는 ORF2p를 포함한다. 일부 실시양태에서, 핵 체류 서열은 Alu 서열이다. 일부 실시양태에서, ORF1p 및/또는 ORF2p는 MS2 코트 단백질에 융합된다. 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 적어도 1개, 2개, 3개 이상의 MS2 헤어핀 서열을 포함한다.

[0122] 본원은 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 포함하는 조성물을 제공하는 것으로, 이때 mRNA는 (i) 인간 LINE-1 트랜스포존 5' UTR 서열, (ii) 인간 LINE-1 트랜스포존 5' UTR 서열의 하류에 있는 ORF1p를 코딩하는 서열, (iii) ORF1p를 코딩하는 서열의 하류에 있는 ORF간 링커 서열, (iv) ORF간 링커 서열의 하류에 있는 ORF2p를 코딩하는 서열, 및 (v) ORF2p를 코딩하는 서열의 하류에 있는 인간 LINE-1 트랜스포존으로부터 유래한 3' UTR 서열을 포함하는 인간 LINE-1 트랜스포존 서열을 포함하고; 이때 3' UTR 서열은 삽입체 서열을 포함하고, 삽입체 서열은 외생성 폴리펩티드를 코딩하는 서열의 역상보체 또는 외생성 조절 요소를 코딩하는 서열의 역상보체이고, 삽입체 서열은 ABCA4, MY07A, CEP290, CDH23, EYS, USH2a, GPR98, ALMS1, GDE, OTOF 및 F8로 구성된 군으로부터 선택된 유전자이다.

[0123] 본원은 (a) 긴 산재된 핵 요소(LINE) 폴리펩티드; 및 (b) 삽입체 서열을 코딩하는 뉴클레오타이드 서열을 포함하는 핵산을 포함하고 실질적으로 면역원성을 나타내지 않는 조성물을 제공하는 것으로, 이때 상기 LINE 폴리펩티드는 인간 ORF1p 및 인간 ORF2p를 포함하고, 상기 삽입체 서열은 외생성 폴리펩티드를 코딩하는 서열의 역상보체 또는 외생성 조절 요소를 코딩하는 서열의 역상보체이고, 상기 삽입체 서열은 ABCA4, MY07A, CEP290, CDH23, EYS, USH2a, GPR98, ALMS1, GDE, OTOF 및 F8로 구성된 군으로부터 선택된 유전자이다.

[0124] 식세포를 사용한 면역요법은 암 세포 또는 감염된 세포와 같은 병든 세포를 공격하고 사멸시키는 대식세포 또는 다른 식세포와 같은 조작된 골수 세포를 만들고 사용하는 단계를 포함한다. 조작된 골수 세포, 예컨대, 대식세포 및 다른 식세포는 암 세포와 같은 표적 세포와 같은 표적의 표면의 특정 항원에 결합하도록 디자인된 표적화된 항원 결합 세포외 도메인을 포함하는 조작된 단백질, 예컨대, 키메라 항원 수용체를 코딩하는 합성 재조합 핵산을, 재조합 핵산 기술을 통해 이 세포 내로 도입함으로써 제조된다. 조작된 키메라 수용체와 표적의 항원, 예컨대, 암 항원(또는 마찬가지로 질환 표적)의 결합은 표적의 식세포작용을 시작한다. 이것은 2배 작용을 유발한다: 첫 번째는 식세포 삼킴과 표적의 용해가 제1차 면역 방어로서 표적을 파괴하고 이를 제거하는 것이고; 두 번째는 표적의 항원이 골수 세포의 포식라이소좀(phagolysosome)에서 분해되고 골수 세포의 표면에 제시된 후, T 세포의 활성화와 면역 반응의 추가 활성화 및 면역학적 기억의 발생을 유발하는 것이다. 키메라 수용체는 그 자신이 도입되고 발현되는 골수 세포의 식세포작용 및 면역 활성화를 향상시키도록 조작된다. 본 개시내용의 키메라 항원 수용체는 본원에서 키메라 융합 단백질, CFP, 식세포 수용체(PR) 융합 단백질(PFP), 또는 식세포작용을 위한 키메라 항원 수용체(CAR-P)로서 다양하게 지칭되지만, 각각의 용어는 재조합 키메라 및/또는 융합 수용

체 단백질의 개념에 관한 것이다. 일부 실시양태에서, 비-수용체 단백질을 코딩하는 유전자도 전형적으로 키메라 항원 수용체 기능의 증강을 위해 골수 세포에서 공-발현된다. 요약하건대, 질환 표적에 대한 골수 세포의 식세포작용 및/또는 면역 반응을 증강시키도록 디자인된 다양한 조작된 수용체들 및 비-수용체 재조합 단백질들; 및 방법과 조성물이 면역요법을 위한 조작된 골수 세포의 생성에 적합하도록 조작된 수용체 또는 비-수용체 재조합 단백질을 코딩하는 재조합 핵산을 생성하고 통합시키는 방법 및 조성물이 본원에서 예상된다.

[0125] 한 측면에서, 본 개시내용은 유전자를 세포 내로 안정적으로 전달하는 조성물 및 방법을 제공하는 것으로, 이때 상기 세포는 임의의 체세포일 수 있다. 일부 실시양태에서, 상기 조성물 및 방법은 세포 특이적 또는 조직 특이적 전달용으로 디자인된다. 일부 경우, 본원에 기재된 방법은 예를 들어, 단백질 대체 요법을 위해 생체내에서 부재하거나 결함이 있는(돌연변이된) 단백질을 보상하도록 기능성 단백질 또는 이의 단편을 제공하는 것에 관한 것이다.

[0126] 세포 내로의 재조합 핵산의 도입은 최신기술에서 이용될 수 있는 하나 이상의 유전자 전달 기법에 의해 달성될 수 있다. 그러나, 치료 목적을 위해 외생성 유전(예를 들어, 핵산) 요소를 게놈 내로 통합시키는 것은 여전히 여러 과제들에 직면한다. 안전하고 신뢰할 수 있는 방식으로 안정한 통합을 달성하는 것과 효율적인 연장된 발현이 이 과제들 중 일부이다. 카고 핵산 서열의 게놈 통합을 목적으로 하는 대다수의 성공적인 유전자 전달 시스템들은 일부 내재된 안전성 및 효능 문제들을 가진 바이러스 전달 기작에 의존한다. 긴 핵산 서열의 전달 및 통합은 현재 유전자 편집 시스템에 의해 달성될 수 없다.

[0127] 지금까지 형질전이유전자의 안정한 장기간 유전자 전달 및 발현을 위해 조작된 골수 세포를 제조하고 사용하는 데 거의 관심을 기울이지 않았다. 예를 들어, 세포 요법을 위해 생체외에서 유전자를 분화된 포유동물 세포에게 전달하는 것은 바이러스 유전자 전달 기작을 통해 달성될 수 있다. 그러나, 시간 경과에 따른 원치 않는 형질전이유전자 침묵 가능성, 다른 유전자(예를 들어, 발암유전자)의 원치 않는 활성화와 관련된 게놈의 전사 활성 부위 내로의 우선적인 통합 및 유전독성을 비롯한, 바이러스 유전자 전달 벡터의 사용과 관련된 몇 가지 전략적 단점이 있다. 안전성 문제 이외에, 증가된 비용과 통합 바이러스의 제조, 저장 및 취급의 번거로운 노력은 종종 치료 적용에 있어서 바이러스 벡터 매개 유전자 변형 세포의 대규모 사용을 방해한다. 안전성에 관한 바이러스 벡터와 관련된 이 지속된 우려뿐만 아니라, 벡터 제조의 비용 및 규모도 효과적인 요법을 위한 대안적인 방법을 필요로 한다.

[0128] 면역요법에 사용되는 세포의 게놈 내로의 형질전이유전자의 통합은 이것이 안정하고 요법 동안 전달을 위해 더 적은 수의 세포가 요구된다는 의미에서 유리할 수 있다. 다른 한편으로, 비-분열 세포 내로의 형질전이유전자의 통합은 세포의 건강 및 기능뿐만 아니라 생체내에서의 세포의 궁극적인 수명에도 영향을 미침으로써, 치료제로서의 그의 전체 유용성에 영향을 미친다는 점에서 어려울 수 있다. 일부 실시양태에서, 면역요법을 위해 골수 세포를 생성하는 본원에 기재된 방법은 예를 들어, 변형시킬 골수 세포의 선택; 재조합 핵산을 골수 세포 내로 도입하는 방법 및 조성물; 재조합 핵산의 발현을 향상시키는 방법 및 조성물; 벡터를 선택하고 변형시키는 방법 및 조성물; 생체내에서 골수 세포에 의한 재조합 핵산의 흡수 및 도입을 위한 생체내 투여에 적합한 재조합 핵산을 제조함으로써 요법을 위한 골수 세포를 생성하는 방법을 포함하나 이들로 제한되지 않는 다수의 단계들과 조성물들의 누적 결과일 수 있다. 일부 측면에서, 본원에 기재된 다양한 발명들의 하나 이상의 실시양태는 서로 간에 전달될 수 있으며, 당업자는 과도한 실험을 필요로 하지 않으면서 이들을 대안으로, 조합으로 또는 교환 가능하게 사용할 것으로 예상된다. 개시된 요소의 모든 이러한 변경이 예상되고 본원에 완전히 포함된다.

[0129] 한 측면에서, 트랜스포존 또는 전위 요소(TE)는 본원에서 관심 있는 형질전이유전자를 코딩하는 이종, 합성 또는 재조합 핵산을 골수 세포 내로 도입하는 수단으로서 간주된다. 트랜스포존 또는 전위 요소는 전위효소(transposase)로서 공지되어 있는 효소를 사용하여 유전 물질의 단편을 게놈 내로 전위시키는 능력을 가진 유전 요소이다. 포유동물 게놈은 많은 수의 전위 요소(TE) 유래 서열들을 함유하고, 본 발명자들의 게놈의 최대 70%는 TE 유래 서열을 나타낸다(de Koning et al. 2011; Richardson et al. 2015). 이 요소는 유전 물질을 세포의 게놈 내로 도입하는 데 이용될 수 있다. TE 요소는 종종 게놈 내에서 유전 물질의 "점핑"으로서 지칭되는 이동을 할 수 있다. TE는 일반적으로 가역적 불활성 후성적 침묵된 형태로 진핵생물 게놈에 존재한다. 본 개시내용에서, 형질전이유전자를 대식세포 및 다른 식세포 내로 효율적 및 안정적으로 통합시키는 방법 및 조성물. 상기 방법은 전위효소 및 전위 요소 mRNA-코딩된 전위효소의 사용에 기반한다. 일부 실시양태에서, 긴 산재된 요소-1(L1) RNA는 형질전이유전자를 세포(예를 들어, 대식세포 또는 식세포) 내로 안정적으로 통합시키고/시키거나 역전위시키는 데 사용된다.

[0130] 본원은 세포 게놈 내로의 외생성 핵산 서열의 레트로트랜스포존 매개 안정한 통합 방법을 예상한다. 이 방법은

불리한 효과를 생성하지 않으면서 레트로트랜스포존을 세포 내로 무작위 게놈 통합시키는 기구를 이용할 수 있다. 본원에 기재된 방법은 외생성 핵산이 게놈 내의 안전한 유전자좌에서 통합되고 세포의 내재된 방어 기작에 의해 침묵되지 않으면서 발현되도록 외생성 핵산 서열을 세포 내로 강력하고 다양하게 도입하는 데 이용될 수 있다. 본원에 기재된 방법은 크기가 약 1 kb, 약 2 kb, 약 3 kb, 약 4 kb, 약 5 kb, 약 6 kb, 약 7 kb, 약 8 kb, 약 9 kb, 약 10 kb 이상인 외생성 핵산을 도입하는 데 이용될 수 있다. 일부 실시양태에서, 외생성 핵산은 리보좀 유전자좌 내로 통합되지 않는다. 일부 실시양태에서, 외생성 핵산은 ROSA26 유전자좌 또는 또 다른 안전 피난처(safe harbor) 유전자좌 내로 통합되지 않는다. 일부 실시양태에서, 본원에 기재된 방법 및 조성물은 외생성 핵산 서열을 세포의 게놈 내의 임의의 위치에서 통합시킬 수 있다. 나아가, 본원은 불리한 효과를 생성하지 않으면서 외생성 핵산 서열을 세포 게놈 내의 특정 예정된 부위 내로 통합시키도록 개발된 역전위 시스템을 예상한다. 개시된 방법 및 조성물은 높은 신뢰도로 외생성 핵산을 세포 내로 고도 특이적으로 도입하기 위해 레트로트랜스포존을 조작하는 여러 기작들을 포함한다.

[0131] 본원에 기재된 방법 및 조성물은 세포의 게놈을 조작하는 분자 시스템 및 기작에 있어서 핵심적인 돌파구를 대표한다. 본원은 외생성 핵산 서열의 큰 단편(적어도 100개 초과의 핵염기, 적어도 1 kb 초과, 적어도 2 kb 초과, 적어도 3 kb 초과 등)을, rDNA 또는 리보좀 유전자좌 또는 지정된 안전 피난처 유전자좌, 예컨대, ROSA 26 유전자좌가 아닌 게놈의 비-보존된 영역 내로 비-바이러스적으로 전달하고 안정적으로 통합시키는 데 있어서 인간 레트로트랜스포존 시스템을 활용하는 방법을 처음으로 제시한다.

[0132] 일부 실시양태에서, 역전위 시스템을 이용하여 비-내생성 핵산을 게놈 내로 안정적으로 통합시키고 발현시키며, 이때 비-내생성 핵산은 핵산 서열 내에 역전위 요소를 포함한다. 일부 실시양태에서, 세포의 내생성 역전위 시스템(예를 들어, 단백질 및 효소)을 이용하여 세포에서 비-내생성 핵산을 안정적으로 발현시킨다. 일부 실시양태에서, 세포의 내생성 역전위 시스템(예를 들어, 단백질 및 효소, 예컨대, LINE-1 역전위 시스템)이 이용되나, 세포에서 비-내생성 핵산을 안정적으로 발현시키기 위해 역전위 시스템의 하나 이상의 성분을 추가로 발현할 수 있다.

[0133] 일부 실시양태에서, 본원은 형질전이유전자를 코딩하고 게놈 통합 및/또는 역전위를 위해 하나 이상의 성분을 코딩하는 합성 핵산을 제공한다.

[0134] 한 측면에서, 본원은 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 세포 내로 도입하는 단계를 포함하는, 핵산 서열을 세포의 게놈 내로 통합시키는 방법을 제공하는 것으로, 이때 mRNA는 외생성 서열, 또는 외생성 서열의 역상보체인 서열을 포함하는 삽입체 서열; 5' UTR 서열, 및 5' UTR 서열의 하류에 있는 3' UTR 서열을 포함하고; 이때 5' UTR 서열 또는 3' UTR 서열은 인간 ORF 단백질에 대한 결합 부위를 포함하고, 상기 삽입체 서열은 세포의 게놈 내로 통합된다. 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 인간 ORF2p에 대한 결합 부위를 포함한다.

[0135] 한 측면에서, 본원은 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 도입하는 단계를 포함하는, 핵산 서열을 면역세포의 게놈 내로 통합시키는 방법을 제공하는 것으로, 이때 mRNA는 (i) 외생성 서열 또는 (ii) 외생성 서열의 역상보체인 서열을 포함하는 삽입체 서열; 5' UTR 서열, 및 5' UTR 서열의 하류에 있는 3' UTR 서열을 포함하고, 이때 5' UTR 서열 또는 3' UTR 서열은 엔도뉴클레아제 결합 부위 및/또는 역전사효소 결합 부위를 포함하고, 형질전이유전자 서열은 면역 세포의 게놈 내로 통합된다.

[0136] 한 측면에서, 본원은 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 도입하는 단계를 포함하는, 핵산 서열을 세포의 게놈 내로 통합시키는 방법을 제공하는 것으로, 이때 mRNA는 (i) 외생성 서열 또는 (ii) 외생성 서열의 역상보체인 서열을 포함하는 삽입체 서열; 5' UTR 서열, 5' UTR 서열의 하류에 있는 인간 레트로트랜스포존의 서열, 및 인간 레트로트랜스포존의 서열의 하류에 있는 3' UTR 서열을 포함하고, 이때 5' UTR 서열 또는 3' UTR 서열은 엔도뉴클레아제 결합 부위 및/또는 역전사효소 결합 부위를 포함하고, 인간 레트로트랜스포존의 서열은 2개의 ORF를 함유하는 단일 RNA로부터 번역된 2개의 단백질을 코딩하고, 상기 삽입체 서열은 세포의 게놈 내로 통합된다.

[0137] 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 ORF2p 결합 부위를 포함한다. 일부 실시양태에서, ORF2p 결합 부위는 3' UTR 서열 내의 폴리 A 서열이다.

[0138] 일부 실시양태에서, mRNA는 인간 레트로트랜스포존의 서열을 포함한다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 5' UTR 서열의 하류에 있다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 3' UTR 서열의 상류에 있다. 일부 실시양태에서, 세포의 게놈 내로 전달되고 통합되기를 원하는 폴리뉴클레오타이드 서열(예를

들어, 삽입체)은 재조합 핵산 구축물에서 ORF1을 코딩하는 서열의 3' 부위에서 삽입된다. 일부 실시양태에서, 세포의 게놈 내로 전달되고 통합되기를 원하는 폴리뉴클레오타이드 서열은 재조합 핵산 구축물에서 ORF2를 코딩하는 서열의 3' 부위에서 삽입된다. 일부 실시양태에서, 세포의 게놈 내로 전달되고 통합되기를 원하는 서열은 ORF1 또는 ORF2, 또는 이들 둘 다의 3' UTR 내에 삽입된다. 일부 실시양태에서, 세포의 게놈 내로 전달되고 통합되기를 원하는 서열인 폴리뉴클레오타이드 서열은 재조합 핵산 구축물에서 ORF2의 폴리 A 꼬리의 상류에서 삽입된다.

[0139] 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 2개의 ORF를 함유하는 단일 RNA로부터 번역되는 2개의 단백질을 코딩한다. 일부 실시양태에서, 2개의 ORF는 비-중첩 ORF이다. 일부 실시양태에서, 2개의 ORF는 ORF1 및 ORF2이다. 일부 실시양태에서, ORF1은 ORF1p를 코딩하고, ORF2는 ORF2p를 코딩한다.

[0140] 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 비-LTR 레트로트랜스포존의 서열을 포함한다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 LINE-1 레트로트랜스포존을 포함한다. 일부 실시양태에서, LINE-1 레트로트랜스포존은 인간 LINE-1 레트로트랜스포존이다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 서열을 포함한다. 일부 실시양태에서, 엔도뉴클레아제 및/또는 역전사효소는 ORF2p이다. 일부 실시양태에서, 역전사효소는 II 군 인트론 역전사효소 도메인이다. 일부 실시양태에서, 엔도뉴클레아제 및/또는 역전사효소는 밍크 고래 엔도뉴클레아제 및/또는 역전사효소이다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 ORF2p를 코딩하는 서열을 포함한다. 일부 실시양태에서, 삽입체 서열은 ORF2p의 엔도뉴클레아제 도메인의 특이성을 이용함으로써 폴리 T 부위에서 게놈 내로 통합된다. 일부 실시양태에서, 폴리 T 부위는 서열 TTTTTA를 포함한다.

[0141] 일부 실시양태에서, 본원은 인간 레트로트랜스포존을 코딩하는 서열을 포함하는 mRNA를 포함하는 폴리뉴클레오타이드 구축물을 제공하는 것으로, 이때 (i) 인간 레트로트랜스포존의 서열은 ORF1p를 코딩하는 서열을 포함하거나, (ii) 상기 mRNA는 ORF1p를 코딩하는 서열을 포함하지 않거나, (iii) 상기 mRNA는 상보체 유전자의 5' UTR 서열에 의한 ORF1p 코딩 서열의 대체를 포함한다. 일부 실시양태에서, 상기 mRNA는 ORF1p를 코딩하는 제1 mRNA 분자, 및 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 mRNA 분자를 포함한다. 일부 실시양태에서, 상기 mRNA는 ORF1p를 코딩하는 제1 서열, 및 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 서열을 포함하는 mRNA 분자이다. 일부 실시양태에서, ORF1p를 코딩하는 제1 서열과 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 서열은 링커 서열에 의해 분리된다.

[0142] 일부 실시양태에서, 링커 서열은 내부 리보좀 진입 서열(IRES)을 포함한다. 일부 실시양태에서, IRES는 CVB3 또는 EV71의 IRES이다. 일부 실시양태에서, 링커 서열은 자가절단 펩티드 서열을 코딩한다. 일부 실시양태에서, 링커 서열은 T2A, E2A 또는 P2A 서열을 코딩한다.

[0143] 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 추가 단백질 서열에 융합된 ORF1p를 코딩하는 서열 및/또는 추가 단백질 서열에 융합된 ORF2p를 코딩하는 서열을 포함한다. 일부 실시양태에서, ORF1p 및/또는 ORF2p는 핵 체류 서열에 융합된다. 일부 실시양태에서, 핵 체류 서열은 Alu 서열이다. 일부 실시양태에서, ORF1p 및/또는 ORF2p는 MS2 코트 단백질에 융합된다. 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 적어도 1개, 2개, 3개 이상의 MS2 헤어핀 서열을 포함한다. 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 mRNA의 폴리 A 꼬리와 엔도뉴클레아제 및/또는 역전사효소의 상호작용을 촉진하거나 향상시키는 서열을 포함한다. 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 폴리 A 결합 단백질(예를 들어, PABP)과 엔도뉴클레아제 및/또는 역전사효소의 상호작용을 촉진하거나 향상시키는 서열을 포함한다. 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 상기 mRNA에 대한 엔도뉴클레아제 및/또는 역전사효소의 특이성을 세포에 의해 발현된 또 다른 mRNA에 비해 증가시키는 서열을 포함한다. 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 Alu 요소 서열을 포함한다.

[0144] 일부 실시양태에서, ORF1p를 코딩하는 제1 서열 및 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 서열은 동일한 프로모터를 가진다. 일부 실시양태에서, 삽입체 서열은 ORF1p를 코딩하는 제1 서열의 프로모터와 상이한 프로모터를 가진다. 일부 실시양태에서, 삽입체 서열은 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 서열의 프로모터와 상이한 프로모터를 가진다. 일부 실시양태에서, ORF1p를 코딩하는 제1 서열 및/또는 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 서열은 유도성 프로모터, CMV 프로모터 또는 전사 시작 부위, T7 프로모터 또는 전사 시작 부위, EF1a 프로모터 또는 전사 시작 부위, 및 이들의 조합으로 구성된 군으로부터 선택된 프로모터 또는 전사 시작 부위를 가진다. 일부 실시양태에서, 삽입체 서열은 유도성 프로모터, CMV 프로모터 또는 전사 시작 부위, T7 프로모터 또는 전사 시작 부위, EF1a 프로모터 또는 전사 시작 부위, 및 이들의 조합

으로 구성된 군으로부터 선택된 프로모터 또는 전사 시작 부위를 가진다.

[0145] 일부 실시양태에서, ORF1p를 코딩하는 제1 서열 및 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 서열은 인간 세포에서 발현되도록 코돈 최적화된다.

[0146] 일부 실시양태에서, mRNA는 WPRE 요소를 포함한다. 일부 실시양태에서, mRNA는 선택 마커를 포함한다. 일부 실시양태에서, mRNA는 친화성 태그를 코딩하는 서열을 포함한다. 일부 실시양태에서, 친화성 태그는 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 서열에 연결된다.

[0147] 일부 실시양태에서, 3' UTR은 폴리 A 서열을 포함하거나, 폴리 A 서열은 시험관내에서 mRNA에 추가된다. 일부 실시양태에서, 폴리 A 서열은 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 서열의 하류에 있다. 일부 실시양태에서, 삽입체 서열은 폴리 A 서열의 상류에 있다.

[0148] 일부 실시양태에서, 3' UTR 서열은 삽입체 서열을 포함한다. 일부 실시양태에서, 삽입체 서열은 외생성 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함한다. 일부 실시양태에서, 삽입체 서열은 폴리아데닐화 부위를 포함한다. 일부 실시양태에서, 삽입체 서열은 SV40 폴리아데닐화 부위를 포함한다. 일부 실시양태에서, 삽입체 서열은 외생성 폴리펩티드를 코딩하는 서열의 역상보체인 서열의 상류에서 폴리아데닐화 부위를 포함한다. 일부 실시양태에서, 삽입체 서열은 리보좀 유전자좌가 아닌 유전자좌에서 게놈 내로 통합된다. 일부 실시양태에서, 삽입체 서열은 rDNA 유전자좌가 아닌 유전자좌에서 게놈 내로 통합된다. 일부 실시양태에서, 삽입체 서열은 유전자 또는 유전자의 조절 영역 내로 통합됨으로써, 유전자를 파괴하거나 유전자의 발현을 하향조절한다. 일부 실시양태에서, 삽입체 서열은 유전자 또는 유전자의 조절 영역 내로 통합됨으로써, 유전자의 발현을 상향조절한다. 일부 실시양태에서, 삽입체 서열은 게놈 내로 통합되고 유전자를 대체한다. 일부 실시양태에서, 삽입체 서열은 게놈 내로 안정적으로 통합된다. 일부 실시양태에서, 삽입체 서열은 게놈 내로 역전위된다. 일부 실시양태에서, 삽입체 서열은 mRNA에 의해 코딩된 엔도뉴클레아제에 의한 표적 부위의 DNA 가닥의 절단에 의해 게놈 내로 통합된다. 일부 실시양태에서, 삽입체 서열은 표적 프라이밍 역전사(TPRT)를 통해 게놈 내로 통합된다. 일부 실시양태에서, 삽입체 서열은 게놈의 DNA 표적 부위 내로의 mRNA의 역스플라이싱을 통해 게놈 내로 통합된다.

[0149] 일부 실시양태에서, 세포는 면역 세포이다. 일부 실시양태에서, 면역 세포는 T 세포 또는 B 세포이다. 일부 실시양태에서, 면역 세포는 골수 세포이다. 일부 실시양태에서, 면역 세포는 단핵구, 대식세포, 수지상 세포, 수지상 전구체 세포 및 대식세포 전구체 세포로 구성된 군으로부터 선택된다.

[0150] 일부 실시양태에서, mRNA는 자가통합 mRNA이다. 일부 실시양태에서, 방법은 mRNA를 세포 내로 도입하는 단계를 포함한다. 일부 실시양태에서, 방법은 mRNA를 코딩하는 벡터를 세포 내로 도입하는 단계를 포함한다. 일부 실시양태에서, 방법은 생체외에서 mRNA 또는 mRNA를 코딩하는 벡터를 세포 내로 도입하는 단계를 포함한다. 일부 실시양태에서, 방법은 세포를 인간 대상체에게 투여하는 단계도 포함한다. 일부 실시양태에서, 방법은 mRNA 또는 mRNA를 코딩하는 벡터를 인간 대상체에게 투여하는 단계를 포함한다. 일부 실시양태에서, 면역 반응은 인간 대상체에서 유발되지 않는다. 일부 실시양태에서, 상기 mRNA 또는 상기 벡터는 실질적으로 면역원성을 나타내지 않는다.

[0151] 일부 실시양태에서, 벡터는 플라스미드 또는 바이러스 벡터이다. 일부 실시양태에서, 벡터는 비-LTR 레트로트랜스포존을 포함한다. 일부 실시양태에서, 벡터는 인간 L1 요소를 포함한다. 일부 실시양태에서, 벡터는 L1 레트로트랜스포존 ORF1 유전자를 포함한다. 일부 실시양태에서, 벡터는 L1 레트로트랜스포존 ORF2 유전자를 포함한다. 일부 실시양태에서, 벡터는 L1 레트로트랜스포존을 포함한다. 일부 실시양태에서, 본원은 인간 LINE 1 역전위 요소를 코딩하는 서열을 포함하는 mRNA, 및 상기 mRNA를 포함하는 세포의 게놈 내로 역전위되고 통합될 수 있는 핵산 서열을 포함하는 페이로드를 제공한다. 일부 실시양태에서, 본원은 살아있는 세포, 예를 들어, 인간 세포 내로 전달될 수 있는 mRNA를 제공하는 것으로, 이 mRNA는 인간 LINE 1 역전위 요소를 코딩하는 서열, 및 상기 세포의 게놈 내로 역전위되고 통합될 수 있는 핵산 서열을 포함하는 페이로드를 포함한다. 일부 실시양태에서, 인간 LINE 1 역전위 요소를 코딩하는 서열은 L1 레트로트랜스포존 ORF1 서열 또는 이의 단편을 포함한다. 일부 실시양태에서, 인간 LINE 1 역전위 요소를 코딩하는 서열은 L1 레트로트랜스포존 ORF2 서열 또는 이의 단편을 포함한다. 일부 실시양태에서, 인간 LINE 1 역전위 요소를 코딩하는 서열은 L1 레트로트랜스포존 ORF1 서열 또는 이의 단편 및 L1 레트로트랜스포존 ORF2 서열 또는 이의 단편, 및 역전위에 의해 세포의 게놈 내로 통합되는 이종 서열인 핵산 "페이로드" 서열을 포함한다(예를 들어, 도 1b 참조).

[0152] 일부 실시양태에서, mRNA는 적어도 약 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9 또는 3 킬로염기이다. 일부 실시양태에서, mRNA는 최대 약 2.5, 2.6, 2.7, 2.8,

2.9, 3, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9 또는 5 킬로염기이다. 일부 실시양태에서, mRNA는 적어도 약 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9 또는 6 킬로염기이다. 일부 실시양태에서, mRNA는 적어도 약 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9 또는 7 킬로염기이다. 일부 실시양태에서, mRNA는 적어도 약 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9 또는 8 킬로염기이다. 일부 실시양태에서, mRNA는 적어도 약 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9 또는 9 킬로염기이다. 일부 실시양태에서, mRNA는 적어도 약 9.1, 9.2, 9.3, 9.4, 9.5, 9.6, 9.7, 9.8, 9.9 또는 10 킬로염기이다.

[0153] 일부 실시양태에서, mRNA는 mRNA의 분해를 억제하거나 방해하는 서열을 포함한다. 일부 실시양태에서, mRNA의 분해를 억제하거나 방해하는 서열은 엑소뉴클레아제(exonuclease) 또는 RNAse에 의한 mRNA의 분해를 억제하거나 방해한다. 일부 실시양태에서, mRNA의 분해를 억제하거나 방해하는 서열은 G 사중체, 슈도노트 또는 삼중체 서열이다. 일부 실시양태에서, mRNA의 분해를 억제하거나 방해하는 서열은 플라비바이러스 RNA의 엑소리보뉴클레아제(exoribonuclease) 내성 RNA 구조 또는 KSV의 ENE 요소이다. 일부 실시양태에서, mRNA의 분해를 억제하거나 방해하는 서열은 데아데닐라제(deadenylase)에 의한 mRNA의 분해를 억제하거나 방해한다. 일부 실시양태에서, mRNA의 분해를 억제하거나 방해하는 서열은 mRNA의 폴리 A 꼬리의 내부 또는 말단에서 비-아데노신 뉴클레오타이드를 포함한다. 일부 실시양태에서, mRNA의 분해를 억제하거나 방해하는 서열은 mRNA의 안정성을 증가시킨다. 일부 실시양태에서, 외생성 서열은 외생성 폴리펩티드를 코딩하는 서열을 포함한다. 일부 실시양태에서, 외생성 폴리펩티드를 코딩하는 서열은 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 서열과 인 프레임으로 있지 않다. 일부 실시양태에서, 외생성 폴리펩티드를 코딩하는 서열은 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 서열과 인 프레임으로 있지 않다. 일부 실시양태에서, 외생성 서열은 인트론을 포함하지 않는다. 일부 실시양태에서, 외생성 서열은 효소, 수용체, 수송 단백질, 구조 단백질, 호르몬, 항체, 수축 단백질 및 저장 단백질로 구성된 군으로부터 선택된 외생성 폴리펩티드를 코딩하는 서열을 포함한다. 일부 실시양태에서, 외생성 서열은 키메라 항원 수용체(CAR), 리간드, 항체, 수용체 및 효소로 구성된 군으로부터 선택된 외생성 폴리펩티드를 코딩하는 서열을 포함한다. 일부 실시양태에서, 외생성 서열은 조절 서열을 포함한다. 일부 실시양태에서, 조절 서열은 시스(cis) 작용 조절 서열을 포함한다. 일부 실시양태에서, 조절 서열은 인핸서(enhancer), 사일런서(silencer), 프로모터 또는 반응 요소로 구성된 군으로부터 선택된 시스 작용 조절 서열을 포함한다. 일부 실시양태에서, 조절 서열은 트랜스 작용 조절 서열을 포함한다. 일부 실시양태에서, 조절 서열은 전사 인자를 코딩하는 트랜스 작용 조절 서열을 포함한다.

[0154] 일부 실시양태에서, 삽입체 서열의 통합은 세포 건강에 불리한 영향을 미치지 않는다. 일부 실시양태에서, 엔도뉴클레아제, 역전사효소 또는 이들 둘 다는 삽입체 서열을 부위 특이적으로 통합시킬 수 있다.

[0155] 일부 실시양태에서, 본원에서 사용된 레트로트랜스포존 시스템은 정확한 부위 특이적 통합을 위해 추가로 조작된다. 일부 실시양태에서, 본원에서 사용된 레트로트랜스포존 시스템은 특이성을 증가시키기 위해 CRISPR-Cas 시스템과 페어링된다. 일부 실시양태에서, ORF 폴리펩티드 결합 서열, 예를 들어, TTTTTA는 세포의 게놈 서열 내로 부위 특이적으로 조작될 수 있다.

[0156] 일부 실시양태에서, mRNA는 추가 뉴클레아제 도메인, 또는 ORF2로부터 유래하지 않은 뉴클레아제 도메인을 코딩하는 서열을 포함한다. 일부 실시양태에서, mRNA는 메가TAL 뉴클레아제 도메인, TALEN 도메인, Cas9 도메인, R2 역요소의 징크 핑거 결합 도메인, 또는 AAV의 Rep78과 같은 반복 서열에 결합하는 DNA 결합 도메인을 코딩하는 서열을 포함한다. 일부 실시양태에서, 엔도뉴클레아제는 돌연변이를 갖지 않은 엔도뉴클레아제에 비해 엔도뉴클레아제의 활성을 감소시키는 돌연변이를 포함한다. 일부 실시양태에서, 엔도뉴클레아제는 ORF2p 엔도뉴클레아제이고, 돌연변이는 S228P이다. 일부 실시양태에서, mRNA는 역전사효소의 신뢰도 및/또는 진행성을 증가시키는 도메인을 코딩하는 서열을 포함한다. 일부 실시양태에서, 역전사효소는 ORF2 이외의 역요소의 역전사효소, 또는 ORF2p의 역전사효소에 비해 더 높은 신뢰도 및/또는 진행성을 가진 역전사효소이다. 일부 실시양태에서, 역전사효소는 II 군 인트론 역전사효소이다. 일부 실시양태에서, II 군 인트론 역전사효소는 IIA 군 인트론 역전사효소, IIB 군 인트론 역전사효소, 또는 IIC 군 인트론 역전사효소이다. 일부 실시양태에서, II 군 인트론 역전사효소는 TGIRT-II 또는 TGIRT-III이다.

[0157] 일부 실시양태에서, mRNA는 Alu 요소 및/또는 리보좀 결합 앱타머를 포함하는 서열을 포함한다. 일부 실시양태에서, mRNA는 DNA 결합 도메인을 포함하는 폴리펩티드를 코딩하는 서열을 포함한다. 일부 실시양태에서, 3' UTR 서열은 바이러스 3' UTR 또는 베타-글로빈 3' UTR로부터 유래한다.

[0158] 한 측면에서, 본원은 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 포함하는 조성물을 제공하는 것으로, 이때 mRNA

는 인간 LINE-1 트랜스포존 5' UTR 서열, 인간 LINE-1 트랜스포존 5' UTR 서열의 하류에 있는, ORF1p를 코딩하는 서열, ORF1p를 코딩하는 서열의 하류에 있는 ORF간 링커 서열, ORF간 링커 서열의 하류에 있는, ORF2p를 코딩하는 서열, 및 ORF2p를 코딩하는 서열의 하류에 있는, 인간 LINE-1 트랜스포존으로부터 유래한 3' UTR 서열을 포함하는 인간 LINE-1 트랜스포존 서열을 포함하고; 이때 3' UTR 서열은 삽입체 서열을 포함하고, 이 삽입체 서열은 외생성 폴리펩티드를 코딩하는 서열의 역상보체 또는 외생성 조절 요소를 코딩하는 서열의 역상보체이다.

[0159] 일부 실시양태에서, 삽입체 서열은 세포 내로 도입될 때 세포의 게놈 내로 통합된다. 일부 실시양태에서, 삽입체 서열은 병태 또는 질환과 관련된 유전자 내로 통합됨으로써, 상기 유전자를 파괴하거나 상기 유전자의 발현을 하향조절한다. 일부 실시양태에서, 삽입체 서열은 유전자 내로 통합됨으로써, 상기 유전자의 발현을 상향조절한다. 일부 실시양태에서, 재조합 mRNA 또는 mRNA를 코딩하는 벡터는 단리되거나 정제된다.

[0160] 한 측면에서, 본원은 (a) 인간 ORF1p 및 인간 ORF2p를 포함하는 긴 산재된 핵 요소(LINE) 폴리펩티드; 및 (b) 외생성 폴리펩티드를 코딩하는 서열의 역상보체 또는 외생성 조절 요소를 코딩하는 서열의 역상보체인 삽입체 서열을 코딩하는 뉴클레오타이드 서열을 포함하는 핵산을 포함하는 조성물로서, 실질적으로 면역원성을 나타내지 않는 조성물을 제공한다. 일부 실시양태에서, 삽입체 서열의 통합은 세포 건강에 불리한 영향을 미치지 않는다.

[0161] 일부 실시양태에서, 상기 조성물은 인간 ORF1p 및 인간 ORF2p 단백질을 포함한다. 일부 실시양태에서, 상기 조성물은 핵산과 복합체를 형성한 인간 ORF1p 및 인간 ORF2p를 포함하는 리보핵단백질(RNP)을 포함한다. 일부 실시양태에서, 핵산은 mRNA이다.

[0162] 한 측면에서, 본원은 본원에 기재된 조성물을 포함하는 세포를 포함하는 조성물을 제공한다. 일부 실시양태에서, 세포는 면역 세포이다. 일부 실시양태에서, 면역 세포는 T 세포 또는 B 세포이다. 일부 실시양태에서, 면역 세포는 골수 세포이다. 일부 실시양태에서, 면역 세포는 단핵구, 대식세포, 수지상 세포, 수지상 전구체 세포 및 대식세포 전구체 세포로 구성된 군으로부터 선택된다. 일부 실시양태에서, 삽입체 서열은 외생성 폴리펩티드를 코딩하는 서열의 역상보체이고, 외생성 폴리펩티드는 키메라 항원 수용체(CAR)이다.

[0163] 한 측면에서, 본원은 본원에 기재된 조성물 및 약학적으로 허용되는 부형제를 포함하는 약학 조성물을 제공한다. 일부 실시양태에서, 약학 조성물은 유전자 요법에 사용하기 위한 것이다. 일부 실시양태에서, 약학 조성물은 질환 또는 병태를 치료하기 위한 의약의 제조에 사용하기 위한 것이다. 일부 실시양태에서, 약학 조성물은 질환 또는 병태를 치료하는 데 사용하기 위한 것이다. 한 측면에서, 본원은 본원에 기재된 약학 조성물을, 질환 또는 병태를 가진 대상체에게 투여하는 단계를 포함하는, 대상체에서 질환을 치료하는 방법을 제공한다. 일부 실시양태에서, 상기 방법은 대상체에서 단백질 또는 기능성 RNA의 양 또는 활성을 증가시킨다. 일부 실시양태에서, 대상체는 단백질 또는 기능성 RNA의 결핍된 양 또는 활성을 가진다. 일부 실시양태에서, 단백질 또는 기능성 RNA의 결핍된 양 또는 활성은 상기 질환 또는 병태과 관련되어 있거나 이러한 질환 또는 병태를 야기한다.

[0164] 일부 실시양태에서, 상기 방법은 인간 침묵 허브(HUSH) 복합체를 억제하는 작용제, FAM208A를 억제하는 작용제, 또는 TRIM28을 억제하는 작용제를 투여하는 단계를 추가로 포함한다. 일부 실시양태에서, 인간 침묵 허브(HUSH) 복합체를 억제하는 작용제는 페리필린(Periphilin), TASOR 및/또는 MPP8을 억제하는 작용제이다. 일부 실시양태에서, 인간 침묵 허브(HUSH) 복합체를 억제하는 작용제는 HUSH 복합체의 어셈블리를 억제한다. 일부 실시양태에서, 상기 작용제는 판코니(Fanconi) 빈혈 복합체를 억제한다. 일부 실시양태에서, 상기 작용제는 FANCD2-FANC1 이종이량체 모노유비퀴틸화를 억제한다. 일부 실시양태에서, 상기 작용제는 FANCD2-FANC1 이종이량체 형성을 억제한다. 일부 실시양태에서, 상기 작용제는 판코니 빈혈(FA) 코어 복합체를 억제한다. FA 코어 복합체는 예를 들어, 화학요법에 의해 유도된 DNA 가닥간 가교결합에서 판코니 빈혈 DNA 손상 복구 경로의 성분이다. FA 코어 복합체는 RING 핑거 서브유닛인 FANCL의 두 카피에 의해 플랭킹된, FANCB의 2개 중심 이량체 및 100 kDa 서브유닛의 FA 관련 단백질(FAAP100)을 포함한다. 이 2개의 이종삼량체는 남은 5개의 서브유닛들을 어셈블링하기 위한 스캐폴드로서 작용하여, 확장된 비대칭 구조를 생성한다. 이 스캐폴드의 불안정화는 전체 복합체를 파괴하여, 비-기능성 FA 경로를 초래할 것이다. FA 코어 복합체를 억제할 수 있는 작용제의 예는 보르테조밉(Bortezomib) 및 쿠르쿠민(curcumin) 유사체 EF24 및 4H-TTD를 포함한다.

[0165] 따라서, 본 발명의 목적은 유전자 요법을 동물에게 제공하는 데 유용한 신규 트랜스포존 기반 벡터를 제공하는 것이다. 본 발명의 목적은 유전자 요법을 동물 또는 인간에게 제공하는 데 유용한 의약의 제조에 사용하기 위한 신규 트랜스포존 기반 벡터를 제공하는 것이다. 본 발명의 또 다른 목적은 세포에서 원하는 단백질 또는 펩티드의 생성을 코딩하는 신규 트랜스포존 기반 벡터를 제공하는 것이다. 본 발명의 또 다른 목적은 세포에서 원하는

핵산의 생성을 코딩하는 신규 트랜스포존 기반 벡터를 제공하는 것이다. 본 발명의 추가 목적은 선택된 유전자를 동물의 특정 세포 또는 조직으로 표적화하는 단계를 포함하는, 트랜스포존 기반 DNA 또는 RNA 구축물을 세포 및 조직 특이적으로 통합시키는 방법을 제공하는 것이다. 본 발명의 또 다른 목적은 전위효소에 의한 선택된 유전자의 안정한 통합을 향상시키는 세포 특이적 프로모터를 가진 DNA 또는 RNA 구축물을 디자인하는 단계 및 세포에서 상기 선택된 유전자를 발현시키는 단계를 포함하는, 트랜스포존 기반 DNA 또는 RNA 구축물을 세포 및 조직 특이적으로 발현시키는 방법을 제공하는 것이다. 본 발명의 목적은 트랜스포존 기반 벡터의 생식세포주 투여를 통해 유전자 요법을 세대에게 제공하는 것이다. 본 발명의 또 다른 목적은 트랜스포존 기반 벡터의 비-생식세포주 투여를 통해 동물에서 유전자 요법을 제공하는 것이다. 본 발명의 또 다른 목적은 트랜스포존 기반 벡터의 투여를 통해 동물에서 유전자 요법을 제공하는 것으로, 이때 상기 동물은 원하는 단백질, 펩티드 또는 핵산을 생성한다. 본 발명의 또 다른 목적은 트랜스포존 기반 벡터의 투여를 통해 동물에서 유전자 요법을 제공하는 것으로, 이때 상기 동물은 표적 세포의 수용체에 의해 인식되는 원하는 단백질 또는 펩티드를 생성한다. 본 발명의 또 다른 목적은 트랜스포존 기반 벡터의 투여를 통해 동물에서 유전자 요법을 제공하는 것으로, 이때 상기 동물은 원하는 융합 단백질 또는 융합 펩티드의 다른 단백질 또는 펩티드 성분을 세포에게 전달하여 생물학적 반응을 유도하기 위해 표적 세포의 수용체에 의해 인식되는 부분을 가진 원하는 융합 단백질 또는 융합 펩티드를 생성한다. 본 발명의 또 다른 목적은 관심 있는 유전자의 조직 특이적 통합 및 발현을 용이하게 하여 원하는 단백질, 펩티드 또는 핵산을 생성하기 위해 조직 특이적 프로모터 및 관심 있는 유전자를 포함하는 트랜스포존 기반 벡터의 투여를 통한 동물의 유전자 치료 방법을 제공하는 것이다. 본 발명의 또 다른 목적은 관심 있는 유전자의 세포 특이적 통합 및 발현을 용이하게 하여 원하는 단백질, 펩티드 또는 핵산을 생성하기 위해 세포 특이적 프로모터 및 관심 있는 유전자를 포함하는 트랜스포존 기반 벡터의 투여를 통한 동물의 유전자 치료 방법을 제공하는 것이다. 본 발명의 또 다른 목적은 관심 있는 유전자의 세포 특이적 통합 및 발현을 용이하게 하여 원하는 단백질, 펩티드 또는 핵산을 생성하기 위해 세포 특이적 프로모터 및 관심 있는 유전자를 포함하는 트랜스포존 기반 벡터의 투여를 통한 동물의 유전자 치료 방법을 제공하는 것으로, 이때 상기 원하는 단백질, 펩티드 또는 핵산은 동물에서 원하는 생물학적 효과를 가진다.

[0166] 한 측면에서, 본원은 하나 이상의 단백질을 코딩하는 핵산 서열을 포함하는 하나 이상의 핵산을 세포, 예를 들어, 골수 세포 내부로 전달하고 안정적으로 도입하는 방법 및 조성물을 제공하는 것으로, 이때 안정적 도입은 비-바이러스 기작을 통해 일어날 수 있다. 일부 실시양태에서, 골수 세포 내로의 핵산 조성물의 전달은 비-바이러스 기작을 통해 일어난다. 일부 실시양태에서, 핵산의 전달은 플라스미드 매개 전달을 우회할 수도 있다. 본원에서 사용된 바와 같이, "플라스미드"는 비-바이러스 발현 벡터, 예를 들어, 유전자 및/또는 유전자의 발현에 필요한 조절 요소를 코딩하는 핵산 분자를 지칭한다. 본원에서 사용된 바와 같이, "바이러스 벡터"는 또 다른 핵산을 세포 내로 수송할 수 있는 바이러스 유래 핵산을 지칭한다. 바이러스 벡터는 적절한 환경에 존재할 때 벡터에 의해 운반되는 하나 이상의 유전자에 의해 코딩된 단백질 또는 단백질들의 발현을 유도할 수 있다. 바이러스 벡터의 예는 레트로바이러스 벡터, 아데노바이러스 벡터, 렌티바이러스 벡터 및 아데노 관련 바이러스 벡터를 포함하나, 이들로 제한되지 않는다.

[0167] 일부 실시양태에서, 본원은 하나 이상의 단백질을 코딩하는 하나 이상의 핵산 서열을 포함하는 조성물을 골수 세포와 같은 세포의 내부로 전달하는 방법을 제공하는 것으로, 이때 상기 하나 이상의 핵산 서열은 RNA이다. 일부 실시양태에서, RNA는 mRNA이다. 일부 실시양태에서, 하나 이상의 핵산 서열을 포함하는 하나 이상의 mRNA가 전달된다. 일부 실시양태에서, 하나 이상의 mRNA는 적어도 하나의 변형 뉴클레오타이드를 포함할 수 있다. 본원에서 사용된 바와 같이, 용어 "뉴클레오타이드"는 염기-당-포스페이트 조합을 지칭한다. 뉴클레오타이드는 합성 뉴클레오타이드를 포함할 수 있다. 뉴클레오타이드는 합성 뉴클레오타이드 유사체를 포함할 수 있다. 뉴클레오타이드는 핵산 서열(예를 들어, 데옥시리보핵산(DNA) 또는 리보핵산(RNA))의 단량체 유닛일 수 있다. 용어 뉴클레오타이드는 리보뉴클레오사이드 트리포스페이트 아데노신 트리포스페이트(ATP), 우리딘 트리포스페이트(UTP), 사이토신 트리포스페이트(CTP), 구아노신 트리포스페이트(GTP) 및 데옥시리보뉴클레오사이드 트리포스페이트, 예컨대, dATP, dCTP, dITP, dUTP, dGTP 또는 이들의 유도체를 포함할 수 있다. 이러한 유도체는 예를 들어, [aS]dATP, 7-데아자-dGTP 및 7-데아자-dATP, 및 그 자신을 함유하는 핵산 분자에 뉴클레아제 내성을 부여하는 뉴클레오타이드 유도체를 포함할 수 있다. 본원에서 사용된 용어 뉴클레오타이드는 디데옥시리보뉴클레오사이드 트리포스페이트(ddNTP) 및 이의 유도체를 지칭할 수 있다. 디데옥시리보뉴클레오사이드 트리포스페이트의 예시적인 예는 ddATP, ddCTP, ddGTP, ddITP 및 ddTTP를 포함할 수 있으나, 이들로 제한되지 않는다. 뉴클레오타이드는 표지부착되지 않을 수 있거나 잘 공지되어 있는 기법에 의해 검출 가능하게 표지부착될 수 있다. 표지부착은 양자 점을 사용함으로써 수행될 수도 있다. 검출 가능한 표지는 예를 들어, 방사성 동위원소, 형광 표지, 화학발광 표지, 생체발광 표지 및 효소 표지를 포함할 수 있다. 뉴클레오타이드의 형광 표지는 플루오레세인, 5-카

르복시플루오레세인(FAM), 2'7'-디메톡시-4'5-디클로로-6-카르복시플루오레세인(JOE), 로다민, 6-카르복시로다민(R6G), N,N,NcN'-테트라메틸-6-카르복시로다민(TAMRA), 6-카르복시-X-로다민(ROX), 4-(4'-디메틸아미노페닐아조)벤조산(DABCYL), 캐스케이드 블루(Cascade Blue), 오레곤 그린(Oregon Green), 텍사스 레드(Texas Red), 시아닌(Cyanine) 및 5-(2'-아미노에틸)아미노나프탈렌-1-설폰산(EDANS)을 포함할 수 있으나, 이들로 제한되지 않는다. 형광 표지부착된 뉴클레오타이드의 구체적인 예는 퍼킨 엘머(Perkin Elmer)(캘리포니아주 포스터 시티 소재)로부터 입수될 수 있는 [R6G]dUTP, [TAMRA]dUTP, [R110]dCTP, [R6G]dCTP, [TAMRA]dCTP, [JOE]ddATP, [R6G]ddATP, [FAM]ddCTP, [R110]ddCTP, [TAN1RA]ddGTP, [ROX]ddTTP, [dR6G]ddATP, [dR110]ddCTP, [dTAMRA]ddGTP 및 [dROX]ddTTP; 아머샴(Amersham)(일리노이주 알링톤 하이츠 소재)으로부터 입수될 수 있는 플루오로링크(FluoroLink) 데옥시뉴클레오타이드, 플루오로링크 Cy3-dCTP, 플루오로링크 Cy5-dCTP, 플루오로링크 플루오르 X-dCTP, 플루오로링크 Cy3-dUTP 및 플루오로링크 Cy5-dUTP; 뵈링거 만하임(Boehringer Mannheim)(인디애나주 인디애나폴리스 소재)으로부터 입수될 수 있는 플루오레세인-15-dATP, 플루오레세인-12-dUTP, 테트라메틸-로다민-6-dUTP, TR770-9-dATP, 플루오레세인-12-ddUTP, 플루오레세인-12-UTP 및 플루오레세인-15-2'-dATP; 및 몰레큘라 프로브스(Molecular Probes)(오레곤주 유진 소재)로부터 입수될 수 있는 염색체 표지부착된 뉴클레오타이드, BODIPY-FL-1 4-UTP, BODIPY-FL-4-UTP, BODIPY-TMR-14-UTP, BODIPY-TMR-14-dUTP, BODIPY-TR-14-UTP, BODIPY-TR-14-dUTP, 캐스케이드 블루-7-UTP, 캐스케이드 블루-7-dUTP, 플루오레세인-12-UTP, 플루오레세인-12-dUTP, 오레곤 그린 488-5-dUTP, 로다민 그린-5-UTP, 로다민 그린-5-dUTP, 테트라메틸로다민-6-UTP, 테트라메틸로다민-6-dUTP, 텍사스 레드-5-UTP, 텍사스 레드-5-dUTP 및 텍사스 레드-12-dUTP를 포함할 수 있다. 뉴클레오타이드는 화학적 변형에 의해 표지부착될 수도 있거나 표식될 수도 있다. 화학적으로 변형된 단일 뉴클레오타이드는 바이오틴-dNTP일 수 있다. 바이오티닐화된 dNTP의 일부 비제한적 예는 바이오틴-dATP(예를 들어, 바이오-N6-ddATP, 바이오틴-14-dATP), 바이오틴-dCTP(예를 들어, 바이오틴-11-cICTP, 바이오틴-14-dCTP) 및 바이오틴-dUTP(예를 들어, 바이오틴-11-dUTP, 바이오틴-1.6-dUTP, 바이오틴-20-dUTP)를 포함할 수 있다.

[0168] 용어 "폴리뉴클레오타이드", "올리고뉴클레오타이드" 및 "핵산"은 단일, 이중 또는 다중 가닥 형태로 존재하는, 데옥시리보뉴클레오타이드 또는 리보뉴클레오타이드인 임의의 길이의 뉴클레오타이드 또는 이의 유사체의 중합체 형태를 지칭하기 위해 교환 가능하게 사용된다. 폴리뉴클레오타이드는 세포에 대한 외생성 또는 내생성 폴리뉴클레오타이드일 수 있다. 폴리뉴클레오타이드는 무세포 환경에 존재할 수 있다. 폴리뉴클레오타이드는 유전자 또는 이의 단편일 수 있다. 폴리뉴클레오타이드는 DNA일 수 있다. 폴리뉴클레오타이드는 RNA일 수 있다. 폴리뉴클레오타이드는 임의의 3차원 구조를 가질 수 있고, 공지되어 있거나 공지되어 있지 않은 임의의 기능을 수행할 수 있다. 폴리뉴클레오타이드는 하나 이상의 유사체(예를 들어, 변경된 골격, 당 또는 핵염기)를 포함할 수 있다. 존재하는 경우, 뉴클레오타이드 구조에 대한 변형은 중합체의 어셈블리 전 또는 후에 부여될 수 있다. 변형 뉴클레오타이드 또는 유사체의 일부 비제한적 예는 슈도우리딘, 5-브로모우라실, 5-메틸사이토신, 퀩티드 핵산, 제노(xeno) 핵산, 모르폴리노, 잠긴 핵산, 글리콜 핵산, 트레오스 핵산, 디데옥시뉴클레오타이드, 코르디셉핀(cordycepin), 7-데아자-GTP, 형광단(예를 들어, 당에 연결된 로다민 또는 플루오레세인), 티올 함유 뉴클레오타이드, 바이오틴 연결된 뉴클레오타이드, 형광 염기 유사체, CpG 아일랜드, 메틸-7-구아노신, 메틸화된 뉴클레오타이드, 이노신, 티오우리딘, 슈도우리딘, 디하이드로우리딘, 퀴오신 및 와이오신을 포함한다. 폴리뉴클레오타이드의 비제한적 예는 유전자 또는 유전자 단편의 코딩 또는 비코딩 영역, 연관 분석으로부터 정의된 유전자좌들(유전자좌), 엑손, 인트론, 메신저 RNA(mRNA), 전달 RNA(tRNA), 리보좀 RNA(rRNA), 짧은 간섭 RNA(siRNA), 짧은 헤어핀 RNA(shRNA), 마이크로-RNA(miRNA), 리보자임, eDNA, 재조합 폴리뉴클레오타이드, 분지된 폴리뉴클레오타이드, 플라스미드, 벡터, 임의의 서열의 단리된 DNA, 임의의 서열의 단리된 RNA, 무세포 DNA(cfDNA) 및 무세포 RNA(cfRNA)를 포함하는 무세포 폴리뉴클레오타이드, 핵산 프로브, 및 프라이머를 포함한다. 뉴클레오타이드의 서열은 비-뉴클레오타이드 성분에 의해 불연속될 수 있다.

[0169] 일부 실시양태에서, 핵산 조성물은 인간 대식세포 또는 수지상 세포 또는 적합한 골수 세포 또는 골수 전구체 세포에서 면역 반응에 관여하는 막횡단 수용체(예를 들어, 식세포 수용체 또는 합성 키메라 항원 수용체)를 코딩하는 적어도 하나의 mRNA를 포함하는 하나 이상의 mRNA를 포함할 수 있다. 일부 실시양태에서, 핵산 조성물은 하나 이상의 mRNA, 및 상기 핵산을 조혈 유래의 세포, 예컨대, 골수 세포 또는 골수 세포 전구체 세포 내로 전달하기 위한 하나 이상의 지질을 포함한다. 일부 실시양태에서, 하나 이상의 지질은 리포좀 복합체를 형성할 수 있다.

[0170] 본원에서 사용된 바와 같이, 본원에 기재된 조성물은 세포 내부로의 전달을 위해 사용될 수 있다. 세포는 하나 이상의 세포를 가진 임의의 유기체로부터 유래할 수 있다. 일부 비제한적 예는 원핵 세포, 진핵 세포, 박테리아 세포, 고세균 세포, 단세포 진핵 유기체의 세포, 원생동물 세포, 식물의 세포(예를 들어, 식물 작물, 과일, 채소, 곡물, 대두, 옥수수, 메이즈(maize), 밀, 종자, 토마토, 쌀, 카사바, 사탕수수, 호박, 건초, 감자, 목화,

대마초, 담배, 개화 식물, 침엽수, 겉씨식물, 양치류, 석송, 붕어마름, 우산이끼, 이끼의 세포), 조류 세포(예를 들어, 보트리오코커스 브라우니이(Botryococcus braunii), 클라미도모나스 레인하르드티이(Chlamydomonas reinhardtii), 난노클로롭시스 가디타나(Nannochloropsis gaditana), 클로렐라 피레노이도사(Chlorella pyrenoidosa), 사르가숨 파텐스 씨. 아가르드(Sargassum patens C. Agardh) 등), 해조류(예를 들어, 다시마), 진균 세포(예를 들어, 효모 세포, 버섯의 세포), 동물 세포, 무척추 동물(예를 들어, 초파리, 자포동물, 극피동물, 선충류 등)의 세포, 척추동물(예를 들어, 어류, 양서류, 파충류, 조류, 포유동물)의 세포, 포유동물(예를 들어, 돼지, 소, 염소, 양, 설치류, 래트, 마우스, 비인간 영장류, 인간 등)의 세포 등을 포함한다. 종종, 세포는 천연 유기체로부터 유래하지 않을 수 있다(예를 들어, 세포는 합성에 의해 제조될 수 있고, 이것은 종종 인공 세포로서 지칭된다). 일부 실시양태에서, 본원에서 언급된 세포는 포유동물 세포이다. 일부 실시양태에서, 세포는 인간 세포이다. 본원에 기재된 방법 및 조성물은 유전 물질을 세포, 보다 구체적으로 인간 세포 내로 도입하는 것에 관한 것으로, 이때 인간 세포는 임의의 인간 세포일 수 있다. 본원에서 사용된 바와 같이, 인간 세포는 임의의 기원의 세포, 예를 들어, 체세포, 뉴런, 섬유모세포, 근육 세포, 상피 세포, 심장 세포 또는 조혈 세포일 수 있다. 본원에 기재된 방법 및 조성물은 외생성 핵산을 형질감염되기 어려운 인간 세포 내로 도입하는 데에도 적용될 수 있고 이러한 도입에 유용할 수 있다. 상기 방법은 일단 적합한 외생성 핵산 구축물이 디자인되고 개발되면 단순하고 보편적으로 적용될 수 있다. 본원에 기재된 방법 및 조성물은 생체외에서 외생성 핵산을 세포 내로 도입하는 데 적용될 수 있다. 일부 실시양태에서, 상기 조성물은 유기체에게 전신 투여하는 데 적용될 수 있고, 이때 상기 조성물 중의 핵산 물질은 생체내에서 세포에 의해 흡수될 수 있고, 그 결과 생체내에서 세포 내로 도입된다.

[0171] 일부 실시양태에서, 본원에 기재된 방법 및 조성물은 외생성 핵산을 인간 조혈 세포, 예를 들어, 조혈 기원의 인간 세포, 예컨대, 인간 골수 세포 또는 골수 세포 전구체 내로 도입하는 것에 관한 것일 수 있다. 그러나, 본원에 기재된 방법 및 조성물은 최소 변형을 가진 임의의 생물학적 세포에 사용될 수 있거나 이 세포에 사용되기에 적합하도록 만들어질 수 있다. 그러므로, 세포는 살아있는 유기체의 기본 구조적, 기능적 및/또는 생물학적 유닛인 임의의 세포를 지칭할 수 있다.

[0172] 한 측면에서, 본원은 하나 이상의 핵산을 세포의 게놈 내로 안정적으로 통합시키기 위해 전위 요소를 사용하는 방법 및 조성물을 제공하는 것으로, 이때 상기 세포는 조혈 세포의 구성원, 예를 들어, 골수 세포이다. 일부 실시양태에서, 하나 이상의 핵산은 면역 반응에서 역할을 하는 막횡단 수용체 단백질을 코딩하는 적어도 하나의 핵산 서열을 포함한다. 일부 실시양태에서, 상기 방법 및 조성물은 하나 이상의 핵산 서열을 골수 세포 내로 도입하기 위한 역전위 요소의 사용에 관한 것이다. 상기 핵산 조성물은 유전자와 같은 하나 이상의 핵산 서열을 포함할 수 있고, 이때 상기 유전자는 형질전이유전자이다. 본원에서 사용된 바와 같이, 용어 "유전자"는 RNA 전사체를 코딩하는 데 관여하는 핵산(예를 들어, DNA, 예컨대, 게놈 DNA 및 cDNA) 및 이의 상응하는 뉴클레오타이드 서열을 지칭한다. 게놈 DNA에 대하여 본원에서 사용된 용어는 조절 영역뿐만 아니라 개재 비코딩 영역도 포함하고 5' 말단 및 3' 말단을 포함할 수 있다. 일부 사용에서, 상기 용어는 5' 비번역 영역 및 3' 비번역 영역(5'-UTR 및 3'-UTR), 엑손 및 인트론을 포함하는 전사된 서열을 포괄한다. 일부 유전자에서, 전사된 영역은 폴리펩티드를 코딩하는 "오픈 리딩 프레임"을 함유할 것이다. 상기 용어의 일부 사용에서, "유전자"는 폴리펩티드를 코딩하는 데 필요한 코딩 서열(예를 들어, "오픈 리딩 프레임" 또는 "코딩 영역")만을 포함한다. 일부 경우, 유전자, 예를 들어, 리보좀 RNA 유전자(rRNA) 및 전달 RNA(tRNA) 유전자는 폴리펩티드를 코딩하지 않는다. 일부 경우, 용어 "유전자"는 전사된 서열을 포함할 뿐만 아니라, 추가로 상류 및 하류 조절 영역, 인핸서 및 프로모터를 비롯한 비전사 영역도 포함한다. 유전자는 유기체 게놈 내의 그의 천연 위치에 있는 "내생성 유전자" 또는 천연 유전자를 지칭할 수 있다. 유전자는 "외생성 유전자" 또는 비천연 유전자를 지칭할 수 있다. 비천연 유전자는 숙주 유기체에서 정상적으로 발견되지 않으나 유전자 전달에 의해 숙주 유기체 내로 도입된 유전자를 지칭할 수 있다. 비천연 유전자는 유기체의 게놈 내의 그의 천연 위치에 있지 않은 유전자도 지칭할 수 있다. 비천연 유전자는 돌연변이, 삽입 및/또는 결실을 포함하는 천연 생성 핵산 또는 폴리펩티드 서열(예를 들어, 비천연 서열)도 지칭할 수 있다.

[0173] 용어 "형질전이유전자"는 본원에서 종종 수용자 세포로서 불릴 수 있는 세포 내로 도입된 임의의 핵산 분자를 지칭한다. 형질전이유전자를 받은 후 생성된 세포는 형질전환 세포로서 지칭될 수 있다. 형질전이유전자는 형질전환 유기체 또는 세포에 대한 부분적 또는 전체적 이종성을 가진(즉, 외래) 유전자를 포함할 수 있거나, 유기체 또는 세포의 내생성 유전자에 대한 동종 유전자를 나타낼 수 있다. 일부 경우, 형질전이유전자는 임의의 폴리뉴클레오타이드, 예컨대, 폴리펩티드 또는 단백질을 코딩하는 유전자, 억제 폴리뉴클레오타이드로 전사되는 폴리뉴클레오타이드, 또는 전사되지 않는(예를 들어, 발현 조절 요소, 예컨대, 전사를 유도하는 프로모터를 결여하는) 폴리뉴클레오타이드를 포함한다. 전사체 및 코딩된 폴리펩티드는 일괄적으로 "유전자 생성물"로서 지칭

될 수 있다. 폴리뉴클레오타이드가 게놈 DNA로부터 유래한 경우, 발현은 진핵 세포에서 mRNA의 스플라이싱을 포함할 수 있다. 발현과 관련하여 "상향조절"은 야생형 상태에서의 그의 발현 수준에 비해 폴리뉴클레오타이드(예를 들어, mRNA와 같은 RNA) 및/또는 폴리펩티드 서열의 증가된 발현 수준을 의미하는 반면, "하향조절"은 야생형 상태에서의 그의 발현에 비해 폴리뉴클레오타이드(예를 들어, mRNA와 같은 RNA) 및/또는 폴리펩티드 서열의 감소된 발현 수준을 의미한다. 형질감염된 유전자의 발현은 세포에서 일시적 또는 안정적으로 일어날 수 있다. "일시적 발현" 동안, 형질감염된 유전자는 세포 분열 동안 딸세포에게 전달되지 않는다. 그의 발현이 형질감염된 세포로 제한되기 때문에, 유전자의 발현은 시간 경과에 따라 상실된다. 대조적으로, 형질감염된 유전자의 안정한 발현은 유전자가 선택 장점을 형질감염된 세포에 부여하는 또 다른 유전자와 공-형질감염될 때 일어날 수 있다. 이러한 선택 장점은 세포에게 제공된 특정 독소에 대한 내성일 수 있다. 형질감염된 유전자가 발현될 것이 요구되는 경우, 본원은 코돈 최적화된 서열의 사용을 예상한다. 코돈 최적화된 서열의 일례는 진핵생물, 예를 들어, 인간에서의 발현을 위해 최적화된 서열(즉, 인간에서의 발현을 위해 최적화됨), 또는 또 다른 진핵생물, 동물 또는 포유동물에서의 발현을 위해 최적화된 서열일 수 있다. 인간 이외의 숙주 종을 위한 코돈 최적화 또는 특정 장기를 위한 코돈 최적화는 공지되어 있다. 일부 실시양태에서, 단백질을 코딩하는 코딩 서열은 진핵 세포와 같은 특정 세포에서 발현되도록 코돈 최적화될 수 있다. 진핵 세포는 인간, 또는 본원에서 논의된 비인간 진핵생물 또는 동물 또는 포유동물, 예를 들어, 마우스, 래트, 토끼, 개, 가축 또는 비인간 포유동물 또는 영장류를 포함하나 이들로 제한되지 않는 특정 유기체, 예컨대, 식물 또는 포유동물의 세포일 수 있거나 이러한 유기체로부터 유래할 수 있다. 코돈 최적화는 천연 아미노산 서열을 유지하면서 천연 서열의 적어도 하나의 코돈(예를 들어, 약 1개, 2개, 3개, 4개, 5개, 10개, 15개, 20개, 25개 또는 50개 이상의 코돈)을 그 숙주 세포의 유전자에서 더 자주 또는 가장 자주 사용되는 코돈으로 대체함으로써 관심 있는 숙주 세포에서 발현을 향상시키도록 핵산 서열을 변형시키는 과정을 지칭한다. 다양한 종들은 특정 아미노산의 특정 코돈에 대한 특정 편향을 나타낸다. 코돈 편향(유기체들 사이에 코돈 사용의 차이)은 종종 메신저 RNA(mRNA)의 번역 효율과 상관관계를 갖고, 이 번역 효율은 특히 번역되는 코돈의 성질과 특정 전달 RNA(tRNA) 분자의 이용 가능성에 의해 좌우되는 것으로 여겨진다. 세포에서 선택된 tRNA의 우세는 일반적으로 펩티드 합성에서 가장 자주 사용되는 코돈을 반영할 수 있다. 따라서, 코돈 최적화를 기반으로 주어진 유기체에서의 최적 유전자 발현을 위해 유전자를 다듬을 수 있다. 코돈 사용 표는 예를 들어, 웹사이트(www.kazusa.orjp/codon/)에서 이용될 수 있는 "코돈 사용 데이터베이스"에서 용이하게 이용될 수 있고, 이 표는 다수의 방식들에 의해 개조될 수 있다. 특정 숙주 세포에서 발현시키기 위해 특정 서열을 코돈 최적화하는 컴퓨터 알고리즘도 이용될 수 있고, 예컨대, Gene Forge(Aptagen; 펜실베니아주 제이코버스 소재)도 이용될 수 있다.

[0174] 본원에서 사용된 "멀티시스트론 전사체"는 하나 초과의 단백질 코딩 영역 또는 시스트론을 함유하는 mRNA 분자를 지칭한다. 2개의 코딩 영역을 포함하는 mRNA는 "바이시스트론 전사체"로서 표기된다. "5'-근위" 코딩 영역 또는 시스트론은 번역 시작 코돈(통상적으로 AUG)이 멀티시스트론 mRNA 분자의 5' 말단에 가장 가까운 코딩 영역이다. "5'-원위" 코딩 영역 또는 시스트론은 번역 시작 코돈(통상적으로 AUG)이 mRNA의 5' 말단에 가장 가까운 시작 코돈이 아닌 코딩 영역이다.

[0175] 용어 "형질감염" 또는 "형질감염된"은 핵산이 비-바이러스 또는 바이러스 기반 방법에 의해 세포 내로 도입되는 것을 의미한다. 핵산 분자는 완전한 단백질 또는 이의 기능성 부분을 코딩하는 유전자 서열일 수 있다. 예를 들어, 문헌[Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 18.1-18.88]을 참조한다.

[0176] 본원에서 사용된 용어 "프로모터"는 세포에서 코딩 서열의 전사를 유도할 수 있는 폴리뉴클레오타이드 서열을 지칭한다. 따라서, 본 개시내용의 폴리뉴클레오타이드 구축물에 사용된 프로모터는 시스 작용 전사 조절 요소, 및 유전자의 전사 시간 및/또는 속도를 제어하거나 조절하는 데 관여하는 조절 서열을 포함한다. 예를 들어, 프로모터는 전사 조절에 관여하는 인핸서, 프로모터, 전사 터미네이터, 복제 기점, 염색체 통합 서열, 5' 및 3' 비번역 영역, 또는 인트론 서열을 비롯한 시스 작용 전사 조절 요소일 수 있다. 이 시스 작용 서열은 전형적으로 단백질 또는 다른 생체분자와 상호작용하여 유전자 전사를 수행한다(켜기/끄기, 제어, 조절 등). "항시성 프로모터"는 거의 모든 조직 유형에서 전사를 시작할 수 있는 프로모터인 반면, "조직 특이적 프로모터"는 하나 또는 소수의 특정 조직 유형에서만 전사를 시작한다. "유도성 프로모터"는 특정 환경 조건, 발생 조건, 또는 약물 또는 화학적 조건 하에서만 전사를 시작하는 프로모터이다. 예시적인 유도성 프로모터는 독시사이클린 또는 테트라사이클린 유도성 프로모터일 수 있다. 테트라사이클린에 의해 조절되는 프로모터는 tet-on 및 tet-off 시스템으로서 지칭되는 테트라사이클린 유도성 또는 테트라사이클린 억제성 프로모터일 수 있다. *tet*에 의해 조절되는 시스템은 두 가지 성분, 즉 테트라사이클린에 의해 조절되는 조절제(트랜스활성화제로서도 지칭됨)(tTA 또는 rtTA) 및 테트라사이클린 의존적 방식으로 하류 cDNA의 발현을 조절하는 tTA/rtTA 의존적 프로모터에 의존한다. tTA는 에스케리키아 콜라이(*Escherichia coli*)의 Tn10 테트라사이클린 내성 오페론의 리프레서(repressor)

및 헤르페스 심플렉스 바이러스의 단백질 16(VP16)의 카르복실 말단 부분을 함유하는 융합 단백질이다. tTA 의존적 프로모터는 *tet* 오퍼레이터(*tetO*) 서열(7개의 동족 오퍼레이터 서열의 어레이)에 융합된 최소 RNA 중합효소 II 프로모터로 구성된다. 이 융합은 진핵 세포에서 *tet* 리프레서를 강한 전사 활성화제로 전환시킨다. 테트라사이클린 또는 이의 유도체(예컨대, 독시사이클린)의 부재 하에서, tTA는 *tetO* 서열에 결합하여, tTA 의존적 프로모터의 전사 활성화를 허용한다. 그러나, 독시사이클린의 존재 하에서, tTA는 그의 표적과 상호작용할 수 없고 전사는 일어나지 않는다. tTA를 사용하는 *tet* 시스템은 테트라사이클린 또는 독시사이클린이 전사 하향조절을 허용하기 때문에 *tet*-OFF로서 지칭된다. 대조적으로, tet-ON 시스템에서, rtTA로서 지칭되는 tTA의 돌연변이체 형태는 무작위 돌연변이유발을 이용함으로써 단리되었다. tTA와 대조적으로, rtTA는 독시사이클린의 부재 하에서 작용하지 않으나, 트랜스활성화를 위해 리간드의 존재를 요구한다. 용어 "엑손"은 인접 서열을 성숙 mRNA 전사체에 기여하는 것으로 생물정보학적으로 예측되고/되거나 실험적으로 확인된, 게놈 DNA에서 발견된 핵산 서열을 지칭한다. 용어 "인트론"은 발현된 단백질의 일부 또는 전부를 코딩하지 않는 것으로 생물정보학적으로 예측되고/되거나 실험적으로 확인되고 내생성 조건에서 RNA(예를 들어, 전구-mRNA) 분자로 전사되나, RNA가 단백질로 번역되기 전에 내생성 RNA(예를 들어, 전구-mRNA)로부터 스플라이싱되는, 게놈 DNA에 존재하는 서열을 의미한다.

[0177] 용어 "스플라이스 수용자 부위"는 전구-mRNA의 스플라이싱 동안 수용자 부위인 것으로 생물정보학적으로 예측되고/되거나 실험적으로 확인된, 게놈 DNA에 존재하는 서열을 지칭하고, 확인된 천연 스플라이스 수용자 부위 및 인위적으로 유도된 또는 유도될 수 있는 스플라이스 수용자 부위, 및 미확인된 천연 스플라이스 수용자 부위 및 인위적으로 유도된 또는 유도될 수 있는 스플라이스 수용자 부위를 포함할 수 있다.

[0178] "내부 리보좀 진입 부위" 또는 "IRES"는 5'-말단/캡 독립적 번역 시작을 허용함으로써, 단일 메신저 RNA(mRNA) 분자로부터 2개의 단백질을 발현시킬 가능성을 높이는 뉴클레오타이드 서열을 의미한다. IRES는 일반적으로 캡핑되지 않은 게놈을 가진 양성 가닥 RNA 바이러스의 5' UTR에 위치한다. 단일 mRNA 분자로부터 2개의 단백질을 발현시키는 또 다른 수단은 2A 펩티드(유사) 서열을 그들의 코딩 서열 사이에 삽입하는 것이다. 2A 펩티드(유사) 서열은 "리보좀 건너뛰기(skipping)", "정지-진행(stop-go)" 번역 및 "정지 계속(stop carry-on)" 번역으로서 다양하게 지칭되는 과정에 의한 1차 번역 생성물의 자가처리를 매개한다. 2A 펩티드(유사) 서열은 피코르나비리대(Picornaviridae), 플라비비리대(Flaviviridae), 테트라비리대(Tetraviridae), 디시스트로비리대(Dicistroviridae), 레오비리대(Reoviridae) 및 토티비리대(Totiviridae)를 포함하는 다양한 양성 및 이중 가닥 RNA 바이러스 군에 존재한다.

[0179] 용어 "2A 펩티드"는 진핵 세포에서 번역하는 동안 폴리펩티드의 "절단"을 매개하는 한 클래스의 18개 내지 22개 아미노산(AA) 길이 바이러스 올리고펩티드를 지칭한다. "2A"라는 명칭은 바이러스 게놈의 특정 영역을 지칭하고 상이한 바이러스 2A들은 일반적으로 이들의 기원인 바이러스의 명칭을 따서 명명되었다. 최초로 발견된 2A는 F2A(구제역 바이러스)이고, 그 후 E2A(말 비염 A 바이러스), P2A(돼지 테스코바이러스(teschovirus)-1 2A), 및 T2A(토세아 아시그나(thosea asigna) 바이러스 2A)도 확인되었다. 2A 매개 "자가 절단"의 기작은 2A 서열의 C-말단에서 글리실-프롤릴 펩티드 결합의 형성을 건너뛰는 리보좀인 것으로 여겨진다. 2A 펩티드(유사) 서열은 "리보좀 건너뛰기", "정지-진행" 번역 및 "정지 계속" 번역으로서 다양하게 지칭되는 과정에 의한 1차 번역 생성물의 자가처리를 매개한다. 2A 펩티드(유사) 서열은 피코르나비리대, 플라비비리대, 테트라비리대, 디시스트로비리대, 레오비리대 및 토티비리대를 포함하는 다양한 양성 및 이중 가닥 RNA 바이러스 군에 존재한다.

[0180] 본원에서 사용된 바와 같이, 용어 "작동 가능하게 연결된"은 핵산 분절 또는 폴리펩티드 분절과 같은 2개 이상의 분절 사이의 기능적 관계를 의미한다. 전형적으로, 이 용어는 전사 조절 서열과 전사된 서열의 기능적 관계를 의미한다.

[0181] 용어 "종결 서열"은 숙주 세포의 중합효소에 의해 인식되고 전사를 종결시키는 핵산 서열을 의미한다. 종결 서열은 천연 또는 합성 유전자의 3' 말단에서 mRNA 전사의 종결 또는 상류 오픈 리딩 프레임의 mRNA 전사 및 리보좀 번역 둘 다를 제공하는 DNA 서열이다. 원핵생물 종결 서열은 통상적으로 AT 풍부 서열이 뒤따르는 2중 대칭을 가진 GC 풍부 영역을 포함한다. 통상적으로 사용되는 종결 서열은 T7 종결 서열이다. 박테리오파지 람다로부터 유래한 TINT3, TL13, TL2, TR1, TR2 및 T6S 종결 신호, 및 박테리아 유전자, 예컨대, 이. 콜라이의 trp 유전자로부터 유래한 종결 신호를 비롯한 다양한 종결 서열들이 당분야에 공지되어 있으며 본 발명의 핵산 구축물에 사용될 수 있다.

[0182] 용어 "폴리아데닐화 서열"("폴리 A 부위" 또는 "폴리 A 서열"로서도 지칭됨)은 신생 RNA 전사체의 종결 및 폴리아데닐화 둘 다를 유도하는 DNA 서열을 지칭한다. 폴리 A 꼬리를 결여한 전사체는 전형적으로 불안정하고 빠르

게 분해되기 때문에, 재조합 전사체의 효율적인 폴리아데닐화가 바람직하다. 발현 벡터에 사용되는 폴리 A 신호는 "이종" 또는 "내생성" 폴리 A 신호일 수 있다. 내생성 폴리 A 신호는 게놈에서 소정의 유전자의 코딩 영역의 3' 말단에서 천연적으로 발견되는 폴리 A 신호이다. 이종 폴리 A 신호는 한 유전자로부터 단리되고 또 다른 유전자, 예를 들어, 단백질에 대한 코딩 서열의 3'에 위치하는 폴리 A 신호이다. 통상적으로 사용되는 이종 폴리 A 신호는 SV40 폴리 A 신호이다. SV40 폴리 A 신호는 237 bp BamHI/BclI 제한 단편에 함유되고 종결 및 폴리아데닐화 둘 다를 유도하고; 다수의 벡터들이 SV40 폴리 A 신호를 함유한다. 또 다른 통상적으로 사용되는 이종 폴리 A 신호는 소 성장 호르몬(BGH) 유전자로부터 유래하고; BGH 폴리 A 신호는 다수의 상업적으로 입수 가능한 벡터들에서도 이용될 수 있다. 헤르페스 심플렉스 바이러스 타이미딘 키나제(HSV tk) 유전자의 폴리 A 신호도 다수의 상업적 발현 벡터들에서 폴리 A 신호로서 사용된다. 폴리아데닐화 신호는 RNA를 세포 핵 내부로부터 세포질로 수송하는 것을 용이하게 할 뿐만 아니라 이러한 RNA의 세포 반감기를 증가시킨다. 폴리아데닐화 신호는 mRNA의 3' 말단에 존재한다.

[0183] 본원에서 사용된 용어 "상보체", "상보체들", "상보적" 및 "상보성"은 소정의 서열에 상보적이고 하이브리드화될 수 있는 서열을 지칭한다. 일부 경우, 소정의 핵산과 하이브리드화된 서열은 예를 들어, A-T, A-U, G-C 및 G-U 염기쌍이 형성되도록 소정의 영역에 걸쳐 그의 염기 서열이 그의 결합 파트너의 염기 서열에 상보적으로 결합할 수 있는 경우 소정의 분자의 "상보체" 또는 "역상보체"로서 지칭된다. 일반적으로, 제2 서열에 하이브리드화될 수 있는 제1 서열은 제2 서열에 특이적 또는 선택적으로 하이브리드화될 수 있으므로, 하이브리드화 반응 동안 제2 서열 또는 제2 서열 세트에의 하이브리드화는 비-표적 서열에의 하이브리드화보다 더 선호된다(예를 들어, 소정의 조건 세트, 예컨대, 당분야에서 통상적으로 사용되는 엄격한 조건 하에서 열역학적으로 더 안정하다). 전형적으로, 하이브리드화될 수 있는 서열들은 이들 각각의 길이의 전체 또는 일부에 걸쳐 어느 정도의 서열 상보성, 예컨대, 적어도 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 및 100% 서열 상보성을 포함하는 25% 내지 100% 상보성을 공유한다. 예컨대, 퍼센트 상보성을 평가하기 위한 서열 동일성은 니들만-분슈(Needleman-Wunsch) 알고리즘(예를 들어, 웹사이트(www.ebi.ac.uk/Tools/psa/embossneedle/nucleotide.html)에서 이용될 수 있는 EMBOSS Needle 정렬기 참조), BLAST 알고리즘(예를 들어, 임의로 디폴트 설정을 가진, 웹사이트(blast.ncbi.nlm.nih.gov/Blast.cgi)에서 이용될 수 있는 BLAST 정렬 수단 참조), 또는 스미스-워터만(Smith-Waterman) 알고리즘(예를 들어, 임의로 디폴트 설정을 가진, 웹사이트(www.ebi.ac.ukaools/psa/emboss_water/nucleotide.html)에서 이용될 수 있는 EMBOSS Water 정렬기 참조)을 포함하나 이들로 제한되지 않는 임의의 적합한 정렬 알고리즘에 의해 측정될 수 있다. 최적 정렬은 디폴트 파라미터를 포함하는, 선택된 알고리즘의 임의의 적합한 파라미터를 사용함으로써 평가될 수 있다.

[0184] 상보성은 완벽할 수 있거나 상당/충분할 수 있다. 두 핵산들 사이의 완벽한 상보성은 두 핵산들이 이중체를 형성할 수 있고 이중체의 모든 염기가 왓슨-크릭(Watson-Crick) 페어링에 의해 상보적 염기에 결합되어 있음을 의미할 수 있다. 상당한 또는 충분한 상보성은 한 가닥의 서열이 반대 가닥의 서열에 완전히 및/또는 완벽하게 상보적이지 않으나, 두 가닥의 염기들 사이에 충분한 결합이 일어나서 하이브리드화 조건 세트(예를 들어, 염 농도 및 온도)에서 안정한 하이브리드 복합체를 형성함을 의미할 수 있다. 이러한 조건은 서열과 표준 수학적 계산을 이용하여 하이브리드화된 가닥의 용융 온도(T_m)를 예측함으로써 예측될 수 있거나, 관용적인 방법을 이용하여 T_m을 실험적으로 측정함으로써 예측될 수 있다.

[0185] 본원에서 사용된 "트랜스포존"은 "점핑 유전자"로서도 공지되어 있는, 게놈 내에서 전위할 수 있는 염색체 내의 분절이다. 두 가지 상이한 클래스의 트랜스포존이 있다: RNA 중간체 및 "복사 및 붙여넣기" 기작을 통해 이동하는 클래스 1 또는 레트로트랜스포존, 및 절단 통합 또는 "절단 및 붙여넣기" 기작을 통해 이동하는 클래스 II 또는 DNA 트랜스포존(Ivics Nat Methods 2009). 박테리아, 하등 진핵생물(예를 들어, 효모) 및 무척추동물 트랜스포존은 대체로 종 특이적인 것으로 보이며, 척추동물 세포에서 DNA의 효율적인 전위를 위해 사용될 수 없다. "슬리핑 뷰티(Sleeping Beauty)"(Ivics Cell 1997)는 어류의 불활성 TE의 서열 셔플링에 의해 인위적으로 재구성된 첫 번째 활성 트랜스포존이었다. 이것은 인간 세포를 포함하는 척추동물 세포 내로의 전위에 의한 DNA 통합을 성공적으로 달성하는 것을 가능하게 만들었다. 슬리핑 뷰티는 트랜스포존의 Tc1/마리너(mariner) 패밀리에 속하는 클래스 II DNA 트랜스포존이다(Ni Genomics Proteomics 2008). 한편, 추가 기능성 트랜스포존은 드로소필라(Drosophila), 개구리 및 심지어 인간 게놈을 포함하는 상이한 종으로부터 확인되었거나 재구성되었고, 이들 모두가 척추동물 및 인간 숙주 세포 게놈 내로의 DNA 전위를 허용하는 것으로 밝혀졌다. 이들 트랜스포존 각각은 전위 효율, 발현의 안정성, 유전적 페이로드 용량 등과 관련된 장점 및 단점을 가진다. 생성된 예시적인 클래스 II 전위효소는 슬리핑 뷰티, PiggyBac, 프로그 프린스(Frog Prince), Himar1, 패스포트(Passport), 미

노스(Minos), hAT, Tol1, Tol2, AciDs, PIF, 하빙거(Harbinger), 하빙거3-DR 및 Hsmar1을 포함한다.

[0186] 본원에서 사용된 "이종"은 분자가 삽입되는 세포에서 천연적으로 발견되지 않을 수 있는 DNA 및 RNA와 같은 분자를 포함한다. 예를 들어, 마우스 또는 박테리아 DNA가 인간 세포의 게놈 내로 삽입될 때, 이러한 DNA는 본원에서 이종 DNA로서 지칭된다. 대조적으로, 본원에서 사용된 용어 "동종"은 분자가 삽입되는 세포에서 천연적으로 발견되는 DNA 및 RNA와 같은 분자를 의미한다. 예를 들어, 마우스 DNA를 마우스 세포의 게놈 내로 삽입하는 것은 동종 DNA를 그 세포 내로 삽입하는 것을 구성한다. 후자의 경우, 동종 DNA는 그 자신이 천연적으로 발견되는 세포 게놈의 부위 내로 삽입될 필요가 없고; 오히려, 동종 DNA는 그 자신이 천연적으로 발견되는 부위 이외의 부위에서 삽입됨으로써, 삽입된 부위에서 유전적 변경(돌연변이)을 생성할 수 있다.

[0187] "전위효소"는 트랜스포존 말단 함유 조성물(예를 들어, 트랜스포존, 트랜스포존 말단)과 함께 기능성 복합체를 형성할 수 있고 시험관내 트랜스포존 반응과 함께 인큐베이션되는 이중 가닥 DNA 내로의 트랜스포존 말단 함유 조성물의 삽입 또는 전위를 촉매작용하는 효소이다. 용어 "트랜스포존 말단"은 시험관내 전위 반응에서 작용하는 전위효소 또는 통합효소(integrase) 효소와 함께 복합체를 형성하는 데 필요한 뉴클레오타이드 서열("트랜스포존 말단 서열")을 함유하는 이중 가닥 DNA를 의미한다.

[0188] 트랜스포존 말단은 트랜스포존 말단을 인식하고 이에 결합하는 전위효소 또는 통합효소와 함께 복합체 또는 시냅스 복합체 또는 트랜스포존 복합체 또는 트랜스포존 조성물을 형성하고, 이 복합체는 트랜스포존 말단을 시험관내 전위 반응에서 그와 함께 인큐베이션되는 표적 DNA 내로 삽입할 수 있거나 전위시킬 수 있다. 트랜스포존 말단은 전달된 트랜스포존 말단 서열 또는 전달된 가닥 및 전달되지 않는 트랜스포존 말단 서열 또는 전달되지 않는 가닥으로 구성된 2개의 상보적 서열을 나타낸다. 예를 들어, 시험관내 전위 반응에서 활성을 나타내는 과활성 Tn5 전위효소와 함께 복합체를 형성하는 하나의 트랜스포존 말단은 전달된 트랜스포존 말단 서열 5' AGATGTGTATAAGAGACAG 3'(서열번호 55)를 나타내는 전달된 가닥, 및 "전달되지 않는 트랜스포존 말단 서열" 5' CTGTCTCTTATACACATCT 3'(서열번호 56)을 나타내는 전달되지 않는 가닥을 포함한다. 전달된 가닥의 3' 말단은 시험관내 전위 반응에서 표적 DNA에 연결되거나 전달된다. 전달된 트랜스포존 말단 서열에 상보적인 트랜스포존 말단 서열을 나타내는 전달되지 않는 가닥은 시험관내 전위 반응에서 표적 DNA에 연결되지 않거나 전달되지 않는다.

[0189] 일부 실시양태에서, 전달된 가닥과 전달되지 않는 가닥은 공유결합된다. 예를 들어, 일부 실시양태에서, 전달된 가닥 서열과 전달되지 않는 가닥 서열은 예를 들어, 헤어핀 구조로 단일 올리고뉴클레오타이드에 제공된다. 따라서, 전달되지 않는 가닥의 자유 말단은 전위 반응에 의해 직접적으로 표적 DNA에 연결되지 않지만, 전달되지 않는 가닥이 헤어핀 구조의 루프에 의해 전달된 가닥에 연결되기 때문에, 전달되지 않는 가닥은 간접적으로 DNA 단편에 부착되게 된다. 본원에서 사용된 바와 같이, "절단 도메인"은 작용제, 예를 들어, 효소에 의한 절단에 민감한 핵산 서열을 의미한다.

[0190] "제한 부위 도메인"은 제한 엔도뉴클레아제를 사용한 절단을 용이하게 하기 위한 서열을 나타내는 태그 도메인을 의미한다. 예를 들어, 일부 실시양태에서, 제한 부위 도메인은 2개의 태그가 부착된 선형 ssDNA 단편을 생성하는 데 사용된다. 일부 실시양태에서, 제한 부위 도메인은 이중 가닥 5' 말단이 주형 의존적 DNA 리가제(ligase)의 사용에 의해 또 다른 DNA 분자에 라이게이션될 수 있도록 태그 도메인에서 호환 가능한 이중 가닥 5' 말단을 생성하는 데 사용된다. 일부 실시양태에서, 태그 내의 제한 부위 도메인은 존재한다 하더라도 표적 DNA에 거의 존재하지 않는 제한 부위(예를 들어, NotI 또는 AscI과 같은 희귀 절단 제한 엔도뉴클레아제에 대한 제한 부위)의 서열을 나타낸다.

[0191] 본원에서 사용된 바와 같이, 용어 "재조합 핵산 분자"는 재조합 DNA 분자 또는 재조합 RNA 분자를 지칭한다. 재조합 핵산 분자는 상이한 최초 공급원으로부터 유래하고 천연적으로 함께 부착되지 않는 연결된 핵산 분자들을 함유하는 임의의 핵산 분자이다. 재조합 RNA 분자는 재조합 DNA 분자로부터 전사된 RNA 분자를 포함한다. 재조합 핵산은 실험실에서 합성될 수 있다. 재조합 핵산은 효소적 제한 분해, 라이게이션 및 DNA 클로닝과 같은 DNA의 효소적 변형을 이용함으로써 재조합 DNA 기술의 이용에 의해 제조될 수 있다. 재조합 DNA는 시험관내에서 전사되어 메신저 RNA(mRNA)를 생성할 수 있고, 재조합 mRNA는 단리 및 정제될 수 있고 세포를 형질감염시키는 데 사용될 수 있다. 재조합 핵산은 단백질 또는 폴리펩티드를 코딩할 수 있다. 적합한 조건 하에서 재조합 핵산은 살아있는 세포 내로 도입될 수 있고, 살아있는 세포 내부에서 발현될 수 있다. 본원에서 사용된 바와 같이, 핵산의 "발현"은 일반적으로 핵산의 전사 및/또는 번역을 의미한다. 핵산 발현의 생성물은 일반적으로 단백질이나, mRNA일 수도 있다. 재조합 핵산이 도입된 세포에서 재조합 핵산에 의해 코딩된 mRNA의 검출은 상기 핵산이 상기 세포에서 "발현"된다는 양성 증거로서 간주된다. 핵산을 세포 내로 삽입하거나 도입하는 과정은 형

질전환, 형질감염 또는 형질도입을 통해 이루어질 수 있다. 형질전환은 박테리아 세포에 의한 외래 핵산의 흡수 과정이다. 이 과정은 플라스미드 DNA의 증식, 단백질 생성 및 기타 적용을 위해 개조된다. 형질전환은 재조합 플라스미드 DNA를, 환경으로부터 세포외 DNA를 흡수하는 형질전환능(competent) 박테리아 세포 내로 도입한다. 일부 박테리아 종은 특정 환경 조건 하에서 천연적으로 형질전환능을 갖지만, 형질전환능은 실험실 환경에서 인위적으로 유도된다. 형질감염은 DNA, RNA 또는 항체와 같은 소분자를 진핵 세포 내로 강제 도입하는 것이다. 다소 혼동되지만, '형질감염'은 박테리아 세포 내로의 박테리오파지의 도입도 지칭한다. '형질도입'은 재조합 바이러스 벡터 입자를 표적 세포 내로 도입하는 것을 설명하는 데 주로 사용되는 반면, '감염'은 야생형 바이러스에 의한 인간 또는 동물의 천연 감염을 지칭한다.

[0192] "줄기-루프(stem-loop)" 서열은 하이브리드화하여, 루프로 돌출되는 비-상보성 영역과 줄기를 형성하기에 충분한 자가상보성을 가진 핵산 서열(예를 들어, RNA 서열)을 지칭한다. 줄기는 불일치 또는 돌출부를 포함할 수 있다.

[0193] 용어 "벡터"는 이종 핵산을 운반하거나 이종 핵산의 발현을 매개할 수 있는 핵산 분자를 지칭한다. 본원에서 사용된 "벡터 서열"은 적어도 하나의 복제 기점 및 적어도 하나의 선택 마커 유전자를 포함하는 핵산 서열을 지칭한다. 그 자신에 작동 가능하게 연결된 유전자 및/또는 핵산 서열의 발현을 유도할 수 있는 벡터는 본원에서 "발현 벡터"로서 지칭된다.

[0194] 플라스미드는 용어 "벡터"에 포함되는 속의 종이다. 일반적으로, 유용한 발현 벡터는 종종 "플라스미드"의 형태로 존재하고, 이 플라스미드는 그의 벡터 형태에서 염색체에 결합되지 않는 원형 이중 가닥 DNA 분자를 지칭하고 전형적으로 코딩된 DNA의 안정한 또는 일시적인 발현을 위한 독립체를 포함한다. 본원에 개시된 방법에 사용될 수 있는 다른 발현 벡터는 플라스미드, 에피좀, 박테리아 인공 염색체, 효모 인공 염색체, 박테리오파지 또는 바이러스 벡터를 포함하나 이들로 제한되지 않으며, 이러한 벡터는 숙주의 게놈 내로 통합될 수 있거나 세포에서 자율적으로 복제할 수 있다. 벡터는 DNA 또는 RNA 벡터일 수 있다. 동등한 기능을 제공하는, 당업자에 의해 공지된 다른 형태의 발현 벡터, 예를 들어, 자가복제 염색체외 벡터 또는 숙주 게놈 내로 통합될 수 있는 벡터도 사용될 수 있다. 예시적인 벡터는 벡터에 연결된 핵산의 자율 복제 및/또는 발현을 할 수 있는 벡터이다. 안전 피난처 유전자좌는 추가 외생성 또는 이종 핵산 서열이 삽입될 수 있는 게놈 내의 영역이고, 숙주 게놈은 삽입된 유전 물질을 수용할 수 있다. 예시적인 안전 피난처 부위는 AAVS1 부위, GGTA1 부위, CMAH 부위, B4GALNT2 부위, B2M 부위, ROSA26 부위, COLA1 부위 및 TIGRE 부위를 포함하나, 이들로 제한되지 않는다. 예를 들어, 본 개시내용에 기재된 이종 핵산은 세포의 게놈 내의 하나 이상의 부위에서 통합될 수 있고, 이때 하나 이상의 부위는 AAVS1 부위, GGTA1 부위, CMAH 부위, B4GALNT2 부위, B2M 부위, ROSA26 부위, COLA1 부위 및 TIGRE 부위로 구성된 군으로부터 선택된다. 일부 실시양태에서, 형질전이유전자를 포함하는 핵산 카고는 R2D 유전자좌로 전달될 수 있다.

[0195] 일부 실시양태에서, 형질전이유전자를 포함하는 핵산 카고는 유전자간 또는 유전자내 영역의 게놈 내로 전달될 수 있다. 일부 실시양태에서, 형질전이유전자를 포함하는 핵산 카고는 내생성 활성 유전자의 0.1 kb, 0.25 kb, 0.5 kb, 0.75 kb, 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 7.5 kb, 10 kb, 15 kb, 20 kb, 25 kb, 50 kb, 75 kb 또는 100 kb 내에서 게놈 5' 또는 3' 내로 통합된다. 일부 실시양태에서, 형질전이유전자를 포함하는 핵산 카고는 내생성 프로모터 또는 인핸서의 0.1 kb, 0.25 kb, 0.5 kb, 0.75 kb, 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 7.5 kb, 10 kb, 15 kb, 20 kb, 25 kb, 50 kb, 75 kb 또는 100 kb 내에서 게놈 5' 또는 3' 내로 통합된다. 일부 실시양태에서, 형질전이유전자를 포함하는 핵산 카고는 50 내지 50,000 염기쌍(bp), 예를 들어, 50 내지 40,000 bp, 500 내지 30,000 bp, 500 내지 20,000 bp, 100 내지 15,000 bp, 500 내지 10,000 bp, 50 내지 10,000 bp, 50 내지 5,000 bp이다. 일부 실시양태에서, 형질전이유전자를 포함하는 핵산 카고는 길이가 1,000개, 1,300개, 1500개, 2,000개, 3,000개, 4,000개, 5,000개 또는 7,500개의 뉴클레오타이드 미만이다.

[0196] L1 레트로트랜스포존 시스템 및 비-L1 레트로트랜스포존 시스템

[0197] 레트로트랜스포존은 그가 존재하는 게놈을 재구성하는 데 있어서 능동적으로 참여하는 전위 요소를 함유할 수 있다. 광범위하게, 레트로트랜스포존은 RNA로 전사되고 단백질로 번역되고 그 자신을 DNA로 다시 역전사하는 능력을 가진 DNA 서열을 지칭할 수 있다. 인간 게놈의 대략 45%는 전위 이벤트로부터 비롯된 서열로 구성된다. 역전위는 종종 표적 부위 결실을 생성하거나 5'-형질도입 및 3'-형질도입으로서 지칭되는 과정으로 비-레트로트랜스포존 DNA를 게놈에 추가한다. 비-동종 레트로트랜스포존 사이의 재조합은 유전자 서열의 결실, 중복 또는 재배열을 야기한다. 지속적인 역전위는 신규 스플라이스 부위, 폴리아데닐화 신호 및 프로모터를 생성할 수 있으므로, 새로운 전사 모듈을 구축한다.

[0198] 일반적으로, 레트로트랜스포존은 두 클래스, 즉 레트로바이러스 유사 LTR 레트로트랜스포존과, 비-LTR 요소, 예컨대, 인간 L1 요소, 뉴로스포라(*Neurospora*) TAD 요소(Kinsey, 1990, Genetics 126:317-326), 드로소필라의 I 인자(Bucheton et al., 1984, Cell 38:153-163) 및 봄빅스 모리(*Bombyx mori*)의 R2Bm(Luan et al., 1993, Cell 72: 595-605)으로 분류될 수 있다. 이 두 가지 유형의 레트로트랜스포존은 구조적으로 상이하고 근본적으로 상이한 기작을 이용하여 역전위한다. LINE-코딩된 폴리펩티드의 예시적인 비제한적 예는 진뱅크(GenBank) 수탁번호 AAC51261, AAC51261, AAC51263, AAC51264, AAC51265, AAC51266, AAC51267, AAC51268, AAC51269, AAC51270, AAC51271, AAC51272, AAC51273, AAC51274, AAC51275, AAC51276, AAC51277, AAC51278 및 AAC51279에서 발견된다.

[0199] 여러 가지 이유로 본 개시내용에 기재된 시스템으로 개발하기 위해 LINE-1에 초점을 맞추기로 결정하였고, 이 이유 중 적어도 일부는 이하에 예시되어 있다: (a) LINE-1(또는 L1-) 요소는 이 역전사 및 통합 과정을 완료하기 위해 단독으로 기구의 전부를 코딩하기 때문에 자율적이고; (b) L1 요소는 이 요소가 게놈의 천연화된 요소로서 간주될 수 있을 정도로 인간 게놈에 풍부하고; (c) L1 레트로트랜스포존은 세포에서 주변을 떠다니는 다른 mRNA에 비해 고도의 특이성으로 그 자신의 mRNA를 역전위한다.

[0200] L1은 필수적이나 불확실한 기능의 40 kDa 오픈 리딩 프레임-1 RNA 결합 단백질(ORF1p), 및 엔도뉴클레아제 및 역전사효소(RT) 활성을 가진 150 kDa ORF2 단백질을 코딩하는 6 kb 바이시스트론 RNA를 발현한다. L1 역전위는 통합 부위에서 L1의 전사, 세포질로의 그의 RNA의 수송, 바이시스트론 RNA의 번역, 리보핵단백질(RNP) 입자의 형성, 핵으로의 그의 재이입 및 표적 프라이밍 역전사를 수반하는 복잡한 과정이다. L1과 상호작용하는 몇몇 전사 인자가 확인되었다. 전사된 L1 RNA는 전사체로부터 번역된 단백질과 시스로 RNP를 형성한다. L1은 5'-TTTT-3'에서 ORF2p 절단에 의한 표적 부위 프라이머 역전사(TPRT)에 의해 게놈 DNA 내로 통합되고, 이때 L1 RNA의 폴리 A 서열은 어닐링되고 역전사효소(RT) 활성을 프라이밍하여 L1 cDNA를 만든다.

[0201] 게놈의 다른 이동 요소는 역전위를 위해 L1 ORF를 "하이재킹"할 수 있다. 예를 들어, *Alu* 요소는 비-자율적 레트로트랜스포존이고 통합을 위해 트랜스 인자를 획득하는 짧은 산재된 요소(SINE)의 클래스에 속하는 이동 DNA 요소이다. Alu 요소 및 SINE-1 요소는 ORF1p 및 ORF2p에 의해 역전위되기 위해 트랜스로 L1 리보핵단백질과 회합할 수 있다. L1 RNA와 다소 유사하게, *Alu* 요소는 종종 A-꼬리로서도 지칭되는 긴 A-런(run)으로 끝나고, 분기된 이량체 구조의 두 절반을 분리하는 더 작은 A 풍부 영역(AA로 표시됨)도 가진다. *Alu* 요소는 RNA 중합효소 III 프로모터(예컨대, 통상적으로 A 박스 및 B 박스 프로모터로서 명명됨)의 내부 구성요소를 가질 가능성이 높으나, RNA 중합효소 III에 대한 터미네이터를 코딩하지 않는다. 이 요소는 전사를 종결하기 위해 *Alu* 요소의 하류에서 다양한 거리에 있는 T 뉴클레오타이드의 스트레치를 사용할 수 있다. 전형적인 *Alu* 전사체는 A-꼬리를 포함하는 전체 *Alu*를 포함하고, 각각의 유전자좌에 대해 고유한 3' 영역을 가진다. *Alu* RNA는 각각의 단량체 유닛에 대해 별도의 구조로 접힌다. RNA는 7SL RNA SRP9 및 14 이종이량체뿐만 아니라 폴리 A 결합 단백질(PABP)에도 결합하는 것으로 밝혀졌다. Alu의 폴리 A 꼬리는 게놈의 T 풍부(TTTT) 영역으로 프라이밍하고 프라이밍된 영역에 결합하기 위해 ORF2p를 유인하여, 그의 엔도뉴클레아제 활성을 통해 T 풍부 영역에서 절단한다. T 풍부 영역은 *Alu* 요소의 3' A-꼬리 영역에서 ORF2p로 역전사를 프라이밍한다. 이것은 *Alu* 요소 본체의 cDNA 카피를 생성한다. 닉(nick)은 제2 가닥에서 비공지된 기작에 의해 일어나고 제2 가닥 합성은 프라이밍된다. 그 다음, 새로운 *Alu* 요소는 제1 닉과 제2 닉 사이의 DNA 서열의 중복체인 짧은 직접 반복부에 의해 플랭킹된다. *Alu* 요소는 유전자 풍부 영역을 선호하기 때문에 RNA 분자 내에서 매우 널리 퍼져 있다. 전체 길이 *Alu*(약 300 bp)는 신호 인식 입자 RNA 7SL로부터 유래하고, A 풍부 링커가 중간에 위치된 2개의 유사한 단량체, 5' 단량체에 존재하는 A 박스와 B 박스, 및 앞서는 폴리아데닐화 신호를 결여하여 연장된 꼬리(길이가 최대 100 bp)를 야기하는 폴리 A 꼬리로 구성된다. *Alu*는 A 박스 및 B 박스 내의 내부 프로모터를 사용함으로써 RNA 중합효소 III에 의해 전사될 수 있으나; *Alu*는 ORF를 함유하지 않으므로 단백질 생성물을 코딩하지 않는다.

[0202] 다른 비-L1 트랜스포존은 SVA 및 HERV-K를 포함한다. 전체 길이 SVA(SINE-VNTR-*Alu*) 요소(약 2 내지 3 kb)는 CCCTCT 반복부, 2개의 *Alu* 유사 서열, VNTR, *env*(외피) 유전자를 가진 SINE-R 영역, HERV-K10의 3' LTR, 및 폴리 A 꼬리가 뒤따르는 폴리아데닐화 신호를 함유하는 복합 유닛이다. SVA 요소가 내부 프로모터를 운반하는지 여부는 알려져 있지 않지만, SVA가 RNA 중합효소 II에 의해 전사될 가능성은 가장 높다.

[0203] 전체 길이 HERV-K 요소(약 9 내지 10 kb)는 내생성 레트로바이러스 서열의 고대 잔해물로 구성되며 3개의 레트로바이러스 ORF를 둘러싸는 2개의 플랭킹 LTR 영역을 포함한다: (1) 레트로바이러스 캡시드의 구조 단백질을 코딩하는 *gag*; (2) 효소를 코딩하는 *pol-pro*: 프로테아제, RT 및 통합효소; 및 (3) 수평 전달을 허용하는 단백질을 코딩하는 *env*. HERV-K의 LTR은 RNA 중합효소 II의 전사 조절 하에 있는 것으로 보이는 내부 양방향 프로모터

를 함유한다.

[0204] L1 역전위 및 RNA 결합은 폴리 A 꼬리에서 또는 이 근처에서 일어날 수 있다. 3' UTR은 ORF1 단백질(ORF1p)의 엄격한 유형의 LINE RNA를 인식하는 역할을 한다. 엄격한 유형의 LINE은 3' UTR의 말단에 위치한 줄기-루프 구조를 함유할 수 있다. 트랜스포존 3' 말단 cDNA와 표적 DNA 사이의 연접, 및 ORF2 단백질(ORF2p) 내에서의 L1 RNA의 특이적 위치선정으로 구성된 분지된 분자는 시험관내에서 L1 역전위의 초기 단계 동안 검출되었다. L1과 Alu가 공유하는 2차 또는 3차 RNA 구조는 아마도 폴리 A 꼬리와 함께 ORF2에 의한 인식 및 ORF2의 결합을 담당할 가능성이 높다. 일부 실시양태에서, 폴리 A 서열의 하류에 위치한 줄기-루프 구조는 절단 강도와 상관관계를 가진다.

[0205] L1 통합을 제한하거나 해소하는 기작도 게놈의 유전적 무결성 및 안정성을 유지하기 위해 진화되었다. XRCC1, Ku70 및 DNA-PK와 같은 비-동종 말단 연결 복구 단백질은 삽입 시 L1 통합의 해소에 관여한다. 또한, 세포는 APOBEC3 패밀리의 사이토신 데아미나제, 아데노신 데아미나제 ADAR1, 염색질 리모델링 인자, 및 남성 생식세포주에서 작용하는 전사 후 유전자 침묵을 위한 piRNA 경로의 구성원을 포함하는, 제한되지 않은 역전위에 저항하는 다수의 단백질들을 진화시켰다.

[0206] **I. 코딩된 단백질의 안정한 발현을 위한 핵산 구축물을 포함하는 조성물 및 관련 방법**

[0207] 본원은 골수 세포와 같은 세포에서 발현될 하나 이상의 단백질을 코딩하는 재조합 핵산을 제공한다. 한 실시양태에서, 재조합 핵산은 재조합 핵산에 의해 코딩된 하나 이상의 단백질 또는 폴리펩티드를 안정적으로 발현하도록 디자인된다. 일부 실시양태에서, 안정한 발현은 세포의 게놈 내로의 재조합 핵산의 통합에 의해 달성된다.

[0208] 당업자는 본원에 기재된 조성물 및 방법을 이용하여 생성물을 디자인할 수 있고, 이때 재조합 핵산이 단백질 또는 폴리펩티드 성분으로서 번역되지 않는 하나 이상의 서열을 포함할 수 있으나, 조절 핵산, 예컨대, 억제제 올리고뉴클레오타이드 생성물, 예컨대, 활성화제 올리고뉴클레오타이드일 수 있는 올리고뉴클레오타이드를 코딩할 수 있음을 용이하게 이해할 수 있다.

[0209] 한 측면에서, 본원은 관심 있는 유전자를 코딩하는 핵산 서열, 및 비-내생성 핵산을 세포 내로 안정적으로 도입하는 하나 이상의 역전위 요소를 포함하는 합성 핵산을 포함하는 조성물을 제공한다. 일부 실시양태에서, 세포는 조혈 세포이다. 일부 실시양태에서, 세포는 골수 세포이다. 일부 실시양태에서, 세포는 전구체 세포이다. 일부 실시양태에서, 세포는 미분화된 세포이다. 일부 실시양태에서, 세포는 추가 분화력을 가진다. 일부 실시양태에서, 세포는 줄기 세포가 아니다.

[0210] A. LINE/Alu 레트로트랜스포존 구축물

[0211] 일부 실시양태에서, 본 개시내용은 역전위 시스템을 이용하여 비-내생성 핵산을 게놈 내로 안정적으로 통합시키고 발현시킬 수 있고, 이때 상기 비-내생성 핵산은 핵산 서열 내에 역전위 요소를 포함한다. 일부 실시양태에서, 본 개시내용은 세포의 내생성 역전위 시스템(예를 들어, 단백질 및 효소)을 이용하여, 세포에서 비-내생성 핵산을 안정적으로 발현시킬 수 있다. 일부 실시양태에서, 본 개시내용은 세포의 내생성 역전위 시스템(예를 들어, 단백질 및 효소, 예컨대, LINE1 역전위 시스템)을 이용할 수 있으나, 세포에서 비-내생성 핵산을 안정적으로 발현시키기 위해 역전위 시스템의 하나 이상의 성분을 추가로 발현시킬 수 있다.

[0212] 일부 실시양태에서, 본원은 형질전이유전자를 코딩하고 역전위를 위한 하나 이상의 성분을 코딩하는 합성 핵산을 제공한다. 본원에 기재된 합성 핵산은 핵산 구축물, 형질전이유전자 또는 외생성 핵산으로서 교환 가능하게 지칭된다.

[0213] 한 측면에서, 본원은 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 세포 내로 도입하는 단계를 포함하는, 핵산 서열을 세포의 게놈 내로 통합시키는 방법을 제공하는 것으로, 이때 mRNA는 외생성 서열, 또는 외생성 서열의 역상보체인 서열을 포함하는 삽입체 서열; 5' UTR 서열, 및 5' UTR 서열의 하류에 있는 3' UTR 서열을 포함하고, 이때 5' UTR 서열 또는 3' UTR 서열은 인간 ORF 단백질에 대한 결합 부위를 포함하고, 삽입체 서열은 세포의 게놈 내로 통합된다.

[0214] 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 인간 ORF2p에 대한 결합 부위를 포함한다.

[0215] 한 측면에서, 본원은 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 도입하는 단계를 포함하는, 핵산 서열을 면역세포의 게놈 내로 통합시키는 방법을 제공하는 것으로, 이때 mRNA는 (i) 외생성 서열 또는 (ii) 외생성 서열의 역상보체인 서열을 포함하는 삽입체 서열; 5' UTR 서열, 및 5' UTR 서열의 하류에 있는 3' UTR 서열을 포함하고, 이때 5' UTR 서열 또는 3' UTR 서열은 엔도뉴클레아제 결합 부위 및/또는 역전사효소 결합 부위를 포

함하고, 형질전이유전자 서열은 면역 세포의 게놈 내로 통합된다.

[0216] 한 측면에서, 본원은 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 도입하는 단계를 포함하는, 핵산 서열을 세포의 게놈 내로 통합시키는 방법을 제공하는 것으로, 이때 mRNA는 (i) 외생성 서열 또는 (ii) 외생성 서열의 역상보체인 서열을 포함하는 삽입체 서열; 5' UTR 서열, 5' UTR 서열의 하류에 있는 인간 레트로트랜스포존의 서열, 및 인간 레트로트랜스포존의 서열의 하류에 있는 3' UTR 서열을 포함하고, 이때 5' UTR 서열 또는 3' UTR 서열은 엔도뉴클레아제 결합 부위 및/또는 역전사효소 결합 부위를 포함하고, 인간 레트로트랜스포존의 서열은 2개의 ORF를 함유하는 단일 RNA로부터 번역되는 2개의 단백질을 코딩하고, 삽입체 서열은 세포의 게놈 내로 통합된다.

[0217] 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 ORF2p 결합 부위를 포함한다. 일부 실시양태에서, ORF2p 결합 부위는 3' UTR 서열 내의 폴리 A 서열이다.

[0218] 일부 실시양태에서, mRNA는 인간 레트로트랜스포존의 서열을 포함한다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 5' UTR 서열의 하류에 있다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 3' UTR 서열의 상류에 있다.

[0219] 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 2개의 ORF를 함유하는 단일 RNA로부터 번역되는 2개의 단백질을 코딩한다. 일부 실시양태에서, 2개의 ORF는 비-중첩 ORF이다. 일부 실시양태에서, 2개의 ORF는 ORF1 및 ORF2이다. 일부 실시양태에서, ORF1은 ORF1p를 코딩하고, ORF2는 ORF2p를 코딩한다.

[0220] 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 비-LTR 레트로트랜스포존의 서열을 포함한다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 LINE-1 레트로트랜스포존을 포함한다. 일부 실시양태에서, LINE-1 레트로트랜스포존은 인간 LINE-1 레트로트랜스포존이다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 서열을 포함한다. 일부 실시양태에서, 엔도뉴클레아제 및/또는 역전사효소는 ORF2p이다. 일부 실시양태에서, 역전사효소는 II 군 인트론 역전사효소 도메인이다. 일부 실시양태에서, 엔도뉴클레아제 및/또는 역전사효소는 밍크 고래 엔도뉴클레아제 및/또는 역전사효소이다. 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 ORF2p를 코딩하는 서열을 포함한다. 일부 실시양태에서, 삽입체 서열은 ORF2p의 엔도뉴클레아제 도메인의 특이성을 이용함으로써 폴리 T 부위에서 게놈 내로 통합된다. 일부 실시양태에서, 폴리 T 부위는 서열 TTTTTA를 포함한다.

[0221] 일부 실시양태에서, (i) 인간 레트로트랜스포존의 서열은 ORF1p를 코딩하는 서열을 포함하거나, (ii) mRNA는 ORF1p를 코딩하는 서열을 포함하지 않거나, (iii) mRNA는 ORF1p를 코딩하는 서열을 상보체 유전자의 5' UTR 서열로 대체하는 것을 포함한다. 일부 실시양태에서, mRNA는 ORF1p를 코딩하는 제1 mRNA 분자, 및 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 mRNA 분자를 포함한다. 일부 실시양태에서, mRNA는 ORF1p를 코딩하는 제1 서열, 및 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 서열을 포함하는 mRNA 분자이다. 일부 실시양태에서, ORF1p를 코딩하는 제1 서열과 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 서열은 링커 서열에 의해 분리된다.

[0222] 일부 실시양태에서, 링커 서열은 내부 리보좀 진입 서열(IRES)을 포함한다. 일부 실시양태에서, IRES는 CVB3 또는 EV71의 IRES이다. 일부 실시양태에서, 링커 서열은 자가절단 펩티드 서열을 코딩한다. 일부 실시양태에서, 링커 서열은 T2A, E2A 또는 P2A 서열을 코딩한다.

[0223] 일부 실시양태에서, 인간 레트로트랜스포존의 서열은 추가 단백질 서열에 융합된 ORF1p를 코딩하는 서열 및/또는 추가 단백질 서열에 융합된 ORF2p를 코딩하는 서열을 포함한다. 일부 실시양태에서, ORF1p 및/또는 ORF2p는 핵 체류 서열에 융합된다. 일부 실시양태에서, 핵 체류 서열은 Alu 서열이다. 일부 실시양태에서, ORF1p 및/또는 ORF2p는 MS2 코트 단백질에 융합된다. 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 적어도 1개, 2개, 3개 이상의 MS2 헤어핀 서열을 포함한다. 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 mRNA의 폴리 A 꼬리와 엔도뉴클레아제 및/또는 역전사효소의 상호작용을 촉진하거나 향상시키는 서열을 포함한다. 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 폴리 A 결합 단백질(PABP)과 엔도뉴클레아제 및/또는 역전사효소의 상호작용을 촉진하거나 향상시키는 서열을 포함한다. 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 상기 mRNA에 대한 엔도뉴클레아제 및/또는 역전사효소의 특이성을 세포에 의해 발현된 또 다른 mRNA에 비해 증가시키는 서열을 포함한다. 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 Alu 요소 서열을 포함한다.

[0224] 일부 실시양태에서, ORF1p를 코딩하는 제1 서열 및 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 서열은 동일한 프로모터를 가진다. 일부 실시양태에서, 삽입체 서열은 ORF1p를 코딩하는 제1 서열의 프로모터와 상이한

프로모터를 가진다. 일부 실시양태에서, 삽입체 서열은 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 서열의 프로모터와 상이한 프로모터를 가진다. 일부 실시양태에서, ORF1p를 코딩하는 제1 서열, 및/또는 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 서열은 유도성 프로모터, CMV 프로모터 또는 전사 시작 부위, T7 프로모터 또는 전사 시작 부위, EF1a 프로모터 또는 전사 시작 부위 및 이들의 조합으로 구성된 군으로부터 선택된 프로모터 또는 전사 시작 부위를 가진다. 일부 실시양태에서, 삽입체 서열은 유도성 프로모터, CMV 프로모터 또는 전사 시작 부위, T7 프로모터 또는 전사 시작 부위, EF1a 프로모터 또는 전사 시작 부위 및 이들의 조합으로 구성된 군으로부터 선택된 프로모터 또는 전사 시작 부위를 가진다.

[0225] 일부 실시양태에서, ORF1p를 코딩하는 제1 서열 및 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 제2 서열은 인간 세포에서 발현되도록 코돈 최적화된다.

[0226] 일부 실시양태에서, mRNA는 WPRE 요소를 포함한다. 일부 실시양태에서, mRNA는 선택 마커를 포함한다. 일부 실시양태에서, mRNA는 친화성 태그를 코딩하는 서열을 포함한다. 일부 실시양태에서, 친화성 태그는 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 서열에 연결된다.

[0227] 일부 실시양태에서, 3' UTR은 폴리 A 서열을 포함하거나, 폴리 A 서열은 시험관내에서 mRNA에 추가된다. 일부 실시양태에서, 폴리 A 서열은 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 서열의 하류에 있다. 일부 실시양태에서, 삽입체 서열은 폴리 A 서열의 상류에 있다.

[0228] 일부 실시양태에서, 3' UTR 서열은 삽입체 서열을 포함한다. 일부 실시양태에서, 삽입체 서열은 외생성 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함한다. 일부 실시양태에서, 삽입체 서열은 폴리아데닐화 부위를 포함한다. 일부 실시양태에서, 삽입체 서열은 SV40 폴리아데닐화 부위를 포함한다. 일부 실시양태에서, 삽입체 서열은 외생성 폴리펩티드를 코딩하는 서열의 역상보체인 서열의 상류에 있는 폴리아데닐화 부위를 포함한다. 일부 실시양태에서, 삽입체 서열은 리보좀 유전자좌가 아닌 유전자좌에서 게놈 내로 통합된다. 일부 실시양태에서, 삽입체 서열은 유전자 또는 유전자의 조절 영역 내로 통합됨으로써, 유전자를 파괴하거나 유전자의 발현을 하향조절한다. 일부 실시양태에서, 삽입체 서열은 유전자 또는 유전자의 조절 영역 내로 통합됨으로써, 유전자의 발현을 상향조절한다. 일부 실시양태에서, 삽입체 서열은 게놈 내로 통합되고 유전자를 대체한다. 일부 실시양태에서, 삽입체 서열은 게놈 내로 안정적으로 통합된다. 일부 실시양태에서, 삽입체 서열은 게놈 내로 역전위된다. 일부 실시양태에서, 삽입체 서열은 mRNA에 의해 코딩된 엔도뉴클레아제에 의한 표적 부위의 DNA 가닥의 절단에 의해 게놈 내로 통합된다. 일부 실시양태에서, 삽입체 서열은 표적 프라이밍 역전사(TPRT)를 통해 게놈 내로 통합된다. 일부 실시양태에서, 삽입체 서열은 게놈의 DNA 표적 부위 내로의 mRNA의 역스플라이싱을 통해 게놈 내로 통합된다.

[0229] 일부 실시양태에서, 세포는 면역 세포이다. 일부 실시양태에서, 면역 세포는 T 세포 또는 B 세포이다. 일부 실시양태에서, 면역 세포는 골수 세포이다. 일부 실시양태에서, 면역 세포는 단핵구, 대식세포, 수지상 세포, 수지상 전구체 세포 및 대식세포 전구체 세포로 구성된 군으로부터 선택된다.

[0230] 일부 실시양태에서, mRNA는 자가통합 mRNA이다. 일부 실시양태에서, 방법은 mRNA를 세포 내로 도입하는 단계를 포함한다. 일부 실시양태에서, 방법은 mRNA를 코딩하는 벡터를 세포 내로 도입하는 단계를 포함한다. 일부 실시양태에서, 방법은 생체외에서 mRNA 또는 mRNA를 코딩하는 벡터를 세포 내로 도입하는 단계를 포함한다. 일부 실시양태에서, 방법은 세포를 인간 대상체에게 투여하는 단계도 포함한다. 일부 실시양태에서, 방법은 mRNA 또는 mRNA를 코딩하는 벡터를 인간 대상체에게 투여하는 단계를 포함한다. 일부 실시양태에서, 면역 반응은 인간 대상체에서 유발되지 않는다. 일부 실시양태에서, mRNA 또는 벡터는 실질적으로 면역원성을 나타내지 않는다.

[0231] 일부 실시양태에서, 벡터는 플라스미드 또는 바이러스 벡터이다. 일부 실시양태에서, 벡터는 비-LTR 레트로트랜스포존을 포함한다. 일부 실시양태에서, 벡터는 인간 L1 요소를 포함한다. 일부 실시양태에서, 벡터는 L1 레트로트랜스포존 ORF1 유전자를 포함한다. 일부 실시양태에서, 벡터는 L1 레트로트랜스포존 ORF2 유전자를 포함한다. 일부 실시양태에서, 벡터는 L1 레트로트랜스포존을 포함한다.

[0232] 일부 실시양태에서, mRNA는 적어도 약 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9 또는 3 킬로염기이다. 일부 실시양태에서, mRNA는 최대 약 2.5, 2.6, 2.7, 2.8, 2.9, 3, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9 또는 5 킬로염기이다.

[0233] 일부 실시양태에서, mRNA는 적어도 약 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9 또는 3 킬로염기인 페이로드를 포함한다. 일부 실시양태에서, mRNA는 최대 약

2.5, 2.6, 2.7, 2.8, 2.9, 3, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9 또는 5 킬로염기이다. 일부 실시양태에서, mRNA는 적어도 약 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9 또는 6 킬로염기이다. 일부 실시양태에서, mRNA는 적어도 약 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9 또는 7 킬로염기이다. 일부 실시양태에서, mRNA는 적어도 약 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9 또는 8 킬로염기이다. 일부 실시양태에서, mRNA는 적어도 약 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9 또는 9 킬로염기이다. 일부 실시양태에서, mRNA는 적어도 약 9.1, 9.2, 9.3, 9.4, 9.5, 9.6, 9.7, 9.8, 9.9 또는 10 킬로염기이다. 일부 실시양태에서, mRNA는 적어도 약 10.1, 10.2, 10.3, 10.4, 10.5, 10.6, 10.7, 10.8, 10.9 또는 11 킬로염기이다. 일부 실시양태에서, mRNA는 적어도 약 11.1, 11.2, 11.3, 11.4, 11.5, 11.6, 11.7, 11.8, 11.9 또는 12 킬로염기이다. 일부 실시양태에서, mRNA는 약 6.8 kB의 페이로드, 예를 들어, ABCA4 유전자 생성물을 코딩하는 서열을 포함한다. 일부 실시양태에서, mRNA는 약 6.7 kB의 페이로드, 예를 들어, MYO7A 유전자 생성물을 코딩하는 서열을 포함한다. 일부 실시양태에서, mRNA는 약 7.5 kB의 페이로드, 예를 들어, CEP290 유전자 생성물을 코딩하는 서열을 포함한다. 일부 실시양태에서, mRNA는 약 10.1 kB의 페이로드, 예를 들어, CDH23 유전자 생성물을 코딩하는 서열을 포함한다. 일부 실시양태에서, mRNA는 약 9.4 kB의 페이로드, 예를 들어, EYS 유전자 생성물을 코딩하는 서열을 포함한다. 일부 실시양태에서, mRNA는 약 15.6 kB의 페이로드, 예를 들어, USH2a 유전자 생성물을 코딩하는 서열을 포함한다. 일부 실시양태에서, mRNA는 약 12.5 kB의 페이로드, 예를 들어, ALMS1 유전자 생성물을 코딩하는 서열을 포함한다. 일부 실시양태에서, mRNA는 약 4.6 kB의 페이로드, 예를 들어, GDE 유전자 생성물을 코딩하는 서열을 포함한다. 일부 실시양태에서, mRNA는 약 6 kB의 페이로드, 예를 들어, OTOF 유전자 생성물을 코딩하는 서열을 포함한다. 일부 실시양태에서, mRNA는 약 7.1 kB의 페이로드, 예를 들어, F8 유전자 생성물을 코딩하는 서열을 포함한다.

[0234] 역전위를 이용하여 핵산을 게놈 내로 통합시키는 방법을 이용하는 장점 중 하나는 이 방법이 임의의 다른 기존 방법을 이용하는 것보다 훨씬 더 큰 핵산 카고를 전달하도록 본원에 기재된 바와 같이 디자인될 수 있다는 점이다. 예를 들어, 렌티바이러스 및 아데노 관련 바이러스(AAV) 유전자 전달 방법은 4 kB보다 더 큰 핵산 카고를 전달할 것으로 예상되지 않는다. 또한, 렌티바이러스 전달은 삽입 돌연변이유발 및 기타 독성의 위험을 수반한다. AAV 매개 전달은 해결되지 않은 간 및 CNS 독성을 수반한다. 반면, 본원에 기재된 바와 같이 mRNA를 사용하는 역전위 매개 방법(Retro-T)은 이 바이러스 방법들보다 더 빠르고 더 안전하며 덜 복잡하다.

[0235] 일부 실시양태에서, mRNA는 mRNA의 분해를 억제하거나 방해하는 서열을 포함한다. 일부 실시양태에서, mRNA의 분해를 억제하거나 방해하는 서열은 엑소뉴클레아제 또는 RNAse에 의한 mRNA의 분해를 억제하거나 방해한다. 일부 실시양태에서, mRNA의 분해를 억제하거나 방해하는 서열은 G 사중체, 슈도노트 또는 삼중체 서열이다. 일부 실시양태에서, mRNA의 분해를 억제하거나 방해하는 서열은 플라비바이러스 RNA의 엑소리보뉴클레아제 내성 RNA 구조 또는 KSV의 ENE 요소이다. 일부 실시양태에서, mRNA의 분해를 억제하거나 방해하는 서열은 데아데닐라제에 의한 mRNA의 분해를 억제하거나 방해한다. 일부 실시양태에서, mRNA의 분해를 억제하거나 방해하는 서열은 mRNA의 폴리 A 꼬리의 내부 또는 말단에서 비-아데노신 뉴클레오타이드를 포함한다. 일부 실시양태에서, mRNA의 분해를 억제하거나 방해하는 서열은 mRNA의 안정성을 증가시킨다. 일부 실시양태에서, 외생성 서열은 외생성 폴리펩티드를 코딩하는 서열을 포함한다. 일부 실시양태에서, 외생성 폴리펩티드를 코딩하는 서열은 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 서열과 인 프레임으로 있지 않다. 일부 실시양태에서, 외생성 폴리펩티드를 코딩하는 서열은 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 서열과 인 프레임으로 있지 않다. 일부 실시양태에서, 외생성 서열은 인트론을 포함하지 않는다. 일부 실시양태에서, 외생성 서열은 효소, 수용체, 수송 단백질, 구조 단백질, 호르몬, 항체, 수축 단백질 및 저장 단백질로 구성된 군으로부터 선택된 외생성 폴리펩티드를 코딩하는 서열을 포함한다. 일부 실시양태에서, 외생성 서열은 키메라 항원 수용체(CAR), 리간드, 항체, 수용체 및 효소로 구성된 군으로부터 선택된 외생성 폴리펩티드를 코딩하는 서열을 포함한다. 일부 실시양태에서, 외생성 서열은 조절 서열을 포함한다. 일부 실시양태에서, 조절 서열은 시스 작용 조절 서열을 포함한다. 일부 실시양태에서, 조절 서열은 인핸서, 사일런서, 프로모터 또는 반응 요소로 구성된 군으로부터 선택된 시스 작용 조절 서열을 포함한다. 일부 실시양태에서, 조절 서열은 트랜스 작용 조절 서열을 포함한다. 일부 실시양태에서, 조절 서열은 전사 인자를 코딩하는 트랜스 작용 조절 서열을 포함한다.

[0236] 일부 실시양태에서, 삽입체 서열의 통합은 세포 건강에 불리한 영향을 미치지 않는다. 일부 실시양태에서, 엔도뉴클레아제, 역전사효소 또는 이들 둘 다는 삽입체 서열을 부위 특이적으로 통합시킬 수 있다.

[0237] 일부 실시양태에서, mRNA는 추가 뉴클레아제 도메인, 또는 ORF2로부터 유래하지 않은 뉴클레아제 도메인을 코딩하는 서열을 포함한다. 일부 실시양태에서, mRNA는 메가TAL 뉴클레아제 도메인, TALEN 도메인, Cas9 도메인, R2 역요소의 징크 핑거 결합 도메인, 또는 AAV의 Rep78과 같은 반복 서열에 결합하는 DNA 결합 도메인을 코딩하는

서열을 포함한다. 일부 실시양태에서, 엔도뉴클레아제는 돌연변이를 갖지 않은 엔도뉴클레아제에 비해 엔도뉴클레아제의 활성을 감소시키는 돌연변이를 포함한다. 일부 실시양태에서, 엔도뉴클레아제는 ORF2p 엔도뉴클레아제이고, 돌연변이는 S228P이다. 일부 실시양태에서, mRNA는 역전사효소의 신뢰도 및/또는 진행성을 증가시키는 도메인을 코딩하는 서열을 포함한다. 일부 실시양태에서, 역전사효소는 ORF2 이외의 역요소의 역전사효소, 또는 ORF2p의 역전사효소에 비해 더 높은 신뢰도 및/또는 진행성을 가진 역전사효소이다. 일부 실시양태에서, 역전사효소는 II 군 인트론 역전사효소이다. 일부 실시양태에서, II 군 인트론 역전사효소는 IIA 군 인트론 역전사효소, IIB 군 인트론 역전사효소 또는 IIC 군 인트론 역전사효소이다. 일부 실시양태에서, II 군 인트론 역전사효소는 TGIRT-II 또는 TGIRT-III이다.

[0238] 일부 실시양태에서, mRNA는 Alu 요소 및/또는 리보좀 결합 앱타머를 포함하는 서열을 포함한다. 일부 실시양태에서, mRNA는 DNA 결합 도메인을 포함하는 폴리펩티드를 코딩하는 서열을 포함한다. 일부 실시양태에서, 3' UTR 서열은 바이러스 3' UTR 또는 베타-글로빈 3' UTR로부터 유래한다.

[0239] 한 측면에서, 본원은 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 포함하는 조성물을 제공하는 것으로, 이때 mRNA는 인간 LINE-1 트랜스포존 5' UTR 서열, 인간 LINE-1 트랜스포존 5' UTR 서열의 하류에 있는, ORF1p를 코딩하는 서열, ORF1p를 코딩하는 서열의 하류에 있는 ORF간 링커 서열, ORF간 링커 서열의 하류에 있는, ORF2p를 코딩하는 서열, 및 ORF2p를 코딩하는 서열의 하류에 있는, 인간 LINE-1 트랜스포존으로부터 유래한 3' UTR 서열을 포함하는 인간 LINE-1 트랜스포존 서열을 포함하고; 이때 3' UTR 서열은 삽입체 서열을 포함하고, 이 삽입체 서열은 외생성 폴리펩티드를 코딩하는 서열의 역상보체 또는 외생성 조절 요소를 코딩하는 서열의 역상보체이다.

[0240] 일부 실시양태에서, 삽입체 서열은 세포 내로 도입될 때 세포의 게놈 내로 통합된다. 일부 실시양태에서, 삽입체 서열은 병태 또는 질환과 관련된 유전자 내로 통합됨으로써, 상기 유전자를 파괴하거나 상기 유전자의 발현을 하향조절한다. 일부 실시양태에서, 삽입체 서열은 유전자 내로 통합됨으로써, 상기 유전자의 발현을 상향조절한다. 일부 실시양태에서, 재조합 mRNA 또는 이 mRNA를 코딩하는 벡터는 단리되거나 정제된다.

[0241] 한 측면에서, 본원은 (a) 인간 ORF1p 및 인간 ORF2p를 포함하는 긴 산재된 핵 요소(LINE) 폴리펩티드; 및 (b) 외생성 폴리펩티드를 코딩하는 서열의 역상보체 또는 외생성 조절 요소를 코딩하는 서열의 역상보체인 삽입체 서열을 코딩하는 뉴클레오타이드 서열을 포함하는 핵산을 포함하고 실질적으로 면역원성을 나타내지 않는 조성물을 제공한다.

[0242] 일부 실시양태에서, 상기 조성물은 인간 ORF1p 및 인간 ORF2p 단백질을 포함한다. 일부 실시양태에서, 상기 조성물은 핵산과 복합체를 형성한 인간 ORF1p 및 인간 ORF2p를 포함하는 리보핵단백질(RNP)을 포함한다. 일부 실시양태에서, 핵산은 mRNA이다.

[0243] 한 측면에서, 본원은 본원에 기재된 조성물을 포함하는 세포를 포함하는 조성물을 제공한다. 일부 실시양태에서, 세포는 면역 세포이다. 일부 실시양태에서, 면역 세포는 T 세포 또는 B 세포이다. 일부 실시양태에서, 면역 세포는 골수 세포이다. 일부 실시양태에서, 면역 세포는 단핵구, 대식세포, 수지상 세포, 수지상 전구체 세포 및 대식세포 전구체 세포로 구성된 군으로부터 선택된다. 일부 실시양태에서, 삽입체 서열은 외생성 폴리펩티드를 코딩하는 서열의 역상보체이고, 외생성 폴리펩티드는 키메라 항원 수용체(CAR)이다.

[0244] 한 측면에서, 본원은 본원에 기재된 조성물 및 약학적으로 허용되는 부형제를 포함하는 약학 조성물을 제공한다. 일부 실시양태에서, 약학 조성물은 유전자 요법에 사용하기 위한 것이다. 일부 실시양태에서, 약학 조성물은 질환 또는 병태를 치료하기 위한 의약의 제조에 사용하기 위한 것이다. 일부 실시양태에서, 약학 조성물은 질환 또는 병태를 치료하는 데 사용하기 위한 것이다. 한 측면에서, 본원은 본원에 기재된 약학 조성물을, 질환 또는 병태를 가진 대상체에게 투여하는 단계를 포함하는, 대상체에서 질환을 치료하는 방법을 제공한다. 일부 실시양태에서, 상기 방법은 대상체에서 단백질 또는 기능성 RNA의 양 또는 활성을 증가시킨다. 일부 실시양태에서, 대상체는 단백질 또는 기능성 RNA의 결핍된 양 또는 활성을 가진다. 일부 실시양태에서, 단백질 또는 기능성 RNA의 결핍된 양 또는 활성은 질환 또는 병태과 관련되어 있거나 질환 또는 병태를 야기한다.

[0245] 일부 실시양태에서, 상기 방법은 인간 침묵 허브(HUSH) 복합체를 억제하는 작용제, FAM208A를 억제하는 작용제, 또는 TRIM28을 억제하는 작용제를 투여하는 단계를 추가로 포함한다. 일부 실시양태에서, 인간 침묵 허브(HUSH) 복합체를 억제하는 작용제는 페리필린, TASOR 및/또는 MPP8을 억제하는 작용제이다. 일부 실시양태에서, 인간 침묵 허브(HUSH) 복합체를 억제하는 작용제는 HUSH 복합체의 어셈블리를 억제한다.

[0246] 일부 실시양태에서, 상기 작용제는 판코니 빈혈 복합체를 억제한다. 일부 실시양태에서, 상기 작용제는 FANCD2-FANC1 이종이량체 모노유비퀴틸화를 억제한다. 일부 실시양태에서, 상기 작용제는 FANCD2-FANC1 이종이량체 형

성을 억제한다. 일부 실시양태에서, 상기 작용제는 판코니 빈혈(FA) 코어 복합체를 억제한다. FA 코어 복합체는 예를 들어, 화학요법에 의해 유도된 DNA 가닥간 가교결합에서 판코니 빈혈 DNA 손상 복구 경로의 성분이다. FA 코어 복합체는 RING 핑거 서브유닛인 FANCL의 두 카피에 의해 플랭킹된, FANCB의 2개 중심 이량체 및 100 kDa 서브유닛의 FA 관련 단백질(FAAP100)을 포함한다. 이 2개의 이종삼량체는 남은 5개의 서브유닛들을 어셈블링하기 위한 스캐폴드로서 작용하여, 확장된 비대칭 구조를 생성한다. 이 스캐폴드의 불안정화는 전체 복합체를 파괴하여, 비-기능성 FA 경로를 야기할 것이다. FA 코어 복합체를 억제할 수 있는 작용제의 예는 보르테조밉 및 쿠르쿠민 유사체 EF24 및 4H-TTD를 포함한다.

[0247] 일부 실시양태에서, 삽입되는 서열은 조직 특이적 요소가 활성을 나타내는 세포에서만 전체 삽입된 DNA가 작용하도록 조직 특이적 요소의 조절 하에 놓일 수 있다.

[0248] 한 측면에서, 본원은 세포의 게놈 내로의 이종 핵산 서열의 역전위를 야기하는 서열에 의해 플랭킹된 관심 있는 이종 핵산 또는 유전자(예를 들어, 형질전이유전자, 조절 서열, 예를 들어, 억제 핵산에 대한 서열, siRNA, miRNA)를 세포 내로 도입함으로써 유전자를 세포에게 안정적으로 전달하는 방법 및 조성물을 제공한다. 일부 실시양태에서, 이종 핵산은 본 명세서에서 설명하기 위해 삽입체로서 지칭되고, 이때 삽입체는 본원에 기재된 구축물의 의도된 디자인에 의해 역전사되고 세포의 게놈 내로 삽입될 핵산 서열이다. 일부 실시양태에서, 이종 핵산은 본 명세서에서 설명하기 위해 카고 또는 카고 서열로서도 지칭된다. 카고는 게놈 내로 삽입되는 이종 핵산의 서열을 포함할 수 있다. 일부 실시양태에서, 세포는 포유동물 세포일 수 있다. 포유동물 세포는 상피, 중피 또는 내피 유래의 포유동물 세포일 수 있다. 일부 실시양태에서, 세포는 줄기 세포일 수 있다. 일부 실시양태에서, 세포는 전구체 세포일 수 있다. 일부 실시양태에서, 세포는 말기 분화된 세포일 수 있다. 일부 실시양태에서, 세포는 근육 세포, 심장 세포, 상피 세포, 조혈 세포, 점액 세포, 표피 세포, 편평 세포, 연골 세포, 골 세포, 또는 포유동물 유래의 임의의 세포일 수 있다. 일부 실시양태에서, 세포는 조혈 계통의 세포이다. 일부 실시양태에서, 세포는 골수 계통의 세포 또는 식세포, 예를 들어, 단핵구, 대식세포, 수지상 세포 또는 골수 전구체 세포이다. 일부 실시양태에서, 형질전이유전자를 코딩하는 핵산은 mRNA이다.

[0249] 일부 실시양태에서, 역전위 요소는 비-LTR 레트로트랜스포존으로부터 유래할 수 있다.

[0250] 본원은 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 세포 내로 도입하는 단계를 포함하는, 핵산 서열을 세포의 게놈 내로 통합시키는 방법을 제공하는 것으로, 이때 mRNA는 삽입체 서열을 포함하고, 삽입체 서열은 세포의 게놈 내로 통합된다. 일부 실시양태에서, 삽입체 서열은 (i) 외생성 서열 또는 (ii) 외생성 서열의 역상보체인 서열; 5' UTR 서열, 및 5' UTR 서열의 하류에 있는 3' UTR 서열을 포함하고; 이때 5' UTR 서열 또는 3' UTR 서열은 인간 ORF 단백질에 대한 결합 부위를 포함한다. 일부 실시양태에서, ORF 단백질은 인간 LINE 1 ORF2 단백질이다. 일부 실시양태에서, ORF 단백질은 비인간 ORF 단백질이다. 일부 실시양태에서, ORF 단백질은 키메라 단백질, 재조합 단백질 또는 조작된 단백질이다.

[0251] 본원은 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 도입하는 단계를 포함하는, 핵산 서열을 면역 세포의 게놈 내로 통합시키는 방법을 제공하는 것으로, 이때 mRNA는 (a) (i) 외생성 서열 또는 (ii) 외생성 서열의 역상보체인 서열을 포함하는 삽입체 서열; (b) 5' UTR 서열, 및 5' UTR 서열의 하류에 있는 3' UTR 서열을 포함하고, 이때 5' UTR 서열 또는 3' UTR 서열은 엔도뉴클레아제 결합 부위 및 역전사효소 결합 부위를 포함하고, 형질전이유전자 서열은 면역 세포의 게놈 내로 통합된다.

[0252] 일부 실시양태에서, RNA 통합 또는 전위를 매개하는 구조 요소는 합성 구축물에 코딩될 수 있고 세포에게 전달할 관심 있는 이종 유전자에 의해 좌우된다. 일부 실시양태에서, 합성 구축물은 관심 있는 이종 유전자를 코딩하는 핵산, 및 게놈 내로의 관심 있는 이종 유전자의 통합 또는 역전위를 야기하는 구조 요소를 포함할 수 있다. 일부 실시양태에서, 통합 또는 역전위를 야기하는 구조 요소는 5' L1 RNA 영역 및 3' L1 영역을 포함할 수 있고, 이때 3' L1 영역은 프라이밍을 위해 폴리 A 3' 영역을 포함한다. 일부 실시양태에서, 5' L1 RNA 영역은 하나 이상의 줄기 루프 영역을 포함할 수 있다. 일부 실시양태에서, 3' L1 영역은 하나 이상의 줄기 루프 영역을 포함할 수 있다. 일부 실시양태에서, 5' L1 영역 및 3' L1 영역은 관심 있는 이종 유전자(형질전이유전자)를 코딩하는 핵산 서열을 플랭킹하는 것으로서 구축된다. 일부 실시양태에서, 구조 유전자는 리보좀 결합 부위인 A 박스 및 B 박스 요소를 포함하는 헤어핀 루프 구조를 포함하는 L1 또는 Alu RNA의 영역을 포함할 수 있다. 일부 실시양태에서, 합성 핵산은 L1-Ta 프로모터를 포함할 수 있다.

[0253] ORF2p에 의한 두 가지 유형의 LINE RNA 인식이 있을 수 있다 - 엄격 및 완화. 엄격 유형에서 RT는 그 자신의 3' UTR 꼬리를 인식하고, 완화 유형에서 RT는 폴리 A 꼬리를 제외하고 어떠한 특이적 인식도 요구하지 않는다. 엄격 유형과 완화 유형으로 나누는 것은 일부 LINE/SINE 쌍들이 동일한 3' 말단을 공유한다는 관찰로부터 유래하

였다. 엄격 유형의 경우, 실험 연구는 3' UTR 줄기-루프가 역전위를 촉진함을 보여주었다. LINE 역전위 서열의 5' UTR은 3개의 보존된 줄기 루프 영역을 함유하는 것으로 확인되었다.

[0254] 일부 실시양태에서, 관심 있는 형질전이유전자 또는 전사체는 5' 및 3' 말단에서 L1 또는 Alu 서열의 전위 요소에 의해 플랭킹될 수 있다. 일부 실시양태에서, 레트로트랜스포존의 5' 영역은 Alu 서열을 포함한다. 일부 실시양태에서, 레트로트랜스포존의 3' 영역은 Alu 서열을 포함한다. 일부 실시양태에서, 레트로트랜스포존의 5' 영역은 L1 서열을 포함한다. 일부 실시양태에서, 레트로트랜스포존의 3' 영역은 L1 서열을 포함한다. 일부 실시양태에서, 관심 있는 형질전이유전자 또는 전사체는 SVA 트랜스포존 서열에 의해 플랭킹된다.

[0255] 일부 실시양태에서, 관심 있는 전사체는 ORF2p에 대한 결합 영역 및 3' 폴리 A 프라이밍 영역을 코딩하는 L1 또는 Alu 서열을 포함할 수 있다. 일부 실시양태에서, 관심 있는 형질전이유전자를 코딩하는 이종 핵산은 ORF1p에 대한 결합 영역 및 3' 폴리 A 프라이밍 영역을 코딩하는 L1 또는 Alu 서열에 의해 플랭킹될 수 있다. 3' 영역은 하나 이상의 줄기 루프 구조를 포함할 수 있다. 일부 실시양태에서, 관심 있는 전사체는 시스 통합 또는 역전위를 위해 구조화된다. 일부 실시양태에서, 관심 있는 전사체는 트랜스 통합 또는 역전위를 위해 구조화된다.

[0256] 일부 실시양태에서, 레트로트랜스포존은 인간 레트로트랜스포존이다. 인간 레트로트랜스포존의 서열은 엔도뉴클레아제 및/또는 역전사효소를 코딩하는 서열을 포함할 수 있다. 인간 레트로트랜스포존의 서열은 2개의 비-중첩 ORF를 함유하는 단일 RNA로부터 번역되는 2개의 단백질을 코딩할 수 있다. 일부 실시양태에서, 2개의 ORF는 ORF1 및 ORF2이다.

[0257] 따라서, 본원은 형질전이유전자를 코딩하는 이종 핵산을 골수 세포와 같은 세포의 게놈 내로 안정적으로 통합시키는 방법으로서, 형질전이유전자; 레트로트랜스포존의 5' 영역을 포함하는, 형질전이유전자를 코딩하는 영역을 플랭킹하는 하나 이상의 5' 핵산 서열; 및 레트로트랜스포존의 3' 영역을 포함하는, 형질전이유전자를 코딩하는 영역을 플랭킹하는 하나 이상의 3' 핵산 서열을 코딩하는 핵산을 세포 내로 도입하는 단계를 포함하는 방법을 제공하는 것으로, 이때 레트로트랜스포존의 3' 영역은 각각 엔도뉴클레아제 및 역전사효소(RT) 활성을 가진 게놈 DNA 프라이밍 서열 및 LINE 전위효소 결합 서열을 포함한다.

[0258] 본원은 재조합 mRNA 또는 mRNA를 코딩하는 벡터를 도입하는 단계를 포함하는, 핵산 서열을 세포의 게놈 내로 통합시키는 방법을 제공하는 것으로, 이때 mRNA는 (a) (i) 외생성 서열 또는 (ii) 외생성 서열의 역상보체인 서열을 포함하는 삽입체 서열; (b) 5' UTR 서열, 5' UTR 서열의 하류에 있는 인간 레트로트랜스포존의 서열, 및 인간 레트로트랜스포존의 서열의 하류에 있는 3' UTR 서열을 포함하고; 이때 5' UTR 서열 또는 3' UTR 서열은 엔도뉴클레아제 결합 부위 및 역전사효소 결합 부위를 포함하고, 인간 레트로트랜스포존의 서열은 2개의 ORF를 함유하는 단일 RNA로부터 번역되는 2개의 단백질을 코딩하고, 삽입체 서열은 세포의 게놈 내로 통합된다.

[0259] 일부 실시양태에서, 상기 방법은 삽입체 서열을 세포의 게놈 내로 전달하고 통합시키기 위해 단일 핵산 분자를 사용하는 단계를 포함한다. 단일 핵산 분자는 플라스미드 벡터일 수 있다. 단일 핵산은 DNA 또는 RNA 분자일 수 있다. 단일 핵산은 mRNA일 수 있다.

[0260] 일부 실시양태에서, 인간 레트로트랜스포존 및 이종 핵산 서열을 포함하는 하나 이상의 폴리뉴클레오타이드를 세포 내로 도입하는 단계를 포함한다. 일부 실시양태에서, 하나 이상의 폴리뉴클레오타이드는 (i) ORF1p를 코딩하는 제1 핵산 분자; (ii) ORF2p를 코딩하는 제2 핵산 분자 및 카고를 코딩하는 서열을 포함한다. 일부 실시양태에서, 제1 핵산 및 제2 핵산은 mRNA이다. 일부 실시양태에서, 제1 핵산 및 제2 핵산은 예를 들어, 별도의 플라스미드 벡터에 코딩된 DNA이다.

[0261] 본원은 세포의 게놈 내로 삽입되는 서열을 포함하는 자가통합 폴리뉴클레오타이드를 제공하고, 삽입체는 자가통합 네이키드(naked) 폴리뉴클레오타이드에 의해 게놈 내로 안정적으로 통합된다. 일부 실시양태에서, 상기 폴리뉴클레오타이드는 RNA이다. 일부 실시양태에서, 상기 폴리뉴클레오타이드는 mRNA이다. 일부 실시양태에서, 상기 폴리뉴클레오타이드는 변형을 가진 mRNA이다. 일부 실시양태에서, 변형은 세포내 환경에서 RNase에 대한 보호를 보장한다. 일부 실시양태에서, 변형은 치환된 변형 뉴클레오타이드, 예를 들어, 5-메틸사이티딘, 슈도우리딘 또는 2-티오우리딘을 포함한다.

[0262] 일부 실시양태에서, 단일 폴리뉴클레오타이드는 삽입체(또는 카고) 핵산의 전달 및 게놈 통합을 위해 사용된다. 일부 실시양태에서, 단일 폴리뉴클레오타이드는 바이시스트론이다. 일부 실시양태에서, 단일 폴리뉴클레오타이드는 트리시스트론이다. 일부 실시양태에서, 단일 폴리뉴클레오타이드는 멀티시스트론이다. 일부 실시양태에서, 2개 이상의 폴리뉴클레오타이드 분자는 삽입체(또는 카고) 핵산의 전달 및 게놈 통합을 위해 사용된다.

[0263] 일부 실시양태에서, 역전위 유전 요소를 생성할 수 있고, 역전위 유전 요소는 (i) 세포의 게놈 내로 삽입될 형

질전이유전자 또는 비코딩 서열을 코딩하는 이종 핵산(삽입체); (ii) 하나 이상의 레트로트랜스포존 ORF 코딩 서열을 코딩하는 핵산 서열; (iii) 삽입될 형질전이유전자 또는 비코딩 서열을 코딩하는 이종 핵산이 UTR 서열 내에 포함되도록 ORF 코딩 서열의 하나 이상의 UTR 영역을 포함하고; 이때 레트로트랜스포존 ORF 코딩 서열의 3' 영역은 게놈 DNA 프라이밍 서열을 포함한다.

[0264] 일부 실시양태에서, 역전위 유전 요소는 형질전이유전자를 게놈 DNA 내로 안정적으로 통합시키기 위해 세포 내로 도입될 수 있다. 일부 실시양태에서, 역전위 유전 요소는 (a) 레트로트랜스포존 단백질 코딩 서열 및 3' UTR; 및 (b) 세포의 게놈 내로 삽입될(예를 들어, 통합될) 이종 핵산을 포함하는 서열을 포함한다. 레트로트랜스포존 단백질 코딩 서열 및 3' UTR은 이종 핵산 서열을 세포의 게놈 내로 전달하기에 완전하고 충분한 유닛일 수 있고, 역전사를 시작하고 이종 핵산을 통합시키기 위해 역전위 요소, 예컨대, 엔도뉴클레아제, 역전사효소, 및 엔도뉴클레아제에 의해 절단된 영역에서 게놈 DNA에 결합하고 이를 프라이밍하기 위한 3' UTR 내의 서열을 포함한다.

[0265] 일부 실시양태에서, 삽입체의 코딩 서열은 하나 이상의 ORF의 코딩 서열에 대하여 정방향으로 존재한다. 일부 실시양태에서, 삽입체의 코딩 서열은 하나 이상의 ORF의 코딩 서열에 대하여 역방향으로 존재한다. 삽입체의 코딩 서열 및 하나 이상의 ORF의 코딩 서열은 5' UTR, 3' UTR, 프로모터, 인핸서 등을 비롯한 상이한 조절 요소들을 포함할 수 있다. 일부 실시양태에서, 삽입체의 3' UTR 또는 5' UTR은 하나 이상의 ORF의 코딩 서열을 포함할 수 있고, 마찬가지로, 삽입체의 코딩 서열은 하나 이상의 ORF의 코딩 서열의 3' UTR 내에 위치할 수 있다.

[0266] 일부 실시양태에서, (a) (i) 외생성 서열, 또는 외생성 서열의 역상보체인 서열을 포함하는 삽입체 서열; (b) 5' UTR 서열, 및 5' UTR 서열의 하류에 있는 3' UTR 서열을 포함하는 역전위 유전 요소를 생성할 수 있고; 이때 5' UTR 서열 또는 3' UTR 서열은 인간 ORF 단백질에 대한 결합 부위를 포함한다.

[0267] 일부 실시양태에서, 레트로트랜스포존은 SINE 또는 LINE 요소를 포함할 수 있다. 일부 실시양태에서, 레트로트랜스포존은 SINE 또는 LINE 줄기 루프 구조, 예컨대, Alu 요소를 포함한다.

[0268] 일부 실시양태에서, 레트로트랜스포존은 LINE-1(L1) 레트로트랜스포존이다. 일부 실시양태에서, 레트로트랜스포존은 인간 LINE-1이다. 인간 LINE-1 서열은 인간 게놈에 풍부하다. 대략 총 13,224개의 인간 L1이 존재하고, 이들 중 약 3.6%를 차지하는 480개는 활성을 나타낸다. 따라서, 인간 L1 단백질은 인간에서 잘 용인되고 면역원성을 나타내지 않는다. 더욱이, 인간에서 무작위 전위의 엄격한 조절은 무작위 전위효소 활성이 본원에 기재된 L1 시스템의 도입에 의해 유발되지 않도록 보장한다. 추가로, 본원에서 디자인된 역전위 구축물은 삽입체 서열의 표적화된 특이적 통합을 포함할 수 있다. 일부 실시양태에서, 역전위 유전 요소는 게놈 불안정성을 초래하는 무작위 통합이 시작되지 않도록 주의하면서 인간 세포에서 활발히 우세한 침묵 기구를 극복하도록 의도된 디자인을 포함할 수 있다.

[0269] 따라서, 역전위 구축물은 인간 LINE-1 ORF1 단백질 및 인간 LINE-1 ORF2 단백질을 코딩하는 서열을 포함할 수 있다. 일부 실시양태에서, 상기 구축물은 하기 서열과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100% 서열 동일성을 가진 ORF1p 단백질을 코딩하는 핵산 서열을 포함한다:

MGKKQNRKTGNSKTQSASPPPKERSSSPATEQSWMENDFDELREEGFRRSNYSELREDIQTKGK
EVENFEKNLEECITRITNTEKCLKELMELKTKARELREECRSLRSRCDQLEERVSAMEDEMNEM
KREGKFREKRIKRNEQSLQEIWDYVKRPNLRLIGVPESDVENGTKLENTLQDIIQENFPNLARQA
NVQIQEIQRTPQRYSSRRATPRHIIVRFTKVEMKEKMLRAAREKGRVTLKGKPIRLTVDLSAETL
QARREWGPIFNILKEKNFQPRISYPAKLSFISEGEIKYFIDKQMLRDFVTTRPALKELLKEALNME
RNNRYQPLQNHAKM(서열번호 57).

[0270] 일부 실시양태에서, 상기 구축물은 하기 서열과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100% 서열 동일성을 가진 핵산 서열을 포함한다:

atgggcaagaagcaaaatcgcaagacggggaattccaagacacaatccgctagcccaccacctaaagagcgttctagctcccctgctactgagcagtc
ctggatggaaaacgacttcgatgaactccgggaagagggatttaggcgatccaactattcagaactccgcgaagatatccagacaaaggggaaggaa
gtcgagaatttcgagaagaacctcgaggagtgcatcacccgtatcacaaacactgagaaatgtctcaaagaactcatggaacttaagacaaaagccag
ggagcttcgagaggagtgtcggagtctgagatccaggtgtgaccagctcgaggagcgcgtgagcgcgatggaagacgagatgaacgagatgaaaa
gagagggcaaattcagggagaagcgcattaagaggaacgaacagagtctgcaggagatttgggattacgtcaagaggcctaacctgcggttgatcgg
cgtccccgagagcgacgtagaaaacgggactaaactggagaatacacttcaagacatcattcaagaaaattttccaaacctggctcggcaagctaatgt
gcaaatccaagagatccaacgcacaccccagcggtatagctctcggcgtgccacccctaggcatattatcgtgcgctttactaaggtggagatgaaaga
gaagatgctgcgagccgctcgggaaaagggaagggtgactttgaagggcaaacctattcggctgacggttgaccttagcgccgagacactccaggca
cgccgggaatggggccccatctttaatatcctgaaggagaagaacttccagccacgaatctcttaccctgcaaagttgagttttatctccgagggtgagat
taagtatttcatcgataaacagatgctgcgagacttcgtgacaactcgcccagctctcaaggaactgctcaaagaggctcttaatatggagcgcaataata

[0271] gatatcaacccttgcagaaccacgcaaagatgtga(**서열번호** 58).

[0272] 일부 실시양태에서, 상기 구축물은 하기 서열과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100% 서열 동일성을 가진 ORF2p 단백질을 코딩하는 핵산 서열을 포함한다:

MTGSNSHITILTLNINGLNSAIKRHRLASWIKSQDPSVCCIQETHLTCRDTHRLKIKGWRKIYQAN
GKQKKAGVAILVSDKTDFKPTKIKRDKEGHYIMVKGSIQQEELTILNIYAPNTGAPRFIKQVLSDL
QRDLDSHTLIMGDFNTPLSTLDRSTRQKVNKDTQELNSALHQADLIDIYRTLHPKSTEYTFFSAP
HHTYSKIDHIVGSKALLSKCKRTEIITNYLSDHSAIKLELRIKNLTQSRSTTWKLNNLLLNDYWV
HNEMKAEIKMFFETNENKDTTYQNLWDAFKAVCRGKFIALNAYKRKQERSKIDTLTSQLKELE
KQEQTHSKASRRQEITKIRAELKEIETQKTLQKINESRSWFFERINKIDRPLARLIKKKREKNQIDTI
KNDKGDITTDPTEIQTTIREYYKHLYANKLENLEEMDTFLDTYTLPRLNQEEVESLNRPITGSEIV
AIINSLPTKKSPGPDGFTAEFYQRYMEELVPFLLKLFQSIEKEGILPNSFYEASIILIPKPGRDTTKKE
NFRPISLMNIDAKILNKILANRIQQHIKKLIHHDQVGFIPGMQGWFNIRKSINVIQHINRAKDKNH
MIISIDAEKAFDKIQQPFMLKTLNKLGIDGTYFKIIRAIYDKPTANIILNGQKLEAFPLKTGTRQGC
PLSPLLFNIVLEVLARAIRQEKEIKGIQLGKEEVKLSLFADDMIVYLENPIVSAQNLLKLISNFSKV
SGYKINVQKSQAFLYTNNRQTESQIMGELPFVIASKRIKYLGIQLTRDVKDLFKENYKPLLKEIKE
DTNKWKNIPCSWVGRINIVKMAILPKVIYRFNAIPIKLPMTFFTELEKTTLKFIWNQKRARIAKSIL
SQKNKAGGITLPDFKLYYKATVTKTAWYWYQNRDIDQWNRTEPSEIMPHIYNYLIFDKPEKNK
QWGKDSLFNKWCWENWLAICRKLKLDPFLTPYTKINSRWIKDLNVKPKTIKTLEENLGITIQDIG
VGKDFMSKTPKAMATKDKIDKWDLIKLKSFCTAKETTIRVNRQPTTWEKIFATYSSDKGLISRIY
NELKQIYKKKTNNPIKKWAKDMNRHFSKEDIYAAKKHMKKCSSSLAIREMQIKTTMRYHLTPV
RMAIIKKSGNNRCWRGCGEIGTLLHCWWDCKLVQPLWKSVWRFLRDLELEIPFDPAIPLLGIYP
NEYKSCCYKDTCTRMFIAALFTIAKTWNQPKCPTMIDWIKKMWHIYTMEYYAAIKNDEFISFVG

[0273] TWMKLETIILSKLSQEQKTKHRIFSLIGGN(**서열번호** 59). 일부 실시양태에서, 상기 구축물은 하기 서열과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100% 서열 동일성을 가진 핵산 서열을 포함한다:

atgaccggctctaactcacatatcaccatccttacacttaacattaacggcctcaactcagctatcaagcgccatcggctggccagctggatcaaatcaca
ggatccaagcgtttgttgcatccaagagacccacctgacctgtagagatactcaccgcctcaagatcaagggatggcgaaagatttatcaggcgaacgg
taagcagaagaaagccggagtcgcaattctggtctcagacaagacggatttcaagcccaccaaaattaagcgtgataaggaaggtcactatattatggtg
aaaggcagcatacagcaggaagaacttaccatattgaacatctacgcgccaaacaccggcgcacctcgctttatcaaacaggtcctgtccgatctgcag
cgagatctggattctcatacgttgattatgggtgatttcaatacaccattgagcaccctggatcgcagcaccaggcaaaaggtaaataaagacacgcaag
agctcaatagcgcactgcatcaggcagatctcattgatatttatcgcactcttcatcctaagagtaccgagtacacattcttcagcgccccacatcatacata
ctcaaagatcgatcatatcgtcggctcaaaggctctgctgtcaaagtgcaagcgcacagagataattacaaattacctgtcagatcatagcgcgatcaag
ctcgagctgagaatcaagaacctgacccagagccggagtaccacttggaagcttaataacctgctgctcaacgattattgggtccacaatgagatgaag
gcagagattaaaatgttcttcgaaacaaatgagaataaggatactacctatcaaaaccttgggatgcctttaaggccgtctgcagaggcaagttcatcgc
cctcaacgcctataaaagaaaacaagagagatctaagatcgatactctcacctctcagctgaaggagttggagaaacaggaacagacccactccaagg
cgtcaagacggcaggagatcacaaagattcgcgccgagttgaaagagatcgaaacccaaaagactcttcagaaaattaacgagtctcgtagttggttctt
cgagcggattaataagatagacagacctctggcacgactgattaagaagaagcgcgaaaagaaccagattgataccatcaagaacgacaagggcgac
atcactactgacccgaccgagatccagaccactattcgggagtattataagcatttgtatgctaacaagcttgagaacctggaagagatggacacttttctg
gatacctatactctgccacggcttaatcaagaggaagtcgagtccctcaaccgcccaattacaggaagcgagattgtggccataattaactccctgccga
caaagaaatctcctggtccggacgggtttacagctgagttttatcaacggtatatggaagagcttgtaccgtttctgctcaagctctttcagtctatagaaaa
ggaaggcatcttgcccaattccttctacgaagcttctataatacttattcccaaaccaggacgcgataccacaaagaaggaaaacttccggcccattagtct
catgaatatcgacgctaaaatattgaacaagattctcgccaacagaatccaacaacatattaagaaattgatacatcacgaccaggtggggtttatacctgg
catgcagggctggtttaacatccggaagagtattaacgtcattcaacacattaatagagctaaggataagaatcatatgatcatctctatagacgcggaaaa
ggcattcgataagattcagcagccatttatgctcaagactctgaacaaactcggcatcgacggaacatattttaagattattcgcgcaatttacgataagcc
gactgctaacattatccttaacggccaaaagctcgaggcctttccgctcaagactggaacccgccaaggctgtcccctctcccgctttgtttaatattgta
ctcgaggtgctggctagggctattcgtcaagagaaagagattaaagggatacagctcgggaaggaagaggtcaagctttccttgttcgccgatgatatg
attgtgtacctggagaatcctattgtgtctgctcagaaccttcttaaacttatttctaactttagcaaggtcagcggctataagattaacgtccagaaatctcag
gcctttctgtacacaaataatcgacagaccgaatcccagataatgggtgagcttccgtttgtcatagccagcaaaaggataaagtatctcggaatccagct
gacacgagacgttaaagatttgtttaaggaaaattacaagcctctcctgaaagagattaaggaagatactaataagtggaagaatatcccctgttcatgggt
tggcagaatcaacatagtgaagatggcaatacttcctaaagtgatatatcgctttaacgccatcccaattaaactgcctatgaccttctttacggagctcgag
aaaacaaccctaaatttatatggaatcaaaagagagcaagaatagcgaagtccatcttgagccagaagaataaggccggtgggattactttgcctgattt
taagttgtattataaagccacagtaactaagacagcctggtattggtatcagaatagagacatcgaccagtggaatcggaccgaaccatcagagataatg
ccccacatctataattaccttatattcgataagccagaaaagaataaacagtggggcaaagacagcctcttcaacaagtggtgttgggagaattggctggc
catatgccggaaactcaagctcgaccccttttcttacacctacactaaaatcaacagtaggtggatcaaggacttgaatgtcaagccaaagactataaaga
cactggaagagaatcttgggatcacaatacaagatataggcgtcggcaaagattttatgtcaaagacgcccaaggccatggccactaaggataagattg
ataagtgggaccttattaagctcaaaagcttctgtactgccaaggagaccacgatcagagttaataggcagcccactacatgggaaaagattttcgccact
tattcatcagataaggggttgataagcagaatatataacgagctgaagcagatctacaagaagaaaacgaataatcccatcaagaagtgggcaaaagat
atgaacaggcattttagcaaagaggatatctacgccgcgaagaagcatatgaagaagtgtagttcaagcttggccattcgtgagatgcagattaagacga
ccatgcgataccaccttaccccagtgaggatggcaatttatcaagaaatctggcaataatagatgttggcggggctgtggcgagattggcaccctgctcca
ttgctggtgggattgcaagctggtgcagccgctttggaaatcagtctggcgctttctgagggacctcgagcttgagattcccttcgatcccgcaattccctt
gctcggaatctatcctaacgaatacaagagctgttgttacaaggatacgtgtacccggatgttcatcgcggccttgtttacgatagctaagacgtggaatca
gcctaagtgccccacaatgatcgattggatcaagaaaatgtggcatatttataccatggagtattacgcagcaattaagaatgacgaatttatttccttcgttg
ggacctggatgaagctggagactattattctgagcaagctgtctcaggagcaaaagacaaagcatagaatcttctctctcattggtggtaactaa(서열
번호 60).

[0274]

[0275] 일부 실시양태에서, 상기 구축물은 하기 서열과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100% 서열 동일성을 가진 ORF2p 단백질을 코딩하는 핵산 서열을 포함한다:

MVIGTYISIITLNVNGLNAPTKRHRLAEWIQKQDPYICCLQETHFRPRDTYRLKVRGWKKIFHAN
GNQKKAGVAILISDKIDFKIKNVTRDKEGHYIMIQGSIQEEDITIINIYAPNIGAPQYIRQLLTAIKE
EIDSNTIIVGDFNTSLTPMDRSSKMKINKETEALNDTIDQIDLIDIYRTFHPKTADYTFFSSAHGTFS
RIDHILGHKSSLSKFKKIEIISSIFSDHNAMRLEMNHREKNVKKTNTWRLNNTLLNNQEITEEIKQ
EIKKYLETNDNENTTTQNLWDAAKAVLRGKFIAIQAYLKKQEKSQVNNLTLHLKKLEKEEQTK
PKVSRRKEIIKIRAEINEIETKKTIAKINKTKSWFFEKINKIDKPLARLIKKKRERTQINKIRNEKGE
VTTDTAEIQNILRDYYKQLYANKMDNLEEMDKFLERYNLPRLNQEETENINRPITSNEIETVIKNL
PTNKSPGPDGFTGEFYQTFREELTPILLKLFQKIAEEGTLPNSFYEATITLIPKPDKDTTKKENYRPI
SLMNIDAKILNKILANRIQQHIKRIIHHDQVGFIPGMQGFFNIRKSINVIHHINKLKKKNHMIISIDA
EKAFDKIQHPFMIKTLQKVGIEGTYLNIIKAIYDKPTANIILNGEKLKAFPLRSGTRQGCPLSPLLF
NIVLEVLATAIREEKEIKGIQIGKEEVKLSLFADDMILYIENPKTATRKLLELINEYGKVAGYKINA
QKSLAFLYTNDEKSEREIMETLPFTIATKRIKYLGINLPKETKDLYAENYKTLMKEIKDDTNRWR
DIPCSWIGRINIVKMSILPKAIYRFNAIPIKLPMAFFTELEQIILKFVWRHKRPRIAKAVLRQKNGA
GGIRLPDFRLYYKATVIKTIWYWHKNRNIDQWNKIESPEINPRTYGQLIYDKGGKDIQWRKDSLF
NKWCWENWTATCKRMKLEYSLTPYTKINSKWIRDLNIRLDTIKLLEENIGRTLFDINHSKIFFDPP
PRVMEIKTKINKWDLMKLQSFCTAKETINKTKRQPSEWEKIFANESTDKGLISKIYKQLIQLNIKE
TNTPIQKWAEDLNRHFSKEDIQTATKHMKRCSTSLIIREMQIKTTMRYHLTPVRMGIIRKSTNNK
CWRGCGEKGTLLHCWWECKLIQPLWRTIWRFLKKLKIELPYDPAIPLLGIYPEKTVIQKDTCTR
MFIAALFTIARSWKQPKCPSTDEWIKKMWYIYTMEYYSAIKRNEIGSFLETWMDLETVIQSEVSQ
KEKNKYRILTHICGTWKNGTDEPVCRTEIETQM(**서열번호** 61).

[0276] 일부 실시양태에서, 상기 구축물은 하기 서열과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100% 서열 동일성을 가진 핵산 서열을 포함한다:

atggtcataggaacatacatatcgataattaccttaaacgtgaatggattaaatgccccaaccaaaagacatagactggctgaatggatacaaaaacaag
acccatatatatgctgtctacaagagacccacttcagacctagggacacatacagactgaaagtgaggggatggaaaaagatattccatgcaaatggaa
atcaaaagaaagctggagtagctatactcatatcagataaaatagactttaaaataaagaatgttacaagagacaaggaaggacactacataatgatcca
gggatcaatccaagaagaagatataacaattataaatatatatgcacccaacataggagcacctcaatacataaggcaactgctaacagctataaaagag
gaaatcgacagtaacacaataatagtgggggactttaacacctcacttacaccaatggacagatcatccaaaatgaaaataaataaggaaacagaagctt
taaatgacacaatagaccagatagatttaattgatatatataggacattccatccaaaaacagcagattacacgttcttctcaagtgcgcacggaacattctc
caggatagatcacatcttgggtcacaaatcaagcctcagtaaatttaagaaaattgaaatcatatcaagcatcttttctgaccacaacgctatgagattagaa
atgaatcacagggaaaaaaacgtaaaaaagacaaacacatggaggctaaacaatacgttactaaataaccaagagatcactgaagaaatcaaacagga
aataaaaaaatacctagagacaaatgacaatgaaaacacgacgacccaaaacctatgggatgcagcaaaagcggttctaagagggaagtttatagctat
acaagcctacctaaagaaacaagaaaaatctcaagtaaacaatctaaccttacacctaaagaaactagagaaagaagaacaaacaaaacccaaagtta
gcagaaggaaagaaatcataaagatcagagcagaaataaatgaaatagaaacaaagaaaacaatagcaaagatcaataaaactaaaagttggttctttg
agaagataaacaaaattgataagccattagccagactcatcaagaaaaagagggagaggactcaaatcaataaaatcagaaatgaaaaaggagaagtt
acaacagacaccgcagaaatacaaaacatcctaagagactactacaagcaactttatgccaataaaatggacaacctggaagaaatggacaaattctta
gaaaggtataaccttccaagactgaaccaggaagaaacagaaaatatcaacagaccaatcacaagtaatgaaattgaaactgtgattaaaaatcttccaa
caaacaaaagtccaggaccagatggcttcacaggtgaattctatcaaacatttagagaagagctaacacccatccttctcaaactcttccaaaaaattgca
gaagaaggaacactcccaaactcattctatgaggccaccatcaccctgataccaaaaccagacaaagacactacaaaaaaagaaaattacagaccaat
atcactgatgaatatagatgcaaaaatcctcaacaaaatactagcaaacagaatccaacaacacattaaaaggatcatacaccacgatcaagtgggattta
tcccagggatgcaaggattcttcaatatacgcaaatcaatcaatgtgatacaccatattaacaaattgaagaagaaaaaccatatgatcatctcaatagatg
cagaaaaagcttttgacaaaattcaacacccatttatgataaaaactctccagaaagtgggcatagagggaacctacctcaacataataaaggccatatat
gacaaacccacagcaaacatcattctcaatggtgaaaaactgaaagcatttcctctaagatcaggaacgagacaaggatgtccactctcaccactattatt
caacatagttctggaagtcctagccacggcaatcagagaagaaaaagaaataaaaggaatacaaattggaaaagaagaagtaaaactgtcactgtttgc
ggatgacatgatactatacatagagaatcctaaaactgccaccagaaaactgctagagctaattaatgaatatggtaaagttgcaggttacaaaattaatgc
acagaaatctcttgcattcctatacactaatgatgaaaaatctgaaagagaaattatggaaacactcccatttaccattgcaacaaaaagaataaaatacccta
ggaataaacctacctaaggagacaaaagacctgtatgcagaaaactataagacactgatgaaagaaattaaagatgataccaacagatggagagatata
ccatgttcttggattggaagaatcaacattgtgaaaatgagtatactacccaaagcaatctacagattcaatgcaatccctatcaaattaccaatggcatttttt
acggagctagaacaaatcatcttaaaatttgtatggagacacaaaagaccccgaatagccaaagcagtcttgaggcaaaaaaatggagctggaggaat
cagactccctgacttcagactatactacaaagctacagtaatcaagacaatatggtactggcacaaaaacagaaacatagatcaatggaacaagatagaa
agcccagagattaacccacgcacctatggtcaactaatctatgacaaaggaggcaaagatatacaatggagaaaagacagtctcttcaataagtggtgct
gggaaaactggacagccacatgtaaaagaatgaaattagaatactccctaacaccatacacaaaaataaactcaaaatggattagagacctaaatataag
actggacactataaaactcttagaggaaaacataggaagaacactctttgacataaatcacagcaagatctttttcgatccacctcctagagtaatggaaat

[0277]

aaaaacaaaaataaacaagtgggacctaatgaaacttcaaagcttttgcacagcaaaggaaaccataaacaagacgaaaagacaaccctcagaatggg
agaaaatatttgcaaatgaatcaacggacaaaggattaatctccaaaatatataaacagctcattcagctcaatatcaaagaaacaaacaccccaatccaa
aaatgggcagaagacctaaatagacatttctccaaagaagacatacagacggccacgaagcacatgaaaagatgctcaacatcactaattattagagaa
atgcaaatcaaaactacaatgaggtatcacctcactcctgttagaatgggcatcatcagaaaatctacaaacaacaaatgctggagagggtgtggagaaa
agggaaccctcttgcactgttggtgggaatgtaaattgatacagccactatggagaacaatatggaggttccttaaaaaactaaaaatagaattaccatatg
acccagcaatcccactactgggcatatacccagagaaaaccgtaattcaaaaagacacatgcacccgaatgttcattgcagcactatttacaatagccag
gtcatggaagcaacctaaatgcccatcgacagacgaatggataaagaagatgtggtacatatatacaatggaatattactcagccataaaaaggaacga
aattgggtcatttttagagacgtggatggatctagagactgtcatacagagtgaagtaagtcagaaagagaaaaacaaatatcgtatattaacgcatatatg
tggaacctggaaaaatggtacagatgaaccggtctgcaggacagaaattgagacacaaatgtaa(**서열번호** 62).

[0278]

[0279] 일부 실시양태에서, 상기 구축물은 PAAKRVKLD(서열번호 63)과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100% 서열 동일성을 가진 핵 국소화 서열을 코딩하는 핵산 서열을 포함한다. 일부 실시양태에서, 핵 국소화 서열은 ORF2p 서열에 융합된다. 일부 실시양태에서, 상기 구축물은 서열 DYKDDDDK(서열번호 64)를 가진 플래그 태그를 코딩하는 핵산 서열을 포함한다. 일부 실시양태에서, 플래그 태그는 ORF2p 서열에 융합된다. 일부 실시양태에서, 플래그 태그는 핵 국소화 서열에 융합된다.

[0280] 일부 실시양태에서, 상기 구축물은 하기 서열과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100% 서열 동일성을 가진 MS2 코트 단백질을 코딩하는 핵산 서열을 포함한다:

ASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV
PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAMASNFTQFVLVD
NGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLN
MELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY(**서열번호** 65).

[0281] 일부 실시양태에서, MS2 코트 단백질 서열은 ORF2p 서열에 융합된다.

[0282] 일부 실시양태에서, 형질전이유전자는 Alu ORF2p 인식 서열을 포함하는 플랭킹 서열을 포함할 수 있다.

[0283] 일부 실시양태에서, 추가 요소를 mRNA 내로 도입할 수 있다. 일부 실시양태에서, 추가 요소는 IRES 요소 또는 T2A 요소일 수 있다. 일부 실시양태에서, mRNA 전사체는 3' 말단에서 1개, 2개, 3개 이상의 정지 코돈을 포함한다.

[0284] 일부 실시양태에서, 1개, 2개, 3개 이상의 정지 코돈은 직렬로 존재하도록 디자인된다. 일부 실시양태에서, 1개, 2개, 3개 이상의 정지 코돈은 모든 3개의 리딩 프레임으로 존재하도록 디자인된다. 일부 실시양태에서, 1개, 2개, 3개 이상의 정지 코돈은 다중 리딩 프레임 및 직렬로 존재하도록 디자인될 수 있다.

[0285] 일부 실시양태에서, 하나 이상의 표적 특이적 뉴클레오타이드를 L1 또는 Alu RNA 프라이밍 영역의 프라이밍 말단에 추가할 수 있다.

[0286] 일부 실시양태에서, 5' UTR 서열 또는 3' UTR 서열은 ORF 단백질에 결합할 수 있는 것 이외에 유전자 역전위 및/또는 안정한 통합을 조절하는 하나 이상의 내생성 단백질에 결합할 수도 있다. 일부 실시양태에서, 플랭킹 서열은 PABP 단백질에 결합할 수 있다.

[0287] 일부 실시양태에서, 전사체를 플랭킹하는 5' 영역은 강한 프로모터를 포함할 수 있다. 일부 실시양태에서, 프로모터는 CMV 프로모터이다.

[0288] 일부 실시양태에서, L1 ORF2p를 코딩하는 추가 핵산을 세포 내로 도입한다. 일부 실시양태에서, L1 ORF1을 코딩하는 서열은 누락되고, L1-ORF2만이 포함된다. 일부 실시양태에서, 플랭킹 요소를 가진 형질전이유전자를 코딩하는 핵산은 mRNA이다. 일부 실시양태에서, 내생성 L1-ORF1p 기능은 억제될 수 있거나 저해될 수 있다.

[0289] 일부 실시양태에서, 역전위 플랭킹 요소와 함께 형질전이유전자를 코딩하는 핵산은 하나 이상의 핵산 변형을 포함한다. 일부 실시양태에서, 역전위 플랭킹 요소와 함께 형질전이유전자를 코딩하는 핵산은 형질전이유전자에서 하나 이상의 핵산 변형을 포함한다. 일부 실시양태에서, 변형은 형질전이유전자 서열의 코돈 최적화를 포함한다. 일부 실시양태에서, 코돈 최적화는 인간 세포에서의 더 효율적인 발현으로 이어지는, 인간 번역 기구에 의한 더 효율적인 인식을 위한 것이다. 일부 실시양태에서, 하나 이상의 핵산 변형은 하나 이상의 줄기-루프 영역을 포함하는 5' 플랭킹 서열 또는 3' 플랭킹 서열에서 수행된다. 역전위 플랭킹 요소와 함께 형질전이유전자를 코딩하는 핵산은 1개, 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개, 10개 이상의 핵산 변형을 포함한다.

[0290] 일부 실시양태에서, 역전위된 형질전이유전자는 세포의 수명 동안 안정적으로 발현된다. 일부 실시양태에서, 세포는 골수 세포이다. 일부 실시양태에서, 골수 세포는 단핵구 전구체 세포이다. 일부 실시양태에서, 골수 세포는 미성숙 단핵구이다. 일부 실시양태에서, 단핵구는 미분화된 단핵구이다. 일부 실시양태에서, 골수 세포는 CD14+ 세포이다. 일부 실시양태에서, 골수 세포는 CD16 마커를 발현하지 않는다. 일부 실시양태에서, 골수 세포는 적합한 조건 하에서 3일 초과, 4일 초과, 5일 초과, 6일 초과, 7일 초과, 8일 초과, 9일 초과, 10일 초과, 11일 초과, 12일 초과, 13일 초과, 14일 초과 또는 그 이상의 원하는 기간 동안 기능적 활성을 유지할 수 있다. 적합한 조건은 시험관내 조건 또는 생체내 조건, 또는 이들의 조합을 의미할 수 있다.

[0291] 일부 실시양태에서, 역전위된 형질전이유전자는 약 2일, 약 3일, 약 4일, 약 5일, 약 6일, 약 7일, 약 8일, 약 9일 또는 약 10일 동안 세포에서 안정적으로 발현될 수 있다. 일부 실시양태에서, 역전위된 형질전이유전자는 10일 이상 동안 세포에서 안정적으로 발현된다. 일부 실시양태에서, 역전위된 형질전이유전자는 2주 이상 동안 세포에서 안정적으로 발현된다. 일부 실시양태에서, 역전위된 형질전이유전자는 약 1개월 동안 세포에서 안정적으로 발현된다.

[0292] 일부 실시양태에서, 역전위된 형질전이유전자는 안정한 발현을 위해 변형될 수 있다. 일부 실시양태에서, 역전위된 형질전이유전자는 생체내 침묵에 대한 내성을 갖도록 변형될 수 있다.

[0293] 일부 실시양태에서, 역전위된 형질전이유전자의 발현은 강한 프로모터에 의해 조절될 수 있다. 일부 실시양태에서, 역전위된 형질전이유전자의 발현은 적절히 강한 프로모터에 의해 조절될 수 있다. 일부 실시양태에서, 역전위된 형질전이유전자의 발현은 생체내 환경에서 조절될 수 있는 강한 프로모터에 의해 조절될 수 있다. 일부 실시양태에서, 프로모터는 CMV 프로모터이다. 일부 실시양태에서, 프로모터는 L1-Ta 프로모터이다.

[0294] 일부 실시양태에서, ORF1p는 과발현될 수 있다. 일부 실시양태에서, ORF2는 과발현될 수 있다. 일부 실시양태에서, ORF1p 또는 ORF2p, 또는 이들 둘 다는 과발현된다. 일부 실시양태에서, ORF1의 과발현 시, ORF1p는 ORF1을 과발현하지 않는 세포보다 적어도 1.1배, 1.5배, 2배, 3배, 4배, 5배, 6배, 7배, 8배, 9배, 10배, 12배, 14배, 16배, 18배, 20배, 30배, 40배, 50배, 60배, 70배, 80배, 90배 또는 적어도 100배 더 높다.

[0295] 일부 실시양태에서, ORF2 서열의 과발현 시, ORF2p는 ORF2p를 과발현하지 않는 세포보다 적어도 1.1배, 1.5배, 2배, 3배, 4배, 5배, 6배, 7배, 8배, 9배, 10배, 12배, 14배, 16배, 18배, 20배, 30배, 40배, 50배, 60배, 70배, 80배, 90배 또는 적어도 100배 더 높다.

[0296] *역전위 신뢰도 및 표적 특이성*

[0297] LINE-1 요소는 역전위를 시작하기 위해 그 자신의 mRNA 폴리 A 꼬리에 결합할 수 있다. LINE-1 요소는 바람직하게는 무작위 mRNA에 비해 그 자신의 mRNA를 역전위한다(Dewannieux et al., 2013, 3,000-fold higher LINE-1 retrotransposition as compared to random mRNAs). 추가로, LINE-1 요소는 비특이적 폴리 A 서열을 게놈 내로 통합시킬 수도 있다.

[0298] 한 측면에서, 본원은 증가된 역전위 특이성을 가진 역전위 조성물 및 이를 사용하는 방법을 제공한다. 예를 들어, 높은 특이성을 가진 역전위 조성물은 매우 특이적이고 효율적인 역전사, 및 뒤따르는 표적 세포, 예를 들어, 골수 세포의 게놈 내로의 통합을 위해 사용될 수 있다. 일부 실시양태에서, 본원에서 제공된 역전위 조성물은 통합 또는 역전위 특이성을 증가시키는 하나 이상의 추가 성분을 포함하는 역전위 카세트를 포함한다. 예를 들어, 레트로트랜스포존 카세트는 폴리 A 서열과 ORF2 사이의 비특이적 결합을 능가하기 위해 고친화성 RNA-단백질 상호작용을 허용하는 하나 이상의 추가 요소를 코딩할 수 있다.

[0299] 따라서, 본원은 통합 또는 역전위 효율을 향상시키기 위한 여러 조치를 개시한다.

[0300] 통합 또는 역전위 효율을 향상시키기 위한 한 예시적인 조치는 세포의 외부 조작이다. 세포 내로 전달된 역전위 기구의 엔도뉴클레아제 기능은 세포의 전위 침묵 기구, 예컨대, DNA 복구 경로에 의해 억제될 가능성이 높다. 예를 들어, 핵산을 도입하기 전에 소분자를 사용하여 세포에서 DNA 복구 경로를 조절하거나 억제할 수 있다. 예를 들어, 세포 주기 동기화된 세포 집단이 세포로의 유전자 전달을 증가시키는 것으로 밝혀졌기 때문에, 예컨대, 전기천공으로 핵산을 도입하기 전에 세포 분류 및/또는 동기화를 이용할 수 있다. 세포 분류를 이용하여 세포 유형을 동기화하거나 균질화하고 외생성 핵산의 균일한 전달 및 발현을 증가시킬 수 있다. 균일성은 비-줄기 세포로부터 줄기 세포를 분류함으로써 달성될 수 있다. 통합 또는 역전위 효율을 향상시키기 위한 또 다른 예시적인 조치는 생화학적 활성을 향상시키는 것이다. 예를 들어, 이것은 역전사효소 진행성 또는 DNA 절단(엔도뉴클레아제) 활성을 증가시킴으로써 달성될 수 있다. 통합 또는 역전위 효율을 향상시키기 위한 또 다른 예시적인 조치는 내생성 침묵 기작을 와해시키는 것이다. 예를 들어, 이것은 전체 LINE-1 서열을 상이한 유기체의 LINE-1로 대체함으로써 달성될 수 있다. 통합 또는 역전위 효율을 향상시키기 위한 또 다른 예시적인 조치는 번역 및 리보좀 결합을 향상시키는 것이다. 예를 들어, 이것은 LINE-1 단백질의 발현을 증가시키거나, LINE 단백질 결합 LINE-1 mRNA를 증가시키거나, 리보좀에 결합하는 LINE-1 복합체를 증가시킴으로써 달성될 수 있다. 통합 또는 역전위 효율을 향상시키기 위한 또 다른 예시적인 조치는 핵 이입 또는 체류를 증가시키는 것이다. 예를 들어, 이것은 LINE-1 서열을 핵 체류 신호 서열에 융합시킴으로써 달성될 수 있다. 통합 또는 역전위 효율을 향상시키기 위한 또 다른 예시적인 조치는 서열 특이적 삽입을 향상시키는 것이다. 예를 들어, 이것은 표적화 도메인을 ORF2에 융합시켜 서열 특이적 역전위를 증가시킴으로써 달성될 수 있다.

[0301] 한 실시양태에서, 방법은 LINE-1 ORF의 UTR 서열을 변형시킴으로써 카고 발현의 특이성 및 강건성을 증가시키기 위해 레트로트랜스포존을 향상시키는 단계를 포함한다. 일부 실시양태에서, ORF1 또는 ORF2 코딩 서열의 5' UTR 상류는 ORF 뉴클레아제가 작용할 수 있고 역전위가 일어날 수 있는 특정 부위에서 상동 재조합을 돕는, 게놈 내의 표적 영역의 서열에 상보적인 서열을 포함하도록 더 변형될 수 있다. 일부 실시양태에서, 상동성에 의해 표적 서열에 결합할 수 있는 서열은 길이가 2개 내지 15개의 뉴클레오타이드이다. 일부 실시양태에서, ORF1 mRNA

의 5' UTR에 포함되는, 게놈 표적에 대해 상동성을 가진 서열은 길이가 약 3개의 뉴클레오타이드, 약 4개의 뉴클레오타이드, 약 5개의 뉴클레오타이드, 약 6개의 뉴클레오타이드, 약 7개의 뉴클레오타이드, 약 8개의 뉴클레오타이드, 약 9개의 뉴클레오타이드 또는 약 10개의 뉴클레오타이드일 수 있다. 일부 실시양태에서, 게놈 표적에 대해 상동성을 가진 서열은 길이가 약 12개 또는 약 15개의 뉴클레오타이드이다. 일부 실시양태에서, 게놈 표적에 대해 상동성을 가진 서열은 길이가 적어도 약 10개, 15개, 20개, 25개, 30개, 35개, 40개, 45개, 50개, 55개, 60개, 65개, 70개, 75개, 80개, 85개, 90개, 95개, 100개, 105개, 110개, 115개, 1120개 또는 125개의 뉴클레오타이드이다. 일부 실시양태에서, 게놈 표적에 대해 상동성을 가진 서열은 게놈 내의 각각의 표적 영역과 상보성을 공유하는 약 2개 내지 5개, 약 2개 내지 6개, 약 2개 내지 8개, 약 2개 내지 10개, 또는 약 2개 내지 12개의 연속 뉴클레오타이드를 포함한다. 일부 실시양태에서, 게놈 표적에 대해 상동성을 가진 서열은 게놈 내의 각각의 표적 영역과 상보성을 공유하는 적어도 약 또는 약 10개, 15개, 20개, 25개, 30개, 35개, 40개, 45개, 50개, 55개, 60개, 65개, 70개, 75개, 80개, 85개, 90개, 95개, 100개, 105개, 110개, 115개, 1120개 또는 125개의 연속 뉴클레오타이드이다.

[0302] 일부 실시양태에서, ORF2는 RNA 결합 활성을 포함하는 추가 단백질 도메인과 회합되거나 융합된다. 일부 실시양태에서, 레트로트랜스포존 카세트는 ORF2와 회합되거나 융합된 추가 단백질 도메인에 대한 친화성을 포함하는 동족 RNA 서열을 포함한다. 일부 실시양태에서, ORF2는 MS2-MCP 코트 단백질과 회합되거나 융합된다. 일부 실시양태에서, 레트로트랜스포존 카세트는 MS2-MCP 코트 단백질과 상호작용하는, 3' 또는 5' UTR 서열 내의 MS2 헤어핀 RNA 서열을 추가로 포함한다. 일부 실시양태에서, ORF2는 PP7 코트 단백질과 회합되거나 융합된다. 일부 실시양태에서, 레트로트랜스포존 카세트는 MS2-MCP 코트 단백질과 상호작용하는, 3' 또는 5' UTR 서열 내의 PP7 헤어핀 RNA 서열을 추가로 포함한다. 일부 실시양태에서, 하나 이상의 추가 요소는 하나 이상의 추가 요소를 갖지 않은 레트로트랜스포존 카세트에 비해 역전위 특이성을 적어도 1.5배, 적어도 2배, 적어도 3배, 적어도 4배, 적어도 5배, 적어도 10배, 적어도 20배, 적어도 30배, 적어도 50배, 적어도 100배, 적어도 200배, 적어도 300배, 적어도 500배, 적어도 1000배, 적어도 1500배, 적어도 2000배, 적어도 3000배, 적어도 5000배 이상 증가시킨다.

[0303] DNA 엔도뉴클레아제 도메인은 일련의 피리미딘이 뒤따르는 표적 부위의 일련의 퓨린 3', 즉 $(Py)_n$↓$(Pu)_n$에 대한 특이성을 가진 듯하다. 예시적인 서열은 $(아데노신)_n$↓$(타이미딘)_n$일 수 있다.

[0304] 한 측면에서, 본원은 높은 표적 특이성을 가진 역전위를 사용하는 방법을 제공한다. 일부 실시양태에서, CRISPR-Cas 가이드 RNA 시스템은 부위 지정 역전위의 정밀도를 증가시키기 위해 본원에서 사용된 LINE-레트로트랜스포존 시스템과 조합된다. 일부 실시양태에서, 상기 시스템은 하나 이상의 ORF 결합 서열을 특정 게놈 유전자좌 내로 통합시키기 위해 프라임 편집 가이드 RNA(pegRNA)를 통합시킨다. 일부 실시양태에서, pegRNA는 인간 ORF에 결합하는 서열, 예를 들어, TTTTTA를 부위 특이적 방식으로 통합시킨다. 일부 실시양태에서, CRISPR-Cas 시스템은 Cas9 효소를 포함한다. 일부 실시양태에서, CRISPR-Cas는 Cfp1 효소를 포함한다. 일부 실시양태에서, Cas9는 닉카제(nickase) 시스템과 페어링된 dCas9이다.

[0305] 결과적으로, 본원은 형질전이유전자를 단핵구 또는 대식세포와 같은 골수 세포의 게놈 내로 안정적으로 통합시키는 방법 및 조성물을 제공하는 것으로, 이때 상기 방법은 비-LTR 레트로트랜스포존 시스템을 이용하여 형질전이유전자를 통합시키는 단계를 포함하고, 역전위는 표적 특이성, 높은 정확도 및 신뢰도로 특정 게놈 유전자좌에서 일어난다. 따라서, 일부 실시양태에서, 상기 방법은 하나 이상의 전위 요소에 의해 플랭킹된 적어도 하나의 형질전이유전자, 및 전위 특이성을 증가시키는 하나 이상의 단백질을 코딩하는 하나 이상의 핵산을 가진 시스템을 포함하는 조성물을 세포에게 투여하는 단계를 포함하고/하거나, 역전위와 관련된 하나 이상의 유전자를 변형시키는 단계를 추가로 포함한다.

[0306] 역전위 요소의 3' UTR 영역에 위치한, 형질전이유전자를 포함하는 핵산은 종종 역전위 카세트로서 지칭된다. 따라서, 일부 실시양태에서, 역전위 카세트는 형질전이유전자 및 플랭킹 Alu 전위 요소를 코딩하는 핵산을 포함한다. 역전위 요소는 레트로트랜스포존, 예를 들어, L1-트랜스포존, 예컨대, L1-ORF 단백질, ORF1p 및 ORF2p에 결합하기 위한 서열을 포함한다. ORF 단백질은 역전위를 위해 그 자신의 mRNA 서열에 결합하는 것으로 알려져 있다. 따라서, 역전위 카세트는 형질전이유전자 서열 외부에서 L1-ORF1p 코딩 서열 및 L1-ORF2p 코딩 서열을 코딩하는 서열을 포함하는, 형질전이유전자; 플랭킹 L1-ORF2p 결합 서열, 및/또는 L1-ORF1p 결합 서열을 코딩하는 핵산을 포함한다. 일부 실시양태에서, L1-ORF1 및 L1-ORF2는 ORF1-ORF2 영역간으로서도 지칭되는 스페이서 영역에 의해 산재된다. 일부 실시양태에서, L1-ORF1 및 L1-ORF2 코딩 서열은 형질전이유전자의 코딩 서열에 대하여 반대 방향으로 존재한다. 역전위 카세트는 L1-ORF2 코딩 서열의 하류에서 폴리 A 영역을 포함할 수 있고, 형질

전이유전자 서열은 폴리 A 서열의 하류에 위치한다. L1-ORF2는 엔도뉴클레아제(EN) 및 역전사효소(RT)를 코딩하는 핵산 서열에 이어 폴리 A 서열을 포함한다. 일부 실시양태에서, 본원에 기재된 역전위 카세트 내의 L1-ORF2 서열은 완전한(온전한) 서열이다, 즉 전체 길이 천연(WT) L1-ORF2 서열을 코딩한다. 일부 실시양태에서, 본원에 기재된 역전위 카세트 내의 L1-ORF2 서열은 부분적 또는 변형된 서열을 포함한다.

[0307] 본원에 기재된 시스템은 L1-ORF1p 및 L1-ORF2p를 발현시키기 위해 프로모터를 포함할 수 있다. 일부 실시양태에서, 형질전이유전자 발현은 별도의 프로모터에 의해 유도된다. 일부 실시양태에서, 형질전이유전자 및 ORF는 직렬 방향으로 존재한다. 일부 실시양태에서, 형질전이유전자 및 ORF는 반대 방향으로 존재한다.

[0308] 일부 실시양태에서, 방법은 레트로트랜스포존 카세트 이외에 하나 이상의 요소를 통합시키는 단계를 포함한다. 일부 실시양태에서, 하나 이상의 추가 요소는 이종 단백질의 하나 이상의 도메인을 코딩하는 핵산 서열을 포함한다. 이종 단백질은 서열 특이적 핵산 결합 단백질, 예를 들어, 서열 특이적 DNA 결합 단백질 도메인(DBD)일 수 있다. 일부 실시양태에서, 이종 단백질은 뉴클레아제 또는 이의 단편이다. 일부 실시양태에서, 추가 요소는 이종 단백질의 하나 이상의 뉴클레아제 도메인 또는 이의 단편을 코딩하는 핵산 서열을 포함한다. 일부 실시양태에서, 이종 뉴클레아제 도메인은 감소된 뉴클레아제 활성을 가진다. 일부 실시양태에서, 이종 뉴클레아제 도메인은 불활성 상태가 된다. 일부 실시양태에서, ORF2 뉴클레아제는 불활성 상태가 되는 반면; 이종 단백질의 하나 이상의 뉴클레아제 도메인은 역전위에 대한 특이성을 갖도록 구성된다. 일부 실시양태에서, 이종 단백질의 하나 이상의 뉴클레아제 도메인 또는 이의 단편은 관심 있는 폴리뉴클레오타이드의 역전위 및 통합이 일어나야 하는 게놈 내의 특정 원하는 폴리뉴클레오타이드를 표적화한다. 일부 실시양태에서, 이종 단백질의 하나 이상의 뉴클레아제 도메인은 메가-TAL 뉴클레아제 도메인, TALEN 또는 징크 핑거 뉴클레아제 도메인, 예를 들어, 뉴클레아제 도메인, 예를 들어, FokI 뉴클레아제 도메인에 융합되거나 회합된 메가-TAL, TALE 또는 징크 핑거 도메인을 포함한다. 일부 실시양태에서, 이종 단백질의 하나 이상의 뉴클레아제 도메인은 특정 표적 유전자좌에 대한 특정 가이드 핵산, 예를 들어, 가이드 RNA(gRNA)가 로딩된 CRISPR-Cas 단백질 도메인을 포함한다. 일부 실시양태에서, CRISPR-Cas 단백질은 Cas9, Cas12a, Cas12b, Cas13, CasX 또는 CasY 단백질 도메인이다. 일부 실시양태에서, 이종 단백질의 하나 이상의 뉴클레아제 도메인은 표적 특이성을 가진다.

[0309] 일부 실시양태에서, 추가 뉴클레아제 도메인은 ORF2 도메인 내로 통합될 수 있다. 일부 실시양태에서, 추가 뉴클레아제는 ORF2p 도메인과 융합될 수 있다. 일부 실시양태에서, 추가 뉴클레아제 도메인은 ORF2p에 융합될 수 있고, 이때 ORF2p는 ORF2p 엔도뉴클레아제 도메인 내에 돌연변이를 포함한다. 일부 실시양태에서, 돌연변이는 ORF2p 엔도뉴클레아제 도메인을 불활성화시킨다. 일부 실시양태에서, 돌연변이는 점 돌연변이이다. 일부 실시양태에서, 돌연변이는 결실이다. 일부 실시양태에서, 돌연변이는 삽입이다. 일부 실시양태에서, 돌연변이는 ORF2 엔도뉴클레아제(닉카제) 활성을 제거한다. 일부 실시양태에서, 돌연변이는 ORF2p 엔도뉴클레아제의 DNA 표적 인식을 불활성화시킨다. 일부 실시양태에서, 돌연변이는 ORF2p 뉴클레아제-DNA 인식과 관련된 영역을 커버한다. 일부 실시양태에서, 돌연변이는 ORF2p 엔도뉴클레아제의 DNA 표적 인식을 감소시킨다. 일부 실시양태에서, ORF2p 엔도뉴클레아제 도메인 돌연변이는 단백질의 N-말단 영역에 존재한다. 일부 실시양태에서, ORF2p 엔도뉴클레아제 도메인 돌연변이는 단백질의 보존된 영역에 존재한다. 일부 실시양태에서, ORF2p 엔도뉴클레아제 도메인 돌연변이는 단백질의 보존된 N-말단 영역에 존재한다. 일부 실시양태에서, 돌연변이는 L1 엔도뉴클레아제 도메인 내에 N14 아미노산을 포함한다. 일부 실시양태에서, 돌연변이는 L1 엔도뉴클레아제 도메인 내에 N14 아미노산을 비롯한 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개, 10개, 11개, 12개, 13개, 14개, 15개, 16개, 17개, 18개, 19개, 20개 이상의 보존적 아미노산을 포함한다. 일부 실시양태에서, 돌연변이는 L1 엔도뉴클레아제 내에 E43 아미노산을 포함한다. 일부 실시양태에서, 돌연변이는 L1 엔도뉴클레아제 도메인 내에 E43 아미노산을 비롯한 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개, 10개, 11개, 12개, 13개, 14개, 15개, 16개, 17개, 18개, 19개, 20개 이상의 보존적 아미노산을 포함한다. 일부 실시양태에서, 돌연변이는 L1 엔도뉴클레아제 도메인 내에 N14 또는 E43, 또는 이들의 조합을 비롯한 2개 이상의 아미노산을 포함한다. 일부 실시양태에서, 돌연변이는 L1 엔도뉴클레아제 도메인의 D145를 포함한다. 일부 실시양태에서, 돌연변이는 D145A일 수 있다. 일부 실시양태에서, 돌연변이는 L1 엔도뉴클레아제 도메인의 D205를 포함할 수 있다. 일부 실시양태에서, 돌연변이는 D205G일 수 있다. 일부 실시양태에서, 돌연변이는 L1 엔도뉴클레아제 도메인의 H230을 포함할 수 있다. 일부 실시양태에서, 돌연변이는 L1 엔도뉴클레아제 도메인의 S228을 포함할 수 있다. 일부 실시양태에서, 돌연변이는 S228P일 수 있다.

[0310] 일부 실시양태에서, 돌연변이는 ORF2p 엔도뉴클레아제의 DNA 표적 인식을 적어도 50% 감소시킨다. 일부 실시양태에서, 돌연변이는 ORF2p 엔도뉴클레아제의 DNA 표적 인식을 적어도 60% 감소시킨다. 일부 실시양태에서, 돌연변이는 ORF2p 엔도뉴클레아제의 DNA 표적 인식을 적어도 70% 감소시킨다. 일부 실시양태에서, 돌연변이는 ORF2p

엔도뉴클레아제의 DNA 표적 인식을 80% 감소시킨다. 일부 실시양태에서, 돌연변이는 ORF2p 엔도뉴클레아제의 DNA 표적 인식을 90% 감소시킨다. 일부 실시양태에서, 돌연변이는 ORF2p의 DNA 표적 인식을 95% 감소시킨다. 일부 실시양태에서, 돌연변이는 ORF2p의 DNA 표적 인식을 100% 감소시킨다.

[0311] 일부 실시양태에서, 돌연변이는 결실이다. 일부 실시양태에서, 결실은 완전하다, 즉 L1 엔도뉴클레아제 도메인의 100%가 결실된다. 일부 실시양태에서, 결실은 부분적이다. 일부 실시양태에서, ORF2 엔도뉴클레아제 도메인의 약 98%, 약 95%, 약 94%, 약 93%, 약 92%, 약 91%, 약 90%, 약 85%, 약 80%, 약 75%, 약 70%, 약 65%, 약 60% 또는 약 50%가 결실된다.

[0312] 일부 실시양태에서, 추가 뉴클레아제 도메인은 ORF2 단백질 서열 내로 삽입된다. 일부 실시양태에서, ORF2 엔도뉴클레아제 도메인은 결실되고 이종 단백질의 엔도뉴클레아제 도메인으로 대체된다. 일부 실시양태에서, ORF2 엔도뉴클레아제는 부분적으로 결실되고 이종 단백질의 엔도뉴클레아제 도메인으로 대체된다. 이종 단백질의 엔도뉴클레아제 도메인은 메가-TAL 뉴클레아제 도메인일 수 있다. 이종 단백질의 엔도뉴클레아제 도메인은 TALEN일 수 있다. 이종 단백질의 엔도뉴클레아제 도메인은 유전자좌에 대한 특이적 gRNA가 로딩된 Cas9일 수 있다.

[0313] 일부 실시양태에서, 엔도뉴클레아제는 (i) 게놈 상의 특정 표적을 갖고 (ii) 절단 부위에서 5'-P 및 3'-OH 말단을 생성하는 엔도뉴클레아제이다.

[0314] 일부 실시양태에서, 이종 단백질의 추가 엔도뉴클레아제 도메인은 관련된 레트로트랜스포존의 엔도뉴클레아제 도메인이다.

[0315] 일부 실시양태에서, 이종 단백질의 엔도뉴클레아제 도메인은 특정 부위를 표적화하도록 조작된 박테리아 엔도뉴클레아제를 포함할 수 있다. 일부 실시양태에서, 이종 단백질의 엔도뉴클레아제 도메인은 귀소(homing) 엔도뉴클레아제의 도메인 또는 이의 단편을 포함할 수 있다. 일부 실시양태에서, 엔도뉴클레아제는 귀소 엔도뉴클레아제이다. 일부 실시양태에서, 귀소 엔도뉴클레아제는 조작된 LAGLIDADG 귀소 엔도뉴클레아제(LHE)(서열번호 66으로서 개시된 "LAGLIDADG") 또는 이의 단편이다. 일부 실시양태에서, 추가 엔도뉴클레아제는 제한 엔도뉴클레아제, Cre, Cas TAL 또는 이의 단편일 수 있다. 일부 실시양태에서, 엔도뉴클레아제는 II 군 인트론 코딩된 단백질(리보자임(ribozyme)) 또는 이의 단편을 포함할 수 있다.

[0316] 추가/이종 엔도뉴클레아제로 인해 특이적 DNA 표적화 능력이 부여된, 이전 단락에서 논의된 바와 같은 조작 또는 변형된 L1-ORF2p는 게놈 내로의 형질전이유전자의 표적화된 안정한 통합을 유도하는 데 매우 유리할 것으로 예상된다. 조작된 L1-ORF2p는 세포에서 발현될 때 조작되지 않은 천연 L1-ORF2p를 사용한 경우보다 훨씬 더 감소된 오프-표적 효과를 생성할 수 있다. 일부 실시양태에서, 조작된 L1-ORF2p는 오프-표적 효과를 생성하지 않는다.

[0317] 일부 실시양태에서, 조작 또는 변형된 L1-ORF2p는 통상의 $(Py)_n$↓$(Pu)_n$ 부위 이외의 인식 부위를 표적화한다. 일부 실시양태에서, 조작된 L1-ORF2p는 $(Py)_n$↓$(Pu)_n$ 부위, 예를 들어, TTTT/AA 부위, 예컨대, 하이브리드 표적 부위를 포함하는 인식 부위를 표적화한다. 일부 실시양태에서, 조작된 L1-ORF2p는 통상의 L1-ORF2 $(Py)_n$↓$(Pu)_n$ 부위, 예를 들어, TTTT/AAG, 또는 TTTT/AAC, 또는 TTTT/AAT, TTTT/AAA, GTTTT/AA, CTTTT/AA, ATTTT/AA, 또는 TTTTT/AA 이외에 적어도 하나의 뉴클레오타드를 가진 인식 부위를 표적화한다. 일부 실시양태에서, 조작된 L1-ORF2p는 통상의 L1-ORF2p $(Py)_n$↓$(Pu)_n$ 부위 이외의 인식 부위를 표적화한다. 일부 실시양태에서, 조작된 L1-ORF2p는 통상의 L1-ORF2p $(Py)_n$↓$(Pu)_n$ 부위 이외의 인식 부위를 표적화한다. 일부 실시양태에서, 조작된 L1-ORF2p는 길이가 4개, 5개, 6개, 7개, 8개, 9개, 10개 이상의 뉴클레오타이드인 인식 부위를 표적화한다. 일부 실시양태에서, 조작 또는 변형된 L1-ORF2p 인식 부위는 11개, 12개, 13개, 14개, 15개, 16개, 17개, 18개, 19개, 20개 이상의 뉴클레오타이드일 수 있다.

[0318] 조작된 L1-ORF2p는 번역 후 그 자신의 mRNA에 결합하고 고효율로 역전사하는 그의 능력을 보유하도록 조작될 수 있다. 일부 실시양태에서, 조작된 L1-ORF2p는 천연(WT) L1-ORF2p에 비해 향상된 역전사 효율을 가진다.

[0319] 일부 실시양태에서, 역전위 요소를 포함하는 시스템은 비특이적 역전위를 감소시키는 유전자 변형을 추가로 포함한다. 일부 실시양태에서, 유전자 변형은 L1-ORF2p를 코딩하는 서열을 포함할 수 있다. 일부 실시양태에서, 변형은 ORF2p가 표적 게놈 DNA에 결합하는 것을 돕는 단백질에 결합하는 데 필수적인 하나 이상의 아미노산의 돌연변이를 포함할 수 있다. ORF2p가 표적 게놈 DNA에 결합하는 것을 돕는 단백질은 염색질-ORF 상호작용체의 일부일 수 있다. 일부 실시양태에서, 변형은 ORF2p DNA 엔도뉴클레아제 활성을 돕는 단백질에 결합하는 데 필수적인 하나 이상의 아미노산을 포함할 수 있다. 일부 실시양태에서, 변형은 ORF2p RT 활성을 돕는 단백질에 결합

하는 데 필수적인 하나 이상의 아미노산을 포함할 수 있다. 일부 실시양태에서, 변형은 단백질과 ORF2p의 회합이 변경되도록 ORF2p 상의 단백질 결합 부위에서의 변형을 포함할 수 있고, 이때 상기 단백질과 ORF2p의 결합은 염색질에의 결합을 위해 요구된다. 일부 실시양태에서, 변형은 단백질과 ORF2p의 회합이 변형의 부재 하에서의 회합보다 더 엄격하고/하거나 특이적이도록 ORF2p 상의 단백질 결합 부위에서의 변형을 포함할 수 있다. 일부 실시양태에서, 단백질 결합 부위에서 ORF2p 코딩 서열의 변형으로 인한 ORF2p와 단백질의 변경된 회합의 결과로서, ORF2p와 표적 DNA의 결합은 증가된 특이성을 가진다. 일부 실시양태에서, 변형은 ORF2p 염색질 상호작용체의 일부인 하나 이상의 단백질과 ORF2의 결합을 감소시킬 수 있다.

[0320] 일부 실시양태에서, 유전자 변형은 ORF2p의 PIP 도메인에 있을 수 있다.

[0321] 일부 실시양태에서, 유전자 변형은 ORF2p에 결합하고 ORF2p의 인식, 결합, 엔도뉴클레아제 또는 RT 활성을 돕는 단백질을 코딩하는 하나 이상의 유전자에 있을 수 있다. 일부 실시양태에서, 유전자 변형은 각각의 단백질에 대한 ORF2p 상호작용 부위, 또는 단백질과 ORF2p의 상호작용 또는 ORF2p와 표적 DNA의 상호작용에 영향을 미치는 부위에서 PCNA, PARP1, PABP, MCM, TOP1, RPA, PURA, PURB, RUVBL2, NAP1, ZCCHC3, UPF1 또는 MOV10 단백질을 코딩하는 하나 이상의 유전자에 있을 수 있다. 일부 실시양태에서, 변형은 ORF2p 상호작용 부위, 또는 단백질과 ORF2p의 상호작용 또는 ORF2p와 표적 DNA의 상호작용에 영향을 미치는 부위에서 PCNA의 ORF2p 결합 도메인에 있을 수 있다. 일부 실시양태에서, 변형은 TOP1의 ORF2p 결합 도메인에 있을 수 있다. 일부 실시양태에서, 변형은 RPA의 ORF2p 결합 도메인에 있을 수 있다. 일부 실시양태에서, 변형은 ORF2p 상호작용 부위, 또는 단백질과 ORF2p의 상호작용 또는 ORF2p와 표적 DNA의 상호작용에 영향을 미치는 부위에서 PARP1의 ORF2p 결합 도메인에 있을 수 있다. 일부 실시양태에서, 변형은 ORF2p 상호작용 부위, 또는 단백질과 ORF2p의 상호작용 또는 ORF2p와 표적 DNA의 상호작용에 영향을 미치는 부위에서 PABP(예를 들어, PABPC1)의 ORF2p 결합 도메인에 있을 수 있다. 일부 실시양태에서, 유전자 변형은 MCM 유전자에 있을 수 있다. 일부 실시양태에서, 유전자 변형은 ORF2p 상호작용 부위, 또는 단백질과 ORF2p의 상호작용 또는 ORF2p와 표적 DNA의 상호작용에 영향을 미치는 부위에서 MCM3 단백질을 코딩하는 유전자에 있을 수 있다. 일부 실시양태에서, 유전자 변형은 ORF2p 상호작용 부위, 또는 단백질과 ORF2p의 상호작용 또는 ORF2p와 표적 DNA의 상호작용에 영향을 미치는 부위에서 MCM5 단백질을 코딩하는 유전자에 있을 수 있다. 일부 실시양태에서, 유전자 변형은 ORF2p 상호작용 부위, 또는 단백질과 ORF2p의 상호작용 또는 ORF2p와 표적 DNA의 상호작용에 영향을 미치는 부위에서 MCM6 단백질을 코딩하는 유전자에 있을 수 있다. 일부 실시양태에서, 유전자 변형은 ORF2p 상호작용 부위, 또는 단백질과 ORF2p의 상호작용 또는 ORF2p와 표적 DNA의 상호작용에 영향을 미치는 부위에서 MEPCE 단백질을 코딩하는 유전자에 있을 수 있다. 일부 실시양태에서, 유전자 변형은 ORF2p 상호작용 부위, 또는 단백질과 ORF2p의 상호작용 또는 ORF2p와 표적 DNA의 상호작용에 영향을 미치는 부위에서 RUVBL1 또는 RUVBL2 단백질을 코딩하는 유전자에 있을 수 있다. 일부 실시양태에서, 유전자 변형은 ORF2p 상호작용 부위, 또는 단백질과 ORF2p의 상호작용 또는 ORF2p와 표적 DNA의 상호작용에 영향을 미치는 부위에서 TROVE 단백질을 코딩하는 유전자에 있을 수 있다.

[0322] 일부 실시양태에서, 본원에 개시된 역전위 시스템은 역전사의 신뢰도를 증가시키는 하나 이상의 요소를 포함한다.

[0323] 일부 실시양태에서, L1-ORF2 RT 도메인은 변형된다. 일부 실시양태에서, 변형은 신뢰도 증가, 진행성 증가, DNA-RNA 기질 친화성 증가; 또는 RNase H 활성의 불활성화 중 하나 이상을 포함한다.

[0324] 일부 실시양태에서, 변형은 RT의 신뢰도가 증가하도록 하나 이상의 돌연변이를 L1-ORF2의 RT 도메인 내로 도입하는 것을 포함한다. 일부 실시양태에서, 돌연변이는 점 돌연변이를 포함한다. 일부 실시양태에서, 돌연변이는 L1-ORF2p RT 도메인에서 1개, 2개, 3개, 4개, 5개, 6개 이상의 아미노산의 치환과 같은 변경을 포함한다. 일부 실시양태에서, 돌연변이는 L1-ORF2p RT 도메인에서 하나 이상의 아미노산, 예를 들어, 1개, 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개, 10개 이상의 아미노산의 결실을 포함한다. 일부 실시양태에서, 돌연변이는 삽입결실 돌연변이를 포함할 수 있다. 일부 실시양태에서, 돌연변이는 프레임시프트 돌연변이를 포함할 수 있다.

[0325] 일부 실시양태에서, 변형은 제2 단백질의 추가 RT 도메인 또는 이의 단편의 포함을 포함할 수 있다. 일부 실시양태에서, 제2 단백질은 바이러스 역전사효소이다. 일부 실시양태에서, 제2 단백질은 비-바이러스 역전사효소이다. 일부 실시양태에서, 제2 단백질은 역전위 요소이다. 일부 실시양태에서, 제2 단백질은 비-LTR 역전위 요소이다. 일부 실시양태에서, 제2 단백질은 II 군 인트론 단백질이다. 일부 실시양태에서, II 군 인트론은 TGIRTII이다. 일부 실시양태에서, 제2 단백질은 Cas 닉카제이고, 이때 역전위 시스템은 가이드 RNA를 도입하는 것을 추가로 포함한다. 일부 실시양태에서, 제2 단백질은 Cas9 엔도뉴클레아제이고, 이때 역전위 시스템은 가이드 RNA를 도입하는 것을 추가로 포함한다. 일부 실시양태에서, 제2 단백질 또는 이의 단편은 L1-ORF2 RT 도메인 또는

변형된 L1-ORF2 RT 도메인의 N-말단에 융합된다. 일부 실시양태에서, 제2 단백질 또는 이의 단편은 L1-ORF2 RT 도메인 또는 변형된 L1-ORF2 RT 도메인의 C-말단에 융합된다.

[0326] 일부 실시양태에서, 전체 길이 WT L1-ORF2p RT 도메인 이외에 제2 단백질의 추가 RT 도메인 또는 이의 단편이 역전위 시스템에 포함된다. 일부 실시양태에서, 제2 단백질의 추가 RT 도메인 또는 이의 단편은 변형된(조작된) L1-ORF2p RT 도메인 또는 이의 단편의 존재 하에서 포함되고, 이때 변형(또는 조작)은 천연 또는 WT ORF2p에 비해 변형된 L1-ORF2p RT의 L1-ORF2p RT 진행성, 안정성 및/또는 신뢰도를 향상시키기 위한 돌연변이를 포함할 수 있다.

[0327] 일부 실시양태에서, 역전사효소 도메인은 다른 역요소 또는 II 군 인트론, 예컨대, TGIRTII의 다른 더 높은 진행성 및 높은 신뢰도 RT 도메인으로 대체될 수 있다.

[0328] 일부 실시양태에서, 변형은 제2 단백질의 추가 RT 도메인 또는 이의 단편과의 융합을 포함할 수 있다. 일부 실시양태에서, 제2 단백질은 역요소를 포함할 수 있다. 제2 단백질의 추가 RT 도메인 또는 이의 단편은 융합된 L1-ORF2p RT 도메인의 역전사의 신뢰도를 증가시키도록 구성된다. 일부 실시양태에서, 상기 추가 RT 도메인 또는 이의 단편을 코딩하는 핵산은 천연 또는 WT L1-ORF2 코딩 서열에 융합된다. 일부 실시양태에서, 제2 단백질의 추가 RT 도메인 또는 이의 단편을 코딩하는 핵산은 변형된 L1-ORF2 코딩 서열에 융합된다. 일부 실시양태에서, 변형은 융합된 RT의 신뢰도가 증가되도록 하나 이상의 돌연변이를 L1-ORF2의 RT 도메인 또는 이의 단편 내로 도입하는 것을 포함한다. 일부 실시양태에서, L1-ORF2의 RT 도메인 또는 이의 단편 내의 돌연변이는 점 돌연변이를 포함한다. 일부 실시양태에서, 돌연변이는 L1-ORF2p RT 도메인에서 1개, 2개, 3개, 4개, 5개, 6개 이상의 아미노산의 변경, 예컨대, 치환을 포함한다. 일부 실시양태에서, 돌연변이는 L1-ORF2p RT 도메인에서 하나 이상의 아미노산, 예를 들어, 1개, 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개, 10개 이상의 아미노산의 결실을 포함한다. 일부 실시양태에서, 돌연변이는 삽입결실 돌연변이를 포함할 수 있다. 일부 실시양태에서, 돌연변이는 프레임시프트 돌연변이를 포함할 수 있다.

[0329] 일부 실시양태에서, 변형된 L1-ORF2p RT 도메인은 WT L1-ORF2p RT 도메인보다 증가된 진행성을 가진다.

[0330] 일부 실시양태에서, 변형된 L1-ORF2p RT 도메인은 WT L1-ORF2p RT 도메인에 비해 적어도 10% 더 높은 진행성 및/또는 신뢰도를 가진다. 일부 실시양태에서, 변형된 L1-ORF2p RT 도메인은 WT L1-ORF2p RT 도메인에 비해 적어도 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 110%, 120%, 150%, 200%, 300%, 400%, 500%, 1000% 이상 더 높은 진행성 및/또는 신뢰도를 가진다. 일부 실시양태에서, 변형된 RT는 6 kb 초과의 핵산 스트레치를 처리할 수 있다. 일부 실시양태에서, 변형된 RT는 7 kb 초과의 핵산 스트레치를 처리할 수 있다. 일부 실시양태에서, 변형된 RT는 8 kb 초과의 핵산 스트레치를 처리할 수 있다. 일부 실시양태에서, 변형된 RT는 9 kb 초과의 핵산 스트레치를 처리할 수 있다. 일부 실시양태에서, 변형된 RT는 10 kb 초과의 핵산 스트레치를 처리할 수 있다.

[0331] B. II 군 인트론 및 리보자임

[0332] II 군 효소는 전구체 RNA를 자가스플라이싱하여, 절제된 인트론 올가미(lariat) RNA를 생성하는 이동성 리보자임이다. 인트론은 역전사효소를 코딩한다. 역전사효소는 정방향 및 역방향 스플라이싱 및 그 후 DNA로의 통합된 인트론 RNA의 전환을 위해 RNA를 안정화시킬 수 있다.

[0333] II 군 RNA는 400개 내지 800개의 뉴클레오타이드에 걸쳐 있는 보존된 2차 구조를 특징으로 한다. 2차 구조는 6개의 도메인 DI 내지 VI에 의해 형성되고, 바퀴와 유사한 구조로 조직화되고, 이때 도메인은 중심점으로부터 방사된다. 도메인은 상호작용하여 보존된 3차 구조를 형성하고, 이 3차 구조는 멀리 떨어진 서열과 함께 활성 부위를 형성한다. 활성 부위는 스플라이스 부위 및 분기점 잔기 뉴클레오타이드에 결합하고 Mg2+ 양이온과 함께 스플라이싱의 촉매작용을 활성화시킨다. DV 도메인은 보존된 촉매 AGC 및 AY 돌출부를 가진 활성 부위 내에 있고, 이 영역들 둘 다가 촉매작용에 필요한 Mg2+ 이온에 결합한다. DI는 카파 및 제타 모티프에 의해 분리된 상부 및 하부 절반을 가진 가장 큰 도메인이다. 하부 절반은 활성 부위와 관련된 ε' 모티브를 함유한다. 상부 절반은 활성 부위에서 5' 및 3' 엑손에 결합하는 서열 요소를 함유한다. DIV는 인트론-코딩된 단백질(IEP)에 대한 고친화성 결합 부위를 함유하는 5' 말단 근처에서 서브도메인 IVa를 가진 IEP를 코딩한다. II 군 인트론은 보존된 5' 말단 서열 및 3' 말단 서열, 즉 각각 GUGYG 및 AY를 가진다.

[0334] II 군 RNA 인트론은 표적 프라이밍 역전사를 통해 관심 있는 서열을 DNA 내로 역전위시키는 데 이용될 수 있다. II 군 RNA 인트론에 의한 이 전위 과정은 종종 역귀소(retrohoming)로서 지칭된다. II 군 인트론은 인트론 RNA와 DNA 표적 서열의 염기 페어링으로 DNA 표적 부위를 인식하고, 인트론 내에 담지된 특정 서열을 원하는 DNA

부위로 재표적화하도록 변형될 수 있다.

[0335] 일부 실시양태에서, 본원에 기재된 역전위 방법 및 조성물은 II 군 인트론 서열, 변형된 II 군 인트론 서열 또는 이의 단편을 포함할 수 있다. 예시적인 II 군 IEP(성숙효소)는 인간 세포에서 작용하는 박테리아, 진균, 효모 IEP를 포함하나, 이들로 제한되지 않는다. 특히, 뉴클레아제는 프라이밍 및 역전사를 위해 또 다른 RT에 의해 활용될 수 있는 DNA의 절단 부위에서 3'-OH를 남긴다. 예시적인 II 군 성숙효소는 TGIRT(열적으로 안정한 II 군 인트론 성숙효소)일 수 있다.

[0336] 본원에 기재된 여러 측면의 하나 이상의 실시양태에서, 핵산 구축물은 RNA를 포함한다. 본 개시내용의 여러 측면의 하나 이상의 실시양태에서, 핵산 구축물은 RNA이다. 본 개시내용의 여러 측면의 하나 이상의 실시양태에서, 핵산 구축물은 mRNA이다. 한 측면에서, mRNA는 이종 유전자 또는 이의 일부의 서열을 포함하고, 이때 이종 유전자 또는 이의 일부는 폴리펩티드 또는 단백질을 코딩한다. 일부 실시양태에서, mRNA는 융합 단백질을 코딩하는 서열을 포함한다. 일부 실시양태에서, mRNA는 재조합 단백질을 코딩하는 서열을 포함한다. 일부 실시양태에서, mRNA는 합성 단백질을 코딩하는 서열을 포함한다. 일부 실시양태에서, 핵산은 하나 이상의 이종 단백질, 하나 이상의 재조합 단백질 또는 하나 이상의 합성 단백질, 또는 이들의 조합을 코딩하는 하나 이상의 서열을 포함한다. 일부 실시양태에서, 핵산은 합성 단백질 또는 재조합 단백질을 포함하는 하나 이상의 이종 단백질을 코딩하는 하나 이상의 서열을 포함한다. 일부 실시양태에서, 합성 또는 재조합 단백질은 재조합 융합 단백질이다.

[0337] C. 부위 지정 편집 및/또는 통합효소를 포함하는 레트로트랜스포존 시스템

[0338] 한 측면에서, 본원은 가이드 RNA 및 Cas 단백질의 도움을 받아 LINE 결합 서열을 게놈 내로 부위 특이적으로 pegRNA 매개 통합시키는 것에 대해 변형 후 높은 표적 특이성을 가진 역전위를 이용하는 방법을 제공한다. 일부 실시양태에서, CRISPR-Cas 가이드 RNA 시스템은 부위 지정 역전위의 정밀도를 증가시키기 위해 본원에서 사용된 LINE-레트로트랜스포존 시스템과 조합되고; 예를 들어, 상기 시스템은 하나 이상의 ORF 결합 서열을 특정 게놈 유전자좌 내로 통합시키기 위해 프라임 편집 가이드 RNA(pegRNA)를 통합시킨다. 일부 실시양태에서, pegRNA는 인간 ORF에 결합하는 서열, 예를 들어, TTTTTA를 부위 특이적 방식으로 통합시킨다. 일부 실시양태에서, CRISPR-Cas 시스템은 Cas9 효소를 포함한다. 일부 실시양태에서 CRISPR-Cas는 Cfp1 효소를 포함한다. 일부 실시양태에서 Cas9는 닉카제 시스템과 페어링된 dCas9이다.

[0339] 일부 실시양태에서, 본원에 설명된 레트로트랜스포존 시스템은 (i) LINE1 레트로트랜스포존 요소, 및 (ii) 통합효소 시스템 또는 이의 일부를 포함한다. 일부 통합효소 시스템은 이중 가닥 DNA를 부위 특이적으로 통합시킬 수 있다. 이중 가닥 DNA 전달 및/또는 게놈 내로의 통합을 우회하기 위해, 본원은 통합효소 또는 이의 단편이 재조합 ORF 단백질 내에 통합되어 있거나, 특정 게놈 부위를 인식하는 통합효소 또는 이의 단편을 코딩하는 별도의 핵산(예를 들어, mRNA)로서 별도로 전달되고 통합효소의 특이성에 의해 안내되는 정확한 위치에서 세포 또는 유기체의 게놈 내로의 카고 서열의 LINE 1 역전사 및 삽입과 커플링되는 재조합 하이브리드 시스템을 제공한다. 이것은 몇 가지 대안적인 방식으로 달성될 수 있다. 일부 실시양태에서, 카고 서열은 마찬가지로 동일한 통합효소에 의해 인식되는, 게놈 내의 랜딩 부위로 카고를 끌어들이기 위해 통합효소에 의해 인식되고 사용되는 부착 부위를 포함한다. 통합효소는 단일 가닥을 절단할 수 있다. 통합효소 DNA 인식 부위, 즉 게놈 랜딩 서열은 10개의 뉴클레오타이드 길이, 예를 들어, 12개, 14개, 16개, 18개, 20개, 22개, 24개, 26개, 28개, 30개 이상의 뉴클레오타이드 길이를 가짐으로써, 임의의 다른 시스템보다 더 큰 특이성을 부여할 수 있다. 통합효소는 ORF가 통합효소 특정 게놈 부위에서 카고 서열을 역전사하고 통합시킬 수 있도록 절두될 수 있거나 다른 방식으로 돌연변이될 수 있다. 반대로, ORF 단백질은 또한 통합효소가 통합효소에 의해 우선적으로 인식되는 게놈 통합 서열("게놈 랜딩 서열 또는 부위"로서도 지칭됨)을 인식할 수 있도록 RNA 인식 부위에서 돌연변이될 수 있다. 대안적인 실시양태에서, 통합효소는 별도의 폴리뉴클레오타이드에 의해 코딩되고, CRISPR Cas 시스템 및 가이드 RNA에 의해 닉킹될 수 있는 부위로 안내될 수 있고, 4개의 뉴클레오타이드를 포함하는 ORF 결합 부위를 추가로 포함하는 통합효소 랜딩 서열이 도입될 수 있고, 그 후 통합효소가 랜딩 서열에 대한 부착 서열을 포함하는 카고 서열에 끌려들어온 후, LINE1 활성이 통합효소 시스템에 의해 특정된 부위에서 게놈 통합을 유발한다. 게놈 내로의 이중 가닥 DNA 통합을 유발하는 통합효소의 임의의 촉매 활성은 돌연변이되거나 절두되거나, 다른 방식으로 침묵된다.

[0340] 본 개시내용의 여러 측면의 하나 이상의 실시양태에서, 핵산 구축물은 진핵 세포에서 발현되도록 개발된다. 일부 실시양태에서, 핵산 구축물은 인간 세포에서 발현되도록 개발된다. 일부 실시양태에서, 핵산 구축물은 조혈 세포에서 발현되도록 개발된다. 일부 실시양태에서, 핵산 구축물은 골수 세포에서 발현되도록 개발된다. 일부

실시양태에서, 골수 세포는 인간 세포이다.

[0341] **II. 코딩된 단백질의 발현을 향상시키는 방법을 위한 핵산 구축물의 변형**

[0342] 본 개시내용의 일부 측면에서, 재조합 핵산은 핵산의 서열에 의해 코딩된 단백질의 향상된 발현을 위해 변형된다. 코딩된 단백질의 향상된 발현은 핵산 안정성, 번역 효율 및 번역된 단백질의 안정성의 함수일 수 있다. 합성 재조합 단백질 또는 이의 단편일 수 있는 외생성 또는 이종 단백질을 코딩하는 메신저 RNA의 안정성과 같은 핵산 안정성을 부여할 수 있는 다수의 변형이 핵산 구축물의 디자인에 포함될 것으로 본원에서 예상된다.

[0343] 일부 실시양태에서, 핵산은 합성 또는 재조합 융합 단백질을 포함하는 하나 이상의 이종 단백질을 코딩하는 하나 이상의 서열을 포함하는 mRNA이다.

[0344] 일부 실시양태에서, mRNA 반감기를 증가시키기 위해 재조합 또는 융합 단백질을 코딩하는 서열을 포함하는 mRNA에서 하나 이상의 변형을 만든다.

[0345] *엑소뉴클레아제에 의한 5' 및 3' 분해를 차단하는 구조 요소: 5'-캡 및 3' UTR 변형*

[0346] 적절한 5'-캡 구조는 기능성 메신저 RNA의 합성에 중요하다. 일부 실시양태에서, 5'-캡은 핵산의 5' 말단에서 GpppG로서 배열된 구아노신 트리포스페이트를 포함한다. 일부 실시양태에서, mRNA는 5' 7-메틸구아노신 캡인 m7-GpppG를 포함한다. 5' 7-메틸구아노신 캡은 mRNA 번역 효율을 증가시키고 mRNA 5'-3' 엑소뉴클레아제의 분해를 방해한다. 일부 실시양태에서, mRNA는 "항-역" 캡 유사체(ARCA, $m^{7,3'-O}$GpppG)를 포함한다. 그러나, 번역 효율은 ARCA의 사용에 의해 현저히 증가될 수 있다. 일부 실시양태에서, 구아노신 캡은 Cap 0 구조이다. 일부 실시양태에서, 구아노신 캡은 Cap 1 구조이다. mRNA 캡은 단백질 합성의 캡 의존적 시작이라는 그의 필수적인 역할 이외에 5'에서 3'으로의 엑소뉴클레아제 절단의 보호기, 및 전구-mRNA 스플라이싱, 폴리아데닐화 및 핵 이출을 위해 단백질 인자를 동원하기 위한 고유 식별자로서도 작용한다. 이것은 단백질 합성을 시작하는 시작 인자의 동원 및 번역 동안 mRNA의 5'에서 3'으로의 루핑을 위한 앵커로서 작용한다. Cap 0 구조를 생성하기 위해 세 가지 효소 활성, 즉 RNA 트리포스파타제(TPase), RNA 구아닐릴트랜스퍼라제(GTase) 및 구아닌-N7 메틸트랜스퍼라제(구아닌-N7 MTase)가 요구된다. 이 효소 활성들 각각은 신생 RNA의 5' 트리포스페이트를 Cap 0 구조로 전환시키는 데 있어서 필수적인 단계를 수행한다. RNA TPase는 5' 트리포스페이트로부터 γ-포스페이트를 제거하여 5' 디포스페이트 RNA를 생성한다. GTase는 라이신-GMP 공유 중간체를 통해 GMP 기를 GTP로부터 5' 디포스페이트로 전달한다. 그 다음, 구아닌-N7 MTase는 메틸 기를 구아닌 캡의 N7 아민에 추가하여 캡 0 구조를 형성한다. Cap 1 구조의 경우, m7G 특이적 2'O 메틸트랜스퍼라제(2'O MTase)는 리보스의 2'O 위치에서 +1 리보뉴클레오타이드를 메틸화하여 캡 1 구조를 생성한다. 핵 RNA 캡핑 효소는 C-말단 헵타드 반복부의 인산화된 Ser5에서 RNA 중합효소 II 복합체의 중합효소 서브유닛과 상호작용한다. RNA 구아닌-N7 메틸트랜스퍼라제는 RNA 중합효소 II 인산화된 헵타드 반복부와도 상호작용한다. 일부 실시양태에서, 캡은 G-사중체 캡이다.

[0347] 일부 실시양태에서, mRNA는 시험관내 전사(IVT)에 의해 합성된다. 일부 실시양태에서, mRNA 합성과 캡핑은 한 단계에서 수행될 수 있다. 캡핑은 IVT와 동일한 반응 혼합물에서 일어날 수 있다. 일부 실시양태에서, mRNA 합성과 캡핑은 별도의 단계에서 수행될 수 있다. 따라서, IVT에 의해 형성된 mRNA는 정제된 후 캡핑된다.

[0348] 일부 실시양태에서, 관심 있는 단백질 또는 폴리펩티드를 코딩하는 하나 이상의 서열을 포함하는 핵산 구축물, 예를 들어, mRNA 구축물은 내생성 5'-3' 엑소리보뉴클레아제, 예를 들어, Xrn1에 의한 mRNA의 분해를 보호하거나, 방지하거나, 억제하거나 감소시키는 요소를 포함하도록 디자인될 수 있다. Xrn1은 5' 일인산화된 RNA를 분해하는 정상적인 RNA 붕괴 경로의 세포 효소이다. 그러나, 일부 바이러스 RNA 구조 요소는 특히 이러한 RNase에 대한 내성을 나타내는 것으로 밝혀졌다(예를 들어, 'xrRNA'로서 지칭되는, 플라비바이러스 sfRNA의 Xrn1 내성 구조). 예를 들어, 모기 매개 플라비바이러스(MBFV) 게놈은 그의 3'-비번역 영역(UTR) 내에 Xrn1의 진행을 차단하는 별도의 RNA 구조를 함유한다. 이 RNA 요소는 보조 단백질을 사용하지 않으면서 Xrn1을 차단하기에 충분하다. xrRNA는 xrRNA의 하류에 위치한 바이러스 RNA가 분해로부터 보호되도록 정해진 위치에서 효소를 정지시킨다. 예를 들어, 지카(Zika) 바이러스 또는 머레이 발리(Murray Valley) 뇌염 바이러스의 xrRNA는 MBFV 구조 전체에 걸쳐 보존된 뉴클레오타이드 세트를 요구하는 독특하고 복잡한 접힘을 생성하는 3-방향 연접 및 다중 슈도노트 상호작용을 포함한다. xrRNA는 xrRNA의 하류에 위치한 바이러스 RNA가 분해로부터 보호되도록 정해진 위치에서 효소를 정지시킨다. RNA의 5' 말단은 상기 접힘의 고리 유사 구조를 통과하고 Xrn1 유사 엑소뉴클레아제로부터 보호된 상태를 유지하는 것으로 여겨진다.

[0349] 일부 실시양태에서, 관심 있는 단백질을 코딩하는 하나 이상의 서열을 포함하는 핵산 구축물은 그의 내부에 통합된 하나 이상의 xrRNA 구조를 포함할 수 있다. 일부 실시양태에서, xrRNA는 하나 이상의 바이러스 xrRNA 서열

의 3' UTR의 보존된 영역을 가진 뉴클레오타이드 스트레치이다. 일부 실시양태에서, 1개, 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개, 10개 이상의 xrRNA 요소가 핵산 구축물 내로 통합된다. 일부 실시양태에서, 2개 이상의 xrRNA 요소가 핵산 구축물 내로 직렬로 통합된다. 일부 실시양태에서, xrRNA는 보존된 서열 또는 이의 단편 또는 이의 변형을 포함하는 하나 이상의 영역을 포함한다. 일부 실시양태에서, xrRNA는 레트로트랜스포존 요소의 3' UTR에 배치된다. 일부 실시양태에서, xrRNA는 하나 이상의 단백질 또는 폴리펩티드를 코딩하는 서열의 상류에 배치된다. 일부 실시양태에서, xrRNA는 레트로트랜스포존 요소의 3' UTR, 예컨대, ORF2 서열, 및 하나 이상의 단백질 또는 폴리펩티드를 코딩하는 서열의 상류에 배치된다.

[0350] 일부 실시양태에서, xrRNA 구조는 MBFV xrRNA 서열, 또는 이와 적어도 90% 동일한 서열을 포함한다. 일부 실시양태에서, xrRNA 구조는 진드기 매개 플라비바이러스(TBFV) xrRNA 서열, 또는 이와 적어도 90% 동일한 서열을 포함한다. 일부 실시양태에서, xrRNA 구조는 진드기 매개 플라비바이러스(TBFV) xrRNA 서열, 또는 이와 적어도 90% 동일한 서열을 포함한다. 일부 실시양태에서, xrRNA 구조는 진드기 매개 플라비바이러스(TBFV) xrRNA 서열, 또는 이와 적어도 90% 동일한 서열을 포함한다. 일부 실시양태에서, xrRNA 구조는 알려지지 않은 절지동물 벡터 플라비바이러스(NKVFV) 구성원으로부터의 xrRNA 서열, 또는 이와 적어도 90% 동일한 서열을 포함한다. 일부 실시양태에서, xrRNA 구조는 곤충 특이적 플라비바이러스(ISFV) 구성원으로부터의 xrRNA 서열, 또는 이와 적어도 90% 동일한 서열을 포함한다. 일부 실시양태에서, xrRNA 구조는 지카바이러스 xrRNA 서열, 또는 이와 적어도 90% 동일한 서열을 포함한다. 임의의 공지된 xrRNA 구조 요소 또는 도출될 수 있는 이의 자명하지 않은 변경이 본원에 기재된 목적을 위해 사용될 수 있음이 예상된다.

[0351] 상이한 유기체로부터의 몇몇 메신저 RNA는 5'-3' 엑소뉴클레아제에 대한 내성을 나타내는 하나 이상의 슈도노트 구조를 나타낸다. 슈도노트의 여러 상이한 접힘 위상구조가 존재하지만, 슈도노트는 최소한 단일 가닥 영역 또는 루프에 의해 연결된 2개의 나선형 분절로 구성된 RNA 구조이다.

[0352] *폴리 A 꼬리 변형*

[0353] mRNA의 3' UTR에서 폴리 A 구조는 mRNA 반감기의 중요한 조절제이다. 폴리 A 꼬리의 3' 말단의 탈아데닐화는 세포내 mRNA 분해의 첫 번째 단계이다. 일부 실시양태에서, mRNA 구축물의 폴리 A 꼬리의 길이는 mRNA 코딩 영역에 의해 코딩된 단백질의 발현 및 mRNA 안정성을 최대화하기 위해 핵심적으로 고려되고 디자인된다. 일부 실시양태에서, 핵산 구축물은 하나 이상의 폴리 A 서열을 포함한다. 일부 실시양태에서, 하나 이상의 단백질 또는 폴리펩티드를 코딩하는 서열의 3' UTR에 있는 폴리 A 서열은 20개 내지 200개의 아데노신 핵염기를 포함한다. 일부 실시양태에서, 폴리 A 서열은 30개 내지 200개의 아데노신 핵염기를 포함한다. 일부 실시양태에서, 폴리 A 서열은 50개 내지 200개의 아데노신 핵염기를 포함한다. 일부 실시양태에서, 폴리 A 서열은 80개 내지 200개의 아데노신 핵염기를 포함한다. 일부 실시양태에서, 하나 이상의 단백질 또는 폴리펩티드를 코딩하는 서열을 포함하는 mRNA 분절은 약 180개의 아데노신 핵염기, 또는 약 140개의 아데노신 핵염기, 또는 약 120개의 아데노신 핵염기를 포함하는 폴리 A 꼬리를 가진 3' UTR을 포함한다. 일부 실시양태에서, 폴리 A 꼬리는 약 122개의 아데노신 핵염기를 포함한다. 일부 실시양태에서, 폴리 A 서열은 50개의 아데노신 핵염기를 포함한다. 일부 실시양태에서, 폴리 A 서열은 30개의 아데노신 핵염기를 포함한다. 일부 실시양태에서, 폴리 A 꼬리의 아데노신 핵염기는 개재 비-아데노신 염기와 함께 또는 이러한 비-아데노신 염기 없이 직렬로 배치된다. 일부 실시양태에서, 특정 엑소뉴클레아제에 대한 추가 내성을 부여하는 하나 이상의 비-아데노신 핵염기는 폴리 A 꼬리 내로 통합된다.

[0354] 일부 실시양태에서, 구축물의 폴리 A 꼬리 내의 아데노신 스트레치는 하나 이상의 비-아데노신(A) 핵염기를 포함한다. 일부 실시양태에서, 비-A 핵염기는 폴리 A 3' 말단 영역에서 -3, -2, -1 및/또는 +1 위치에 존재한다. 일부 실시양태에서, 비-A 염기는 구아노신(G) 또는 사이토신(C) 또는 우라실 염기(U)를 포함한다. 일부 실시양태에서, 비-A 염기는 G이다. 일부 실시양태에서, 하나 초과의 비-A 염기, 예를 들어, GG는 직렬로 존재한다. 일부 실시양태에서, 하나 이상의 비-A 염기를 가진 폴리 A 꼬리의 3' 말단에서의 변형은 폴리 A 꼬리에서 A 염기 적층을 파괴하도록 유도된다. 폴리 A 염기 적층은 다양한 탈아데닐화 효소들에 의한 탈아데닐화를 촉진하므로, -AAAG, -AAAGA 또는 -AAAGGA로 끝나는 폴리 A 꼬리의 3' 말단은 탈아데닐화에 대한 안정성을 부여하는 데 효과적이다. 일부 유기체에서, 폴리 A 서열에 개재된 GC 서열은 3'-5' 엑소뉴클레아제 매개 붕괴를 효과적으로 보여주는 것으로 확인된다. 본원에서 고려되는 변형은 개재 비-A 잔기, 또는 3' 말단에서 폴리 A 스트레치에 개재된 비-A 잔기 이중체를 포함한다.

[0355] 일부 실시양태에서, 3' 말단을 이용하여 엑소뉴클레아제 활성을 효과적으로 정지시키거나 늦추는 삼중체 구조가 3' UTR 내로 도입된다.

[0356] 일부 실시양태에서, 전술된 변형을 가진 mRNA는 연장된 반감기를 갖고 변형되지 않은 mRNA보다 더 긴 기간에 걸쳐 안정한 발현을 나타낸다. 일부 실시양태에서, mRNA는 2일, 3일, 4일, 5일, 6일, 7일, 8일, 9일 또는 10일 이상 동안 안정적으로 발현하고, mRNA 또는 이의 단백질 생성물은 생체내에서 검출 가능하다. 일부 실시양태에서, mRNA는 생체내에서 3일, 4일, 5일, 6일, 7일, 8일, 9일, 10일, 11일, 12일, 13일, 14일 또는 15일까지 검출된다. 일부 실시양태에서, mRNA의 단백질 생성물은 생체내에서 3일, 4일, 5일, 6일, 7일, 8일, 9일, 10일, 11일, 12일, 13일, 14일, 15일, 16일, 17일, 18일, 20일, 25일 또는 30일까지 검출된다.

[0357] *circRNA 및 tectoRNA*

[0358] 원형 RNA는 합성 단백질 쇄, 예컨대, 긴 다중 반복 단백질 쇄를 유도하기 위한 메신저 RNA로서 사용되는 안정한 형태의 RNA의 디자인 및 생성에 유용하다. 원형 RNA(circRNA)를 만드는 방법은 거의 없다. 이 방법은 인트론의 두 절반이 전사된 mRNA의 말단에 위치하는 경우 인트론 그 자체가 스플라이싱되고 라이게이션된 생성물을 남기도록 RNA 리가제를 사용하고 분할 자가스플라이싱 인트론을 사용하는 RNA 말단의 단백질 매개 라이게이션을 포함한다(도 3a). 또 다른 기법은 라이게이션될 RNA 말단이 올리고뉴클레오타이드에 의해 함께 유지될 때 RNA 리가제로서 작용하는 T4 DNA 리가제의 능력에 의존한다. 이 기법들 둘 다가 비효율적이고 다량의 효소를 필요로 한다. 세 번째 기법은 반응을 수행하는 대부분의 인트론 서열들이 원의 일부로 남아 있어야 하는 경우 I 군 인트론의 고리화 또는 원형화 활성을 이용한다. I 군 인트론은 촉매 코어를 형성하는 일련의 보존된 RNA 줄기 루프를 함유하는 복잡한 2차 및 3차 구조 세트를 공유한다. 이 인트론들 중 대다수가 시험관내에서 자가스플라이싱하고 보조 단백질 인자 없이 스플라이싱하여 RNA로서 2개의 라이게이션된 엑손을 형성할 수 있다. I 군 자가촉매 반응에 의해 생성된 생성물은 (1) 5' 스플라이스 부위에서 하류 엑손의 3' 스플라이스 부위에서 라이게이션된 상류 엑손 및 (2) 추가 가역적 자가촉매작용을 거쳐 원형 인트론을 형성할 수 있는 선형 인트론이다. 이러한 고도로 구조화된 큰 핵산 서열의 존재는 그 기법에 의해 원형으로 만들 수 있는 RNA 서열의 유형을 심각하게 제한한다. 또한, 인트론의 촉매 활성이 남아 원형 RNA의 구조와 기능을 방해할 수 있다.

[0359] RNA의 말단을 더 가깝게 유지함으로써 반응 속도 및 이에 따른 전체 효율을 증가시키는 것이 유용하다. 선행 연구는 상보적 RNA 서열 3' 및 5'를 mRNA의 말단에 포함시켜, 이 서열들의 하이브리드화 시 mRNA의 말단이 상기 상보적 서열을 갖지 않는 경우에 비해 전반적으로 더 빠른 속도로 라이게이션 또는 자가스플라이싱 반응을 겪을 수 있도록 더 가까이 인접하게 함으로써 이를 달성하였다. 이들은 원형화 반응의 자가스플라이싱 버전의 상동성 아암(도 3a)으로서 지칭된다. 이러한 하이브리드화 전략의 주요 문제점은 상동성 아암 중 어느 하나에 상보적인 서열이 코딩 영역 내에 있는 경우, 하이브리드화가 실제로 스플라이싱 반응을 억제할 것이고 상기 아암이 각각의 새로운 코딩 영역에 대해 최적화될 필요가 있을 것이라는 점이다. 본원에 기재된 이 전략의 대안은 3차원 구조로 접혀 서열과 무관한 안정한 결합 상호작용을 형성하는 RNA 서열의 사용이다.

[0360] 비-왓슨-크릭 RNA 3차 상호작용을 이용하여 자가어셈블리를 할 수 있는 RNA 분자로서 정의된 'tectoRNA' 분자 유닛을 구축할 수 있다. 이러한 유형의 3차 상호작용의 이용은 양이온 농도(예를 들어, Mg^{2+}) 및/또는 적합한 온도를 조작하고 모듈식으로 디자인된 '선택제' RNA 분자를 사용함으로써 어셈블리 과정을 제어하고 조절할 수 있게 한다. 1차원 어레이의 자가어셈블리를 위해, 각각의 나선형 아암에서 상호작용 모듈과 함께 4-방향 연접을 포함하는 기본 모듈식 유닛을 디자인하였다. 일부 실시양태에서, 상호작용 모듈은 GAAA 루프 또는 특정 GAAA 루프 수용체이다. 각각의 tectoRNA는 4개의 루프-수용체 상호작용의 형성을 통해 2개의 다른 tectoRNA와 상호작용할 수 있으며, 이때 2개가 각각의 파트너 분자와 상호작용할 수 있다.

[0361] 일부 실시양태에서, tectoRNA 구조는 적절히 선택되고, 엑손 및 인트론을 포함하는 RNA 내로 통합되어 circRNA를 형성한다. 일부 실시양태에서, 통합은 라이게이션과 같은 잘 알려진 분자생물학 기법에 의해 수행된다. 일부 실시양태에서, tectoRNA는 고온에서 안정한 구조를 형성한다. tectoRNA 구조는 내부 RNA 서열과 경쟁하지 않음으로써, 고효율 원형화 및 스플라이싱을 생성한다.

[0362] circRNA는 이전 단락들 중 임의의 단락에 기재된 코딩 서열을 포함할 수 있다. 예를 들어, 이것은 테더링(tethering) 또는 수용체 분자를 포함하는 융합 단백질을 코딩하는 서열을 포함할 수 있다. 수용체는 식세포 수용체 융합 단백질일 수 있다.

[0363] 일부 실시양태에서, 인트론은 자가스플라이싱 인트론이다.

[0364] 일부 실시양태에서, circRNA에 대한 스캐폴딩 영역으로서도 지칭되는, 3차 구조를 가진 말단 영역은 약 30개의 뉴클레오타이드 내지 약 100개의 뉴클레오타이드 길이를 가진다. 일부 실시양태에서, 3차 구조 모티프는 약 45개의 뉴클레오타이드, 약 50개의 뉴클레오타이드, 약 55개의 뉴클레오타이드, 약 60개의 뉴클레오타이드, 약 65

개의 뉴클레오타이드, 약 70개의 뉴클레오타이드 또는 약 75개의 뉴클레오타이드 길이를 가진다. 일부 실시양태에서, 3차 모티프는 고온에서 형성된다. 일부 실시양태에서, 3차 모티프는 안정하다.

[0365] 일부 실시양태에서, 본원에 기재된 바와 같이 하나 이상의 변형을 갖고 하나 이상의 단백질 또는 폴리펩티드를 코딩하는 하나 이상의 서열을 포함하는 핵산 구축물은 생체내로 투여될 때 안정하다. 일부 실시양태에서, 핵산은 mRNA이다. 일부 실시양태에서, 하나 이상의 단백질 또는 폴리펩티드를 코딩하는 하나 이상의 서열을 포함하는 mRNA는 2일 초과, 3일 초과, 4일 초과, 5일 초과, 6일 초과, 7일 초과, 8일 초과, 9일 초과, 10일 초과, 11일 초과, 12일 초과, 13일 초과, 14일 초과, 15일 초과, 16일 초과, 17일 초과, 18일 초과, 19일 초과 또는 20일 초과의 기간 동안 생체내에서 안정하다. 일부 실시양태에서, mRNA의 서열에 의해 코딩된 단백질은 3일, 4일, 5일, 6일, 7일, 8일, 9일, 10일, 11일, 12일, 13일, 14일, 15일, 16일, 17일, 18일, 19일 또는 20일 초과의 기간 동안 생체내에서 검출될 수 있다. 일부 실시양태에서, mRNA의 서열에 의해 코딩된 단백질은 mRNA가 투여된 후 약 7일 동안 생체내에서 검출될 수 있다. 일부 실시양태에서, mRNA의 서열에 의해 코딩된 단백질은 mRNA가 투여된 후 약 14일 동안 생체내에서 검출될 수 있다. 일부 실시양태에서, mRNA의 서열에 의해 코딩된 단백질은 mRNA가 투여된 후 약 21일 동안 생체내에서 검출될 수 있다. 일부 실시양태에서, mRNA의 서열에 의해 코딩된 단백질은 mRNA가 투여된 후 약 30일 동안 생체내에서 검출될 수 있다. 일부 실시양태에서, mRNA의 서열에 의해 코딩된 단백질은 mRNA가 투여된 후 약 30일 초과의 기간 동안 생체내에서 검출될 수 있다.

[0366] 일부 측면에서, 역전위의 발현을 향상시키기 위해 세포 내로의 핵산 흡수 또는 도입을 향상시키는 것이 고려된다. 방법 중 하나는 예를 들어, 플라스미드 벡터 구축물의 경우 형질감염을 통한, 또는 전기천공 또는 핵산 분자를 세포 내로 전달하는 데 적절하게 이용될 수 있는 임의의 다른 수단을 통한 핵산의 도입을 시작하기 위해 균질한 세포 집단을 수득하는 단계를 포함한다. 일부 실시양태에서, 세포 주기 동기화가 추구될 수 있다. 세포 주기 동기화는 특정 공통 표현형에 대해 세포를 분류함으로써 달성될 수 있다. 일부 실시양태에서, 세포 집단은 특정 단계에서 모든 세포들의 세포 주기 진행을 정지시킬 수 있는 시약으로 처리될 수 있다. 예시적인 시약은 www.tocris.com/cell-biology/cell-cycle-inhibitors 또는 www.scbt.com/browse/chemicals-Other-Chemicals-cell-cycle-arresting-compounds와 같은 상업용 데이터베이스에서 발견될 수 있다. 예를 들어, 몇몇을 나열하자면, G1 기에서 세포 주기를 억제하는 이트라코나졸 또는 노코다졸, 또는 G0/G1 기에서 세포 주기를 정지시키는 시약, 예를 들어, 5-[(4-에틸페닐)메틸렌]-2-티옥소-4-티아졸리디논(화합물 10058-F4)(Tocris Bioscience); 또는 G2M 세포 주기 차단제, 예컨대, G2M, G1 또는 S 기에서 세포 주기를 차단하는 AZD 5438(화학명, 4-[2-메틸-1-(1-메틸에틸)-1*H*-이미다졸-5-일]-*N*-[4-(메틸설포닐)페닐]-2-피리미딘아민). 사이클로스포린, 하이드록시우레아, 타이미딘은 세포 주기 정지를 야기할 수 있는 잘 알려진 시약이다. 일부 시약은 세포 상태를 비가역적으로 변경시킬 수 있거나 세포에 대한 독성을 나타낼 수 있다. 세포 유형에 따라 전기천공 또는 형질감염 전에 약 2시간 내지 16시간 동안 세포의 혈청 박탈도 세포 동기화를 위한 용이한 가역적 전략일 수 있다.

[0367] 일부 실시양태에서, 역전위 효율은 본원에 기재된 바와 같은 역전위 구축물에 의해 형질감염되었거나 전기천공된 세포에서 DNA 이중 가닥 절단의 생성을 촉진하고/하거나 DNA 복구 기구를 조절함으로써 증가될 수 있다. 이 기법들의 적용은 이 방법에 의한 핵산 서열의 안정적인 통합을 위해 생체외에서 유전자 조작될 세포의 최종 용도에 따라 제한될 수 있다. 일부 경우, 통합된 핵산에 의해 코딩된 단백질 또는 전사체의 강력한 발현이 소정의 시간 동안 결과로서 예상되는 경우 이러한 기법의 이용이 고려될 수 있다. 이중 가닥 절단을 세포 내로 도입하는 방법은 세포를 짧은 시간 동안 약 0.1 Gy 이하의 조절된 이온화 방사선에 노출시키는 단계를 포함한다.

[0368] 일부 실시양태에서, LINE-1 매개 역전위의 효율은 역전사효소가 작용할 윈도우를 증가시키기 위해 DNA 복구 단백질의 소분자 억제제로 세포를 처리함으로써 증가될 수 있다. DNA 복구 단백질의 예시적인 소분자 억제제는 벤즈아미드(Benzamide)(CAS 55-21-0), 올라파립(Olaparib)(Lynparza)(CAS 763113-22-0), 루카파립(Rucaparib)(Clovis-AG014699, PF-01367338 Pfizer), 니라파립(Niraparib)(MK-827 Tesaro)(CAS 1038915-60-4); 벨리파립(Veliparib)(ABT-888 Abbvie)(CAS 912444-00-9); 캄프토테신(Camptothecin)(CPT)(CAS 7689-03-4); 이리노테칸(Irinotecan)(CAS 100286-90-6); 토포테칸(Topotecan)(Hycamtin® GlaxoSmithKline)(CAS 123948-87-8); NSC 19630(CAS 72835-26-8); NSC 617145(CAS 203115-63-3); ML216(CAS 1430213-30-1); 6-하이드록시DL-도파(CAS 21373-30-8); D-103; D-G23; DIDS(CAS 67483-13-0); B02(CAS 1290541-46-6); RI-1(CAS 415713-60-9); RI-2(CAS 1417162-36-7); 스트렙토니그린(Streptonigrin)(SN)(CAS 3930-19-6)일 수 있다.

[0369] *III. 핵산 카고:*

[0370] *A. 형질전이유전자*

[0371] 한 측면에서, 세포의 게놈 내로 삽입될 이종 핵산 서열인 형질전이유전자 또는 비코딩 서열은 mRNA로서 전달된

다. mRNA는 약 100개, 200개, 300개, 400개, 500개, 600개, 700개, 800개, 900개, 1000개, 1100개, 1200개, 1300개, 1400개, 1500개, 1600개, 1700개, 1800개, 1900개, 2000개, 3000개, 4000개, 5000개, 6000개, 7000개, 8000개, 9000개, 10,000개 초과의 염기를 포함할 수 있다. 일부 실시양태에서, mRNA는 길이가 10,000개 초과의 염기일 수 있다. 일부 실시양태에서, mRNA는 길이가 약 11,000개의 염기일 수 있다. 일부 실시양태에서, mRNA는 길이가 약 12,000개의 염기일 수 있다. 일부 실시양태에서, mRNA는 융합 단백질을 코딩하는 형질전이유전자 서열을 포함한다. 일부 실시양태에서, 핵산은 플라스미드로서 전달된다.

[0372] 일부 실시양태에서, 핵산은 형질감염에 의해 세포 내로 전달된다. 일부 실시양태에서, 핵산은 전기천공에 의해 세포 내로 전달된다. 일부 실시양태에서, 형질감염 또는 전기천공은 세포 내로의 핵산의 도입을 향상시키기 위해 1회 초과의 빈도로 반복된다.

[0373] 식세포 또는 테더링 수용체(PR) 융합 단백질(CFP)을 코딩하는 재조합 핵산의 레트로트랜스포존 매개 안정한 통합이 본원에서 고려된다. 일부 실시양태에서, CFP는 막횡단 도메인, 및 세포내 신호전달 도메인을 포함하는 세포내 도메인을 포함하는 PR 서브유닛; 및 표적 세포의 항원에 특이적인 항원 결합 도메인을 포함하는 세포외 도메인을 포함하고; 이때 막횡단 도메인과 세포외 도메인은 작동 가능하게 연결된다.

[0374] 일부 실시양태에서, 핵산은 CD5 결합 도메인을 포함하는 세포외 도메인, 및 세포외 도메인에 작동 가능하게 연결된 막횡단 도메인을 포함하는 키메라 융합 단백질(CFP)을 코딩하는 서열을 포함한다. 일부 실시양태에서, CD5 결합 도메인은 항체의 항원 결합 단편, Fab 단편, scFv 도메인 또는 sdAb 도메인과 같은 CD5 결합 단백질이다. 일부 실시양태에서, CD5 결합 도메인은 (i) EIQLVQSGGGLVKPGGSVRISCAASGYTFTNYGMNWVRQAPGKGLEWMGWINTHTGEPTYADSFKGRFTFSLDDSKNTAYLQINSLRAEDTAVYFCTRRGYDWYFDVWGQGTTVTV(서열번호 1)과 적어도 90% 서열 동일성을 가진 가변 중쇄(VH) 서열; 및 (ii) DIQMTQSPSSLSASVGDRVTITCRASQDINSYLSWFQQKPGKAPKTLIYRANRLESGVPSRFSGSGSGTDYTLTISSLQYEDFGIYYCQQYDESPWTFGGGTKLEIK(서열번호 2)와 적어도 90% 서열 동일성을 가진 가변 경쇄(VL) 서열을 포함하는 scFv를 포함한다. 일부 실시양태에서, CFP는 세포내 도메인을 추가로 포함하고, 이때 세포내 도메인은 하나 이상의 세포내 신호전달 도메인을 포함하고, 세포내 도메인을 포함하는 야생형 단백질은 세포외 도메인을 포함하지 않는다. 일부 실시양태에서, 하나 이상의 세포내 신호전달 도메인은 식세포 신호전달 도메인을 포함한다. 일부 실시양태에서, 식세포작용 신호전달 도메인은 Megf10, MerTk, FcαR 및 Bai1 이외의 수용체로부터 유래한 세포내 신호전달 도메인을 포함한다. 일부 실시양태에서, 식세포작용 신호전달 도메인은 FcγR, FcαR 또는 FcεR로부터 유래한 세포내 신호전달 도메인을 포함한다. 일부 실시양태에서, 식세포작용 신호전달 도메인은 LYCRRLKIQVRKAAITSYEKSDGVYTGLSTRNQETYETLKHEKPP(서열번호 67)과 적어도 90% 서열 동일성을 가진 세포내 신호전달 도메인을 포함한다. 일부 실시양태에서, 하나 이상의 세포내 신호전달 도메인은 염증촉진성 신호전달 도메인을 추가로 포함한다. 일부 실시양태에서, 염증촉진성 신호전달 도메인은 PI3-키나제(PI3K) 동원 도메인을 포함한다. 일부 실시양태에서, 염증촉진성 신호전달 도메인은 YEDMRGILYAAPQLRSIRGQPGPNHEEDADSYENM(서열번호 4)와 적어도 90% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, 염증촉진성 신호전달 도메인은 CD40의 세포내 신호전달 도메인으로부터 유래한다. 일부 실시양태에서, 염증촉진성 신호전달 도메인은 KVAKKPTNKAPHPKQEPQEINFPDDLPGSNTAAPVQETLHGCQPVTQEDGKESRISVQERQ(서열번호 68)과 적어도 90% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, 막횡단 도메인은 CD8 막횡단 도메인을 포함한다. 일부 실시양태에서, 막횡단 도메인은 IYIWAPLAGTCGVLLLSLVIT(서열번호 6)과 적어도 90% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, 세포외 도메인은 CD8로부터 유래한 힌지 도메인을 추가로 포함하고, 이때 힌지 도메인은 막횡단 도메인 및 CD5 결합 도메인에 작동 가능하게 연결된다. 일부 실시양태에서, 세포외 도메인은 ALSNSIMYFSHFVPVFLPAKPTTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLD(서열번호 7)과 적어도 90% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, CFP는 CD5에 특이적으로 결합하는 scFv 및 CD8로부터 유래한 힌지 도메인을 포함하는 세포외 도메인; CD28로부터 유래한 힌지 도메인 또는 CD68로부터의 세포외 도메인의 적어도 일부; CD8 막횡단 도메인, CD28 막횡단 도메인 또는 CD68 막횡단 도메인; 및 적어도 2개의 세포내 신호전달 도메인을 포함하는 세포내 도메인을 포함하고, 이때 적어도 2개의 세포내 신호전달 도메인은 FcγR 또는 FcεR로부터 유래한 제1 세포내 신호전달 도메인, 및 PI3K 동원 도메인을 포함하거나 CD40으로부터 유래한 제2 세포내 신호전달 도메인을 포함한다. 일부 실시양태에서, 재조합 폴리핵산은 mRNA 또는 circRNA이다. 일부 실시양태에서, 핵산은 골수 세포 내로 전달된다. 일부 실시양태에서, 핵산은 CD14+ 세포, CD14+CD16- 세포, M0 대식세포, M2 대식세포, M1 대식세포 또는 모자이크 골수 세포/대식세포 내로 전달된다. 일부 실시양태에서, 융합 단백질은 하기 서열과 적어도 90% 서열 동일성을 가진 서열을 포함한다:

EIQLVQSGGGLVKPGGSVRISCAASGYTFTNYGMNWVRQAPGKGLEWMGWINTHTGEPTYAD
SFKGRFTFSLDDSKNTAYLQINSLRAEDTAVYFCTRRGYDWYFDVWGQGTTVTVSSGGGGSGG
GGSGGGGSDIQMTQSPSSLSASVGDRVTITCRASQDINSYLSWFQQKPGKAPKTLIYRANRLESG
VPSRFSGSGSGTDYTLTISSLQYEDFGIYYCQQYDESPWTFGGGTKLEIKSGGGGSGALSNSIMYF
SHFVPVFLPAKPTTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDIYIWAPLAGTCGV
LLLSLVITLYCRRLKIQVRKAAITSYEKSDGVYTGLSTRNQETYETLKHEKPPQGSGSYEDMRGI
LYAAPQLRSIRGQPGPNHEEDADSYENM(**서열번호** 69).

[0375] 일부 실시양태에서, 융합 단백질은 하기 서열과 적어도 90% 서열 동일성을 가진 서열을 포함한다:

EIQLVQSGGGLVKPGGSVRISCAASGYTFTNYGMNWVRQAPGKGLEWMGWINTHTGEPTYAD
SFKGRFTFSLDDSKNTAYLQINSLRAEDTAVYFCTRRGYDWYFDVWGQGTTVTVSSGGGGSGG
GGSGGGGSDIQMTQSPSSLSASVGDRVTITCRASQDINSYLSWFQQKPGKAPKTLIYRANRLESG
VPSRFSGSGSGTDYTLTISSLQYEDFGIYYCQQYDESPWTFGGGTKLEIKSGGGGSGALSNSIMYF
SHFVPVFLPAKPTTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDIYIWAPLAGTCGV
LLLSLVITLYCRLKIQVRKAAITSYEKSDGVYTGLSTRNQETYETLKHEKPPQKKVAKKPTNKAP
HPKQEPQEINFPDDLPGSNTAAPVQETLHGCQPVTQEDGKESRISVQERQ(**서열번호** 70), **또는**
EIQLVQSGGGLVKPGGSVRISCAASGYTFTNYGMNWVRQAPGKGLEWMGWINTHTGEPTYAD
SFKGRFTFSLDDSKNTAYLQINSLRAEDTAVYFCTRRGYDWYFDVWGQGTTVTVSSGGGGSGG
GGSGGGGSDIQMTQSPSSLSASVGDRVTITCRASQDINSYLSWFQQKPGKAPKTLIYRANRLESG
VPSRFSGSGSGTDYTLTISSLQYEDFGIYYCQQYDESPWTFGGGTKLEIKSGGGGSGALSNSIMYF
SHFVPVFLPAKPTTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDIYIWAPLAGTCGV
LLLSLVITLYCRRLKIQVRKAAITSYEKSDGVYTGLSTRNQETYETLKHEKPPQKKVAKKPTNKA
PHPKQEPQEINFPDDLPGSNTAAPVQETLHGCQPVTQEDGKESRISVQERQ(**서열번호** 71).

[0376]

[0377] 일부 실시양태에서, 융합 단백질은 막횡단 단백질, 세포내 단백질 또는 세포내 단백질이다. 한 실시양태에서, 융합 단백질은 단핵구, 대식세포, 수지상 세포 또는 이들의 전구체로부터 선택된 면역 세포, 예를 들어, 골수 세포의 기능을 향상시키도록 유도된다. 한 실시양태에서, 융합 단백질은 면역 세포의 세포 기능, 예컨대, 식세포작용을 증강시킨다. 본 개시내용은 기재된 방법 및 조성물을 사용함으로써 발현될 수 있는 형질전이유전자에 의해 제한되지 않는다. 이 단락에 표시된 형질전이유전자는 예시하기 위한 것이다.

[0378] 본원은 식세포의 게놈 내로의 안정한 통합을 위한 예시적인 형질전이유전자 후보를 제공한다. 한 실시양태에서, 형질전이유전자는 식세포 수용체(PR) 융합 단백질(CFP)을 코딩하는 재조합 핵산이다. 재조합 핵산은 (i) 막횡단 도메인 및 (ii) 식세포 수용체 세포내 신호전달 도메인을 포함하는 세포내 도메인을 포함하는 PR 서브유닛; 및 표적 세포의 항원에 특이적인 세포외 항원 결합 도메인을 갖고; 이때 막횡단 도메인과 세포외 항원 결합 도메인은 융합된 수용체의 세포외 항원 결합 도메인에 의한 표적에의 항원 결합이 식세포 수용체의 세포내 신호전달 도메인에서 활성화되도록 작동 가능하게 연결된다. 일부 실시양태에서, 재조합 핵산은 키메라 항원 수용체를 코딩한다. 일부 실시양태에서, 키메라 항원 수용체는 키메라 항원 수용체(식세포작용)(CAR-P)이다. 일부 실시양태에서, 융합 단백질은 항-식세포 신호를 잠그기 위한 재조합 단백질이다. 일부 실시양태에서, 융합 단백질은 식세포작용을 향상시키는 키메라 단백질이다. 일부 실시양태에서, 키메라 단백질은 활성 식세포작용 신호 전달도입 도메인을 포함하는 세포내 도메인을 가진다. 일부 실시양태에서, 키메라 단백질은 이것이 발현되는 식세포의 염증 잠재력을 향상시킴으로써 식세포 잠재력을 향상시킨다. 일부 실시양태에서, 형질전이유전자는 표적 세포에서 항원과의 접촉에 의해 활성화되는 키메라 단백질을 발현하도록 디자인되고, 그 결과 식세포는 표적 세포를 포식하고 표적 세포를 사멸시킨다.

[0379] 융합 단백질과 관련하여 사용된 용어 "스페이서" 또는 "링커"는 융합 단백질의 단백질 도메인을 연결하는 펩티드 서열을 지칭한다. 일반적으로, 스페이서는 단백질 또는 RNA 서열 사이에 어느 정도의 최소 거리 또는 다른 공간적 관계를 연결하거나 보존하는 것 이외에 특정 생물학적 활성을 갖지 않는다. 그러나, 일부 실시양태에서, 스페이서의 구성 아미노산은 분자의 접힘, 순 전하 또는 소수성과 같은 분자의 일부 성질에 영향을 미치도록 선택될 수 있다. 본 개시내용의 실시양태에서 사용하기에 적합한 링커는 당업자에게 잘 공지되어 있고, 직쇄 또는 분지쇄 탄소 링커, 헤테로환형 탄소 링커 또는 펩티드 링커를 포함하나, 이들로 제한되지 않는다. 링커는 일부

실시양태에서 각각의 항원성 펩티드가 적절히 접히는 것을 보장하기에 충분한 거리로 2개의 항원성 펩티드를 분리하는 데 사용된다. 예시적인 펩티드 링커 서열은 유연한 확장된 입체구조를 채택하고 정돈된 2차 구조를 발생시키는 경향을 나타내지 않는다. 유연한 단백질 영역의 전형적인 아미노산은 Gly, Asn 및 Ser을 포함한다. Gly, Asn 및 Ser을 함유하는 아미노산 서열의 사실상 임의의 순열은 링커 서열에 대한 상기 기준을 충족시킬 것으로 예상된다. Thr 및 Ala과 같은 다른 거의 중성 아미노산도 링커 서열에 사용될 수 있다.

[0380] 식세포의 면역 잠재력을 향상시키기 위해 발현될 수 있는 형질전이유전자에 의해 코딩된 다양한 예시적인 단백질들이 이하에 기재되어 있다. 이것은 완전한 목록이 아니라, 본 개시내용의 범위 내에서 형질전이유전자 디자인을 위한 예시적인 목록으로서 사용된다.

[0381] 일부 실시양태에서, PSP 서브유닛은 식세포 수용체의 막횡단(TM) 도메인을 포함한다.

[0382] 일부 실시양태에서, PSP 서브유닛은 식세포 수용체의 ICD 도메인을 포함한다.

[0383] 일부 실시양태에서, 재조합 핵산에 의해 코딩된 ICD는 렉틴(lectin), 덱틴(dectin) 1, 만노스 수용체(CD206), 스캐빈저 수용체 A1(SRA1), MARCO, CD36, CD163, MSR1, SCARA3, COLEC12, SCARA5, SCARB1, SCARB2, CD68, OLR1, SCARF1, SCARF2, CXCL16, STAB1, STAB2, SRCRB4D, SSC5D, CD205, CD207, CD209, RAGE, CD14, CD64, F4/80, CCR2, CX3CR1, CSF1R, Tie2, HuCRIg(L) 및 CD169 수용체로 구성된 군으로부터 선택된 도메인을 포함한다.

[0384] 일부 실시양태에서, ICD는 렉틴, 덱틴 1, 만노스 수용체(CD206), 스캐빈저 수용체 A1(SRA1), MARCO(콜라겐 구조를 가진 대식세포 수용체, 별칭: SRA6, SCARA2), CD36(트롬보스폰딘 수용체, 별칭: 스캐빈저 수용체 클래스 B, 구성원 3), CD163(스캐빈저 수용체, 시스테인 풍부 1형), MSR1, SCARA3, COLEC12(별칭: C형 렉틴을 가진 스캐빈저 수용체, SCARA4, 또는 콜렉틴 12), SCARA5, SCARB1, SCARB2, CD68(SCARD, 마이크로시알린), OLR1(산화된 저밀도 지단백질 수용체 1, LOX1 또는 C형 렉틴 도메인 패밀리 8 구성원 A), SCARF1, SCARF2, SRCRB4D, SSC5D 및 CD169(별칭: 시알로어드헤신 수용체, SIGLEC1) 중 어느 하나 이상으로부터 유래한 신호전달 도메인을 포함한다.

[0385] 일부 실시양태에서, 재조합 핵산은 예를 들어, 인간 MARCO의 세포내 도메인을 코딩한다. PSR 서브유닛은 아미노산 서열 MRNKKILKEDELLSETQQAAFHQIAMEPFEINVPKPKRRNGVNF(서열번호 72)를 가진 인간 MARCO의 44개 아미노산 ICD를 가진 세포내 도메인을 포함한다. 일부 실시양태에서, PSR 서브유닛은 MARCO의 세포내 도메인과 적어도 70%, 75%, 80%, 85%, 90% 또는 95% 동일한 변이체를 포함한다.

[0386] 일부 실시양태에서, 예를 들어, PSR(식세포 스캐빈저 수용체)은 인간 MARCO의 막횡단 영역을 포함한다.

[0387] 일부 실시양태에서, 재조합 핵산은 인간 SRA1의 세포내 도메인을 코딩한다. PSR 서브유닛은 아미노산 서열 MEQWDHFHNQQEDTDSCSESVKFDARSMTA LLPPNPKNSPSLQEKLKSFK(서열번호 73)을 가진 인간 SRA1의 50개 아미노산 ICD를 가진 세포내 도메인을 포함한다. 일부 실시양태에서, PSR 서브유닛은 인간 SRA1의 세포내 도메인과 적어도 70%, 75%, 80%, 85%, 90% 또는 95% 동일한 변이체를 포함한다. SRA의 세포내 영역은 인산화 부위를 가진다.

[0388] 일부 실시양태에서, PSR은 인간 SRA1의 막횡단 영역을 포함한다.

[0389] 일부 실시양태에서, 예를 들어, 재조합 핵산은 CD36의 세포내 도메인을 포함한다. 일부 실시양태에서, 재조합 핵산은 CD36의 TM 도메인을 포함한다. 천연 생성 전체 길이 CD36은 2개의 TM 도메인과 2개의 짧은 세포내 도메인을 갖고, CD36의 세포외 도메인은 산화된 LDL에 결합한다. 세포내 도메인 둘 다가 지방산 아실화된 시스테인 쌍을 함유한다. 이것은 공지되어 있는 신호전달 도메인(예를 들어, 키나제, 포스파타제, g-단백질 결합 또는 스캐폴딩 도메인)을 결여한다. N-말단 세포질 도메인은 매우 짧고(5개 내지 7개의 아미노산 잔기) 원형질막의 내부 소엽과 밀접하게 연관되어 있다. 카르복시 말단 도메인은 신호전달 분자와 상호작용하는 것으로 알려진, CD4 및 CD8의 세포내 도메인 내의 영역에 상동한 CXCX5K 모티프를 함유하는 13개의 아미노산을 함유한다. CD36의 세포내 도메인은 lyn 키나제, MAP 키나제 및 국소 부착 키나제(FAK)를 활성화시키는 신호전달 복합체를 어셈블링할 수 있고 src 상동성 2-함유 포스포티로신 포스파타제(SHP-2)를 불활성화시킬 수 있다. 구아닌 뉴클레오타이드 교환 인자(GEF)의 구성원은 잠재적 핵심 신호전달 중간체로서 확인되었다.

[0390] 일부 실시양태에서, 재조합 핵산은 예를 들어, 인간 SCARA3의 세포내 도메인을 코딩한다. 일부 실시양태에서, PSR 서브유닛은 인간 SCARA3의 세포내 도메인과 적어도 70%, 75%, 80%, 85%, 90% 또는 95% 동일한 변이체를 포함한다. 일부 실시양태에서, PSR은 SCRA3의 TM 도메인을 포함한다. 일부 실시양태에서, TM 도메인은 길이가 약 20개 내지 30개의 아미노산이다.

[0391] 스캐빈저 수용체는 동종이량체 또는 이종이량체로서 존재할 수 있다. 예를 들어, MARCO는 동종삼량체로서 존재한다.

[0392] 일부 실시양태에서, PSP의 TM 도메인 또는 ICD 도메인은 FcR, Megf10, Bai1 또는 MerTK로부터 유래하지 않는다. 일부 실시양태에서, PSR의 ICD는 CD3 제타 세포내 도메인을 포함하지 않는다.

[0393] 일부 실시양태에서, 세포내 도메인 및 막횡단 도메인은 FcR 베타로부터 유래한다.

[0394] 한 측면에서, 재조합 핵산은 (a) 표적 세포의 항원에 특이적인 세포외 항원 결합 도메인을 포함하는 세포외 도메인, (b) 막횡단 도메인, 및 (c) 재조합 PSR 세포내 신호전달 도메인을 포함하는 식세포 스캐빈저 수용체(PSR) 융합 단백질(CFP)인, 향상된 식세포작용을 위한 키메라 항원성 수용체(CAR-P)를 코딩하고, 이때 재조합 PSR 세포내 신호전달 도메인은 식세포 수용체로부터 유래한 제1 부분 및 비-식세포 수용체로부터 유래한 제2 부분을 포함한다.

[0395] 일부 실시양태에서, 제2 부분은 PI3K 동원 도메인이 아니다. 일부 실시양태에서, 제2 부분은 PI3K 동원 도메인이다.

[0396] 비-식세포 수용체로부터 유래한 제2 부분은 재조합 핵산을 발현하는 조작된 식세포의 식세포작용 및/또는 염증 잠재력을 향상시키는 세포내 신호전달 도메인을 포함할 수 있다. 일부 실시양태에서, 비-식세포 수용체로부터 유래한 제2 부분은 하나 초과의 세포내 도메인(ICD)을 포함한다. 일부 실시양태에서, 비-식세포 수용체로부터 유래한 제2 부분은 제2 ICD를 포함한다. 일부 실시양태에서, 비-식세포 수용체로부터 유래한 제2 부분은 제2 ICD 및 제3 ICD를 포함한다. 일부 실시양태에서, 비-식세포 수용체로부터 유래한 제2 부분은 제2 ICD, 제3 ICD 및 제4 ICD를 포함하고, 이때 제2 부분은 재조합 핵산에 의해 코딩된다. 비-식세포 수용체로부터 유래한 제2, 제3 또는 제4 ICD를 포함하는 각각의 제2 부분은 다음과 같이 기재된다.

[0397] *세포내 신호전달 및 염증 활성화를 향상시키는 키메라 항원 수용체*

[0398] 한 측면에서, 재조합 핵산은 식세포 ICD 이외에, 예컨대, 대식세포가 감염과 싸울 때, 강력한 염증촉진성 면역 활성화 능력을 부여하는 제2 세포내 도메인을 코딩한다. 제2 세포내 도메인(제2 ICD)은 제1 식세포 ICD의 세포질 말단에 융합된다. 제2 세포내 도메인은 인플라마좀(inflammasome) 및 염증촉진성 신호를 유발하는 데 필요한 제2 신호를 제공한다. Nod 유사 수용체(NLR)는 선천성 면역 반응에서 활성화되는 수용체의 서브세트이고, 올리고머화하여 다중단백질 복합체를 형성하고, 이 다중단백질 복합체는 염증촉진성 캐스파제를 동원하고 이의 절단 및 활성화를 유도하는 플랫폼으로서 사용된다. 이것은 ROS의 직접적인 활성화로 이어지고, 종종 파이롭토시스(pyroptosis)로서 공지되어 있는 격렬한 세포 사멸을 초래한다. 4개의 인플라마좀 복합체들인 NLRP1m, NLRP3, IPAF 및 AIM2가 존재한다.

[0399] 종양 미세환경(TME)은 면역억제 환경을 구성한다. IL-10, 글루코코르티코이드 호르몬, 아폽토시스 세포 및 면역복합체의 영향은 선천성 면역 세포 기능을 방해할 수 있다. 식세포를 포함하는 면역 세포는 면역관용 표현형으로 정착된다. 대식세포에서, 통상적으로 M2 표현형으로서도 공지되어 있는 이 표현형은 대식세포가 강력하고 병원체를 사멸시킬 수 있는 M1 표현형과 구별된다. 예를 들어, LPS 또는 IFN-감마에 노출된 대식세포는 M1 표현형을 향해 분극화될 수 있는 반면, IL-4 또는 IL-13에 노출된 대식세포는 M2 표현형을 향해 분극화될 것이다. LPS 또는 IFN-감마는 대식세포 표면의 Toll 유사 수용체 4(TLR4)와 상호작용하여 Trif 및 MyD88 경로를 유도함으로써, 전사 인자 IRF3, AP-1 및 NFKB의 활성화를 유도하여, 염증촉진성 M1 대식세포 반응에 필요한 TNF 유전자, 인터페론 유전자, CXCL10, NOS2, IL-12 등을 활성화시킬 수 있다. 유사하게, IL-4 및 IL-13은 IL-4R에 결합하여, 소염 반응(M2 반응)과 관련된 유전자인 CCL17, ARG1, IRF4, IL-10, SOCS3 등의 발현을 조절하는 Jak/Stat6 경로를 활성화시킨다. CD14, CD80, D206의 발현 및 CD163의 낮은 발현은 M1 표현형을 향한 대식세포 분극화의 지표이다.

[0400] 일부 실시양태에서, 재조합 핵산은 염증 반응을 위한 세포질 도메인을 포함하는 하나 이상의 추가 세포내 도메인을 코딩한다. 일부 실시양태에서, 조작된 대식세포에서 염증 반응을 위한 세포질 도메인을 포함하는 식세포 수용체(PR) 융합 단백질(CFP)을 코딩하는 재조합 핵산의 발현은 M1 표현형과 유사한 강력한 염증촉진성 반응을 부여한다.

[0401] 일부 실시양태에서, 염증 반응을 위한 세포질 도메인은 TLR3, TLR4, TLR9, MYD88, TRIF, RIG-1, MDA5, CD40, IFN 수용체, NLRP-1 내지 NLRP-14, NOD1, NOD2, 파이린(Pyrin), AIM2, NLRC4, CD40의 신호 전달도입 도메인 또는 영역일 수 있다.

[0402] 일부 실시양태에서, 식세포 스캐빈저 수용체(PSR) 융합 단백질(CFP)을 코딩하는 재조합 핵산의 발현은 IL-1 신호전달 캐스케이드의 활성화를 위한 염증촉진성 세포질 도메인을 포함한다.

[0403] 일부 실시양태에서, 키메라 수용체(예를 들어, 식세포 수용체(PR) 융합 단백질(CFP))의 세포질 부분은 Toll 유사 수용체의 세포질 도메인, 예컨대, Toll 유사 수용체 3(TLR3), Toll 유사 수용체 4(TLR4), Toll 유사 수용체 7(TLR7), Toll 유사 수용체 8(TLR8), Toll 유사 수용체 9(TLR9)의 세포내 신호전달 도메인을 포함한다. 일부 실시양태에서, 키메라 수용체의 세포질 부분은 인터류킨-1 수용체 관련 키나제 1(IRAK1)의 적합한 영역을 포함한다. 일부 실시양태에서, 키메라 수용체의 세포질 부분은 분화 1차 반응 단백질(MYD88)의 적합한 영역을 포함한다. 일부 실시양태에서, 키메라 수용체의 세포질 부분은 미엘린 및 림프구 단백질(MAL)의 적합한 영역을 포함한다. 일부 실시양태에서, 키메라 수용체의 세포질 부분은 레티노산 유도성 유전자(RIG-1)의 적합한 영역을 포함한다.

[0404] 일부 실시양태에서, PSR의 막횡단 도메인은 MYD88, TLR3, TLR4, TLR7, TLR8, TLR9, MAL, IRAK1 단백질들 중 어느 한 단백질의 막횡단 도메인을 포함한다.

[0405] 일부 실시양태에서, 재조합 PSR 세포내 신호전달 도메인은 식세포 수용체로부터 유래한 제1 부분 및 비-식세포 수용체로부터 유래한 제2 부분을 포함하고, 이때 비-식세포 수용체로부터 유래한 제2 부분은 인산화 부위를 포함한다. 일부 실시양태에서, 인산화 부위는 자가인산화 부위에 적합한 아미노산 서열을 포함한다. 일부 실시양태에서, 인산화 부위는 Src 패밀리 키나제에 의한 인산화에 적합한 아미노산 서열을 포함한다. 일부 실시양태에서, 인산화 부위는 인산화 시 키나제의 SH2 도메인에 결합할 수 있는 아미노산 서열을 포함한다. 일부 실시양태에서, 수용체 티로신 키나제 도메인은 제1 세포질 부분 이외에 CFP의 세포질 말단에서 융합된다. 일부 실시양태에서, 인산화는 티로신 인산화이다.

[0406] 일부 실시양태에서, 제2 세포내 도메인은 면역 수용체 티로신 활성화 모티프(ITAM)이다. ITAM 모티프는 포유동물 α 및 β 면역글로불린 단백질, TCR ɣ 수용체, FCR ɣ 수용체 서브유닛, CD3 쇄 수용체 및 NFAT 활성화 분자에 존재한다.

[0407] 일부 실시양태에서, CFP 세포내 도메인은 1개의 ITAM 모티프를 포함한다. 일부 실시양태에서, CFP 세포내 도메인은 1개 초과의 ITAM 모티프를 포함한다. 일부 실시양태에서, CFP 세포내 도메인은 2개 이상의 ITAM 모티프를 포함한다. 일부 실시양태에서, CFP 세포내 도메인은 3개 이상의 ITAM 모티프를 포함한다. 일부 실시양태에서, CFP 세포내 도메인은 4개 이상의 ITAM 모티프를 포함한다. 일부 실시양태에서, CFP 세포내 도메인은 5개 이상의 ITAM 모티프를 포함한다. 일부 실시양태에서, CFP 세포내 도메인은 6개 이상의 ITAM 모티프를 포함한다. 일부 실시양태에서, CFP 세포내 도메인은 7개 이상의 ITAM 모티프를 포함한다. 일부 실시양태에서, CFP 세포내 도메인은 8개 이상의 ITAM 모티프를 포함한다. 일부 실시양태에서, CFP 세포내 도메인은 9개 이상의 ITAM 모티프를 포함한다. 일부 실시양태에서, CFP 세포내 도메인은 10개 이상의 ITAM 모티프를 포함한다.

[0408] 일부 실시양태에서, 제1 식세포 ICD 내의 하나 이상의 도메인은 돌연변이를 포함한다.

[0409] 일부 실시양태에서, 제2 ICD 내의 하나 이상의 도메인은 키나제 결합 도메인의 향상, 인산화 부위의 생성, SH2 도킹 부위의 생성 또는 이들의 조합을 위해 돌연변이를 포함한다.

[0410] *염증 유전자의 공-발현*

[0411] 한 측면에서, 재조합 핵산은 조작된 세포에서 CFP와 공-발현되는, 염증촉진성 유전자에 대한 코딩 서열을 포함한다. 일부 실시양태에서, 염증촉진성 유전자는 사이토카인이다. 예는 TNF-α, IL-1α, IL-1β, IL-6, CSF, GMCSF, 또는 IL-12 또는 인터페론을 포함하나, 이들로 제한되지 않는다.

[0412] 염증촉진성 유전자를 코딩하는 재조합 핵산은 모노시스트론일 수 있고, 이때 (a) PSP 및 (b) 염증촉진성 유전자에 대한 2개의 코딩 서열은 독립적인 발현을 위해 전사 후 또는 번역 후 절단된다.

[0413] 일부 실시양태에서, 상기 2개의 코딩 서열은 예를 들어, P2A 서열을 코딩하는 자가절단 도메인을 포함한다.

[0414] 일부 실시양태에서, 상기 2개의 코딩 영역은 IRES 부위에 의해 분리된다.

[0415] 일부 실시양태에서, 상기 2개의 코딩 서열은 바이시스트론 유전 요소에 의해 코딩된다. (a) PSP 및 (b) 염증촉진성 유전자에 대한 코딩 영역은 단방향적일 수 있고, 이때 각각은 별도의 조절 제어 하에 있다. 일부 실시양태에서, 이들 둘 다에 대한 코딩 영역은 양방향적일 수 있고 반대 반향으로 유도된다. 각각의 코딩 서열은 별도의 조절 제어 하에 있다.

[0416] 염증촉진성 유전자의 공-발현은 대식세포의 강한 염증 자극을 부여하고 염증을 위해 주변 조직을 활성화시키도록 디자인된다.

[0417] *인테그린 활성화 도메인*

[0418] 세포-세포 및 세포-기질 부착은 인테그린 세포외 도메인과 다양한 단백질 리간드들의 결합에 의해 매개되나; 이 부착 상호작용 및 동적 세포 반응, 예컨대, 세포 확산 또는 이동으로의 이의 번역의 세포 조절은 인테그린 세포질 꼬리를 요구한다. 이 짧은 꼬리는 수용체를 신호전달 경로 및 세포골격 네트워크에 연결하는 세포내 리간드에 결합한다(Calderwood DA, 2004, Integrin Activation, Journal of Cell Science 117, 657-666). 인테그린은 α 서브유닛과 β 서브유닛의 비-공유 회합에 의해 형성된 이종이량체성 부착 수용체이다. β4 서브유닛을 제외한 각각의 서브유닛은 비교적 큰 세포외 도메인 및 짧은 세포질 꼬리를 가진 I형 막횡단 당단백질이다. 개별 인테그린 패밀리 구성원은 다수의 리간드들을 인식하는 능력을 가진다. 인테그린은 많은 수의 세포외 매트릭스 단백질들(골 매트릭스 단백질, 콜라겐, 피브로넥틴, 피브리노겐, 라미닌, 트롬보스폰딘, 비트로넥틴 및 폰 빌레브란트 인자)에 결합할 수 있고, 이것은 세포외 매트릭스에의 세포 부착에 있어서 인테그린의 주요 기능을 반영한다. 많은 "반대 수용체들"은 리간드이고, 이것은 세포-세포 상호작용을 매개하는 데 있어서 인테그린의 역할을 반영한다. 인테그린은 리간드 친화성을 증가시키기 위해 입체구조적 변화를 겪는다.

[0419] 인테그린 β_2 서브패밀리는 4개의 상이한 인테그린 수용체인 $\alpha_M\beta_2$(CD11b/CD18, Mac-1, CR3, Mo-1), $\alpha_L\beta_2$(CD11a/CD18, LFA-1), $\alpha_X\beta_2$(CD11c/CD18) 및 $\alpha_D\beta_2$(CD11d/CD18)로 구성된다. 이 백혈구 인테그린들은 면역 반응, 내피에의 부착 및 내피 관통, 병원체의 식세포작용 및 백혈구 활성화를 포함하는, 백혈구 기능의 사실상 모든 측면에 관여한다.

[0420] 모든 β_2 인테그린의 α 서브유닛은 I 또는 A 도메인으로서 지칭되는 약 200개 아미노산의 삽입된 영역을 함유한다. 고도로 보존된 I 도메인은 여러 다른 인테그린 α 서브유닛 및 다른 단백질, 예컨대, 특정 응고 및 보체 단백질에서 발견된다. I 도메인은 단백질-단백질 상호작용을 매개하고, 인테그린에서는 단백질 리간드의 결합에 통합적으로 관여한다. I 도메인이 그의 인테그린의 리간드 결합 기능을 지배하지만, α 서브유닛의 다른 영역은 리간드 인식에 영향을 미친다. 예로서, mAb(OKM1)는 $\alpha_M\beta_2$에서 I 도메인 외부의 에피토프를 인식하나, α_M 서브유닛에서는 리간드 결합을 억제하고; α 서브유닛에서 I 도메인을 가진 인테그린인 $\alpha_L\beta_2$ 및 $\alpha_2\beta_1$에서 EF-핸드(hand) 영역은 리간드 인식에 기여한다. α_M 서브유닛 및 아마도 다른 α 서브유닛은 비-단백질 리간드의 개입에 관여하는 렉틴 유사 도메인을 함유하고, 점유는 I 도메인의 기능을 조절할 수 있다.

[0421] 인테그린은 효소 활성을 결여하기 때문에, 원형질막의 세포질 면에서의 신호전달 복합체의 어셈블리에 의해 신호전달이 유도된다. 이 복합체의 형성은 두 가지 방식; 첫째, 분자 상호작용의 결합력을 증가시켜 이펙터 분자의 결합 속도를 증가시키는 수용체 클러스터링, 및 둘째, 이펙터 결합 부위를 생성하거나 노출시키는 수용체의 입체구조적 변화의 유도에 의해 달성된다. ECM 내에서, 인테그린은 피브로넥틴, 라미닌, 콜라겐, 테나신, 비트로넥틴 및 트롬보스폰딘에 결합하는 능력을 가진다. 인테그린/ECM 상호작용의 클러스터는 국소 부착을 형성하여, 세포 내에서 세포골격 성분 및 신호전달 분자를 농축시킨다. 인테그린의 세포질 꼬리는 F-액틴을 막에 고정시키는 데 관여하는 단백질인 빈쿨린(vinculin)을 동원하는 α-액티닌(actinin)과 탈린(talin)에 대한 결합 부위로서 사용된다. 탈린은 단백질 키나제 C(PKCα)와 같은 키나제에 의해 활성화된다.

[0422] 인테그린은 셀렉틴(selectin)에 의해 활성화된다. 백혈구는 L-셀렉틴을 발현하고, 활성화된 혈소판은 P-셀렉틴을 발현하고, 활성화된 내피 세포는 E-셀렉틴 및 P-셀렉틴을 발현한다. P-셀렉틴 매개 부착은 β2 인테그린의 케모카인 또는 혈소판 활성화 인자 유발 활성화를 가능하게 하여, 부착을 안정화시킨다. 또한, 이것은 부착 백혈구로부터의 케모카인의 방출을 용이하게 한다. P-셀렉틴 당단백질 리간드 1의 세포질 도메인은 Nef 관련 인자 1과 함께 항시성 복합체를 형성하였다. P-셀렉틴의 결합 후, Src 키나제는 포스포이노시타이드-3-OH 키나제 p85-p110δ 이종이량체를 동원하고 백혈구 인테그린을 활성화시키는 Nef 관련 인자 1을 인산화하였다. E-셀렉틴 리간드는 β2 인테그린 기능에도 영향을 미치는 신호를 전달도입한다. 셀렉틴은 Src 패밀리 키나제의 활성화를 유발한다. 셀렉틴 개입에 의해 활성화된 SFK는 DAP12 및 FcRγ의 세포질 도메인에서 면역수용체 티로신 기반 활성화 모티프(ITAM)를 인산화한다. 일부 측면에서, CD44는 E-셀렉틴으로부터 신호를 전달도입하기에 충분하다. CD44는 인테그린의 내부-외부 신호전달을 유발한다. 인테그린 활성화의 마지막 공통 단계는 탈린과 β 서브유닛의 세포질 꼬리의 결합이다. 세포질 어댑터의 또 다른 군인 킨들린(Kindlin)은 인테그린 β 꼬리의 상이한 영역에 결합한다. 킨들린은 탈린에 의해 활성화된 인테그린의 클러스터링을 증가시킨다. 킨들린은 셀렉틴 신호전달에 반응하나, 킨들린은 주로 호중구와 같은 조혈 세포에서 발견된다. 케모카인 성분에 의한 인테그린 활성화 시

신호전달뿐만 아니라 셀렉틴 신호전달도 SFK, Syk 및 SLP-76을 비롯한 성분을 공유한다.

[0423] 일부 실시양태에서, 재조합 PSR 융합 단백질의 세포내 도메인은 인테그린 활성화 도메인을 포함한다. 인테그린 활성화 도메인은 셀렉틴, 예를 들어, P-셀렉틴, L-셀렉틴 또는 E-셀렉틴의 세포내 도메인을 포함한다.

[0424] 일부 실시양태에서, 재조합 PSR 융합 단백질의 세포내 도메인은 라미닌의 인테그린 활성화 도메인을 포함한다.

[0425] 일부 실시양태에서, 재조합 PSR 융합 단백질의 세포내 도메인은 탈린의 활성화를 위해 인테그린 활성화 도메인을 포함한다.

[0426] 일부 실시양태에서, 재조합 PSR 융합 단백질의 세포내 도메인은 식세포 수용체 ICD 도메인의 세포질 말단에 융합된 인테그린 활성화 도메인을 포함한다.

[0427] *항원 교차제시를 향상시키는 키메라 수용체*

[0428] 일부 실시양태에서, 재조합 핵산은 항원의 교차제시를 가능하게 하는 도메인을 코딩한다. 일반적으로, MHC 클래스 I 분자는 세포 내에서 합성되는 자가 또는 병원체 유래 항원을 제시하는 반면, 세포내이입 흡수를 통해 유래한 외생성 항원은 CD4+ T 세포에게 제시되기 위해 MHC 클래스 II 분자에 로딩된다. 펩티드가 프로테아좀에 의해 생성되는 내생성 항원의 MHC-I 제한 제시. 그러나, 일부 경우, DC는 CD8+ T 세포에게 제시하기 위해 외생성 항원을 MHC-I 경로로 처리할 수 있다. 이것은 항원의 교차제시로서 지칭된다. 가용성 또는 외생성 항원 성분은 세포내이입 경로를 따르는 대신에 액포 내의 라이소좀 프로테아제에 의해 분해될 수 있고 DC에 의해 교차제시될 수 있다. 일부 경우, 열 충격 단백질 90(Hsp90)과 같은 샤페론은 특정 APC에 의한 항원의 교차제시를 돕는 것으로 밝혀졌다. HSP-펩티드 복합체는 유리 폴리펩티드에 비해 상이한 수용체 군에 의해 내재화되는 것으로 알려져 있다. 이 수용체는 스캐빈저 수용체 패밀리의 수용체이고, LOX-1, SREC-I/SCARF-I 및 FEEL1/스타빌린(Stabilin)-1을 포함한다. SREC-1 및 LOX-1 둘 다가 분자 샤페론 결합 항원의 교차제시를 매개하고 $CD8^+$ T 림프구의 활성화를 유발하는 것으로 확인되었다.

[0429] SREC-1(내피 세포에 의해 발현된 스캐빈저 수용체)은 다른 유형의 스캐빈저 수용체에 대해 유의미한 상동성을 갖지 않지만 독특한 도메인 구조를 가진다. 이것은 세포외 도메인에서 EGF 유사 시스테인 풍부 모티프의 10개 반복부를 함유한다. 최근에, SREC-I의 구조는 아폽토시스 세포를 인식하는 세포 표면 식세포 수용체로서 작용하는, 캐노르하브디티스 엘레간스(*Caenorhabditis elegans*) 유전자 *ced-I*에 의해 코딩된 16개의 EGF 유사 반복부를 가진 막횡단 단백질의 구조와 유사한 것으로 밝혀졌다.

[0430] 클래스-I MHC 경로를 통한 암 항원의 교차제시는 세포독성과 관련되어 있으므로 종양 퇴행에 유리한 $CD8^+$ T 세포 반응을 향상시킨다. 일부 실시양태에서, CFP의 세포내 도메인은 SREC1 세포내 도메인을 포함한다. 일부 실시양태에서, CFP의 세포내 도메인은 SRECII 세포내 도메인을 포함한다.

[0431] 일부 실시양태에서, PSR 서브유닛은 SREC1 또는 SRECII로부터의 PSR 세포내 신호전달 도메인을 포함하는 세포내 도메인을 포함한다.

[0432] 일부 실시양태에서, PSR 서브유닛은 (i) 막횡단 도메인, 및 (ii) SREC1 또는 SRECII로부터의 PSR 세포내 신호전달 도메인을 포함하는 세포내 도메인을 포함한다.

[0433] 일부 실시양태에서, PSR 서브유닛은 (i) 막횡단 도메인, (ii) PSR 세포내 신호전달 도메인을 포함하는 세포내 도메인, 및 (iii) SREC1 또는 SRECII로부터의 세포외 도메인을 포함한다.

[0434] *CFP 융합 단백질의 막횡단 도메인*

[0435] 일부 실시양태에서, 재조합 핵산에 의해 코딩된 TM은 스캐빈저 수용체(SR)의 도메인을 포함한다. 일부 실시양태에서, TM은 렉틴, 덱틴 1, 만노스 수용체(CD206), SRA1, MARCO, CD36, CD163, MSR1, SCARA3, COLEC12, SCARA5, SCARB1, SCARB2, CD68, OLR1, SCARF1, SCARF2, SRCRB4D, SSC5D 및 CD169 중 어느 하나 이상의 수용체의 TM 도메인일 수 있거나 이러한 수용체로부터 유래할 수 있다.

[0436] 일부 실시양태에서, TM 도메인은 길이가 약 20개 내지 30개의 아미노산이다. SR의 TM 도메인은 길이가 약 20개 내지 30개의 아미노산이다.

[0437] PSP의 TM 도메인 또는 ICD 도메인은 Megf10, Bai1 또는 MerTK로부터 유래하지 않는다. PSR의 ICD는 CD3 제타 세포내 도메인을 포함하지 않는다.

[0438] 일부 실시양태에서, TM은 ICD와 동일한 식세포 수용체로부터 유래한다.

[0439] 일부 실시양태에서, TM 영역은 원형질막 단백질로부터 유래한다. TM은 Fc 수용체(FcR)로부터 선택될 수 있다. 일부 실시양태에서, 특정 FcR의 도메인을 코딩하는 핵산 서열은 재조합 구축물의 세포 특이적 발현에 사용된다. TM 도메인을 포함하는 FCR 알파 영역은 구축물의 대식세포 특이적 발현에 사용될 수 있다. FcRβ 재조합 단백질은 비만 세포에서 발현된다.

[0440] 일부 실시양태에서, CFP는 FCR 베타(FcRβ)의 TM을 포함한다.

[0441] 일부 실시양태에서, CFP는 FcRβ TM 및 ICD 도메인 둘 다를 포함한다.

[0442] 일부 실시양태에서, TM 도메인은 CD8로부터 유래한다.

[0443] 일부 실시양태에서, TM은 CD2로부터 유래한다.

[0444] 일부 실시양태에서, TM은 FCR 알파로부터 유래한다.

[0445] *CFP 융합 단백질의 세포외 도메인*

[0446] 세포외 도메인은 표적 세포의 하나 이상의 표적 항원에 결합하는 항원 결합 도메인을 포함한다. 표적 결합 도메인은 표적에 특이적이다. 세포외 도메인은 인트라바디, 펩티바디, 나노바디, 단일 도메인 항체, SMIP 및 다중특이적 항체로부터 선택된 항체 또는 항원 결합 도메인을 포함할 수 있다.

[0447] 일부 실시양태에서, 세포외 도메인은 Fab 결합 도메인을 포함한다. 다른 이러한 실시양태에서, 세포외 도메인은 scFv를 포함한다.

[0448] 일부 실시양태에서, 키메라 항원 수용체는 항원 결합 단편(Fab), 단일 쇄 가변 단편(scFv), 나노바디, VH 도메인, VL 도메인, 단일 도메인 항체(sdAb), VNAR 도메인, 및 VHH 도메인, 이중특이적 항체, 디아바디, 또는 이들 중 임의의 분자의 기능성 단편으로 구성된 군으로부터 유래한 세포외 항원 결합 도메인을 포함한다. 일부 실시양태에서, 항원 결합 단편(Fab), 단일 쇄 가변 단편(scFv), 나노바디, VH 도메인, VL 도메인, 단일 도메인 항체(sdAb), VNAR 도메인, 및 VHH 도메인, 이중특이적 항체, 디아바디, 또는 이들 중 임의의 분자의 기능성 단편은 하나 이상의 항원에 특이적으로 결합한다.

[0449] 일부 실시양태에서, 항원은 암 항원이고, 표적 세포는 표적 암 세포이다. 일부 실시양태에서, 표적 암 세포에 대한 항원은 CD3, CD4, CD5, CD7, CD19, CCR2, CCR4, CD30, CD37, TCRB1/2, TCR □□, TCR □□. CD22, HER2(ERBB2/neu), 메소텔린(Mesothelin), PSCA, CD123, CD30, CD171, CD138, CS-1, CLECL1, CD33, CD79b, EGFRvIII, GD2, GD3, BCMA, PSMA, ROR1, FLT3, TAG72, CD38, CD44v6, CEA, EPCAM, B7H3(CD276), KIT(CD117), CD213A2, IL-1 IRa, PRSS21, VEGFR2, CD24, MUC-16, PDGFR-베타, SSEA-4, CD20, MUC1, EGFR, NCAM, 프로스타제(Prostase), PAP, ELF2M, 에프린(Ephrin) B2, FAP, EphA2, GM3, TEM1/CD248, TEM7R, CLDN6, TSHR, GPRC5D, CD97, CD179a, ALK 및 IGLL1로 구성된 군으로부터 선택된다.

[0450] 다양한 암 항원 표적들은 당업자에게 공지되어 있는 암 항원들로부터 선택될 수 있다. 관련된 암 및 세포 유형에 따라, 암 항원은 돌연변이된 천연 단백질이다. 항원 결합 도메인은 천연 항원이 아닌 돌연변이된/암 항원에 대한 특이성에 대해 스크리닝된다.

[0451] 일부 실시양태에서, 예를 들어, 표적 암 세포에 대한 암 항원은 하기 돌연변이된/암 항원들 중 하나 이상일 수 있다: MUC16, CCAT2, CTAG1A, CTAG1B, MAGEA1, MAGEA2, MAGEA3, MAGEA4, MAGEA6, PRAME, PCA3, MAGEC1, MAGEC2, MAGED2, AFP, MAGEA8, MAGE9, MAGEA11, MAGEA12, IL13RA2, PLAC1, SDCCAG8, LSP1, CT45A1, CT45A2, CT45A3, CT45A5, CT45A6, CT45A8, CT45A10, CT47A1, CT47A2, CT47A3, CT47A4, CT47A5, CT47A6, CT47A8, CT47A9, CT47A10, CT47A11, CT47A12, CT47B1, SAGE1 및 CT55.

[0452] 일부 실시양태에서, 예를 들어, 표적 암 세포에 대한 암 항원은 하기 돌연변이된/암 항원들 중 하나 이상일 수 있고, 이때 암은 T 세포 림프종이다: CD2, CD3, CD4, CD5, CD7, CD8, CD20, CD30, CD45, CD56.

[0453] 일부 실시양태에서, 예를 들어, 표적 암 세포에 대한 암 항원은 하기 돌연변이된/암 항원들 중 하나 이상일 수 있고, 이때 암은 교모세포종이다: IDH1, ATRX, PRL3 또는 ETBR.

[0454] 일부 실시양태에서, 예를 들어, 표적 암 세포에 대한 암 항원은 하기 돌연변이된/암 항원들 중 하나 이상일 수 있고, 이때 암은 난소암이다: CA125, 베타-hCG, 비뇨기 고나도트로핀 단편, AFP, CEA, SCC, 인히빈(inhibin) 또는 엑스트라디올(extradiol).

[0455] 일부 실시양태에서, 표적 암 세포에 대한 암 항원은 HER2일 수 있다.

[0456] 일부 실시양태에서, 표적 암 세포에 대한 암 항원은 EGFR 변이체 III일 수 있다.

[0457] 일부 실시양태에서, 표적 암 세포에 대한 암 항원은 CD19일 수 있다.

[0458] 일부 실시양태에서, SR 서브유닛 영역은 스캐빈저 수용체의 세포외 도메인(ECD)을 포함한다. 일부 실시양태에서, 스캐빈저 수용체의 ECD는 ICD 및 TM 도메인을 포함하는 SR의 ECD 도메인을 포함한다. 일부 실시양태에서, SR-ECD는 식세포와 표적 세포의 결합에 기여하고, 그 다음 활성화되고, 표적 세포의 식세포작용을 활성화시킨다.

[0459] 일부 실시양태에서, PSR 도메인은 PSR에 통합된 ICD 및 TM 도메인을 가진 각각의 스캐빈저 수용체의 ECD 도메인 또는 이의 일부를 임의로 포함한다. 따라서, 일부 실시양태에서, 재조합 핵산에 의해 코딩된 ECD는 렉틴, 덱틴 1, 만노스 수용체(CD206), 스캐빈저 수용체 A1(SRA1), MARCO, CD36, CD163, MSR1, SCARA3, COLEC12, SCARA5, SCARB1, SCARB2, CD68, OLR1, SCARF1, SCARF2, CXCL16, STAB1, STAB2, SRCRB4D, SSC5D, CD205, CD207, CD209, RAGE, CD14, CD64, F4/80, CCR2, CX3CR1, CSF1R, Tie2, HuCRIg(L) 및 CD169 수용체로 구성된 군으로부터 선택된 도메인을 포함한다. 대부분의 대식세포 스캐빈저 수용체들의 세포외 도메인은 외래 물질의 비특이적 항체 독립적 인식에 있어서 자가와 비-자가를 구별하는 데 사용될 수 있는 광범위한 결합 특이성을 가진 스캐빈저 수용체를 함유한다. I형 및 II형 클래스 A 스캐빈저 수용체들(SR-AI1 및 SR-AII)은 작은 NH2-말단 세포내 도메인, 및 짧은 스페이서 도메인, α-나선형 코일드-코일 도메인 및 삼중 나선형 콜라겐성 도메인을 함유하는 세포외 부분을 가진 삼량체성 막 당단백질이다. I형 수용체는 시스테인 풍부 COOH 말단(SRCR) 도메인을 추가로 함유한다. 이 수용체는 신체 전체에 걸쳐 다양한 조직들에서 대식세포에 존재하고 비정상적으로 광범위한 리간드 결합 특이성을 나타낸다. 이들은 변형된 LDL과 같은 화학적으로 변형된 단백질을 비롯한 매우 다양한 다중음이온에 결합하고, 죽상형성 동안 콜레스테롤 침착에 관여한다. 이들은 대식세포 관련 숙주 방어 및 염증성 병태의 세포 부착 과정에서 역할을 할 수도 있다.

[0460] 일부 실시양태에서, SR ECD는 전구아폽토시스 세포에 결합하도록 디자인된다. 일부 실시양태에서, 스캐빈저 수용체 ECD는 암 세포 또는 감염된 세포의 세포 표면 분자에 대한 결합 도메인을 포함한다.

[0461] 일부 실시양태에서, PR 서브유닛의 세포외 도메인은 링커에 의해 표적 세포 결합 도메인, 예컨대, 암 항원에 특이적인 항체 또는 이의 일부에 연결된다.

[0462] 일부 실시양태에서, 세포외 항원 결합 도메인은 하나의 항원 결합 도메인을 포함한다. 일부 실시양태에서, 세포외 항원 결합 도메인은 하나 초과의 결합 도메인을 포함한다. 일부 실시양태에서, 결합 도메인은 scFv이다. 일부 실시양태에서, 결합 도메인은 단일 도메인 항체(sdAb)이다. 일부 실시양태에서, 결합 도메인은 세포외 도메인에서 재조합 PR에 융합된다. 일부 실시양태에서, PR의 결합 도메인(예를 들어, scFv)과 세포외 도메인은 링커를 통해 연결된다.

[0463] 일부 실시양태에서, ECD 항원 결합 도메인은 세포내 항원에 결합할 수 있다. 일부 실시양태에서, 세포내 항원은 암 항원이다.

[0464] 일부 실시양태에서, 세포외 항원 결합 도메인은 1000 nM 미만의 친화성으로 표적 리간드에 결합한다. 일부 실시양태에서, 세포외 항원 결합 도메인은 500 nM 미만의 친화성으로 표적 리간드에 결합한다. 일부 실시양태에서, 세포외 항원 결합 도메인은 450 nM 미만의 친화성으로 표적 리간드에 결합한다. 일부 실시양태에서, 세포외 항원 결합 도메인은 400 nM 미만의 친화성으로 표적 리간드에 결합한다. 일부 실시양태에서, 세포외 항원 결합 도메인은 350 nM 미만의 친화성으로 표적 리간드에 결합한다. 일부 실시양태에서, 세포외 항원 결합 도메인은 250 nM 미만의 친화성으로 표적 리간드에 결합한다. 일부 실시양태에서, 세포외 항원 결합 도메인은 200 nM 미만의 친화성으로 표적 리간드에 결합한다. 일부 실시양태에서, 세포외 항원 결합 도메인은 100 nM 미만의 친화성으로 표적 리간드에 결합한다. 일부 실시양태에서, 세포외 항원 결합 도메인은 200 nM 내지 1000 nM의 친화성으로 표적 리간드에 결합한다. 일부 실시양태에서, 세포외 항원 결합 도메인은 300 nM 내지 1.5 mM의 친화성으로 표적 리간드에 결합한다. 일부 실시양태에서, 항원 결합 도메인은 > 200 nM, > 300 nM 또는 > 500 nM의 친화성으로 표적 리간드에 결합한다.

[0465] *펩티드 링커*

[0466] 일부 실시양태에서, 세포외 항원 결합 도메인인 scfv는 링커에 의해 TM 도메인 또는 다른 세포외 도메인에 연결된다. 일부 실시양태에서, 세포외 항원 결합 도메인에서 하나 초과의 scfv가 있는 경우, 하나 초과의 scfv는 링

커에 의해 서로 연결된다.

[0467] 일부 실시양태에서, 링커는 유연하다. 일부 실시양태에서, 링커는 힌지 영역을 포함한다. 링커는 통상적으로 짧은 펩티드 서열이다. 일부 실시양태에서, 링커는 글리신 및 하나 이상의 세린 잔기의 스트레치이다. 짧은 펩티드 링커에 바람직한 다른 아미노산은 트레오닌(Thr), 세린(Ser), 프롤린(Pro), 글리신(Gly), 아스파르트산(Asp), 라이신(Lys), 글루타민(Gln), 아스파라긴(Asn), 알라닌(Ala), 아르기닌(Arg), 페닐알라닌(Phe) 및 글루탐산(Glu)을 포함하나, 이들로 제한되지 않는다. 이들 중 Pro, Thr 및 Gln은 천연 링커를 위해 자주 사용되는 아미노산이다. Pro은 매우 제한된 입체구조를 야기하는 고리형 측쇄를 가진 독특한 아미노산이다. Pro 풍부 서열은 피루베이트 데하이드로게나제에서 리포일과 E3 결합 도메인 사이의 링커를 비롯한 도메인간 링커로서 사용된다($GA_2PA_3PAKQEA_3PAPA_2KAEAPA_3PA_2KA$(서열번호 75)). 본 개시내용의 목적을 위해, 실험적 링커는 유연성 링커, 강성 링커 및 절단 가능한 링커일 수 있다. (G4S)x(이때, x는 1, 2, 3, 4 등으로서 표기된, 모이어티의 다중 카피임)(서열번호 76)와 같은 서열은 유연성 링커 서열을 포함한다. 본원에서 사용된 다른 유연성 서열은 글리신의 여러 반복부, 예를 들어, (Gly)6(서열번호 77) 또는 (Gly)8(서열번호 78)을 포함한다. 다른 한편으로, 강성 링커, 예를 들어, 링커(EAAAK)x(서열번호 79)가 사용될 수 있고, 이때 x는 강성 링커를 생성하는 1, 2, 3, 4 등의 정수이다.

[0468] 일부 실시양태에서, 링커는 적어도 2개 또는 적어도 3개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 4개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 5개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 6개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 7개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 8개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 9개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 8개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 10개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 11개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 12개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 13개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 14개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 15개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 16개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 17개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 18개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 19개의 아미노산을 포함한다. 일부 실시양태에서, 링커는 20개의 아미노산을 포함한다.

[0469] 본원에서 예상되는 바와 같이, 임의의 적합한 ECD, TM 또는 ICD 도메인은 내생성 수용체에 비해 향상된 식세포작용을 가진 단백질을 수득하기 위해 본 개시내용에 기재된 CARP 수용체들 중 어느 한 수용체의 적합한 부분에서 교환 가능하게 클로닝될 수 있다.

[0470] *융합 단백질의 특성:*

[0471] CFP는 이를 발현하는 세포의 세포막 내로 구조적으로 통합될 수 있다. 핵산 구축물의 특정 리더 서열, 예컨대, 신호 펩티드를 사용하여 코딩된 단백질의 원형질막 발현을 유도할 수 있다. 구축물에 의해 코딩된 막횡단 도메인은 발현된 단백질을 세포의 원형질막 내로 통합시킬 수 있다.

[0472] 일부 실시양태에서, 막횡단 도메인은 대식세포에서 내생성 FcR 감마 수용체와 이량체화하여 대식세포 특이적 발현을 보장하는 FcR 알파 수용체의 TM 도메인을 포함한다.

[0473] CFP는 이를 발현하는 세포를 강력한 식세포로 만들 수 있다. CFP를 코딩하는 재조합 핵산이 세포에서 발현될 때, 세포는 재조합 핵산을 발현하지 않는 세포에 비해 표적 세포의 항원을 가진 표적 세포의 증가된 식세포작용을 나타낼 수 있다. 재조합 핵산이 세포에서 발현될 때, 세포는 재조합 핵산을 발현하지 않는 세포에 비해 표적 세포의 항원을 가진 표적 세포의 증가된 식세포작용을 나타낼 수 있다. 일부 실시양태에서, 재조합 핵산이 세포에서 발현될 때, 세포는 재조합 핵산을 발현하지 않는 세포에 비해 표적 세포의 항원을 가진 표적 세포의 적어도 2배 증가된 식세포작용을 나타낸다. 일부 실시양태에서, 재조합 핵산이 세포에서 발현될 때, 세포는 재조합 핵산을 발현하지 않는 세포에 비해 표적 세포의 항원을 가진 표적 세포의 적어도 3배, 4배, 5배, 6배, 7배, 8배, 9배, 10배, 20배, 30배 또는 적어도 5배 증가된 식세포작용을 나타낸다.

[0474] 일부 실시양태에서, SIRP-ΔICD의 발현은 이를 발현하는 세포의 식세포작용을, SIRP-ΔICD를 발현하지 않는 세포에 비해 1.1배 이상, 1.2배 이상, 1.3배 이상, 1.4배 이상, 1.5배 이상, 1.6배 이상, 1.7배 이상, 1.8배 이상, 1.9배 이상, 2배 이상, 3배 이상, 4배 이상, 5배 이상, 8배 이상, 10배 이상, 15배 이상, 20배 이상, 30배 이상, 40배 이상, 50배 이상, 60배 이상, 70배 이상, 80배 이상, 90배 이상 또는 100배 이상 향상시킨다.

[0475] 일부 실시양태에서, SIRP-ΔICD와 본원에 기재된 식세포 수용체를 코딩하는 CFP를 공-발현하는 세포는 이 단백

질들 중 어느 하나를 발현하지 않는 세포에 비해 증강된 식세포작용을 나타낸다. 일부 실시양태에서, SIRP-ΔICD와 본원에 기재된 식세포 수용체를 코딩하는 CFP의 공-발현은 SIRP-ΔICD 또는 식세포 수용체를 코딩하는 CFP를 발현하지 않는 세포에 비해 (식세포 지수의 배수 변화로 측정된) 식세포 잠재력의 2배 초과, 3배 초과, 4배 초과, 5배 초과, 6배 초과, 7배 초과, 8배 초과, 9배 초과, 10배 초과, 20배 초과, 30배 초과, 40배 초과, 50배 초과, 60배 초과, 70배 초과, 80배 초과, 90배 초과, 100배 초과, 150배 초과 또는 200배 초과의 증가를 나타낸다.

[0476] 일부 실시양태에서, SIRPα의 세포외 도메인 및 식세포 수용체의 세포내 도메인을 차단하는 CD47을 발현하는 CFP 중 어느 하나의 발현은 이를 발현하는 세포의 식세포 활성을, CFP를 발현하지 않는 세포에 비해 또는 SIRP-ΔICD를 발현하는 세포에 비해 적어도 1.5배 이상, 1.6배 이상, 1.7배 이상, 1.8배 이상, 1.9배 이상, 2배 이상, 3배 이상, 4배 이상, 5배 이상, 8배 이상, 10배 이상, 15배 이상, 20배 이상, 30배 이상, 40배 이상, 50배 이상, 60배 이상, 70배 이상, 80배 이상, 90배 이상 또는 100배 이상 증강시킨다.

[0477] 일부 실시양태에서, SIRP-ΔICD를 발현하는 세포에 의한 표적 세포의 식세포작용의 향상은 SIRP-ΔICD를 발현하지 않는 식세포에 비해 크게 증가된다.

[0478] 일부 실시양태에서, SIRPα의 세포외 도메인 및 식세포 수용체의 세포내 도메인을 차단하는 CD47을 포함하는 CFP를 발현하는 세포에 의한 표적 세포의 식세포작용의 향상은 융합 단백질을 발현하지 않는 대조군 식세포 또는 SIRP-ΔICD를 발현하는 대조군 식세포에 비해 크게 증가된다.

[0479] 일부 실시양태에서, 본원에 기재된 재조합 핵산이 세포에서 발현될 때, 세포는 증가된 사이토카인 생성을 나타낸다. 사이토카인은 IL-1, IL-6, IL-12, IL-23, TNF, CXCL9, CXCL10, CXCL11, IL-18, IL-23, IL-27 및 인터페론 중 어느 하나를 포함할 수 있다.

[0480] 일부 실시양태에서, 본원에 기재된 재조합 핵산이 세포에서 발현될 때, 세포는 증가된 세포 이동을 나타낸다.

[0481] 일부 실시양태에서, 본원에 기재된 재조합 핵산이 세포에서 발현될 때, 세포는 증가된 면역 활성을 나타낸다. 일부 실시양태에서, 재조합 핵산이 세포에서 발현될 때, 세포는 MHC II의 증가된 발현을 나타낸다. 일부 실시양태에서, 재조합 핵산이 세포에서 발현될 때, 세포는 CD80의 증가된 발현을 나타낸다. 일부 실시양태에서, 재조합 핵산이 세포에서 발현될 때, 세포는 CD86의 증가된 발현을 나타낸다. 일부 실시양태에서, 재조합 핵산이 세포에서 발현될 때, 세포는 증가된 iNOS 생성을 나타낸다.

[0482] 일부 실시양태에서, 재조합 핵산이 세포에서 발현될 때, 세포는 재조합 핵산을 발현하지 않는 세포에 비해 표적 세포의 항원을 발현하는 표적 세포의 감소된 트로고사이토시스(trogocytosis)를 나타낸다.

[0483] 일부 실시양태에서, 키메라 수용체는 글리코실화될 수 있고/있거나, 페길화될 수 있고/있거나 다른 방식으로 번역 후 변형될 수 있다. 추가 실시양태에서, 글리코실화, 페길화, 및/또는 다른 번역 후 변형은 생체내 또는 시험관내에서 일어날 수 있고/있거나 화학적 기법을 이용함으로써 수행될 수 있다. 추가 실시양태에서, 임의의 글리코실화, 페길화 및/또는 다른 번역 후 변형은 N-연결될 수 있거나 O-연결될 수 있다. 일부 실시양태에서, 키메라 수용체들 중 어느 한 키메라 수용체는 세포외 도메인이 리간드에 의해 결합될 때 신호가 변환되어 대식세포를 분극화하도록 효소적 또는 기능적 활성을 나타낼 수 있다.

[0484] 일부 실시양태에서, 키메라 융합 단백질(CFP)은 예를 들어, 서열번호 1로 기재된 아미노산 서열을 가진 중쇄 가변 영역(VH)을 포함하는, CD5에 결합하도록 표적화된 세포외 도메인(ECD)(CD5 결합 도메인)을 포함한다. 일부 실시양태에서, 키메라 CFP는 서열번호 1과 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 서열 동일성을 가진 아미노산 서열을 포함하는 CD5 결합 중쇄 가변 도메인을 포함한다. 일부 실시양태에서, CD5에 결합하도록 표적화된 세포외 도메인(ECD)(CD5 결합 도메인)은 서열번호 2로 기재된 아미노산 서열을 가진 경쇄 가변 도메인(V_L)을 포함한다. 일부 실시양태에서, 키메라 CFP는 서열번호 2와 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 서열 동일성을 가진 아미노산 서열을 포함하는 CD5 결합 경쇄 가변 도메인을 포함한다.

[0485] 일부 실시양태에서, CFP는 예를 들어, 서열번호 8로 기재된 중쇄 가변 도메인 아미노산 서열 및 서열번호 9로 기재된 경쇄 가변 도메인 아미노산 서열을 가진, HER2에 결합하도록 표적화된 세포외 도메인(HER2 결합 도메인)을 포함한다. 일부 실시양태에서, CFP는 서열번호 8과 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 서열 동일성을 가진 아미노산 서열을 포함하는 HER2 결합 중쇄 가변 도메인을 포함한다. 일부 실시양태에서, CFP는 서열번호 9와 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 서열 동일성을 가진 아미노산 서열을 포함하는 HER2 결합 경쇄 가변 도메인을 포함한다.

[0486] 일부 실시양태에서, CFP는 ECD를 막횡단(TM)에 연결하는 힌지를 포함한다. 일부 실시양태에서, 힌지는 CD8 수용체의 힌지 영역의 아미노산 서열을 포함한다. 일부 실시양태에서, CFP는 서열번호 7로 기재된 아미노산 서열을 가진 힌지(CD8α 쇄 힌지 도메인)를 포함할 수 있다. 일부 실시양태에서, PFP 힌지 영역은 서열번호 7과 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 서열 동일성을 가진 아미노산 서열을 포함한다.

[0487] 일부 실시양태에서, CFP는 예를 들어, 서열번호 6으로 기재된 아미노산 서열을 가진 CD8 막횡단 영역을 포함한다. 일부 실시양태에서, CFP TM 영역은 서열번호 6과 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 서열 동일성을 가진 아미노산 서열을 포함한다.

[0488] 일부 실시양태에서, CFP는 FcR 도메인을 가진 세포내 도메인을 포함한다. 일부 실시양태에서, CFP는 서열번호 3으로 기재된 아미노산 서열, 또는 서열번호 3과 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 서열 동일성을 가진 서열을 포함하는 FcR 도메인 세포내 도메인을 포함한다.

[0489] 일부 실시양태에서, CFP는 PI3K 동원 도메인을 가진 세포내 도메인을 포함한다. 일부 실시양태에서, PI3K 동원 도메인은 서열번호 4로 기재된 아미노산 서열을 포함한다. 일부 실시양태에서, PI3K 동원 도메인은 서열번호 4와 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 서열 동일성을 가진 아미노산 서열을 포함한다.

[0490] 일부 실시양태에서, CFP는 CD40 세포내 도메인을 가진 세포내 도메인을 포함한다. 일부 실시양태에서, CD40 ICD는 서열번호 5로 기재된 아미노산 서열을 포함한다. 일부 실시양태에서, CD40 ICD는 서열번호 5와 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 서열 동일성을 가진 아미노산 서열을 포함한다.

[0491] 일부 실시양태에서, CD5 결합 도메인은 (i) 서열번호 1의 가변 중쇄(V_H) 서열 또는 서열번호 1과 적어도 90% 서열 동일성을 가진 가변 중쇄(V_H) 서열; 및 (ii) 서열번호 2의 가변 경쇄(V_L) 서열 또는 서열번호 2와 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 가변 경쇄(V_L) 서열을 포함하는 scFv를 포함한다. 일부 실시양태에서, CD5 결합 도메인은 서열번호 33의 서열, 또는 서열번호 33과 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함하는 scFv를 포함한다. 일부 실시양태에서, HER2 결합 도메인은 (i) 서열번호 8의 가변 중쇄(V_H) 서열 또는 서열번호 8과 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 가변 중쇄(V_H) 서열; 및 (ii) 서열번호 9의 가변 경쇄(V_L) 서열 또는 서열번호 9와 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 가변 경쇄(V_L) 서열을 포함하는 scFv를 포함한다. 일부 실시양태에서, CD5 결합 도메인은 서열번호 32의 서열, 또는 서열번호 32와 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함하는 scFv를 포함한다. 일부 실시양태에서, CFP는 세포내 도메인을 추가로 포함하고, 이때 세포내 도메인은 하나 이상의 세포내 신호전달 도메인을 포함하고, 세포내 도메인을 포함하는 야생형 단백질은 세포외 도메인을 포함하지 않는다.

[0492] 일부 실시양태에서, 세포외 도메인은 CD8로부터 유래한 힌지 도메인을 추가로 포함하고, 이때 상기 힌지 도메인은 막횡단 도메인 및 항-CD5 결합 도메인에 작동 가능하게 연결된다. 일부 실시양태에서, 세포외 힌지 도메인은 서열번호 7의 서열, 또는 서열번호 7과 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함한다.

[0493] 일부 실시양태에서, CFP는 서열번호 30의 막횡단 도메인, 또는 서열번호 30과 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 막횡단 도메인에 융합된 세포외 도메인을 포함한다. 일부 실시양태에서, CFP는 서열번호 31의 막횡단 도메인, 또는 서열번호 31과 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 막횡단 도메인에 융합된 세포외 도메인을 포함한다.

[0494] 일부 실시양태에서, 막횡단 도메인은 CD8 막횡단 도메인을 포함한다. 일부 실시양태에서, 막횡단 도메인은 서열번호 6 또는 29의 서열, 또는 서열번호 6 또는 29와 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, 막횡단 도메인은 서열번호 18의 서열, 또는 서열번호 18과 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, 막횡단 도메인은 서열번호 34의 서열, 또는 서열번호 34와 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, 막횡단 도메인은 서열번호 19의 서열, 또는 서열번호 19와

적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함한다.

[0495] 일부 실시양태에서, CFP는 식세포 신호전달 도메인을 포함하는 하나 이상의 세포내 신호전달 도메인을 포함한다. 일부 실시양태에서, 식세포작용 신호전달 도메인은 Megf10, MerTk, FcRα 및 Bai1 이외의 수용체로부터 유래한 세포내 신호전달 도메인을 포함한다. 일부 실시양태에서, 식세포작용 신호전달 도메인은 Megf10, MerTk, FcR 및 Bai1 이외의 수용체로부터 유래한 세포내 신호전달 도메인을 포함한다. 일부 실시양태에서, 식세포작용 신호전달 도메인은 CD3ζ 이외의 수용체로부터 유래한 세포내 신호전달 도메인을 포함한다. 일부 실시양태에서, 식세포작용 신호전달 도메인은 FcRγ, FcRα 또는 FcRε으로부터 유래한 세포내 신호전달 도메인을 포함한다. 일부 실시양태에서, 식세포작용 신호전달 도메인은 CD3ζ로부터 유래한 세포내 신호전달 도메인을 포함한다. 일부 실시양태에서, CFP는 서열번호 3, 20, 27 및 28 중 어느 한 서열의 세포내 신호전달 도메인, 또는 서열번호 3, 20, 27 및 28 중 어느 한 서열과 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 세포내 신호전달 도메인을 포함한다. 일부 실시양태에서, 하나 이상의 세포내 신호전달 도메인은 염증촉진성 신호전달 도메인을 추가로 포함한다. 일부 실시양태에서, 염증촉진성 신호전달 도메인은 PI3-키나제(PI3K) 동원 도메인을 포함한다. 일부 실시양태에서, 염증촉진성 신호전달 도메인은 서열번호 4의 서열, 또는 서열번호 4와 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, 염증촉진성 신호전달 도메인은 CD40의 세포내 신호전달 도메인으로부터 유래한다. 일부 실시양태에서, 염증촉진성 신호전달 도메인은 서열번호 5의 서열, 또는 서열번호 5와 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, CFP는 서열번호 21의 세포내 신호전달 도메인, 또는 서열번호 21과 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 세포내 신호전달 도메인을 포함한다. 일부 실시양태에서, CFP는 서열번호 23의 세포내 신호전달 도메인, 또는 서열번호 23과 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 세포내 신호전달 도메인을 포함한다.

[0496] 일부 실시양태에서, CFP는 서열번호 14의 서열, 또는 서열번호 14와 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, CFP는 서열번호 15의 서열, 또는 서열번호 15와 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, CFP는 서열번호 16의 서열, 또는 서열번호 16과 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, CFP는 서열번호 24의 서열, 또는 서열번호 24와 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함한다. 일부 실시양태에서, CFP는 서열번호 25의 서열, 또는 서열번호 25와 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 가진 서열을 포함한다.

[0497] 일부 실시양태에서, CFP는 (a) (i) CD5에 특이적으로 결합하는 scFv, 및 (ii) CD8로부터 유래한 힌지 도메인, CD28로부터 유래한 힌지 도메인 또는 CD68의 세포외 도메인의 적어도 일부를 포함하는 세포외 도메인; (b) CD8 막횡단 도메인, CD28 막횡단 도메인, CD2 막횡단 도메인 또는 CD68 막횡단 도메인; 및 (c) (i) FcRα, FcRγ 또는 FcRε으로부터 유래한 제1 세포내 신호전달 도메인, 및 (ii) (A) PI3K 동원 도메인을 포함하거나 (B) CD40으로부터 유래한 제2 세포내 신호전달 도메인을 포함하는 적어도 2개의 세포내 신호전달 도메인을 포함하는 세포내 도메인을 포함한다. 일부 실시양태에서, CFP는 (i) 식세포 수용체 세포내 도메인으로부터 유래한 제1 세포내 신호전달 도메인, 및 (ii) (A) PI3K 동원 도메인을 포함하거나 (B) CD40으로부터 유래한 스캐빈저 수용체 식세포 수용체 세포내 도메인으로부터 유래한 제2 세포내 신호전달 도메인을 포함하는 적어도 2개의 세포내 신호전달 도메인을 포함하는 세포내 도메인을 상기 (c)에 대한 대안으로서 포함한다. 세포내 신호전달 도메인의 기원이 될 수 있는 예시적인 스캐빈저 수용체는 표 2에서 확인될 수 있다. 일부 실시양태에서, CFP는 선천성 면역 수용체의 세포내 신호전달 도메인으로부터 유래한 세포내 신호전달 도메인을 포함한다.

[0498] 일부 실시양태에서, 재조합 폴리핵산은 mRNA이다. 일부 실시양태에서, 재조합 폴리핵산은 circRNA이다. 일부 실시양태에서, 재조합 폴리핵산은 바이러스 벡터이다. 일부 실시양태에서, 재조합 폴리핵산은 바이러스 벡터를 통해 전달된다.

[0499] 일부 실시양태에서, 골수 세포는 CD14+ 세포, CD14+/CD16- 세포, CD14+/CD16+ 세포, CD14-/CD16+ 세포, CD14-/CD16- 세포, 수지상 세포, M0 대식세포, M2 대식세포, M1 대식세포 또는 모자이크 골수 세포/대식세포/수지상

세포이다.

[0500] 한 측면에서, 본원은 암의 치료를 필요로 하는 인간 대상체에서 암을 치료하는 방법으로서, (a) (i) 항-CD5 결합 도메인을 포함하는 세포외 도메인 및 (ii) 세포외 도메인에 작동 가능하게 연결된 막횡단 도메인을 포함하는 키메라 융합 단백질(CFP)을 코딩하는 서열을 포함하는 재조합 폴리핵산 서열을 포함하는 골수 세포; 및 (b) 약학적으로 허용되는 담체를 포함하는 약학 조성물을 상기 인간 대상체에게 투여하는 단계를 포함하는 방법을 제공하는 것으로; 이때 골수 세포는 CFP를 발현한다.

[0501] 일부 실시양태에서, CFP가 대상체의 표적 암 세포에 의해 발현된 CD5에 결합할 때, 골수 세포의 사멸 또는 식세포작용 활성은 CFP를 발현하지 않는 골수 세포에 비해 20% 이상 증가된다. 일부 실시양태에서, 종양의 성장은 인간 대상체에서 억제된다.

[0502] 일부 실시양태에서, 암은 CD5+ 암이다. 일부 실시양태에서, 암은 백혈병, T 세포 림프종 또는 B 세포 림프종이다. 일부 실시양태에서, CFP는 하기 표 A 및/또는 표 B에 제시된 하나 이상의 서열을 포함한다.

[0503] [표 A]

CFP 및 이의 도메인의 예시적인 서열

서열 번호	PFP/도메인	서열
1	항-CD5 중쇄 가변 도메인	EIQLVQSGGGLVKPGGSVRISCAASGYTFTNYGMNWV RQAPGKGLEWMGWINTHTGEPTYADSFKGRFTFSLDD SKNTAYLQINSLRAEDTAVYFCTRRGYDWYFDVWGQ GTTVTV
2	항-CD5 경쇄 가변 도메인	DIQMTQSPSSLSASVGDRVTITCRASQDINSYLSWFQQK PGKAPKTLIYRANRLESGVPSRFSGSGSGTDYTLTISSLQ YEDFGIYYCQQYDESPWTFGGGTKLEIK
33	항-CD5 scFv	EIQLVQSGGGLVKPGGSVRISCAASGYTFTNYGMNWV RQAPGKGLEWMGWINTHTGEPTYADSFKGRFTFSLDD SKNTAYLQINSLRAEDTAVYFCTRRGYDWYFDVWGQ GTTVTV*SSGGGGSGGGGSGGGGS*DIQMTQSPSSLSASV GDRVTITCRASQDINSYLSWFQQKPGKAPKTLIYRANR LESGVPSRFSGSGSGTDYTLTISSLQYEDFGIYYCQQYD ESPWTFGGGTKLEIK
3	FcRɣ-쇄 세포내 신호전달 도메인	LYCRRLKIQVRKAAITSYEKSDGVYTGLSTRNQETYET LKHEKPPQ
20	FcRɣ-쇄 세포내 신호전달 도메인	LYCRLKIQVRKAAITSYEKSDGVYTGLSTRNQETYETL KHEKPPQ
27	FcRɣ-쇄 세포내 신호전달 도메인	RLKIQVRKAAITSYEKSDGVYTGLSTRNQETYETLKHE KPPQ
28	FcRɣ-쇄 세포내 신호전달 도메인	RLKIQVRKAAITSYEKSDGVYTGLSTRNQETYETLKHE KPPQ
4	PI3K 동원 도메인	YEDMRGILYAAPQLRSIRGQPGPNHEEDADSYENM
5	CD40 세포내 도메인	KKVAKKPTNKAPHPKQEPQEINFPDDLPGSNTAAPVQE TLHGCQPVTQEDGKESRISVQERQ
6	CD8α 쇄 막횡단 도메인	IYIWAPLAGTCGVLLLSLVIT
29	CD8α 쇄 막횡단 도메인	IYIWAPLAGTCGVLLLSLVITLYC
7	CD8α 쇄 힌지 도메인	ALSNSIMYFSHFVPVFLPAKPTTTPAPRPPTPAPTIASQP LSLRPEACRPAAGGAVHTRGLD
8	항-HER2 중쇄 가변 도메인	DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQ KPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDFTLTISSL QPEDFATYYCQQHYTTPPTFGQGTKVEIKRTGSTSGSG KPGSGEGSEVQLVE

[0504]

서열 번호	PFP/도메인	서열
9	항-HER2 경쇄 가변 도메인	LVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEW VARIYPTNGYTRYADSVKGRFTISADTSKNTAYLQMNS LRAEDTAVYYCSRWGGDGFYAMDVWGQGTLVTV
32	항-HER2 scFv	DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQ KPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDFTLTISSL QPEDFATYYCQQHYTTPPTFGQGTKVEIKRTGSTSGSG KPGSGEGSEVQLVE*SSGGGGSGGGGSGGGGS*LVQPGGS LRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTN GYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTA VYYCSRWGGDGFYAMDVWGQGTLVTV
17	GMCSF 신호 펩티드	MWLQSLLLLGTVACSIS
18	CD28 막횡단 도메인	FWVLVVVGGVLACYSLLVTVAFIIFWV
34	CD2 막횡단 도메인	IYLIIGICGGGSLLMVFVALLVFYIT
19	CD68 막횡단 도메인	ILLPLIIGLILLGLLALVLIAFCII
21	TNFR1 세포내 도메인	QRWKSKLYSIVCGKSTPEKEGELEGTTTKPLAPNPSFSP TPGFTPTLGFSPVPSSTFTSSSTYTPGDCPNFAAPRREVA PPYQGADPILATALASDPIPNPLQKWEDSAHKPQSLDTD DPATLYAVVENVPPLRWKEFVRRLGLSDHEIDRLELQN GRCLREAQYSMLATWRRRTPRREATLELLGRVLRDMD LLGCLEDIEEALCGPAALPPAPSLLR
22	TNFR2 세포내 도메인	PLCLQREAKVPHLPADKARGTQGPEQQHLLITAPSSSSS SLESSASALDRRAPTRNQPQAPGVEASGAGEARASTGS SDSSPGGHGTQVNVTCIVNVCSSSDHSSQCSSQASSTM GDTDSSPSESPKDEQVPFSKEECAFRSQLETPETLLGSTE EKPLPLGVPDAGMKPS
23	MDA5 세포내 도메인	MSNGYSTDENFRYLISCFRARVKMYIQVEPVLDYLTFL PAEVKEQIQRTVATSGNMQAVELLLSTLEKGVWHLGW TREFVEALRRTGSPLAARYMNPELTDLPSPSFENAHDE YLQLLNLLQPTLVDKLLVRDVLDKCMEEELLTIEDRNR IAAAENNGNESGVRELLKRIVQKENWFSAFLNVLRQTG NNELVQELTGSDCSESNAEIEN

[0505]

서열 번호	PFP/도메인	서열
30	CD8α 쇄 힌지 도메인 + 막횡단 도메인	ALSNSIMYFSHFVPVFLPAKPTTTPAPRPPTPAPTIASQP LSLRPEACRPAAGGAVHTRGLDIYIWAPLAGTCGVLLL SLVITLYC
31	CD8α 쇄 힌지 도메인 + 막횡단 도메인	ALSNSIMYFSHFVPVFLPAKPTTTPAPRPPTPAPTIASQP LSLRPEACRPAAGGAVHTRGLDIYIWAPLAGTCGVLLL SLVIT
14	CD5-FcRγ-PI3K	MWLQSLLLLGTVACSISEIQLVQSGGGLVKPGGSVRISC AASGYTFTNYGMNWVRQAPGKGLEWMGWINTHTGEP TYADSFKGRFTFSLDDSKNTAYLQINSLRAEDTAVYFC TRRGYDWYFDVWGQGTTVTV*SSGGGGSGGGGSGGGG S*DIQMTQSPSSLSASVGDRVTITCRASQDINSYLSWFQQ KPGKAPKTLIYRANRLESGVPSRF*SGSGSG*TDYTLTISSL QYEDFGIYYCQQYDESPWTFGGGTKLEIK*SGGGGSG*AL SNSIMYFSHFVPVFLPAKPTTTPAPRPPTPAPTIASQPLSL RPEACRPAAGGAVHTRGLDIYIWAPLAGTCGVLLLSLV ITLYCRRLKIQVRKAAITSYEKSDGVYTGLSTRNQETYE TLKHEKPPQ*GSGS*YEDMRGILYAAPQLRSIRGQPGPNHE EDADSYENM
15	HER2-FcRγ-PI3K	MWLQSLLLLGTVACSISDIQMTQSPSSLSASVGDRVTIT CRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVP SRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFG QGTKVEIKRTGSTSGSGKPGSGEGSEVQLVE*SGGG*LVQ PGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARI YPTNGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAE DTAVYYCSRWGGDGFYAMDVWGQGTLVTV*SSSGGGG SG*ALSNSIMYFSHFVPVFLPAKPTTTPAPRPPTPAPTIAS QPLSLRPEACRPAAGGAVHTRGLDIYIWAPLAGTCGVL LLSLVITLYCRRLKIQVRKAAITSYEKSDGVYTGLSTRN QETYETLKHEKPPQ*GSGS*YEDMRGILYAAPQLRSIRGQP GPNHEEDADSYENM

[0506]

[0507]

서열 번호	PFP/도메인	서열
16	CD5-FcRγ-CD40	MWLQSLLLLGTVACSISEIQLVQSGGGLVKPGGSVRISC AASGYTFTNYGMNWVRQAPGKGLEWMGWINTHTGEP TYADSFKGRFTFSLDDSKNTAYLQINSLRAEDTAVYFC TRRGYDWYFDVWGQGTTVTV*SSGGGGSGGGGSGGGG* *S*DIQMTQSPSSLSASVGDRVTITCRASQDINSYLSWFQQ KPGKAPKTLIYRANRLESGVPSRFSGSGSGTDYTLTISSL QYEDFGIYYCQQYDESPWTFGGGTKLEIK*SGGGGSG*AL SNSIMYFSHFVPVFLPAKPTTTPAPRPPTPAPTIASQPLSL RPEACRPAAGGAVHTRGLDIYIWAPLAGTCGVLLLSLV ITLYCRLKIQVRKAAITSYEKSDGVYTGLSTRNQETYET LKHEKPPQKKVAKKPTNKAPHPKQEPQEINFPDDLPGS NTAAPVQETLHGCQPVTQEDGKESRISVQERQ
24	CD5-FcRγ-MDA5	MWLQSLLLLGTVACSISEIQLVQSGGGLVKPGGSVRISC AASGYTFTNYGMNWVRQAPGKGLEWMGWINTHTGEP TYADSFKGRFTFSLDDSKNTAYLQINSLRAEDTAVYFC TRRGYDWYFDVWGQGTTVTVSSGGGGSGGGGSGGGG SDIQMTQSPSSLSASVGDRVTITCRASQDINSYLSWFQQ KPGKAPKTLIYRANRLESGVPSRFSGSGSGTDYTLTISSL QYEDFGIYYCQQYDESPWTFGGGTKLEIKSGGGGSGAL SNSIMYFSHFVPVFLPAKPTTTPAPRPPTPAPTIASQPLSL RPEACRPAAGGAVHTRGLDIYIWAPLAGTCGVLLLSLV ITLYCRLKIQVRKAAITSYEKSDGVYTGLSTRNQETYET LKHEKPPQGSGSMSNGYSTDENFRYLISCFRARVKMYI QVEPVLDYLTFLPAEVKEQIQRTVATSGNMQAVELLLS TLEKGVWHLGWTREFVEALRRTGSPLAARYMNPELTD LPSPSFENAHDEYLQLLNLLQPTLVDKLLVRDVLDKCM EEELLTIEDRNRIAAAENNGNESGVRELLKRIVQKENW FSAFLNVLRQTGNNELVQELTGSDCSESNAEIEN

서열 번호	PFP/도메인	서열
25	CD5-FcRγ-TNFR1	MWLQSLLLLGTVACSISEIQLVQSGGGLVKPGGSVRISC AASGYTFTNYGMNWVRQAPGKGLEWMGWINTHTGEP TYADSFKGRFTFSLDDSKNTAYLQINSLRAEDTAVYFC TRRGYDWYFDVWGQGTTVTVSSGGGGSGGGGSGGGG SDIQMTQSPSSLSASVGDRVTITCRASQDINSYLSWFQQ KPGKAPKTLIYRANRLESGVPSRFSGSGSGTDYTLTISSL QYEDFGIYYCQQYDESPWTFGGGTKLEIKSGGGGSGAL SNSIMYFSHFVPVFLPAKPTTTPAPRPPTPAPTIASQPLSL RPEACRPAAGGAVHTRGLDIYIWAPLAGTCGVLLLSLV ITLYCRLKIQVRKAAITSYEKSDGVYTGLSTRNQETYET LKHEKPPQGSGSQRWKSKLYSIVCGKSTPEKEGELEGT TTKPLAPNPSFSPTPGFTPTLGFSPVPSSTFTSSSTYTPGD CPNFAAPRREVAPPYQGADPILATALASDPIPNPLQKWE DSAHKPQSLDTDDPATLYAVVENVPPLRWKEFVRRLG LSDHEIDRLELQNGRCLREAQYSMLATWRRRTPRREAT LELLGRVLRDMDLLGCLEDIEEALCGPAALPPAPSLLR
26	CD5-FcRγ-TNFR2	MWLQSLLLLGTVACSISEIQLVQSGGGLVKPGGSVRISC AASGYTFTNYGMNWVRQAPGKGLEWMGWINTHTGEP TYADSFKGRFTFSLDDSKNTAYLQINSLRAEDTAVYFC TRRGYDWYFDVWGQGTTVTVSSGGGGSGGGGSGGGG SDIQMTQSPSSLSASVGDRVTITCRASQDINSYLSWFQQ KPGKAPKTLIYRANRLESGVPSRFSGSGSGTDYTLTISSL QYEDFGIYYCQQYDESPWTFGGGTKLEIKSGGGGSGAL SNSIMYFSHFVPVFLPAKPTTTPAPRPPTPAPTIASQPLSL RPEACRPAAGGAVHTRGLDIYIWAPLAGTCGVLLLSLV ITLYCRLKIQVRKAAITSYEKSDGVYTGLSTRNQETYET LKHEKPPQGSGSPLCLQREAKVPHLPADKARGTQGPEQ QHLLITAPSSSSSSLESSASASALDRRAPTRNQPQAPGVEAS GAGEARASTGSSDSSPGGHGTQVNVTCIVNVCSSSDHS SQCSSQASSTMGDTDSSPSESPKDEQVPFSKEECAFRSQ LETPETLLGSTEEKPLPLGVPDAGMKPS

[0508]

[0509] [표 B]

링커 서열

서열번호	서열
10	SSGGGGSGGGGSGGGGS
11	SGGGGSG
12	SGGG
13	GSGS

[0510]

[0511] *IV. 세포의 게놈 내로의 전달 및 통합을 위한 비코딩 외생성 서열*

[0512] 비코딩 서열은 세포 내로 전달될 수 있고 세포의 게놈 내로 통합되도록 디자인될 수 있다. 본원에서 사용된 비코딩 서열은 번역된 단백질 생성물을 생성하지 않으나 조절 요소, 예컨대, 전사된 생성물, 예컨대, 억제 RNA를 가질 수 있는 서열이다. 일부 실시양태에서, 이러한 서열은 miRNA 서열일 수 있다. 일부 실시양태에서, 서열은 siRNA 생성을 위한 서열일 수 있다. 일부 실시양태에서, 서열은 인트론 서열을 포함할 수 있거나, 하나 이상의 DNA 결합 단백질이 결합 부위에 도킹하여 인접 영역의 성질과 거동에 영향을 미치도록 생성된 결합 부위를 포함할 수 있다. 일부 실시양태에서, 서열은 전사 인자 결합 부위일 수 있다. 일부 실시양태에서, 서열은 인핸서 결

합 부위를 포함할 수 있다. 일부 실시양태에서, 서열은 토포이소머라제(topoisomerase), 자이라제(gyrase), 역전사효소, 중합효소, 폴리 A 결합 단백질, 구아닐릴 사이클라제(guanylyl cyclase), 리가제(ligase), 제한효소, DNA 메틸라제(methylase), HDAC 효소 등에 대한 결합 부위를 포함할 수 있다. 일부 실시양태에서, 비코딩 서열은 이종염색질을 조작하도록 유도될 수 있다. 본원에서 비코딩 삽입체 서열로서도 지칭될 수 있는 비코딩 서열은 몇 개의 뉴클레오타이드 내지 5 kB의 길이를 가질 수 있다.

[0513] ***V. 삽입체 서열을 포함하는 플라스미드 디자인 및 재조합 핵산 디자인***

[0514] 하나 이상의 단백질 또는 폴리펩티드를 코딩하는 하나 이상의 서열을 포함하는 핵산 구축물은 전사 및 mRNA 생성을 위해 플라스미드 내로 통합된다. mRNA는 세포 추출물의 합성 시스템을 이용함으로써 시험관내 시스템에서 전사될 수 있다. 대안적으로, mRNA는 세포에서 생성되고 수거될 수 있다. 세포는 박테리아 세포와 같은 원핵 세포일 수 있다. 일부 실시양태에서, 세포는 진핵 세포이다. 일부 실시양태에서, 전사는 합성 시스템에서 일어난다. 본원은 예시적인 플라스미드 디자인을 제공한다.

[0515] 본 개시내용의 다양한 측면의 일부 실시양태에서, 플라스미드는 단백질 또는 폴리펩티드를 코딩하는 관심 있는 이종 서열을 포함하는 mRNA 분자를 발현하도록 디자인된다. 플라스미드는 특히, 단백질 또는 폴리펩티드를 코딩하는 관심 있는 이종 서열의 통합을 위한 게놈 통합 요소에 대한 서열; 숙주 게놈(예를 들어, mRNA가 투여되는 대상체) 내로의 통합 후 숙주에서의 그의 발현을 위해 필요한 그의 별도의 프로모터 및 조절 요소에 작동 가능하게 연결된, 형질전이유전자 또는 이의 단편을 포함하는 서열; 예를 들어, 박테리아 세포 또는 세포 추출물에서 mRNA를 발현시키기 위한 프로모터 및 3' 안정화 요소를 포함하는, mRNA의 전사 및 생성을 위한 하나 이상의 조절 요소; 하나 이상의 검출 마커 및/또는 선택 마커를 위한 서열을 포함한다.

[0516] 당업자에게 공지되어 있는 바와 같이, 플라스미드 골격은 전사된 서열의 최상의 발현을 위해 다양한 방식으로, 예를 들어, 하나 이상의 바람직한 제한 분해 부위를 MCS(다중 클로닝 부위)에 도입하거나, 전체 mRNA 전사를 위한 원하는 프로모터, 예컨대, T7 프로모터를 도입하거나, 플라스미드 벡터 내의 기존 서열을 하나 이상의 원하는 서열로 교체하거나, 선택 마커 서열과 같은 하나 이상의 원하는 분절을 도입함으로써(그러나, 이들로 제한되지 않음) 개선될 수 있는 이용 가능한 벡터, 예컨대, 사내 또는 상업적으로 개발된 벡터일 수 있다.

[0517] 플라스미드는 5' 영역에 있는 프로모터와 같은 전사 조절 요소 및 3' 안정화 요소를 포함한다. 일부 실시양태에서, 프로모터는 이. 콜라이 박테리아 세포와 같은 원하는 세포에서 향상된 mRNA 전사를 위해 선택된다. 일부 실시양태에서, 플라스미드의 전사를 위한 프로모터는 T7 프로모터, Sp6 프로모터, pL(람다) 프로모터, T3 프로모터, trp 프로모터, araBad 프로모터, lac 프로모터 또는 Ptac 프로모터로부터 선택된다. 일부 실시양태에서, 프로모터는 T7 프로모터이다. T7 또는 Sp6 프로모터는 항시성 프로모터이고 고수준의 전사 또는 시험관내 전사에 유용하다. 일부 실시양태에서, 3' 안정화 요소는 BGH 3' 요소, WPRE 3' 요소, SV40 요소, hGH 요소 및 기타 요소로부터의 서열이다. 3' 요소는 필요한 폴리 A 및 전사 종결 서열을 포함한다.

[0518] 예시적인 선택 마커는 항생제 선택 마커 및/또는 발현 검출 마커를 포함한다. 항생제 선택 마커는 앰피실린에 대한 내성을 부여하는 앰피실린 내성 유전자 서열(베타 락타마제(lactamase) 유전자 또는 이의 단편), 예를 들어, G418 선택 마커, 테트라사이클린에 대한 내성을 부여하는 테트라사이클린 내성 유전자 서열, 카나마이신에 대한 내성을 부여하는 카나마이신 내성 유전자 서열, 에리쓰로마이신에 대한 내성을 부여하는 에리쓰로마이신 내성 유전자 서열, 클로람페니콜에 대한 내성을 부여하는 클로람페니콜 내성 유전자 서열, 네오마이신에 대한 내성을 부여하는 네오마이신 내성 유전자 서열 등을 포함하나, 이들로 제한되지 않는다. 예시적인 발현 검출 마커는 FLAG, HA, GFP 등을 포함한다.

[0519] 일부 실시양태에서, 융합되어 있는 원하는 단백질 또는 펩티드의 발현을 위해 대용물로서 작용하는 하나 이상의 코딩 서열에 융합될 수 있는 다른 태그.

[0520] 일부 실시양태에서, 플라스미드는 길이가 20 kb 미만이다. 일부 실시양태에서, 플라스미드는 길이가 19 kb 미만이다. 일부 실시양태에서, 플라스미드는 길이가 20 kb 미만이다. 일부 실시양태에서, 플라스미드는 길이가 18 kb 미만이다. 일부 실시양태에서, 플라스미드는 길이가 20 kb 미만이다. 일부 실시양태에서, 플라스미드는 길이가 17 kb 미만이다. 일부 실시양태에서, 플라스미드는 길이가 20 kb 미만이다. 일부 실시양태에서, 플라스미드는 길이가 16 kb 미만이다. 일부 실시양태에서, 플라스미드는 길이가 15 kb 미만이다. 일부 실시양태에서, 플라스미드는 길이가 14 kb 미만이다. 일부 실시양태에서, 플라스미드는 길이가 13 kb 미만이다. 일부 실시양태에서, 플라스미드는 길이가 12 kb 미만이다. 일부 실시양태에서, 플라스미드는 길이가 약 15 kb, 약 14 kb, 약 13 kb, 약 12 kb 또는 약 10 kb이다.

[0521] 일부 실시양태에서, 코돈은 전사 시스템에 적합한 최대화된 전사를 위해 최적화된다.

[0522] ***VI. 생체내에서의 형질전이유전자의 발현과 관련된 특징***

[0523] *재조합 핵산 구축물(형질전이유전자)의 전사 조절 요소*

[0524] 일부 실시양태에서, 재조합 핵산은 코딩된 단백질의 원하는 발현 프로파일을 위해 조작될 수 있는 비코딩 영역 내에 하나 이상의 조절 요소를 포함한다. 일부 실시양태에서, 비코딩 영역은 적합한 인핸서를 포함할 수 있다. 일부 실시양태에서, 인핸서는 인핸서의 영향 하에서 코딩된 단백질의 발현을 시작하기 위해 세포 또는 세포를 포함하는 시스템에 첨가될 수 있는 조절제 단백질 또는 펩티드에 대한 결합 영역을 포함한다. 대조적으로, 조절 요소는 동족 단백질과 결합된 상태로 남아 있는 단백질 결합 도메인을 포함할 수 있고, 상기 단백질이 결합된 위치로부터 분리되어 단백질 합성의 시작을 허용하도록 세포외 신호가 제공될 때까지 재조합 단백질의 전사 및/또는 번역을 계속 억제할 수 있다. 예는 당업자에게 공지되어 있는 테트라사이클린 유도성(Tet 유도성 또는 Tet-on) 시스템 및 테트라사이클린 억제성(Tet-off) 시스템을 포함하나, 이들로 제한되지 않는다.

[0525] *대사 스위치를 포함하는 구축물:* 일부 실시양태에서, 구축물의 코딩 영역을 플랭킹하는 5' 및 3' 비번역 영역은 전술된 핵산 구축물에 의해 코딩된 재조합 단백질의 발현을 조절하도록 조작될 수 있다. 예를 들어, 3' UTR은 mRNA를 안정화시키기 위해 삽입되는 하나 이상의 요소를 포함할 수 있다. 일부 실시양태에서, mRNA를 안정화시키거나 불안정화시켜 mRNA 반감기를 조절할 수 있게 하는 RNA 결합 단백질의 결합을 야기하는 AU 풍부 요소(ARE) 서열이 3' UTR에 삽입된다.

[0526] 일부 실시양태에서, 3' UTR은 성숙 mRNA 가닥에 결합하여 번역을 방해하는 RNA 결합 단백질(예를 들어, GAPDH)에 대한 보존된 영역을 포함할 수 있다. 일부 실시양태에서, 해당과정은 mRNA 가닥 번역을 허용하는, RNA 결합 단백질(예를 들어, GAPDH)의 분리를 야기한다. 대사 스위치의 원리는 세포가 특정 대사 상태에 들어갈 때 표적 유전자의 발현을 유발하는 것이다. 예를 들어, GAPDH는 휴면 세포에서 RNA 결합 단백질(RBP)이다. 이 단백질은 3' UTR의 ARE 서열에 결합하여, mRNA의 번역을 방해한다. 세포가 해당과정에 들어갈 때, GAPDH는 글루코스를 ATP로 전환시켜, mRNA로부터 빠져나와 단백질 번역이 일어날 수 있게 하는 데 필요하다. 일부 실시양태에서, 재조합 핵산을 포함하는 세포가 존재하는 환경은 유전자 발현에 대한 대사 스위치를 제공한다. 예를 들어, 저산소 상태는 mRNA로부터의 GAPDH의 분리를 유도하는 대사 스위치를 유발할 수 있다. 따라서, mRNA의 발현은 대식세포가 순환계를 떠나 저산소 상태인 종양 환경으로 들어갈 때에만 유도될 수 있다. 이것은 핵산 또는 핵산을 포함하는 세포의 전신 투여를 허용하나, 종양 환경을 특이적으로 표적화하는 국소 발현을 보장한다.

[0527] 일부 실시양태에서, 핵산 구축물은 전술된 바와 같이, 예를 들어, 구축물의 일부가 항시성 발현 시스템의 조절 하에서 발현될 수 있게 하는 반면, 핵산의 또 다른 일부가 대사 스위치의 조절 하에서 발현되게 하는 분할 구축물일 수 있다. 일부 실시양태에서, 핵산은 바이시스트론 조절 하에 있을 수 있다. 일부 실시양태에서, 바이시스트론 벡터는 항시적 조절 하에 있을 수 있는 표적 인식 모이어티의 코딩 서열을 포함하는, 제1 조절 제어 하에 있는 제1 코딩 서열; 및 대사 스위치 하에 있을 수 있는 염증 유전자 발현을 코딩하는 제2 코딩 서열을 포함한다. 일부 실시양태에서, 바이시스트론 벡터는 단방향 벡터일 수 있다. 일부 실시양태에서, 바이시스트론 벡터는 양방향 벡터일 수 있다.

[0528] 일부 실시양태에서, ARE 서열은 ADK, ALDH18A1, ALDH6A1, ALDOA, ASS1, CCBL2, CS, DUT, ENO1, FASN, FDPS, GOT2, HADHB, HK2, HSD17B10, MDH2, NME1, NQO1, PKM2, PPP1CC, SUCLG1, TP11, GAPDH 또는 LDH에 결합하는, ARE 서열 결합을 위한 단백질 결합 모티프를 포함한다.

[0529] *약학 조성물 및 면역요법*

[0530] 한 측면에서, 본원은 (i) 형질전이유전자, 5' 및 3' 플랭킹 전위 또는 역전위 요소, 발현 조절 요소, 예컨대, 프로모터, 인트론; 및 전위효소 또는 역전위효소를 코딩하는 핵산을 포함하는 전위 또는 역전위 시스템 내에 통합된, 형질전이유전자를 코딩하는 핵산, (ii) 핵산 전달 비히클 및 약학적으로 허용되는 염 또는 부형제를 포함하는 약학 조성물을 제공한다.

[0531] 일부 실시양태에서, 약학 조성물은 세포의 게놈 내로 안정적으로 통합된 형질전이유전자를 코딩하는 핵산을 포함하는 세포 및 약학적으로 허용되는 부형제를 포함한다. 핵산 구축물은 양이온성 지질의 사용(Goddard, et al, Gene Therapy, 4:1231-1236, 1997; Gorman, et al, Gene Therapy 4:983-992, 1997; Chadwick, et al, Gene Therapy 4:937-942, 1997; Gokhale, et al, Gene Therapy 4:1289-1299, 1997; Gao, and Huang, Gene Therapy 2:710-722, 1995), 바이러스 벡터의 사용(Monahan, et al, Gene Therapy 4:40-49, 1997; Onodera, et al,

Blood 91:30-36, 1998), "네이키드 DNA"의 흡수 등에 의해 전달될 수 있다. 당분야에서 잘 공지되어 있는 세포 형질전환 기법(상기 논의 참조)은 핵산 구축물의 생체외 투여에 사용될 수 있다. 정확한 제제, 투여 경로 및 용량은 실험적으로 선택될 수 있다(예를 들어, 문헌[Fingl et al., 1975, "The Pharmacological Basis of Therapeutics", Ch. 1 pl] 참조).

[0532] 일부 실시양태에서, 형질전이유전자 및 전위 요소를 포함하는 핵산은 핵산을 세포 내부로 전달하는 공지된 방법, 예컨대, 리포펙타민 또는 인산칼슘의 사용을 통해, 또는 전기천공 또는 핵감염과 같은 물리적 수단을 통해 세포 내로 도입되거나 통합된다. 일부 실시양태에서, 핵산은 리포좀 또는 지질 나노입자 내로 캡슐화된다. LNP는 직경이 100 내지 300 nm이고 mRNA를, 대식세포를 비롯한 다양한 세포 유형에게 효율적으로 전달하는 수단을 제공한다. 일부 실시양태에서, 핵산은 다른 나노입자에 의해 전달된다. 일부 실시양태에서, CFP의 발현을 위한 벡터는 바이러스로부터 유래한 벡터, 즉 렌티바이러스 벡터 또는 아데노바이러스 벡터이다. 일부 실시양태에서, 재조합 핵산을 코딩하는 핵산은 렌티바이러스 벡터에 의해 코딩된다. 일부 실시양태에서, 렌티바이러스 벡터는 사내에서 제조되고 목적을 위해 대규모로 제조된다. 일부 실시양태에서, 당업자에게 공지되어 있는 바와 같이, 상업적으로 입수 가능한 렌티바이러스 벡터가 사용된다.

[0533] 일부 실시양태에서, 바이러스 벡터는 아데노 관련 바이러스(AAV) 벡터이다.

[0534] 이 방법은 외생성 핵산을 표적 세포 내로 도입하는 것이 요구되는 다양한 적용 분야에 유용하고 표적 세포에서 발현 카세트에 의해 코딩된 단백질을 발현하는 것이 요구되는 경우에서 특히 관심을 끌고, 이때 상기 표적 세포 또는 세포들은 다세포 유기체의 일부이다. 전위효소 시스템은 표적화 구축물이 예를 들어, 생체내 또는 생체외 프로토콜을 통해 표적 세포(들)에 들어갈 수 있게 하는 방식으로 유기체 또는 숙주에게 투여될 수 있다. 이러한 세포 또는 장기는 전형적으로 생체로 돌려보내진다.

[0535] 일부 실시양태에서, 면역 기능과 관련된 융합 단백질을 코딩하는 형질전이유전자는 생체외에서 대상체의 살아있는 세포 내로 안정적으로 통합되고, 이어서 상기 형질전이유전자를 포함하는 세포는 대상체로 돌려보내진다. 예시적으로 중요하게는, CFP 형질전이유전자(식세포 수용체 융합 단백질)는 형질전이유전자의 안정한 전달을 위해 재조합 핵산과 생체외에서 접촉되고 대상체의 질환과 싸우기 위해 상기 대상체 내로 재도입되는 면역 세포, 예컨대, 골수 세포, 식세포, 대식세포, 단핵구 또는 수지상 세포 계통의 세포에서 발현하기 위한 것이다. 고려되는 질환은 감염 질환, 암 및 자가면역 질환을 포함한다. 본원에 기재된 융합 단백질(CFP)을 포함하는 PSR 서브유닛을 코딩하는 핵산은 암 치료를 위해 조작된 식세포를 생성하는 데 사용된다.

[0536] 암은 T 세포 림프종, 피부 림프종, B 세포 암(예를 들어, 다발성 골수종, 발덴스트롬 마크로글로불린혈증), 중쇄 질환(예를 들어, 알파 쇄 질환, 감마 쇄 질환 및 뮤 쇄 질환), 양성 단일클론 감마병증 및 면역구성 아밀로이드증, 흑색종, 유방암, 폐암, 기관지암, 대장암, 전립선암(예를 들어, 전이성, 호르몬 불응성 전립선암), 췌장암, 위암, 난소암, 요로방광암, 뇌 또는 중추신경계 암, 말초신경계 암, 식도암, 자궁경부암, 자궁암 또는 자궁내막암, 구강 또는 후두의 암, 간암, 신장암, 고환암, 담도암, 소장암 또는 충수암, 침샘암, 갑상선암, 부신암, 골육종, 연골육종, 혈액 조직의 암 등을 포함하나, 이들로 제한되지 않는다. 본 개시내용에 포함되는 방법에 적용될 수 있는 암의 유형의 다른 비제한적 예는 인간 육종 및 암종, 예를 들어, 섬유육종, 점액육종, 지방육종, 연골육종, 골원성 육종, 척삭종, 혈관육종, 내피육종, 림프관육종, 림프관내피육종, 활막종, 중피종, 유윙 종양, 평활근육종, 횡문근육종, 결장 암종, 대장암, 췌장암, 유방암, 난소암, 편평 세포 암종, 기저 세포 암종, 선암종, 땀샘 암종, 피지샘 암종, 유두상 암종, 유두상 선암종, 낭선암종, 수질성 암종, 기관지원성 암종, 신장 세포 암종, 간암종, 담관 암종, 간암, 융모막암종, 정상피종, 배아 암종, 윌름스 종양, 자궁경부암, 골암, 뇌 종양, 고환암, 폐 암종, 소세포 폐 암종, 방광 암종, 상피 암종, 신경교종, 성상세포종, 수모세포종, 두개인두종, 뇌실막세포종, 송과체종, 혈관모세포종, 청신경종, 희소돌기아교세포종, 뇌수막종, 흑색종, 신경모세포종, 망막모세포종; 백혈병, 예를 들어, 급성 림프구성 백혈병 및 급성 골수구성 백혈병(골수모세포성, 전구골수구성, 골수단핵구성, 단핵구성 및 적백혈병); 만성 백혈병(만성 골수구성(과립구성) 백혈병 및 만성 림프구성 백혈병); 및 진성 적혈구증가증, 림프종(호지킨병 및 비-호지킨병), 다발성 골수종, 발덴스트롬 마크로글로불린혈증, 및 중쇄 질환을 포함한다. 일부 실시양태에서, 암은 상피암, 예컨대, 방광암, 유방암, 자궁경부암, 결장암, 부인과암, 신장암, 후두암, 폐암, 구강암, 두경부암, 난소암, 췌장암, 전립선암 또는 피부암이나, 이들로 제한되지 않는다. 다른 실시양태에서, 암은 유방암, 전립선암, 폐암 또는 결장암이다. 다른 실시양태에서, 상피암은 비-소세포 폐암, 비유두상 신장 세포 암종, 자궁경부 암종, 난소 암종(예를 들어, 장액성 난소 암종) 또는 유방 암종이다. 상피암은 장액성, 자궁내막모양, 점액성, 투명 세포 또는 미분화를 포함하나 이들로 제한되지 않는 다양한 다른 방식들로 특징져질 수 있다. 일부 실시양태에서, 본 개시내용은 맨틀 세포 림프종을 포함하나 이것으로 제한되지 않는 림프종 또는 이의 서브타입의 치료, 진단 및/또는 예후에

사용된다. 림프증식성 장애도 증식성 질환인 것으로 간주된다.

[0537] 일반적으로, 세포 면역요법은 세포가 이식편 대 숙주 질환인 GVHD를 유발하지 않도록 대상체와의 적합성을 위해 HLA 일치되어야 하는 살아있는 세포를 포함하는 의약을 환자에게 제공하는 단계를 포함한다. 전술된 바와 같이 맞춤화된 의약 및 면역요법을 위해 병원에 도착하는 대상체는 대상체에 의해 발현된 HLA 항원을 확인하기 위해 관용적으로 HLA 유형분류된다.

[0538] **mRNA 유도 전달의 치료적 이점**

[0539] 한 실시양태에서, 본원은 불리한 효과 없이 세포에서 지속적인 유전자 발현을 위해 핵산 서열을 세포 내로 도입하는 방법을 제공한다. 일부 실시양태에서, 세포는 살아있는 시스템, 예를 들어, 인간과 같은 숙주 유기체 내에 존재한다. 핵산 서열은 mRNA이다.

[0540] 특히, 레트로트랜스포존을 통한 전달은 수익성이 높은 방식인 것으로 보인다. mRNA 유도 전달은 유전자 전달을 단순화한다. 다른 기술은 비싸고 정교한 디자인과 제조, 핵산을 세포 내로 전달하기 위한 해법, 및 통합을 보조하기 위한 유전자 편집 기술을 요구하는 반면, 레트로트랜스포존 매개 전달은 그 자체가 편집 기구를 코딩하고 전달될 새로운 유전자를 코딩한다. 추가로, 단일 mRNA가 유전자 전달 및 편집에 충분할 수 있다.

[0541] 한 실시양태에서, mRNA 전달은 크기 제한 없이 핵산 카고의 도입을 보장할 수 있다는 점에서 유리하다.

[0542] 표 9는 다른 기존 핵산 전달 방법에 비해 장점 중 일부를 요약한다.

표 9

레트로트랜스포존 매개 유전자 전달의 장점

[0543]

	렌티바이러스 전달	AAV 전달	레트로트랜스포존 전달
페이로드	약 4 kb	약 4 kb	>10 kb
독성	삽입 돌연변이유발	해결되지 않은 간 및 CNS 독성	알려지지 않음, 임상 개발 진행 중
제조	복잡하고, 비싸다	복잡하고, 비싸다	저렴하고, 빠르다

[0544] 레트로트랜스포존은 여러 방식에 걸쳐 적용하는 데 유리하다. 이 방법을 이용한 유전자 조작은 생체내 및 생체외 둘 다에서 용이하게 달성된다. 한 실시양태에서, 레트로트랜스포존의 적용은 생체내에서 이루어질 수 있으며, mRNA에 코딩된 유전 물질의 조각은 전신 또는 국소 도입에 의해 환자 내로 직접 도입될 수 있다. 대조적으로, 대상체로부터 세포를 채취하여 생체외에서 조작한 다음, 동일한 대상체(자가) 또는 또 다른 인간(이종) 내로 도입할 수 있다.

[0545] 한 실시양태에서, 본원에 기재된 레트로트랜스포존 및 관련 방법은 유전자 요법에 도움이 될 수 있다. 큰 페이로드를 도입하는 능력을 이용하여, 다회 도입 또는 다수의 편집 이벤트를 요구하지 않으면서 전체 단백질을 코딩하는 유전자를 보유하는 큰 DNA 구획을 한 번에 도입할 수 있다. 한 실시양태에서, 예를 들어, 결함이 있는 단백질을 코딩하는 유전자를 절제할 수 있고, 레트로트랜스포존 매개 전달을 이용하여 1회 통합 이벤트로 정확한 유전자를 정확한 부위에서 도입할 수 있다. 한 예에서, CRISPR 편집을 이용하여 정확한 유전자좌로부터 유전자를 절제할 수 있으며 역전위를 이용하여 정확한 유전자를 대체할 수 있다. 일부 실시양태에서, 바람직한 레트로트랜스포존 통합 부위는 절제 부위에서 도입될 수 있다.

[0546] 한 실시양태에서, 본원에 기재된 레트로트랜스포존 및 관련 방법은 유전자 편집에 도움이 될 수 있다.

[0547] 한 실시양태에서, 본원에 기재된 레트로트랜스포존 및 관련 방법은 전사 조절에 도움이 될 수 있다.

[0548] 한 실시양태에서, 본원에 기재된 레트로트랜스포존 및 관련 방법은 게놈 조작에 도움이 될 수 있다.

[0549] 한 실시양태에서, 본원은 (i) 게놈 내로의 삽입체 서열의 부위 지정 통합을 위한 통합효소 또는 이의 단편을 코딩하는 서열 및 (ii) 통합효소에 의해 작동될 수 있는 게놈 랜딩 부위 서열을 추가로 포함하는 하나 이상의 폴리핵산을 포함하는 이종 핵산 서열을 게놈 내로 통합시키는 조성물을 제공하는 것으로, 이때 상기 게놈 랜딩 서열은 4개, 5개, 6개, 7개, 8개, 9개 또는 10개 초과의 연속 뉴클레오타이드 길이를 가진다. 일부 실시양태에서,

통합효소는 박테리아 통합효소이다. 일부 실시양태에서, 통합효소는 세린 통합효소이다.

[0550] 일부 실시양태에서, 상기 조성물은 ORF2 및 통합효소 또는 이의 단편을 포함하고, 이때 통합효소 및 ORF2는 별도의 폴리뉴클레오타이드에 존재한다. 일부 실시양태에서, 통합효소는 약 10개의 뉴클레오타이드 길이, 약 20개의 뉴클레오타이드 길이, 약 30개의 뉴클레오타이드 길이 또는 약 40개의 뉴클레오타이드 길이를 가진 게놈 랜딩 서열을 가진 게놈 부위 내로 핵산 서열을 통합시키는 능력을 가진다.

[0551] 일부 실시양태에서, ORF2 및 통합효소는 단일 폴리뉴클레오타이드에 존재한다. 일부 실시양태에서, ORF2는 10개 초과의 연속 뉴클레오타이드 길이를 가진 게놈 랜딩 서열을 인식할 수 있는 통합효소 단백질의 단편을 통합시키도록 변형되고, 이때 LINE 레트로트랜스포존 시스템은 이종 게놈 삽입체를 게놈 내로 통합된 통합효소 단백질의 단편에 의해 인식되는 게놈 랜딩 서열 내로 통합시킨다.

[0552] 일부 실시양태에서, 통합효소는 세포의 게놈 내로 통합되지 않는다. 일부 실시양태에서, 통합효소는 재조합 단백질이다. 일부 실시양태에서, ORF 단백질, 예를 들어, ORF2 단백질은 통합효소, 예를 들어, 박테리아 통합효소, 예를 들어, 박테리아 세린 통합효소의 게놈 랜딩 서열을 인식할 수 있는 박테리아 세린 통합효소의 적어도 단편을 포함하는 재조합(예를 들어, 키메라) 단백질이고, 이때 게놈 랜딩 서열은 통합효소에 의해 작동될 수 있고 20개 초과의 뉴클레오타이드 길이 또는 30개 초과의 뉴클레오타이드 길이를 가진다. 일부 실시양태에서, ORF 단백질 서열 중 하나 이상은 돌연변이를 포함한다. 일부 실시양태에서, 재조합(예를 들어, 키메라) ORF 단백질은 표적 인식 서열의 결실을 포함한다. 일부 실시양태에서, 재조합(예를 들어, 키메라) ORF 단백질은 촉매 도메인 또는 이의 단편, 부위 특이적 게놈 통합 인식 도메인을 포함하는, 통합효소, 예를 들어, 박테리아 통합효소, 예를 들어, 박테리아 세린 통합효소의 단편을 포함한다.

[0553] 본원은 약학 조성물을 제공하는 것으로, 이때 삽입체 서열은 통합효소에 의해 작동될 수 있는 부착 부위를 포함한다.

[0554] 일부 실시양태에서, 게놈 랜딩 부위는 가이드 RNA 및 Cas 시스템의 사용에 의해 게놈 내로 삽입된다. 일부 실시양태에서, 가이드 RNA, CAS 시스템 및 게놈 랜딩 서열은 LINE1-ORF를 코딩하는 서열 및 삽입체 서열을 포함하는 폴리뉴클레오타이드로부터 분리되어 있는 폴리뉴클레오타이드에 존재한다.

[0555] 본원은 이종 게놈 삽입체 서열을 포유동물 세포의 게놈 내로 부위 특이적으로 통합시키는 방법으로서, (i) (a) 이종 삽입체 서열과 관련된 하나 이상의 인간 레트로트랜스포존 요소를 코딩하는 서열을 포함하는 폴리뉴클레오타이드, 및 (b) 가이드 RNA, RNA 가이딩 통합효소 또는 이의 단편, 및 상기 통합효소에 의해 작동될 수 있는 랜딩 서열을 코딩하는 서열을 포함하는 폴리뉴클레오타이드를 세포 내로 도입하는 단계; 및 (ii) 상기 이종 삽입체 서열이 게놈 부위 내로 통합되었는지를 확인하는 단계를 포함하는 방법을 제공한다.

[0556] 본원은 LINE 레트로트랜스포존 시스템을 이용하여 이종 게놈 삽입체를 부위 특이적으로 통합시키는 방법을 제공하는 것으로, 이때 LINE 레트로트랜스포존 시스템은 10개 초과의 연속 뉴클레오타이드 길이를 가진 게놈 랜딩 서열을 인식할 수 있는 통합효소 단백질의 단편을 통합하도록 변형되고, LINE 레트로트랜스포존 시스템은 이종 게놈 삽입체를, 게놈 내로 통합된 통합효소 단백질의 단편에 의해 인식되는 게놈 랜딩 서열 내로 통합시킨다. 일부 실시양태에서, 통합효소는 게놈 랜딩 서열을 인식하여 접촉하고, LINE(L1) 레트로트랜스포존 시스템의 역전사효소 요소는 삽입체 서열(예를 들어, 카고 서열)을 역전사하고 게놈 랜딩 서열의 특정 부위에서 통합시킨다. 일부 실시양태에서, 상기 방법은 10개 초과의 연속 뉴클레오타이드 길이를 가진 게놈 랜딩 서열을 게놈 내로 통합시키는 단계를 포함한다. 일부 실시양태에서, 게놈 랜딩 서열을 게놈 내로 통합시키는 단계는 RNA 가이딩 CRISPR-Cas 시스템에 의해 수행된다. 일부 실시양태에서, RNA 가이딩 CRISPR-Cas 시스템은 10개 초과의 연속 뉴클레오타이드 길이를 가진 서열을 특정 게놈 부위 내로 통합시킬 수 있는 편집 기능을 가진다.

[0557] 따라서, 본원은 부위 특이적 통합효소 DNA 인식 모이어티를 포함하는 변형된 L1 역전위 시스템을 제공하는 것으로; 이때 상기 통합효소 DNA 인식 모이어티는 부위 특이적 통합을 위해 게놈 서열을 인식하고, L1 역전위 시스템은 이종 삽입체를 포함하는 서열을 역전사하고 통합효소에 의해 인식되는 부위 또는 이 부위 근처에서 게놈 부위 내로 통합시킨다. 세린 통합효소와 같은 통합효소를 포함하는 기존 시스템은 예를 들어, 플라스미드를 사용하는 DNA 통합 시스템 및 바이러스 전달 시스템을 포함할 수 있으며, 이들 각각은 카고 크기의 제한 없이 본원에서 사용된 바와 같은 보다 안전하고 확실한 mRNA 시스템의 경우 우회될 수 있다.

[0558] 한 실시양태에서, 본원에 기재된 레트로트랜스포존 및 관련 방법은 NK 세포 요법 또는 골수 세포 요법에서 세포 요법, 예를 들어, 키메라 항원 수용체(CAR) T 세포를 개발하는 데 도움이 될 수 있다. 한 실시양태에서, 본원에 기재된 레트로트랜스포존 및 관련 방법은 기존 기술에 의해 접근되기 어려운 뉴런 내로의 유전자의 전달에 도움

이 될 수 있다.

[0559] 한 측면에서, 본원은 (A) 하나 이상의 돌연변이를 포함하는 핵산 서열을 세포의 게놈으로부터 절제하기 위한 표적화 절제 기구를 포함하는 제1 단백질 복합체를 코딩하는 폴리뉴클레오타이드 서열; 및 (B) 제2 단백질 복합체를 코딩하는 재조합 mRNA를 세포 내로 도입하는 단계를 포함하는, 세포의 게놈 핵산 서열의 표적화 대체 방법을 제공하는 것으로, 이때 상기 재조합 mRNA는 (i) 하나 이상의 돌연변이를 함유하지 않는, (A)의 절제된 핵산 서열을 포함하는 핵산 서열, 및 (ii) 독립적인 프로모터의 영향 하에서 L1 레트로트랜스포존 ORF2 단백질을 코딩하는 서열을 포함한다.

[0560] 한 실시양태에서, 제1 단백질 복합체는 제2 단백질 복합체와 독립적인 엔도뉴클레아제 복합체일 수 있다. 한 실시양태에서, 제1 단백질 복합체는 서열 가이딩 게놈 DNA 절제를 이용하는 CRISPR-CAS 시스템을 포함한다. 한 실시양태에서, 본원에 기재된 방법은 CRISPR CAS 시스템 또는 임의의 다른 유전자 편집 시스템을, 4 kb, 또는 5 kb, 또는 6 kb, 또는 7 kb, 또는 8 kb, 또는 9 kb, 또는 10 kb보다 더 큰 페이로드 용량을 가진 대체 유전자를 전달하는 L1 트랜스포존 기구(예를 들어, 제2 단백질 복합체)와 커플링시킨다. 이 커플링은 게놈 유전자좌로부터 큰 단편(질환을 유발하는 돌연변이된 유전자)을 정확하게 절제하고 돌연변이되지 않은 정확한 서열을 코딩하는 큰 유전자 단편 또는 전체 유전자를 통합시키는 데 이용될 수 있다.

[0561] 다수의 유전 질환들은 종종 기존 방법의 기능적 용량을 초과하는 큰 페이로드의 유전자 전달을 요구할 수 있다. 레트로트랜스포존을 사용하여 이러한 질환에 대한 요법을 추가로 디자인하는 데 도움이 될 수 있는 본원에 개시된 방법 및 조성물이 본원에서 고려된다. 유전 질환의 예시적인 목록은 표 10에 나열된 질환들을 포함하나 이들로 제한되지 않는다.

표 10

잠재적인 유전자 요법 적용의 목록

[0562]

질환	유전자	CDS	발현	유병률
스타가르트	ABCA4	6.8 kb	간상 및 원추 PR	1:8000
어셔 1B	MY07A	6.7 kb	RPE 및 PR	3.2: 100,000
LCA10	CEP290	7.4 kb	PR(범망막)	1:50,000
USH1D, DFNB12	CDH23	10.1 kb	PR	3:100,000
RP	EYS	9.4 kb	PR ECM	1:50,000
USH2A	USH2a	15.6 kb	간상 및 원추 PR	4: 100,000
USH2C	GPR98	18.0 kb	주로 PR	1: 100,000
알스트롬 증후군	ALMS1	12.5 kb	간상 및 원추 PR	1: 1,000,000
글리코겐 저장 질환 III	GDE	4.6 kb	근육, 간	1:8000
비-증후군성 난청	OTOF	6.0 kb	귀	14: 100,000
A 형 혈우병	F8	7.1 kb	간	1:10,000
레버 선천성 흑암시	CEP290	7.5 kb	망막	5:100,000

[0563] 본원은 세포에서 게놈 핵산 서열을 표적화 대체하는 방법을 제공한다. 한 실시양태에서, 이 방법은 (A) 하나 이상의 돌연변이를 포함하는 핵산 서열을 세포의 게놈으로부터 절제하는 단계, 및 (B) (i) (A)의 절제된 서열에 비해 하나 이상의 돌연변이를 함유하지 않는 야생형 서열을 포함하는 핵산 서열, 및 (ii) 독립적인 프로모터의

영향 하에서 L1 레트로트랜스포존 ORF2 단백질을 코딩하는 서열을 코딩하는 재조합 mRNA를 세포 내로 도입하는 단계를 포함한다. 한 실시양태에서, 단계 (A)는 절제 부위에서 적어도 복수의 아데닐레이트 잔기를 포함하는 짧은 서열을 도입하는 단계를 추가로 포함한다. 한 실시양태에서, 야생형 서열을 포함하는 핵산 서열은 ORF2 역전사효소가 돌연변이되지 않은 야생형 서열을 포함하는 서열을 게놈 내로 통합시키는 방식으로 ORF2 코딩 서열과 작동 가능하게 연결된다.

[0564] 한 실시양태에서, 세포는 림프구이다.

[0565] 한 실시양태에서, 세포는 상피 세포이다. 일부 실시양태에서, 세포는 망막 색소 상피 세포(RPE)이다.

[0566] 한 실시양태에서, 세포는 뉴런이다.

[0567] 한 실시양태에서, 세포는 골수 세포이다.

[0568] 한 실시양태에서, 세포는 줄기 세포이다.

[0569] 한 실시양태에서, 세포는 암 세포이다.

[0570] 한 실시양태에서, 유전자는 ABCA4, MY07A, CEP290, CDH23, EYS, USH2a, GPR98, ALMS1, GDE, OTOF 및 F8로 구성된 군으로부터 선택된다.

[0571] 한 실시양태에서, mRNA는 유도성 프로모터에 대한 서열을 포함한다.

[0572] 한 실시양태에서, 돌연변이되지 않은 서열을 포함하는 핵산 서열의 발현은 감염된 지 적어도 20일, 21일, 22일, 23일, 24일, 25일, 26일, 27일, 28일, 29일, 30일, 31일, 32일, 33일, 34일 또는 35일 후에 검출될 수 있다.

[0573] 한 실시양태에서, 상기 방법은 생체내에서 재조합 mRNA를 세포 내로 도입하는 단계를 포함한다.

[0574] 한 실시양태에서, 상기 방법은 생체외에서 재조합 mRNA를 세포 내로 도입하는 단계를 포함한다.

[0575] 본원은 유전 질환의 치료를 필요로 하는 대상체에서 유전 질환을 치료하는 방법으로서, L1 레트로트랜스포존을 코딩하는 서열에 작동 가능하게 연결된, 유전자 또는 이의 단편을 코딩하는 폴리시스트론 mRNA를 포함하는 조성물을 상기 대상체 내로 도입하는 단계를 포함하는 방법을 제공하는 것으로; 이때 유전자 또는 이의 단편은 길이가 적어도 10.1 kb이다.

[0576] 한 실시양태에서, 상기 방법은 mRNA를 직접 전신 도입하는 단계를 포함한다.

[0577] 한 실시양태에서, 상기 방법은 mRNA를 직접 국소 도입하는 단계를 포함한다.

[0578] 한 실시양태에서, 유전 질환은 망막 질환이다. 예를 들어, 상기 질환은 황반 이영양증이다. 한 실시양태에서, 상기 질환은 청소년 황반 변성 또는 노란점 안저(fundus flavimaculatus)로서도 알려진 스타가르트병이다. 이 질환은 황반의 점진적인 변성과 손상을 야기한다. 이 질환은 ATP 결합 카세트(ABC) 수송제 유전자인 (*ABCA4*) 유전자의 돌연변이로 인한 유전적 근거를 갖고, 2차 광수용체 세포 사멸과 함께 망막 색소 상피(RPE)에서의 리포푸신 유사 물질의 침착으로부터 발생한다. 일부 실시양태에서, 상기 방법은 mRNA를 망막으로 직접 전달하는 단계를 포함한다.

[0579] 한 실시양태에서, 상기 방법은 비-증후군성 상염색체 열성 난청(DFNB12), 및 색소성 망막염 및 전정 기능장애와 관련된 난청(USH1D)을 치료하는 단계를 포함한다. 한 실시양태에서, 본원은 비-증후군성 난청(DFNB12) 또는 어셔 증후군(USH1D)을 치료하는 방법으로서, L1 레트로트랜스포존을 코딩하는 서열에 작동 가능하게 연결된, CDH23 또는 이의 단편의 카피를 포함하는 mRNA를 도입하는 단계를 포함한다.

[0580] 외생성 폴리펩티드의 세포 특이적 발현

[0581] 외생성 폴리펩티드의 안정한 발현은 세포(예를 들어, 표적 세포)의 게놈 내로의 폴리뉴클레오타이드 서열(예를 들어, 본원에서 종종 구축물의 삽입체 서열로서 지칭됨)의 통합을 표적화하는 이동성 유전 요소를 사용함으로써 다양한 세포 유형(예를 들어, 표적 세포 유형)에서 달성될 수 있다. 일부 실시양태에서, 표적 세포는 유사분열후 세포, 예를 들어, 포유동물 심근세포 또는 RPE 세포이다. 일부 실시양태에서, 이동성 유전 요소는 인간 LINE 1 서열을 포함한다. 일부 실시양태에서, 이동성 유전 요소는 인간 LINE1 서열이다. 일부 실시양태에서, 이동성 유전 요소는 인간 L1 ORFp1 단백질을 코딩하는 서열을 포함한다. 일부 실시양태에서, 이동성 유전 요소는 인간 L1-ORFp2 단백질을 코딩하는 서열을 포함한다. 일부 실시양태에서, 이동성 유전 요소는 인간 ORFp1 및 인간 ORFp2 폴리펩티드를 코딩하는 서열을 포함한다. 일부 실시양태에서, 폴리뉴클레오타이드 서열은 표적 세포 내로

도입되고, 폴리뉴클레오타이드는 이동성 유전 요소를 코딩하는 서열 및 삽입체 서열을 포함한다. 일부 실시양태에서, 이동성 유전 요소는 표적 세포의 게놈 내로의 삽입체 서열의 안정한 통합을 번역 후 촉진하는 폴리펩티드를 포함한다. 일부 실시양태에서, 이동성 유전 요소는 주어진 삽입체 서열을 표적 세포의 게놈 내로 부위 특이적으로 통합시키도록 추가로 구성된 폴리펩티드를 포함하나, 예를 들어, 본 명세서의 다른 곳에 기재된 실시양태로 제한되지 않는다.

[0582] 일부 실시양태에서, 본원에 기재된 폴리뉴클레오타이드는 mRNA이다. mRNA는 바이시스트론 또는 폴리시스트론일 수 있다. 작업 실시양태에 의해 예시되는 바와 같이, 삽입체 서열 및 이동성 유전 요소를 코딩하는 서열은 서로에 대해 역방향으로 존재할 수 있다. 이동성 유전 요소는 표적 프라이밍 역전사(TPRT)를 통해 삽입체 서열을 통합시킬 수 있다. 일부 실시양태에서, 이동성 유전 요소는 인간 L1 레트로트랜스포존 또는 이의 단편을 포함한다.

[0583] 일부 실시양태에서, 폴리뉴클레오타이드는 세포 유형에 특이적으로 표적화될 수 있다. 일부 실시양태에서, 폴리뉴클레오타이드는 나노입자로 구성될 수 있으며, 이때 나노입자는 당업자에게 알려진 하나 이상의 표적화 모이어티를 포함한다.

[0584] 본원은 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 방법으로서, 폴리핵산을 포함하는 조성물을 표적 세포와 접촉시키는 단계; 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 단계; 및 표적 세포에서 외생성 폴리펩티드를 발현시키는 단계를 포함하는 방법을 제공하는 것으로, 이때 폴리핵산은 외생성 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함하는 삽입체 서열, 및 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소를 포함하고, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 게놈 DNA 내로의 삽입체 서열의 통합을 촉진하고, 상기 표적 세포는 인간 간세포이다. 일부 실시양태에서, 폴리핵산은 mRNA이다. 일부 실시양태에서, mRNA는 세포내 전달을 위해 나노입자에 캡슐화된다. 일부 실시양태에서, 나노입자는 지질을 포함한다. 일부 실시양태에서, 폴리핵산은 예를 들어, 간세포에 의한 폴리핵산의 성공적인 흡수를 야기하도록 폴리핵산 및 전달 비히클(예를 들어, 나노입자는 지질을 포함함)을 포함하는 조성물을 간세포와 접촉시킴으로써 간세포 내로 도입된다. 일부 실시양태에서, 도입은 전기천공을 통해 이루어진다. 일부 실시양태에서, 폴리핵산 조성물은 간세포 내로 전기천공된다. 일부 실시양태에서, 간세포는 폴리핵산에 코딩된 폴리펩티드의 발현에 적합하고 간세포의 생존율에 기여하는 조건 하에서 전기천공된다. 일부 실시양태에서, 간세포의 게놈 내로의 삽입체의 통합은 폴리핵산의 도입 후에 당업자에게 공지된 방법, 예를 들어, 게놈 시퀀싱에 의해 확인된다. 일부 실시양태에서, 통합된 삽입체 서열로부터의 폴리펩티드의 발현은 폴리핵산을 간세포 내로 도입한 후에 적합한 간격에서 확인되고, 이때 적합한 간격은 약 4시간, 6시간, 8시간, 10시간, 12시간, 24시간 또는 48시간이다. 일부 실시양태에서, 간세포를 포함하는 세포 집단의 전기천공 후, 전기천공된 간세포를 포함하는 세포 집단은 간세포에 적합한 조건 하에서 적어도 약 2시간, 약 3시간, 약 4시간, 약 5시간, 약 6시간, 약 8시간, 약 10시간 또는 약 24시간 동안 배양된다. 일부 실시양태에서, 간세포는 생장을 위해 간세포에 적합한 조건 하에서 약 48시간, 72시간, 96시간 이상 동안 배양된다. 일부 실시양태에서, 폴리뉴클레오타이드(예를 들어, 삽입체)에 의해 코딩된 폴리펩티드의 발현은 배양물에서 48시간, 72시간 또는 96시간 후, 또는 4일, 5일, 6일, 7일, 8일, 9일, 10일 이상 후에 확인된다. 일부 실시양태에서, 인간 간세포의 적어도 2%는 도입 후 10일째 날에 외생성 폴리펩티드를 발현한다. 일부 실시양태에서, 전기천공된 세포 집단 내의 인간 간세포의 약 2%, 또는 약 5%, 또는 약 10% 이상은 도입 후 10일째 날에 외생성 폴리펩티드를 발현한다.

[0585] 본원은 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 방법으로서, 폴리핵산을 포함하는 조성물을 표적 세포와 접촉시키는 단계; 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 단계; 및 표적 세포에서 외생성 폴리펩티드를 발현시키는 단계를 포함하는 방법을 제공하는 것으로, 이때 폴리핵산은 외생성 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함하는 삽입체 서열, 및 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소를 포함하고, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 게놈 DNA 내로의 삽입체 서열의 통합을 촉진하고, 상기 표적 세포는 인간 심근세포이다. 일부 실시양태에서, 폴리핵산은 mRNA이다. 일부 실시양태에서, mRNA는 세포내 전달을 위해 나노입자에 캡슐화된다. 일부 실시양태에서, 나노입자는 지질을 포함한다. 일부 실시양태에서, 폴리핵산은 예를 들어, 심근세포에 의한 폴리핵산의 성공적인 흡수를 야기하도록 폴리핵산 및 전달 비히클(예를 들어, 나노입자는 지질을 포함함)을 포함하는 조성물을 심근세포와 접촉시킴으로써 심근세포 내로 도입된다. 일부 실시양태에서, 도입은 전기천공을 통해 이루어진다. 일부 실시양태에서, 폴리핵산 조성물은 심근세포 내로 전기천공된다. 일부 실시양태에서, 심근세포는 폴리핵산에 코딩된 폴리펩티드의 발현에 적합하고 심근세포의 생존율에 기여하는 조건 하에서 전기천공된다. 일부 실시양태에서, 심근세포의 게놈 내로의 삽입체의 통합은 폴리핵산의 도입 후에 당업자에게 공지된 방법, 예를

들어, 게놈 시퀀싱에 의해 확인된다. 일부 실시양태에서, 통합된 삽입체 서열로부터의 폴리펩티드의 발현은 폴리핵산을 심근세포 내로 도입한 후에 적합한 간격에서 확인되고, 이때 적합한 간격은 약 4시간, 6시간, 8시간, 10시간, 12시간, 24시간 또는 48시간이다. 일부 실시양태에서, 심근세포를 포함하는 세포 집단의 전기천공 후, 전기천공된 심근세포를 포함하는 세포 집단은 심근세포에 적합한 조건 하에서 적어도 약 2시간, 약 3시간, 약 4시간, 약 5시간, 약 6시간, 약 8시간, 약 10시간 또는 약 24시간 동안 배양된다. 일부 실시양태에서, 심근세포는 생장을 위해 심근세포에 적합한 조건 하에서 약 48시간, 72시간, 96시간 이상 동안 배양된다. 일부 실시양태에서, 폴리뉴클레오타이드(예를 들어, 삽입체)에 의해 코딩된 폴리펩티드의 발현은 배양물에서 48시간, 72시간 또는 96시간 후, 또는 4일, 5일, 6일, 7일, 8일, 9일, 10일 이상 후에 확인된다. 일부 실시양태에서, 인간 심근세포의 적어도 2%는 도입 후 10일째 날에 외생성 폴리펩티드를 발현한다. 일부 실시양태에서, 전기천공된 세포 집단 내의 인간 심근세포의 약 2%, 또는 약 5%, 또는 약 10% 이상은 도입 후 10일째 날에 외생성 폴리펩티드를 발현한다. 본원은 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 방법으로서, 폴리핵산을 포함하는 조성물을 표적 세포와 접촉시키는 단계; 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 단계; 및 표적 세포에서 외생성 폴리펩티드를 발현시키는 단계를 포함하는 방법을 제공하는 것으로, 이때 폴리핵산은 외생성 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함하는 삽입체 서열, 및 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소를 포함하고, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 게놈 DNA 내로의 삽입체 서열의 통합을 촉진하고, 상기 표적 세포는 인간 망막 색소 상피 세포(RPE)이다. 일부 실시양태에서, 폴리핵산은 mRNA이다. 일부 실시양태에서, mRNA는 세포내 전달을 위해 나노입자에 캡슐화된다. 일부 실시양태에서, 나노입자는 지질을 포함한다. 일부 실시양태에서, 폴리핵산은 예를 들어, RPE에 의한 폴리핵산의 성공적인 흡수를 야기하도록 폴리핵산 및 전달 비히클(예를 들어, 나노입자는 지질을 포함함)을 포함하는 조성물을 RPE와 접촉시킴으로써 RPE 내로 도입된다. 일부 실시양태에서, 도입은 전기천공을 통해 이루어진다. 일부 실시양태에서, 폴리핵산 조성물은 RPE 내로 전기천공된다. 일부 실시양태에서, RPE는 폴리핵산에 코딩된 폴리펩티드의 발현에 적합하고 RPE의 생존율에 기여하는 조건 하에서 전기천공된다. 일부 실시양태에서, RPE의 게놈 내로의 삽입체의 통합은 폴리핵산의 도입 후에 당업자에게 공지된 방법, 예를 들어, 게놈 시퀀싱에 의해 확인된다. 일부 실시양태에서, 통합된 삽입체 서열로부터의 폴리펩티드의 발현은 폴리핵산을 RPE 내로 도입한 후에 적합한 간격에서 확인되고, 이때 적합한 간격은 약 4시간, 6시간, 8시간, 10시간, 12시간, 24시간 또는 48시간이다. 일부 실시양태에서, RPE를 포함하는 세포 집단의 전기천공 후, 전기천공된 RPE를 포함하는 세포 집단은 RPE에 적합한 조건 하에서 적어도 약 2시간, 약 3시간, 약 4시간, 약 5시간, 약 6시간, 약 8시간, 약 10시간 또는 약 24시간 동안 배양된다. 일부 실시양태에서, RPE는 생장을 위해 RPE에 적합한 조건 하에서 약 48시간, 72시간, 96시간 이상 동안 배양된다. 일부 실시양태에서, 폴리뉴클레오타이드(예를 들어, 삽입체)에 의해 코딩된 폴리펩티드의 발현은 배양물에서 48시간, 72시간 또는 96시간 후, 또는 4일, 5일, 6일, 7일, 8일, 9일, 10일 이상 후에 확인된다. 일부 실시양태에서, 인간 RPE의 적어도 2%는 도입 후 10일째 날에 외생성 폴리펩티드를 발현한다. 일부 실시양태에서, 전기천공된 세포 집단 내의 인간 RPE의 약 2%, 또는 약 5%, 또는 약 10% 이상은 도입 후 10일째 날에 외생성 폴리펩티드를 발현한다.

[0586] 본원은 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 방법으로서, 폴리핵산을 포함하는 조성물을 표적 세포와 접촉시키는 단계; 삽입체 서열을 표적 세포의 게놈 DNA 내로 안정적으로 통합시키는 단계; 및 표적 세포에서 외생성 폴리펩티드를 발현시키는 단계를 포함하는 방법을 제공하는 것으로, 이때 폴리핵산은 외생성 폴리펩티드를 코딩하는 서열의 역상보체인 서열을 포함하는 삽입체 서열, 및 폴리펩티드를 코딩하는 서열을 포함하는 이동성 유전 요소를 포함하고, 이동성 유전 요소의 서열에 의해 코딩된 폴리펩티드는 게놈 DNA 내로의 삽입체 서열의 통합을 촉진하고, 상기 표적 세포는 인간 뉴런 세포이다. 일부 실시양태에서, 폴리핵산은 mRNA이다. 일부 실시양태에서, mRNA는 세포내 전달을 위해 나노입자에 캡슐화된다. 일부 실시양태에서, 나노입자는 지질을 포함한다. 일부 실시양태에서, 폴리핵산은 예를 들어, 뉴런 세포에 의한 폴리핵산의 성공적인 흡수를 야기하도록 폴리핵산 및 전달 비히클(예를 들어, 나노입자는 지질을 포함함)을 포함하는 조성물을 뉴런 세포와 접촉시킴으로써 뉴런 세포 내로 도입된다. 일부 실시양태에서, 도입은 전기천공을 통해 이루어진다. 일부 실시양태에서, 폴리핵산 조성물은 뉴런 세포 내로 전기천공된다. 일부 실시양태에서, 뉴런 세포는 폴리핵산에 코딩된 폴리펩티드의 발현에 적합하고 뉴런 세포의 생존율에 기여하는 조건 하에서 전기천공된다. 일부 실시양태에서, 뉴런 세포의 게놈 내로의 삽입체의 통합은 폴리핵산의 도입 후에 당업자에게 공지된 방법, 예를 들어, 게놈 시퀀싱에 의해 확인된다. 일부 실시양태에서, 통합된 삽입체 서열로부터의 폴리펩티드의 발현은 폴리핵산을 뉴런 세포 내로 도입한 후에 적합한 간격에서 확인되고, 이때 적합한 간격은 약 4시간, 6시간, 8시간, 10시간, 12시간, 24시간 또는 48시간이다. 일부 실시양태에서, 뉴런 세포를 포함하는 세포 집단의 전기천공 후, 전기천공된 뉴런 세포를 포함하는 세포 집단은 뉴런 세포에 적합한 조건 하에서 적어도 약 2시간, 약 3시간, 약

4시간, 약 5시간, 약 6시간, 약 8시간, 약 10시간 또는 약 24시간 동안 배양된다. 일부 실시양태에서, 뉴런 세포는 생장을 위해 뉴런 세포에 적합한 조건 하에서 약 48시간, 72시간, 96시간 이상 동안 배양된다. 일부 실시양태에서, 폴리뉴클레오타이드(예를 들어, 삽입체)에 의해 코딩된 폴리펩티드의 발현은 배양물에서 48시간, 72시간 또는 96시간 후, 또는 4일, 5일, 6일, 7일, 8일, 9일, 10일 이상 후에 확인된다. 일부 실시양태에서, 인간 뉴런 세포의 적어도 2%는 도입 후 10일째 날에 외생성 폴리펩티드를 발현한다. 일부 실시양태에서, 전기천공된 세포 집단 내의 인간 뉴런 세포의 약 2%, 또는 약 5%, 또는 약 10% 이상은 도입 후 10일째 날에 외생성 폴리펩티드를 발현한다.

[0587] **실시예**

[0588] 본 발명은 지금부터 기재되는 실시예로 제한되는 것으로 해석되어서는 안 되고; 오히려, 본 발명은 본원에서 제공된 임의의 모든 적용 및 통상의 기술자의 기술 내의 모든 동등한 변경을 포함하는 것으로 해석되어야 함을 인식해야 한다.

[0589] **실시예 1. 예시적인 레트로트랜스포존 디자인 구축물:**

[0590] 본원은 세포의 게놈 내로 통합시키고 예시적인 형질전이유전자를 발현시키기 위한 레트로트랜스포존 구축물을 디자인하는 예시적인 전략을 제공한다. 도 1b 및 도 1c는 형질전이유전자를 코딩하는 mRNA를 세포의 게놈 내로 통합시키기 위한 다양한 전략적 디자인을 예시한다. 여기서 박스로 표시된 GFP는 예시적인 형질전이유전자이다. 형질전이유전자(예를 들어, GFP)를 코딩하는 mRNA는 ORF2p 단백질을 코딩하는 핵산 서열과 센스 또는 안티센스 방향 중 어느 한 방향으로 공-발현될 수 있고; 각각의 코딩 서열은 예시적인 시스 전략 하에 표시된 모노시스트론 또는 바이시스트론 구축물에 존재할 수 있다(도 1b 및 도 1c). CMV/T7은 프로모터이다.

[0591] 한편, 상기 구축물은 트랜스 방식으로 발현되도록 유도될 수 있다. 트랜스 전략은 바이시스트론 서열의 ORF2p 단백질 또는 ORF1p 및 ORF2p 단백질 둘 다를 코딩하는 서열, 및 임의의 유전자의 3' UTR에서 센스 또는 안티센스 방향으로 GFP를 코딩하는 mRNA를 포함할 수 있다. 형질전이유전자는 전위효소 결합 서열, A 박스 및 B 박스, 및 폴리 A 꼬리를 포함하는 역전위 서열에 의해 플랭킹된다. 도 2a는 구축물을 사용하여 GFP를 코딩하는 서열을 안정적으로 통합시킴으로써 예시적인 형질전이유전자 GFP를 발현시키기 위한 세 가지 예시적인 디자인을 보여준다. 첫 번째 구축물은 임의의 전위효소 결합 요소의 부재 하에서 L1 5' UTR에 의해 플랭킹된, GFP를 코딩하는 서열; 및 3' UTR의 폴리 A 서열을 포함한다. 두 번째 구축물 및 세 번째 구축물은 GFP를 코딩하는 서열, 3' UTR A 박스 및 B 박스, 및 3' UTR의 폴리 A 서열을 포함한다. 세 번째 구축물은 ORF2p를 코딩하는 추가 서열을 포함한다. 72시간에서 예상된 GFP 발현 수준은 오른쪽에 표시되어 있다. 도 2b는 구축물을 사용하여 GFP를 코딩하는 서열을 안정적으로 통합시킴으로써 RFP 또는 ORF2p를 코딩하는 mRNA에서 예시적인 형질전이유전자 GFP를 발현시키기 위한 세 가지 예시적인 디자인을 보여준다. 첫 번째 구축물은 임의의 L1 요소의 부재 하에서 RFP를 코딩하는 서열, 및 3' UTR의 폴리 A 서열을 포함한다. 두 번째 및 세 번째 구축물은 A 박스 및 B 박스를 포함하는 3' UTR, 및 3' UTR의 폴리 A 서열을 포함한다. 두 번째 구축물은 RFP를 코딩하는 서열을 포함하고 세 번째 구축물은 ORF2p를 코딩하는 서열을 포함한다. 72시간에서 예상된 RFP 및 GFP 발현 수준은 오른쪽에 표시되어 있다.

[0592] **실시예 2. 예시적인 circRNA 디자인 구축물:**

[0593] 이 실시예에서, circRNA를 제조하기 위해 천연 생성 3차 구조를 가진 약 50개 뉴클레오타이드 길이 RNA의 스트레치를 통합시키는, circRNA에 대한 모듈식 디자인을 입증한다. 3차 구조 형성 RNA의 사용은 원형화를 위해 circRNA 형성 과정이 서열 매개 하이브리드화와 무관하게 만든다. 3차 구조를 가진 이 RNA 모티프는 5' 및 3' 상동성 아암 대신에 엑손 및 인트론을 가진 원하는 RNA 내로 통합됨으로써, 원형화를 위한 말단 RNA 스캐폴드를 형성할 수 있다.

[0594] TectoRNA: RNA-RNA 결합 계면은 GNRA 루프/루프-수용체 상호작용 모티프의 쌍을 조합하여, 높은 친화성 및 높은 특이성을 가진 3차 구조를 생성함으로써 구축된다(도 3b). GNRA 루프/루프-수용체 상호작용 모티프 쌍은 헤어핀 리보자임으로부터의 4-방향 연접을 이용함으로써 융합되어, 폐쇄된 협력적 어셈블링 고리 모양 복합체를 형성하는 데 도움이 되는 2가 자가어셈블링 스캐폴딩 유닛('tectoRNA')을 생성한다. 두 개의 직교 루프/루프-수용체 상호작용 모티프를 사용하여, 평행('위-위') 또는 역평행('위-아래') 어셈블리 방식으로 방향성 어셈블리를 할 수 있는 RNA 단량체를 디자인한다. 상호작용하는 분자의 역평행 어셈블리에서, 각각의 통합된 단량체는 성장하는 쇄의 방향성을 전환하여 그의 고유 굽힘을 보상함으로써, 길고 비교적 곧은 다중유닛 쇄를 생성한다. 대안적인 2차 구조의 발생이 최소화된 tectoRNA 스캐폴드를 선택하기 위해, 주어진 RNA 서열의 열역학적으로 선호된 2차 구조를 예측하는 RNA 폴딩 프로그램 Mfold(bioinfo.math.rpi.edu/~zukerm/rna/mfold)에 서열을 제출하여 서

열을 확인한다. 열역학적으로 선호된 구조는 최소한의 대안적인 2차 구조를 가진 스캐폴딩을 위해 선택된다(배타적이 아니라 전형적으로, 에너지 면에서 최저 에너지 구조에 15% 이상 더 가까운 다른 2차 구조는 없다). RNA 분자는 통상의 방법, 예컨대, T7 RNA 중합효소를 사용한 시험관내 런-오프(run-off) 전사에 의해 제조된다. 도 3b는 RL-GAAA 루프 구조를 보여준다. tectoRNA 이종이량체를 프로파일링하기 위해, 형광 기반 칩-유동 조각 시험 방법을 이용한다. 이 방법에서, 구조화된 RNA(칩 조각)의 잠재적 변이체의 라이브러리를 DNA 주형으로서 합성하고 RNAP 시작을 위해 시퀀싱 어댑터 및 영역을 포함하도록 증폭한다. 각각의 DNA 변이체를 제자리에서 RNA로 전사하여, 서열이 식별된 RNA 클러스터를 시퀀싱 칩의 표면에 표시할 수 있다. 형광 표지부착된 tectoRNA 결합 파트너인 "유동 조각"을 증가하는 농도로 시퀀싱 칩 유동 셀에 도입하여, 평형화 후 각각의 RNA 클러스터에 대해 결합된 형광을 정량한다. 이 형광 값을 이용하여, 해리 상수(K_d) 및 결합 자유 에너지($\Delta G = RT \log(K_d)$)의 관점에서 각각의 칩 조각 변이체에 대한 유동 조각의 친화성을 유도한다(도 3c).

[0595] 3차 구조를 포함하는 선택된 말단 RNA 스캐폴드 분절을, T7 전사를 이용하여 통합시키거나 원형화될 원하는 RNA의 5' 및 3' 말단에서 라이게이션하거나; 임의의 공지되어 있는 분자생물학 기법으로 원하는 RNA 내로 통합시킨다.

[0596] **실시예 3: 향상된 특이성을 가진 예시적인 레트로트랜스포존 디자인**

[0597] 이 실시예에서, 향상된 표적 특이성을 위해 L1 매개 레트로트랜스포존을 위한 핵산 구축물에 대한 디자인을 입증한다. ORF2를 사용하여 관심 있는 유전자를 세포의 게놈 내로 통합시키기 위해 ORF2 코딩 서열 및 관심 있는 유전자를 코딩하는 서열을 포함하는 mRNA를 디자인한다. 한 예시적인 디자인에서, 구축물은 더 변형된 ORF2를 포함한다.

[0598] 도 4a에 나타낸 바와 같이, ORF2 단백질은 그 자신의 폴리 A 서열에 결합함으로써 역전위를 시작한다. 그러나, 폴리 A는 mRNA에 풍부하게 존재하기 때문에, 비특이적 결합 및 통합이 가능해진다. 특이성을 증가시키기 위해, 이종 단백질의 mRNA 결합 도메인을 포함하는 재조합 ORF2를 디자인하고, 이종 mRNA 결합 도메인에 대한 동족 mRNA 서열을 3' UTR의 폴리 A 서열 및 ORF2 폴리 A 결합 부위 근처에 삽입한다.

[0599] 이로써, 도 4b에 나타낸 키메라 ORF2가 생성되고, 이때 이종 단백질 코딩 서열의 고친화성 RNA 결합 도메인은 ORF2 서열에 통합되거나 융합되고, 고친화성 RNA 결합 단백질에 상응하는 동족 RNA 서열은 폴리 A 영역에 근접한 mRNA의 3' UTR 영역 내로 통합된다. 이 실시예에서, MCP 코트 단백질 MS2(도면에서 M으로서 표시됨)로부터 유래한 이종 고친화성 RNA 결합 도메인은 ORF2 서열 내로 통합되고 동족 서열인 MS2 헤어핀은 mRNA의 3' UTR 서열에 포함된다(도 4b). MS2는 동족 서열에 결합하여, 포유동물 세포 게놈에서 ORF2 mRNA와 관련된 각각의 서열의 역전사 및 통합을 위해 그 자신의 mRNA에 대한 키메라 ORF2의 특이성을 증가시킨다(도 4b).

[0600] 다른 예시적인 디자인에서, ORF2에 의한 표적 세포 게놈 내로의 형질전이유전자의 통합의 특이성을 증가시키려는 시도가 착수된다. 한 예시적인 디자인에서, 도 4c에 나타낸 바와 같이 메가 TAL 코딩 서열을 ORF2에 융합에 융합시킨다(상부 패널). 이와 함께, ORF2를 돌연변이시켜, 더 낮은 특이성을 가진 RNA 서열을 인식하고 이에 결합하는 그의 능력을 제거한다. 융합된 단백질은 3' UTR 내로 통합된 TAL 결합 서열로 향하고 엔도뉴클레아제 기능을 수행한다. 메가 TAL DNA 결합 서열은 융합 단백질에 의해 표적화된다. 유사하게, 다른 키메라(도 4c(중간 패널)) 및 특정 DNA 결합 도메인을 가진 융합 단백질(도 4c(하부 패널))을 디자인한다.

[0601] **실시예 4. 외생성 핵산 서열의 LINE-1 매개 역전위를 위한 예시적인 플라스미드 디자인 및 개발**

[0602] 이 실시예에서, 포유동물 세포 내로 전위될 수 있는 형질전이유전자에 작동 가능하게 연결된 ORF2 트랜스포존 요소, 및 mRNA 전사와 안정화를 위한 조절 요소를 포함하는 재조합 LINE-1 구축물의 전달 및 통합을 위한 플라스미드 벡터를 생성한다. 시험관내에서 포유동물 세포 내로 도입하거나, 포유동물, 설치류, 양, 돼지 또는 인간과 같은 유기체에게 투여하기 위해 더 처리될 수 있고/있거나 정제될 수 있는 mRNA는 박테리아 숙주 세포에서 전사될 수 있다.

[0603] 임의의 적합한 벡터 골격을 사용하여, 박테리아 시스템에서 재조합 핵산 서열을 삽입체로서 통합시키고 전사하여 mRNA를 생성하거나; 시험관내 전사 시스템을 이용하여 재조합 핵산 서열을 포함하는 mRNA를 생성할 수 있다. 여러 특징을 플라스미드에 추가한다. 성공적인 확장 가능한 mRNA 생성 및 정제 시, mRNA를 골수 세포와 같은 관심 있는 포유동물 세포 내로 도입할 수 있다.

[0604] 역전위 연구 분야에서 전통적으로 사용되는 플라스미드는 디자이너 유전자인 유전자 블록을 결여하고, 깁슨(Gibson) 어셈블리 방법을 정기적으로 이용하여 상이한 특징을 삽입하였다. 기존 벡터의 특징을 갖되, 새로운

특징을 삽입할 유연성을 가진 신규 벡터는 유전자 전달 시스템으로서 LINE-1 요소의 연구 및 최적화 둘 다에 유리할 수 있다. 플라스미드를 단독으로 사용하거나 플라스미드로부터 전사된 mRNA를 사용하여, 역전위 빈도를 증가시킬 수 있는 기본 특징 및 추가 특징은 아래에 요약되어 있다. 이 특징들 중 임의의 특징을 교체하기 위해 제한 부위 없이 원래의 5' UTR, 3' UTR 및 ORF간 서열을 가진 천연 LINE-1 서열을 함유하는 예시적인 플라스미드 디자인이 도 5(I)에 그래프로 표시되어 있다. 신규 최적화된 플라스미드:

[0605] • CMV 또는 EIF1a 또는 EF1a 프로모터로 대체된, 제거된 Dox 유도성 프로모터

[0606] • mRNA를 만들기 위해 추가된 T7 부위

[0607] • 코돈 최적화된 ORF1 및 ORF2

[0608] • mRNA를 안정화시키기 위해 추가된 WPRE 요소

[0609] • 단백질 보호를 돕기 위해 ORF2에 추가된 FLAG 태그

[0610] • 18 kb에서 14 kb로 감소된 크기

[0611] • 삽입을 용이하게 하기 위해 각각의 특징에서 추가된 블런트 제한 부위(블런트 화살표를 가진 점선)

[0612] • 포함된 G418 선택 마커

[0613] 플라스미드는 도(II)에 표시되어 있다.

[0614] 깁슨을 이용하여, 도 5(Ⅲ)에 나타낸 바와 같이 플라스미드 리포터 유전자용으로 역 분할 GFP를 삽입한다. mRNA 리포터용으로 완전한 역 GFP를 도 5(IV)에서와 같이 삽입한다.

[0615] 도 5(V)의 플라스미드 구축물을 모체로서 사용하여, 핵 이입을 돕기 위해 ORF2의 N-말단에서 핵 국소화 서열(NLS)을 삽입한다(도 5(VI)). ORF2의 발현을 증가시키기 위해 IRES 또는 또 다른 종결/프로모터 서열을 삽입한다(도 5(VII)). ORF2와 mRNA 사이의 더 강한 상호작용을 용이하게 하기 위해, MS2 헤어핀을 3' UTR에 삽입하고 MS2 코트 단백질 서열을 ORF2 단백질의 N-말단에서 삽입한다(도 5(VIII)). 향상된 특이성을 가진 상응하는 예시적인 ORF 및 이의 작용 기작은 이전 실시예 및 도 4b에 개시되어 있다. mRNA와 번역 리보좀의 더 강한 상호작용을 용이하게 하고 신생 ORF2가 mRNA에 더 잘 결합하도록 번역을 지연시키기 위해, Alu 요소를 mRNA의 3' UTR에 삽입한다(도 5(IX)). 활성이 더 높은 ORF2 단백질을 잠재적으로 사용하기 위해, ORF2의 RT 도메인을 II 군 인트론의 역전사효소 도메인으로 대체한다(도 5(X)). 또한, 밍크 고래 게놈은 가장 높은 수 및 퍼센트의 활성 LINE 요소를 가진다(480개(3.6% 활성)를 가진 인간에 비해 약 5,000개(60% 활성)). 두 서열들은 67% 동일하며 고래 서열은 활성 엔도뉴클레아제 및 역전사효소 잔기를 가진다. 각각의 밍크 고래 도메인은 천연 ORF2 엔도뉴클레아제 및/또는 RT 도메인을 대체하거나 키메라 도메인을 디자인하는 데 사용될 수 있다.

[0616] **실시예 5. mRNA 디자인 합성 mRNA 생성**

[0617] 올리고뉴클레오타이드의 올리고합성 및/또는 라이게이션에 의한 합성 제조를 위해 mRNA를 전략적으로 디자인할 수 있다. 추가로, 이러한 디자인은 시험관내 전사(IVT) 매개 mRNA 생성에 유용하다. mRNA 전략은 이전 실시예에서 논의된 플라스미드 전략과 동일한 변이체를 포함할 수 있다. 주요 차이점은 리포터 GFP 서열이 인트론을 포함하지 않고(도 6a) 구축물이 ORF1 코딩 영역 없이 전달될 수 있다는 점(도 6b)이다.

[0618] **실시예 6. 증가된 mRNA 반감기를 위한 구조적 특징**

[0619] 이 실시예에서, 구조적 특징은 mRNA 반감기를 증가시키기 위해 역전위 요소 및/또는 형질전이유전자를 포함하는 mRNA에 도입된다. 목표는 1차 단핵구에서 mRNA로부터의 단백질 발현의 기간을 3일에서 적어도 5일까지 증가시키는 것이고, 이때 궁극적인 목표는 10일이다.

[0620] 도 7b(왼쪽)에 나타낸 바와 같이, CD14+ 골수 세포(단핵구) 내로 도입될 때 형질전이유전자를 코딩하는 서열을 포함하는 mRNA는 번역되고 암 세포의 항원에 결합할 수 있는 키메라 수용체(ATAK 구축물)를 발현한다.

[0621] 다수의 mRNA 디자인은 (i) G-사중체, (ii) 5' UTR 내의 바이러스 슈도노트 구조; 및/또는 3' UTR 내의 (iii) 하나 또는 (iv) 더 많은 xrRNA 루프 구조; (v) 도 7a에 나타낸 삼중체 RNA 구조 중 하나 이상을 단독으로 또는 조합으로 포함하는 다양한 유전자 블록들을 합성함으로써 생성되고; 형질전이유전자의 코딩 서열에 인접한 각각

의 UTR에서 전사 벡터 내로 클로닝된다. 이 구축물들은 외부 판매회사에 의해 개별적으로 제조되고 키메라 수용체의 발현에 의해 측정된 mRNA의 안정성을 측정하기 위해 사내에서 시험된다(예시적인 수용체 및 이의 기능은 도 7b(왼쪽)에 그래프로 도시되어 있다). 공정 순서도는 도 7b(오른쪽)에 표시되어 있다. 요약하건대, 구축물을 코딩 또는 변형된 폴리 A 꼬리를 가진 플라스미드 내로 클로닝한다. mRNA를 전사하고 정제하였다. 한편, 냉동된 단핵구를 해동시키고 수거한다. 수거된 세포를 정제된 mRNA(5 내지 10 μg)로 전기천공하고 1일, 2일, 3일 또는 5일 동안 배양하였다. 키메라 수용체에 대해 양성인 세포(결합제 양성 세포)를 표적 세포, 또는 표적 항원으로 코팅된 기판에 결합하는 그의 능력으로 검출한다. 예상된 결과는 도 7c에 표시되어 있다. (i) 내지 (v)에 요약된 구조적 특징들 중 하나 이상 또는 이들의 조합을 발현하는 벌크 또는 정제된 mRNA(실선 정사각형으로 표시된 데이터)는 (i) 내지 (v)에 요약된 특징들 중 임의의 특징을 함유하지 않는 상업적으로 입수 가능한 대응물(정삼각형으로 표시된 데이터)을 능가한다.

[0622] **실시예 7. GFP 유전자의 LINE-1 레트로트랜스포존 플라스미드 매개 전달**

[0623] 이 시험 실행에서, LINE-1 레트로트랜스포존 시스템을 이용하여 GFP 카고의 게놈 통합 및 GFP 단백질의 발현을 검증하였다. LINE-1-GFP 구축물(LINE-1 플라스미드 GFP)은 도 8a에 예시되어 있다: ORF1p(ORF1)를 코딩하는 LINE-1 서열, ORF2p(ORF2)를 코딩하는 서열, 및 ORF에 대해 역방향으로 LINE-1의 3' UTR에 위치한 CMV 프로모터 유도 분할 GFP 유전자를 가진 플라스미드 구축물. 코딩 서열의 중간에서 비코딩 서열의 통합 및 스플라이싱 매개 제거가 일어난 후에만 GFP가 발현되게 하기 위해, 스플라이스 공여자 부위와 수용자 부위 사이에 삽입된 인트론 서열을 갖도록 분할 GFP를 디자인한다. 이 경우, 카고는 2.1 kb이다. 퓨진(Fugene) 시약을 사용하여 상기 플라스미드로 HEK293T 세포를 형질감염시켰고, 플라스미드 양성 세포를 퓨로마이신으로 선택하였다. 모의 형질감염된 세포에 비해 평균 형광 강도(MFI)(도 8b) 및 GFP 형광 강도를 가진 세포의 비율(도 8c)의 변화로서 표시된 바와 같이, 게놈 내로 통합된 GFP로부터 생성된 mRNA는 성공적으로 번역되고 유세포분석에 의해 측정된다. 모의 형질감염된 세포는 GFP 서열을 결여하는 플라스미드를 받았다.

[0624] **실시예 8. 키메라 수용체 유전자의 LINE-1 레트로트랜스포존 플라스미드 매개 전달**

[0625] 이 실시예는 세포에서 LINE-1 서열을 사용하여 재조합 유전자를 성공적으로 발현시킬 수 있음을 입증한다. CD5 결합 도메인의 인트론 서열에 의해 불연속된 재조합 수용체 단백질 CD5-인트론-fcr-PI3K(ATAK)를 코딩하는 3 kb 카고 서열과 함께 LINE-1 요소를 가진 플라스미드로 HEK 293 세포를 형질감염시켰다. 카고는 CD5 결합 세포외 도메인, FCRɣ 막횡단 도메인, 및 PI3-키나제 동원 도메인을 가진 세포내 도메인을 가진 키메라 수용체이다. 레트로트랜스포존 플라스미드의 도식적 표시는 도 9a에 제시되어 있다. 상기 실험의 디자인에서와 같이, ATAK 수용체는 게놈 내로 통합되고 인트론이 스플라이싱되지 않는 한 발현될 수 없다. HEK293T 세포에서의 형질감염 후, CD5 결합 세포외 도메인에 대한 미끼로서 표지부착된 CD5를 사용하여 수용체 발현을 검출한다. 도 9b 및 9c에 제시된 결과는 수용체의 성공적인 통합 및 발현을 보여준다. 36.5% 세포는 ATAK(CD5 결합제) 양성이었다(도 9c).

[0626] 추가 변형에서, 불연속되지 않은 재조합 수용체 단백질 CD5-인트론-fcr-PI3K를 코딩하는 더 긴 3.7 kb 카고 서열, 및 수용체와 GFP 서열 사이에 T2A 서열을 가진 불연속된 GFP 서열을 가진 LINE-1 구축물(LINE-1플라스미드-cd5_fcr-pi3k_t2a_GFP인트론)(도 10a). 모의 형질감염된 세포에 대해 정규화되었을 때, ATAK 수용체의 10배 초과의 증가가 있었고 GFP 이중 양성 세포가 인지되었다(도 10b). 실험 실행에 대한 GFP 및 형광 태그부착된 CD5 결합 및 게이팅 정량의 예시적인 형광 식별이 도 10c 및 도 10d에 표시되어 있다.

[0627] **실시예 9. 카고 유전자의 전달을 위한 LINE-1 레트로트랜스포존을 코딩하는 mRNA**

[0628] 이 어세이에서, LINE-1 역전위 유전자 서열을 mRNA로서 전달하고 발현하는 능력을 시험하였다. ORF1을 코딩하는 mRNA(ORF1-FLAG-mRNA), 및 CMV 프로모터 서열과 안티센스 방향으로 ORF2 및 GFP를 코딩하는 mRNA(ORF2-FLAG-GFPai)를 도 11a에 표시된 바와 같이 디자인한다. 이 어세이에서 카고 크기는 2.4 kb이었고, GFP는 ORF2 서열에 대해 안티센스 방향으로 존재한다. 상기 mRNA를 293T 세포 내로 전기천공하였고, 리포터 유전자 발현을 도 11b에 나타낸 바와 같이 입증하였다. 이 실험 설정은 ORF1-리드쓰루가 ORF2p의 발현에 필요하지 않으며 상이한 mRNA 분자로부터의 ORF2p의 발현이 ORF2p 및 GFP의 더 높은 발현을 허용할 수 있음을 입증하였다. 이 결과로, mRNA 형태로 LINE-1과 카고의 성공적인 전달을 달성하였다.

[0629] ORF1 및 ORF2 mRNA의 상대적 수준이 GFP 발현에 영향을 미치는지를 확인하기 위해, 다양한 양의 ORF1 및 ORF2 mRNA를 시험할 실험을 설정하였다(도 11a). GFP+ 세포의 증가에 대해 3배 양을 각각 및 함께 시험하고, 결과는 도 12a에 제시되어 있다. 배수 증가는 1x ORF2-GFP 및 1x ORF1 mRNA를 기준으로 한 것이다. GFP 발현은 3x

ORF1을 1x ORF2와 함께 사용하였을 때 더 높았으나, 그 반대의 경우는 그러하지 않은 반면; 3x ORF1 및 3x ORF2 둘 다를 가진 것은 비교된 세트에서 최대 수준의 GFP 발현을 보여주었다. 이때, 카고 크기는 2.4 kb이다. 도 12b는 레트로트랜스포존 mRNA 전기천공 후 GFP+ 세포의 형광 현미경관찰 영상을 보여준다.

[0630] 단일 mRNA 분자에서 안티센스 방향으로 ORF1과 ORF2 둘 다 및 GFP 형질전이유전자를 코딩하는 완전한 LINE-1 mRNA(LINE 1-GFP mRNA 구축물)를 세포 내로의 전달 및 게놈 통합에 대해 시험하였다. mRNA는 3'에서 5'으로 3' UTR 내에 CMV-GFP 서열과 함께 바이시스트론 ORF1 및 ORF2 서열을 함유한다(도 13a). 이 연구에서, 카고 크기는 2.4 kb이다. 도 13b에 나타낸 바와 같이, 전달된 ORF2-cmv-GFP 안티센스(LINE-1 mRNA)의 역전위 시, 왼쪽에서 세 번째 막대로 표시된 세포는 ORF1과 ORF2가 별도의 mRNA 분자에 있는 경우(그래프 막대 1, 2)에 비해 더 높은 GFP를 발현하였다. LINE-1 완전한 mRNA 이외에 별도의 mRNA에 ORF1을 포함시켰을 경우 LINE-1 단독에 비해 GFP 발현이 증가되었다. ORF2+GFP의 포함은 GFP 카고 코딩 mRNA와 추가 ORF2의 기여일 수 있는 더 높은 GFP를 예상대로 보여주었다.

[0631] 후속 전기천공이 역전위 효율을 증가시키는지를 시험하기 위해, 세포를 48시간마다 전기천공하였다. 24시간 내지 72시간 동안 배양한 후 유동을 이용하여 GFP 양성 세포를 평가하였다. 형광 데이터를, 단일 전기천공 이벤트를 이용한 세트의 값으로 정규화하였다. 도 14a 및 14b에 나타낸 바와 같이, 다회 전기천공은 전위된 유전자의 발현이 상향되는 추세를 이끌어냈으나, 변화는 완만하였다.

[0632] **실시예 10. mRNA로 역전위를 향상시키기 위한 ORF2 단백질 서열의 변형**

[0633] 판독값으로서 GFP 리포터를 사용하여 mRNA 전달을 통한 역전위를 향상시키기 위한 LINE-1 서열의 변형을 시험하였다. 다음과 같이 실험을 수행하였다. 모든 변형은 바이시스트론 ORF1 및 ORF2 서열과 관련되었다. (i) C-말단 NLS 서열을 ORF2 서열에 삽입함으로써 ORF2-NLS 융합체를 생성하였다. (ii) 인간 ORF2를 밍크 고래 ORF2로 대체하였다(Ivancevic et al., 2016). (iii) 3' UTR 내로의 Alu 요소의 통합: LINE-1의 3' UTR에서 Alu 요소의 최소 서열(AJL-H33△; Ahl et al., 2015)을 사용한다. (iv) 3' UTR + ORF2-MCP 융합체에서의 MS2 헤어핀: LINE-1 서열의 3' UTR 내의 MS2 헤어핀 및 ORF2 서열에 융합된 MS2 헤어핀 결합 단백질(MCP)(도 15a). 모의 구축물은 야생형 인간 ORF2 서열을 가졌다.

[0634] 모의 구축물로 전기천공한 세포에 비해 GFP 양성 세포의 비율의 배수 증가의 정량은 도 15b에 제시되어 있다.

[0635] **실시예 11. 면역 세포에서의 역전위**

[0636] 이 실험에서, HEK 세포의 형질감염 후 임의의 변화가 GFP 발현을 개선하는지를 확인하기 위해 ORF간 영역을 더 조작한다. LINE-1 플라스미드 GFP를 사용하여 ORF간 영역을 다음과 같이 조작한다: (a) 한 구축물에서, ORF간 영역을 CVB3의 IRES로 대체하고; (b) 또 다른 구축물에서, ORF간 영역을 EV71의 IRES로 대체하고; (c) 3개의 별도의 구축물에서, E2A 또는 P2A 또는 T2A 자가절단 서열을 ORF간 영역에 삽입한다. 결과는 도 16에 표시된 바와 같다. 특히 T2A 서열 삽입 시, GFP 판독값이 LINE-1 플라스미드 GFP(LINE-1 야생형 플라스미드)에 비해 단지 약간 변화되었다. EV71 IRES 서열의 삽입은 GFP 발현을 개선한 반면, CVB3 IRES는 어떠한 개선도 보이지 않았다.

[0637] **실시예 12. 면역 세포에서의 역전위**

[0638] 면역 세포에서 역전위를 시험하기 위해, T 세포 림프종 세포주인 Jurkat 세포 내로 전기천공함으로써 CMV-GFP 안티센스 리포터 카고로 LINE-1 플라스미드 및 mRNA를 시험하였다(도 17a 및 17b). GFP 서열을 갖지 않은 플라스미드로 모의 세트를 전기천공하였다. 형질감염된 세포에서 GFP 발현을 평가하였고, 전기천공 후 4일째 날 대표적인 데이터가 도 17b에 표시되어 있다. 모의 형질감염된 세포를 기준으로 한 배수 증가를 보고한다. 플라스미드 및 mRNA 전달 방식 둘 다가 성공적인 GFP 발현을 야기하였다.

[0639] 다음으로, LINE-1 서열, 및 키메라 HER-2 결합 수용체 및 분할 GFP를 코딩하는 3.7 kb 카고를 가진 플라스미드(LINE-1 플라스미드 Her2-Cd3z-T2A-GFP인트론)로 THP-1 세포(골수, 단핵구 세포주)를 전기천공하였다(도 18a). 상기 카고는 HER2 결합 세포외 도메인, CD3z 막횡단 도메인 및 분할 GFP 리포터를 포함하는 키메라 수용체이다. 도 18b에서 입증된 바와 같이, 상기 플라스미드는 성공적으로 게놈 내로 통합되었고 연장된 발현을 보였다. 형질감염 후 6일째 날 대표적인 발현이 도면에 표시되어 있다. 이 연구로부터, LINE-1 매개 유전자 전달이 상피 세포 유형(HEK-293T 세포); T 세포(예를 들어, Jurkat 세포); 및 골수 계통의 세포(예를 들어, THP-1 세포)를 비롯한 다양한 세포 유형에서 성공적인 안정한 게놈 통합을 야기할 수 있고 발현을 연장시킬 수 있음이 입증되었다. 더욱이, 역전위는 프라임(Prime) 편집과 같은 CRISPR 의존적 기술과 달리 큰 유전 카고의 통합을 야기할 수 있고 이러한 카고를 단일 핵산 구축물로서 전달할 수 있다.

[0640] **실시예 13. 카고 서열의 LINE-1 매개 역전위의 효율을 더 향상시키는 외부 방법**

[0641] 세포 게놈 내로의 카고 서열의 역전위의 효율을 더 향상시키는 방법이 이 단락에 상세히 기재되어 있다.

[0642] 세포 주기의 특정 단계에 있는 집단 내의 세포의 선택 또는 적합한 작용제에 의한 G1 정지에 의한 세포 주기 동기화는 더 높은 핵산 흡수 효율, 예를 들어, 플라스미드 벡터 형질감염 효율 또는 전기천공 효율로 이어질 수 있다. 이 어세이에서, 세포를 미리 분류하고 각각의 군을 따로 전기천공하여 균일한 전기천공을 보장한다. 이 군들 사이에 전기천공의 효율을 비교하고 GFP 시험 플라스미드 또는 mRNA의 발현에 의해 확인된 최대 효율을 야기하는 세포 주기 단계를 선택한다(도 19).

[0643] 이 실험의 또 다른 변경에서, 전기천공 전에 몇 시간 동안 세포 주기 정지 시약으로 세포를 처리함으로써, 세포를 분류하거나 분류하지 않으면서 동기화한다. 세포 주기 정지 시약의 예시적인 목록은 표 1에 제공된다. 이 목록은 완전하지 않고 전구아폽토시스성을 나타낼 수 있는 시약을 포함하므로, 목적, 용량 및 인큐베이션 시간에 적합한 신중한 선택은 특정 상황에서의 사용을 위해 최적화된다.

표 1

세포 주기를 억제하는 데 사용되는 소분자 시약의 예시적인 불완전한 목록

[0644]

작용제	세포 주기	기작
5-[(4-에틸페닐)메틸렌]-2-티옥소-4-티아졸리디논	G0-G1에서 세포 주기 정지	c-Myc-Max 이량체화 억제
이트라코나졸(Itraconazole)	G1에서 세포 주기 억제	SMO 길항제
ABT 751(Tocris Bioscience, 카탈로그 #4138)	G2M에서 세포 주기 차단	마이크로튜불 증식 억제
아르테수네이트(Artesunate)	G2M에서 세포 주기 정지	ROS 유도 NLRP3 억제
AZD 5438	G2M, M, S 및 G1 기에서 세포 주기 차단	Cdk 억제
바이칼레인(Baicalein)	G1 및 G2 기에서 세포 주기 정지	리폭시게나제(lipoxygenase) 억제
CPI 203(별칭: TEN 101)	G1 기에서 세포 주기 정지	BET 브로모도메인 억제제
디아드제인(Diadzein)	G1에서 세포 주기 정지	에스트로겐 수용체 아고니스트
DIM	G2M 에서 세포 주기 차단	EGFR 활성화 유도
에포틸론(Epothilone) B	G2M 에서 세포 주기 정지	튜불린 증식 억제
인디루빈-3'-옥심	항증식	GSK3b 억제
MPC 6827 하이드로클로라이드	세포 주기 정지	마이크로튜불 증식 억제
플라디에놀라이드(Pladienolide)	G1 및 G2/M 억제	mRNA 스플라이싱 감소
플럼바긴(Plumbagin)	G2/M 정지 유도	TOR 신호전달 등의 억제
템시롤리무스(Temsirolimus)	G1/S 유도	mTOR 억제제
토세라닙(Toceranib)	세포 주기 정지	PDGFR 및 VEGFR 억제
WYE 687 디하이드로클로라이드	G1 정지 유도	mTOR 억제제
YC1	G1 정지 유도	구아닐릴 사이클라제 활성화제

[0645] 특정 생체외 용도의 경우, 역전사효소를 위한 상동 재조합 및 프라이밍에 대한 기회를 생성하는 조절된 방사선조사를 이용하여, 상기 실시예들 중 임의의 실시예에 기재된 바와 같은 역전위 기구를 발현하는 세포에서 DNA 이중 가닥 절단(DSB)을 유도함으로써 역전위를 향상시킨다(도 20). 또 다른 예에서, 세포를 LINE-1 플라스미드 GFP 구축물로 형질감염시키고 방사선조사 펄스에 노출시킨다. GFP 발현을 모니터링한다. 더 높은 GFP 발현에 의해 표시된 바와 같이, 최대 이점을 얻기 위해 방사선조사의 강도와 시간을 최적화한다.

[0646] 또 다른 예에서, LINE-1 플라스미드 GFP에 의해 형질감염된 세포를 다음과 같이 처리된 실험 세트로 나누었다: (i) DSB를 유도하기 위한 방사선조사(전술된 바와 같음); (ii) DNA 리가제를 차단하여 DNA 손상 복구 기구를 억

제하는 소분자, 예컨대, SCR7로 이 세트의 세포를 처리함. 보호 복구 기작이 역전위의 진행을 억제하는 것을 방해하는 것은 GFP 발현을 향상시킬 것으로 예상된다: (iii) 세포를 방사선조사한 후, 세포를 SCR7로 처리함, 이들 두 방법의 조합은 더 강력한 효과를 보여줄 것으로 예상된다. GFP 발현을 6일의 기간에 걸쳐 모니터링하고, 가장 긴 기간에 걸쳐 최대 GFP 형광을 보이는 세트는 추가 연구에서 채택되는 조건을 표시한다.

[0647] **실시예 14. 구축물의 추가 변형에 의한 카고 서열의 LINE-1 매개 역전위의 효율 향상**

[0648] *I. 안정성 및 더 높은 발현을 제공하기 위한 구축물의 비코딩 영역의 향상*. 이 실시예에서, 다음과 같이 증가된 GFP 발현에 대해 시험하기 위해 LINE-1 플라스미드-GFP를 더 변형시킨다: (a) 한 구축물에서, 5' UTR을 상보체 유전자의 UTR로 대체하고; (b) 또 다른 구축물에서, 증가된 안정성을 위해 3' UTR을 B-글로빈 유전자의 UTR 서열로 대체하고; (c) 또 다른 구축물에서, ORF간 영역을 CVB3의 IRES로 대체하고; (d) 또 다른 구축물에서, ORF간 영역을 EV71의 IRES로 대체하고; (e) 3개의 별도의 구축물에서, E2A 또는 P2A 또는 T2A 자가절단 서열을 도 21의 도식에 나타낸 바와 같이 ORF간 영역에 삽입한다. 전술된 것 이외에, 전술된 바와 동일한 설정을 이용하여 (a) 내지 (e)의 다양한 조합 및 표 2에 나열된 추가 조합을 시험한다. 병렬 시험 세트에서 구축물을 HEK293T 세포 내로 형질감염시킨 후에 GFP 발현을 모니터링하여, 이 구축물들 중 임의의 구축물이 LINE-1 플라스미드 GFP 단독에 비해 GFP 발현을 증가시켰는지를 확인한다. 개선을 보여주는 조합을 채택한다.

표 2

역전위 효율을 증가시키기 위해 LINE-1 구축물에 포함시킬 5' 및 3' UTR, 및 ORF간 삽입 요소의 예시적인 조합

[0649]

서열들로부터 선택된 5'-UTR 서열	3'-UTR 서열	ORF간 서열
상보체 5' UTR	WPRE	T2A, E2A, P2A
코비드(Covid)-19 5' 리더 서열	B-글로빈 3' UTR	CVB3 IRES
CYBA 5' UTR	RSV RSE	EV71 IRES
CYP2E1 5' UTR	ARE	EMCV IRES
	ER에 대한 RNA 짚코드	PV IRES
	mtRNR1-AES	CSFV IRES
		HRV2 IRES
		AAA(트리알라닌 융합 또는 임의의 융합-링커 서열)

[0650] *II. 핵에서의 ORF의 국소화 및 체류의 향상*. 이 실시예에서, 증가된 GFP 발현에 대해 시험하기 위해 LINE-1 플라스미드-GFP를 다음과 같이 더 변형시킨다: (a) ORF2 코딩 서열을 핵 국소화 서열(NLS)과 융합시키고(도 15a에 그래프로 표시됨, 위에서 두 번째 구축물); (b) ORF1 코딩 서열을 핵 국소화 서열(NLS)과 융합시키고(도 22에 그래프로 표시됨); (c) Alu 결합 서열을 ORF2 역전사효소 코딩 서열의 3'에서 삽입하고(도 15a에 그래프로 표시됨, 위에서 네 번째 구축물); (d) (a)와 (c) 둘 다와 함께(표시되어 있지 않음); (e) (b)와 (c) 둘 다와 함께, NLS 서열을 ORF1 N-말단과 융합시키고, Alu 결합 서열을 ORF2 역전사효소 코딩 서열의 3'에서 삽입하고(도 22); (f) SINE 유래 핵 RNA 국소화(SIRLOIN) 서열을 LINE-1 3' UTR 내로 통합시킨다. HEK-293T 세포를 구축물 (a) 내지 (f) 및 LINE-플라스미드 GFP 구축물로 동시에 형질감염시켰다. HEK293T 세포 내로 형질감염시킨 후에 GFP 발현을 모니터링한다. 가장 긴 기간에 걸쳐 최대 GFP 형광을 보여주는 세트를 채택한다.

[0651] *III. LINE-1-단백질-RNA 복합체와 리보좀의 결합을 증가시키기 위한 구축물의 변형*. 이 실시예에서, LINE-1 단백질 RNA 구축물과 리보좀의 결합을 증가시키기 위해 Alu 요소 또는 리보좀 결합 앱타머인 추가 서열을 LINE-1 구축물의 3' UTR에 삽입한다(도 23).

[0652] LINE-1 단백질-RNA 복합체와 리보좀의 결합을 향상시키기 위해, 하기 요소를 mRNA의 3' UTR에 삽입하고 상기 실

험과 유사하게 시험한다. Alu 요소의 삽입은 앞서 기재되어 있다. 별도의 구축물에서, Alu 요소 절두체, 리보좀 결합 앱타머(109.2-3) 및 리보좀 확장 분절(ES9S) 결합 서열을 삽입하고 GFP 발현의 증가에 대해 각각 시험한다.

[0653] *IV. 역전위를 위한 ORF2와 그 자신의 mRNA의 결합 향상*. 이 실시예에서, MS2 결합 루프 구조를 함유하는 서열을 LINE-1의 3' UTR 내로 도입하고, MS2 RNA 결합 도메인을 코딩하는 서열을 ORF2p-RT의 RNA 결합 도메인과 융합시킨다(도 4a와 4b 및 도 24에 그래프로 표시됨, 구축물 서열번호 15). 융합된 단백질은 3' UTR에서 MS2 결합 구조 모티프에 특이적으로 부착할 것이므로, 임의의 비특이적 결합 및 역전위가 최소화된다(도 24). HEK293T 세포 내로 형질감염시킨 후에 GFP 발현을 모니터링한다. 유사한 디자인에 따라, ORF를 동일한 행의 상응하는 오른쪽 열에 표시된 ORF2의 3' UTR 영역에 삽입된 동족 서열과 조합된, 하기 표 3의 왼쪽 열에 표시된 단백질 결합 서열과 융합시킨다.

표 3

LINE-1 단백질의 번역 효율과 안정성의 향상 및 LINE-1 단백질의 증가된 발현을 위한 요소의 예시적인 목록

[0654]

LINE-1 ORF2 와 융합될 요소	요소에 의해 인식될 수 있는 3' UTR 서열
PP7 코트 단백질	PP7
스트렙타비딘	S1m 앱타머
토브라마이신	토브라마이신 앱타머

[0655] *V. 레트로트랜스포존의 엔도뉴클레아제 기능의 변형*. 이 실시예에서, GFP 발현의 증가를 시험하기 위해 구축물을 다음과 같이 변형시킨다. 첫 번째 실험 세트에서, LINE-1 플라스미드 GFP를 ORF2의 엔도뉴클레아제 코딩 서열의 3' 말단에서 절단하고, 이종 징크 핑거 단백질(ZFP)의 DNA 결합 도메인(DBD)을 코딩하는 서열을 삽입한다. 또 다른 실험 세트에서, 엔도뉴클레아제 도메인을 CRISPR 뉴클레아제와 융합시킨다. 표 4의 불완전한 목록에 표시된 바와 같이, DNA 결합 도메인 및 절단 도메인을 사용하여 융합 단백질을 생성하여 LINE-1 플라스미드 GFP ORF를 변형시킴으로써 다양한 뉴클레아제들을 시험할 수 있다. 또한, 이량체화를 용이하게 하기 위해 2개의 ORF-2 도메인을 한 세트로 코딩한다. ORF2 엔도뉴클레아제보다 더 높은 GFP 발현을 가진 구축물을 추가로 선택할 수 있다. 플라스미드 디자인은 도 25에 그래프로 표시되어 있다. 플라스미드를 HEK293T 세포 내로 형질감염시킨 후에 GFP 발현을 모니터링하고, 가장 좋은 결과를 제공하는 세트를 확인한다.

표 4

LINE-1 엔도뉴클레아제에 융합될 수 있거나 이 엔도뉴클레아제 대신에 삽입될 수 있는 추가 DNA 절단 도메인/효소의 예시적인 불완전한 목록

유전자/효소	설명
FokI	플라보박테리움 오케아노코이테스(Flavobacterium okeanokoites)의 클래스 II 엔도뉴클레아제, 인식 서열과 절단 서열은 몇 개의 뉴클레오타이드에 의해 분리됨; DNA 서열 5-GGATG-3' 인식
제한 효소	예를 들어, HindII, EcoR1, BamH1
LAGLIDADG 패밀리 뉴클레아제 A(서열번호 66 으로서 개시된 "LAGLIDADG")	박테리아를 비롯한 다양한 속들에서 발견된 인트론 코딩된 귀소 단백질
GIY-YIG	이 도메인은 엑소뉴클레아제 abc 서브유닛 c(uvrC), 박테리오파지 T4, 엔도뉴클레아제 segA, segB, seg C, seg D 및 seg E, 진균 및 파지의 I 군 인트론의 아미노 말단 영역에서 발견됨
His-Cys 박스	100 개 아미노산 영역에 걸쳐 보존된 히스티딘과 시스테인 잔기의 두 클러스터를 함유하는 귀소 엔도뉴클레아제
H-N-H	파지 DNA 에 널리 존재하는 뉴클레아제. 이. 콜라이 파지 HK97 의 종결효소(terminase) 팩키징 반응의 중요한 성분.
PD-(D/E)xK	다수의 단백질들, 예를 들어, DUF4420, DUF3883, DUF4263, COG5482, COG1395, Tsp45I, HaeII, Eco47II, ScaI, HpaII 에 존재하는 포스포디에스터라제.
Vsr 유사/EDxHD	박테리아의 매우 짧은 복구(Vsr) 엔도뉴클레아제에 대한 인식 가능한 상동성을 나타내는 C-말단 뉴클레아제 도메인

[0656]

[0657] *VI. 레트로트랜스포존의 역전사효소 기능의 변형*. 이 실시예에서, ORF2의 역전사효소 도메인을 그의 효율을 증가시키도록 변형한다. 한 실험 세트에서, LINE-1 플라스미드 GFP에서 인간 ORF2를 코딩하는 서열을 절제하고 MMLV 또는 TGIRTII를 코딩하는 서열로 대체한다. 또 다른 실험 세트에서, ORF2 역전사효소 도메인을 이종 단백질의 DNA 결합 도메인과 융합시킨다. 역전사효소 도메인 및/또는 DNA 결합 도메인을 표 5A 및 표 5B에 제공된 불완전한 목록으로부터 선택할 수 있다. 구축물은 도 26에 그래프로 예시되어 있다. HEK293T 세포 내로 형질감염시킨 후에 GFP 발현을 모니터링한다.

[0658] [표 5A]

더 높은 효율을 위해 LINE-1 RT 를 대체하는 역전사효소의 선택된 불완전한 목록

역전사효소	설명
M-MLV-RT	뮤린 백혈병 바이러스
TGIRT-II	높은 신뢰도 및 진행성을 가진 열안정성 II 군 인트론 역전사효소
AMV-RT	조류 골수아구증 바이러스 역전사효소
II 군 인트론 성숙효소(maturase) RT	유박테리움 렉테일(*Eubacterium rectale*)로부터 유래함
HIV-RT	HIV 로부터 유래한 효율적인 RT
TERT	6 개 뉴클레오타이드 텔로머성 반복부 유닛인 5'-TTAGGG-3'를 가진 3'-염색체 말단의 RNA 의존적 연장을 촉매작용함.

[0659]

[0660] [표 5B]

더 높은 효율을 위해 RT 에 융합시킬 DNA 결합 도메인의 선택된 불완전한 목록

DNA 결합 도메인(DBD)
징크 핑거 도메인
류신 지퍼(bZip)
나선-회전-나선 도메인
HMG-박스
R2 역요소 DBD
Sso7d
단백질 A(ssDNA)
OB-폴드(ssDNA)

[0661]

[0662] *VII. 다른 유기체의 LINE-1에 의한 인간 LINE-1의 대체*. 이 실시예에서, 인간 LINE-1을 코딩하는 서열을 상이한 유기체의 LINE-1로 대체한다. 한 예에서, 인간 LINE-1 구축물을, 인간 LINE-1이 밍크 고래 LINE-1 서열로 대체되어 있는 구축물과 비교한다(도 27). 동일한 실험 프레임워크를 이용하여, 다수의 ORF들을 시험한다. 예시적인 불완전한 목록이 하기 표 6에 제공된다. 더 포괄적인 목록은 문헌[Ivancevic A. et al., Genome Biol Evol 8(11):3301-3322]에서 이용 가능하다.

표 6

인간 LINE-1 의 대체에 사용하기 위한 유기체의 예시적인 LINE-1 요소

종명	총 LINE-1 의 수/불활성의 수/퍼센트 활성
발라에노프테라 아쿠토로스트라타 스캄모니 (Balaenoptera acutorostrata scammoni)	8,012/5,006/62.4%
리노피테쿠스 록셀라나(Rhinopithecus roxellana)	11,115/2,954/26.5%
무스 무스쿨루스(Mus musculus)	18,280/4,143/22.66%
애데스 애깁티(Aedes aegypti)	519/184/35.4%
제아 메이스(Zea mays)	744/165/22.17%
브라시카 나푸스(Brassica napus)	1,929/565/29.2%
브라시카 라파(Brassica rapa)	543/228/41.9%
다니오 레리오(Danio rerio)	590/268/45.4%

[0663]

[0664] 또 다른 세트에서, 인간 LINE-1을 GFP 플라스미드에서와 같이 유지하되, 인간 LINE-1 사일런서의 억제제를 사용하여 HUSH 복합체 TASOR 단백질과 같은 내생성 단백질에 의한 인식을 방해한다. 이 경우, TASOR 억제제는 miRNA와 같은 억제 RNA이다.

[0665] VIII. 표적 특이성을 위한 LINE-1 융합 단백질. 이 실시예에서, LINE-1 플라스미드 GFP ORF2를 메가TAL 뉴클레아제의 도메인, CRISPR-CAS 뉴클레아제, TALEN, R2 역요소 결합 징크 핑거 결합 도메인, 또는 반복 요소, 예컨대, Rep78 AAV에 결합할 수 있는 DNA 결합 도메인과 융합시킨다. 도 28은 디자인을 예시한다. 표 7은 서열 특이적 역전위를 증가시키기 위해 융합될 수 있는 상이한 요소들의 목록을 제공한다.

표 7

역전위 특이성을 증가시키기 위해 ORF2 에 융합시킬 DNA 결합 도메인을 가진 예시적인 단백질

[0666]

요소
전사 인자
메가 TAL 뉴클레아제
TALEN
다른 역요소의 징크 핑거 결합 도메인
안전한 은신처 결합 단백질
Cfp1

[0667] 각각의 플라스미드를 HEK293 세포 내로 형질감염시키고, GFP 발현을 모니터링한다.

[0668] 판독값으로서 GFP를 사용하여 역전위 효율의 증가에 대해 시험하기 위해 이 단락의 (I) 내지 (VIII)에 기재된 변형을 디자인한다. 이것 다음에, (I) 내지 (VIII)의 많은 유용한 변형들을 단일 역전위 구축물 내로 도입하고, 결과를 위해 삽입체로서 GFP를 사용하여 시험하고, GFP 서열을 원하는 삽입체 서열로 대체한다.

[0669] **실시예 15. 레트로트랜스포존 기술을 이용한 장기간 발현을 위한 큰 페이로드의 전달**

[0670] 본원은 레트로트랜스포존 구축물이 핵산 페이로드를 세포의 게놈 내로 통합시키고 예시적인 형질전이유전자를 발현하는 데 유용함을 입증하는 예를 제공한다. 레트로트랜스포존 구축물을 본 개시내용의 다른 부분에서 상세히 설명된 바와 같이 디자인하였다.

[0671] 요약하건대, 검증 실험의 한 세트에서, GFP 코딩 페이로드를 다음과 같이 구축하였다: 독시사이클린 유도성 조절 하에 있는 안티센스 프로모터 서열에 이어, 센스 방향으로 인트론에 의해 분할된 안티센스 GFP 유전자를 LINE-1 ORF의 하류에 배치하였다(도 29). 스플라이싱 공여자(SD)와 스플라이싱 수용자(SA) 서열은 mRNA가 상부 가닥의 프로모터로부터 생성될 때에만 인식되고 스플라이싱되므로, 스플라이싱된 mRNA로부터 게놈 내로 통합된 GFP 유전자만이 형광 신호를 생성한다. 도 2의 대표적인 유세포분석 데이터에서 확인된 바와 같이, 유세포분석을 이용하여 ORF 발현의 독시사이클린 유도 후 35일째 날에 GFP 발현을 측정하고(녹색 히스토그램) 음성 대조군 플라스미드(회색 히스토그램)와 비교하였다. 이 경우, 카고 크기는 2.4 kb이었다.

[0672] 이전 구축물의 카고 GFP 유전자를 인트론에 의해 불연속된 CD5-FcR-PI3K CAR-M 서열로 대체하였다(Morrissey et al., 2018). 역전위된 세포가 플라스미드로 형질감염시킨 음성 대조군 세포 집단(회색 히스토그램)에 비해 CD5-AF647 양성(적색 히스토그램)을 나타내도록 Alexa647-접합된 CD5 단백질을 사용하여 유세포분석으로 CD5 결합제 발현을 측정하였다(도 30). 도면에 나타낸 바와 같이 3.0 kb 구축물의 성공적인 발현이 입증되었다.

[0673] 인트론에 의해 불연속된 GFP 유전자를 CD5-FcR-PI3K CAR-M 서열의 하류에 있는 T2A 서열 뒤에 추가함으로써 카고 유전자 길이를 연장시켰다(도 31). Alexa647-접합된 CD5 단백질을 사용하여 유세포분석으로 CD5 결합제 발현을 측정하였다. 음성 대조군(회색 히스토그램)에 비해 적색 히스토그램으로 표시된 CD5 결합제 양성 세포. 유세포분석을 이용하여 GFP 발현을 측정하고(녹색 히스토그램) 음성 대조군 플라스미드로 형질감염시킨 세포(회색

히스토그램)와 비교한다. Q2의 유세포분석 신호는 10.8%의 세포가 CAR-M 및 GFP 단백질 둘 다를 발현한다는 것을 보여주었다.

[0674] 도 32에 나타낸 바와 같이, 3.9 kb 페이로드를 사용한 레트로트랜스포존 전달 및 통합(Retro-T 전달)에 의해 페이로드 크기 제한이 도달되지 않았다. 여기에 설명된 전달 기작은 제1세대 CART 구축물 및 GFP(T2A 부위에 의해 분리됨)의 발현에 성공적이었다.

[0675] 이 실시예에서, 삽입체 서열의 역전위 효율에 대해 다양한 구축물들을 시험하였다. 도 33a는 mRNA가 성공적인 통합을 야기할 때 유전자 전달을 보여준다. 이 데이터는 Retro-T가 mRNA로서 전달될 수 있음을 보여주는 최초의 데이터이다. 도 33b(상부 패널)에 그래프로 예시된 바와 같이, 유전자 전달을 위해 3' UTR에서 안티센스 프로모터 및 GFP 카고를 가진 ORF1 및 ORF2(ORF2-GFPai)를 코딩하는 별도의 mRNA를 사용하는 트랜스 전략을 연구하였다. 도 33b 내지 33d는 다수의 대표적인 어세이로부터의 실험 결과를 보여준다. LINE-1 단백질을 발현하는 별도의 mRNA는 역전위에 필요한 RNA-단백질 복합체를 재구성할 수 있다. 시스 전략은 3' UTR에서 안티센스 프로모터 및 GFP 유전자 카고를 가진 단일 바이시스트론 LINE-1 mRNA를 사용한다. 다양한 양 및 비율의 ORF2 및 ORF1을 포함하는 구축물을 도 33b 및 33c에 나타낸 바와 같이 페이로드로서 GFP 코딩 서열과 비교하였다. 도 33d는 단일 mRNA의 도입이 세포당 더 많은 수의 통합을 일으킴을 보여준다. 생화학적 어세이 및 통합 어세이를 위해 역전위된 세포를 농후화하기 위한 293T GFP 세포의 분류. 세포는 도 33b의 세포와 동일하고 하부 패널에서 분류 후 4일째 날에 GFP 발현을 보여준다. 그래프는 다양한 LINE-1 플라스미드에 의해 형질감염된, LINE-1 mRNA(레트로-mRNA), 및 ORF1 및 ORF2-GFP mRNA 전기천공된 세포로부터의 게놈 DNA 통합에 대한 qPCR 어세이를 보여준다. 2개의 qPCR 프라이머-프로브 세트를 사용하였는데, 하나는 하우스킵핑 유전자 RPS30을 위한 qPCR 프라이머-프로브 세트이고 다른 하나는 GFP 유전자를 위한 qPCR 프라이머-프로브 세트이다. 플라스미드로 형질감염시킨 세포는 SV40 유지 서열을 함유하지 않는 플라스미드를 사용한다. 플라스미드 및 게놈 DNA의 표준 곡선의 보간을 통해 샘플당 카피 수를 측정하고 293T 세포당 2개의 RPS30 카피에 대해 정규화함으로써 세포당 통합을 계산한다. 오차 막대는 3회 기술적 재현실험 측정치의 표준 편차를 나타낸다.

[0676] **실시예 16. 다양한 세포 유형으로의 전달**

[0677] 도 34 내지 38에 나타낸 바와 같이, 예를 들어, CAR 단백질 또는, 예를 들어, GFP 단백질을 코딩하는 관심 있는 유전자를 포함하는 mRNA 구축물을 다양한 세포 유형, 예컨대, 상피 세포(예를 들어, HEK 293 세포), 단핵구 세포주(예를 들어, THP-1 세포), 림프모구성 세포주(예를 들어, K562 세포) 및 1차 림프구(T 세포)에서 효율적으로 발현시킬 수 있다. 활성화된 1차 T 세포도 GFP의 게놈 통합 및 발현을 가진 mRNA로 성공적으로 형질감염시켰다(도 36). 1차 T 세포를 단리하고 IL7/IL15를 사용하여 증폭시켰고; 제1세대 CAR 구축물을 활성화 후 2일째 날에 전달하였다. 세포를 분류하고 냉동시켰다. GFP 발현은 냉동-해동 주기 후에 검출될 수 있었다(도 37a 및 37b). 이것은 mRNA 매개 전달 및 L1-트랜스포존 매개 통합의 다양한 성질을 표시한다. 도 38은 293T 세포, K562 세포, THP-1 세포 및 1차 T 세포에서의 GFP mRNA 통합 및 발현의 대표적인 어세이를 보여준다.

[0678] **실시예 17. 역전위 시스템의 변형, 게놈 통합을 위한 통합효소의 부위 특이성의 교환**

[0679] 이 예후적 실시예에서, 세린 통합효소를 사용하여 고도의 위치 특이성으로 레트로트랜스포존 통합 시스템을 이용하여 비분열 포유동물 세포에서 큰 mRNA 구축물을 세포의 게놈 내로 안정적으로 통합시키고 발현시킨다. 레트로트랜스포존 구축물은 본 개시내용의 다른 부분에서 상세히 설명된 바와 같이 디자인된다.

[0680] 이 실시예에서, 제1 구축물인 mRNA 구축물은 통합효소에 의한 이중 가닥 폴리뉴클레오타이드 통합의 가능성을 제거하는 돌연변이를 촉매 도메인 내에 포함하는 세린 통합효소와 융합된 Cas 엔도뉴클레아제, 및 어떤 가이드 RNA가 구체적으로 디자인되는지에 따라 특정 게놈 유전자좌, 예를 들어, 히스톤 유전자 유전자좌, 액틴 유전자 유전자좌 등으로 향하는 프라임 편집 가이드 RNA(pegRNA) 올리고뉴클레오타이드를 포함하도록 디자인된다. 하나 이상의 T2A 절단 부위는 개별 단백질, 예를 들어, Cas9, 및 세린 통합효소 또는 이의 단편의 코딩 서열을 분리한다. pegRNA는 41 bp AttB 랜딩 부위를 삽입하도록 디자인된다. 하나 이상의 대안적인 디자인에서, 통합효소 결합 부위와 관련하여 ORF2 단백질의 결합을 위해 편리하게 가까이 위치하는(예를 들어, 병치된) LINE 1 ORF2 결합 부위가 통합된다.

[0681] 제2 mRNA 구축물은 5, 6, 7, 8, 9 또는 10 킬로염기보다 더 큰 카고 서열과 융합되고 단백질을 코딩하는 LINE1 mRNA; 돌연변이된 엔도뉴클레아제; 및 통합효소가 카고 서열을 유인하여 랜딩 서열로 데려오도록 카고 서열에 융합된 AttP 부착 부위를 포함하도록 디자인된다. AttP 부착 부위는 AttB 랜딩 부위에 상보적이다.

[0682] 시험관내 전사를 통해 mRNA 구축물을 제조한 후 정제한다. 각각의 생성된 정제된 mRNA 구축물을 지질 나노입자

(LNP) 내에 따로 도입한다. 전기천공을 이용하여, LNP를 사용하여 2개의 mRNA 구축물을 세포 내로 공-전달한다. pegRNA에 융합된 Cas9 엔도뉴클레아제는 게놈 랜딩 서열을 특정 유전자 유전자좌 내로 삽입하도록 안내된다. AttB 랜딩 부위는 디자인에 의해 LINE1 mRNA와 연결된 카고 서열 내에 포함되어 있는 AttP 부착 부위를 유인하는 비콘처럼 작용한다. 단백질의 공-번역은 ORF 폴리펩티드의 생성으로 이어지고, 통합효소에 의해 안내되고 부착 부위에 의해 부착되며, ORF는 근접한 관련 ORF 결합 부위와 접촉하고 카고를 특정 부위에서 통합시킨다.

[0683] 또 다른 대안적인 디자인 시스템에서, 통합효소 시스템은 모두 우회되고 PEG-RNA는 가이드 RNA에 의해 안내된 특정 게놈 유전자좌에서 ORF 결합 부위만을 통합시킨다.

[0684] 부위 특이적 통합은 생거(Sanger) 시퀀싱에 의해 확인된다. 세포에서의 삽입체 서열의 발현을 입증하기 위해 유세포분석을 수행한다. mRNA가 세포에 해로운 영향을 덜 미치기 때문에 세포 생존 및 게놈 통합은 플라스미드 또는 벡터 기반 시스템에서 흔히 발견되는 세포 생존 및 게놈 통합보다 더 높은 것으로 발견되었다.

[0685] **실시예 18. 예시적인 서열**

[0686] 다음은 실시예에서 사용된 구축물의 예시적인 서열들이다. 이 서열들은 예시적인 참고 목적을 위한 것이며 당업자가 과도한 실험 없이 도출할 수 있는 서열 변경 및 최적화는 본 개시내용에 의해 예상되고 포함된다. mRNA 서열이 서열 제목에서 언급되는 경우, 구축물은 DNA 주형의 뉴클레오타이드를 언급하고 당업자는 상응하는 mRNA 서열을 쉽게 유도할 수 있다.

표 8

플라스미드 및 mRNA 구축물 서열
ORF1-FLAG-mRNA(코돈 최적화된 인간 ORF1 코딩 서열-FLAG)(서열번호 35): TAATACGACTCACTATAGGGAGAAAGACGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGGA ATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAGT CCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAAC TCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTGC ATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCAG GGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGCG CGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGAG GAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGCGT CCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGAAA ATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCGGT ATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGAGA AGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGACG GTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGAAG GAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGATT AAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAACTG CTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGATG GATTATAAGGATGACGATGATAAATGA (서열번호 35)
ORF2-FLAG-GFP aim RNA(코돈 최적화된 인간 ORF2 코딩 서열)(서열번호 36) TAATACGACTCACTATAGGGAGAAAGACGCCACCATGACAGGTTCAAATAGTCACATTACGATTCT CACTCTGAATATAAATGGGCTGAATTCTGCAATTAAACGGCACAGGCTTGCTTCCTGGATAAAGTC TCAAGACCCCTCAGTGTGCTGTATTCAGGAAACGCATCTCACGTGCAGGGACACCCATCGGCTGAA AATAAAAGGCTGGCGGAAGATCTACCAAGCCAATGGAAAACAAAAGAAGGCTGGGGTGGCGATAC TTGTAAGCGATAAAACAGACTTTAAACCAACTAAGATCAAACGGGACAAAGAGGGCCATTACATC ATGGTAAAGGGTAGTATTCAACAAGAGGAGCTGACTATCCTGAATATTTATGCACCTAATACTGGA GCCCCCAGATTCATAAAGCAAGTGTTGAGTGACCTTCAACGCGACCTCGACTCCCACACTCTGATC ATGGGAGACTTTAACACCCCGCTGTCCACTCTCGACAGATCTACTAGACAGAAAGTCAACAAGGAT ACACAGGAACTGAACAGTGCTCTCCACCAAGCGGACCTTATCGACATCTACAGAACACTCCACCCC AAAAGCACAGAATATACCTTCTTTTCAGCCCCTCACCACACCTATTCCAAAATTGACCACATTGTGG GGAGTAAAGCCCTTCTCTCCAAATGTAAACGGACCGAAATTATCACTAACTATCTCTCCGACCACA GTGCAATAAAACTTGAATTGCGAATTAAGAATCTCACTCAAAGTAGATCCACGACATGGAAACTGA ACAATCTCCTCTTGAATGACTACTGGGTGCATAACGAAATGAAGGCTGAAATAAAGATGTTCTTTG AGACCAACGAAAACAAAGACACCACGTACCAGAATCTCTGGGACGCTTTCAAAGCAGTGTGTCGA GGAAAATTTATTGCACTGAATGCTTACAAGCGGAAGCAGGAAAGATCCAAAATAGACACCCTGAC TAGCCAACTTAAAGAACTGGAAAAGCAAGAGCAAACTCATAGCAAAGCTAGCCGTCGCCAAGAAA TTACGAAAATCAGAGCTGAACTGAAGGAAATTGAGACACAGAAAACCCTGCAAAAGATAAATGAA AGCCGCAGCTGGTTCTTTGAACGCATCAACAAAATCGATAGGCCACTTGCTCGCCTTATCAAGAAG AAAAGGGAGAAGAATCAAATCGACACTATAAAGAATGATAAAGGCGATATAACCACCGATCCCAC AGAAATTCAAACAACCATACGCGAATACTACAAACACCTCTACGCCAATAAACTCGAAAATCTCGA GGAAATGGATACATTCCTCGACACGTACACCCTTCCCAGGCTGAACCAGGAAGAAGTTGAATCACT GAATCGGCCTATCACGGGGAGTGAAATAGTAGCTATCATCAATTCACTCCCTACCAAGAAGTCACC CGGACCTGATGGATTCACCGCCGAATTCTACCAGAGATACATGGAAGAACTGGTGCCCTTCTTGCT GAAACTTTTCCAAAGTATTGAGAAAGAGGGAATACTTCCAAACTCATTTTATGAGGCATCCATCAT TCTGATCCCGAAGCCCGGCAGGGACACGACCAAGAAAGAGAATTTTCGACCAATCTCATTGATGAA CATTGATGCAAAGATCCTCAATAAAATACTGGCAAATCGGATTCAGCAGCACATAAAGAAGCTGAT CCACCATGATCAAGTAGGCTTCATCCCCGGTATGCAAGGTTGGTTCAATATACGAAAATCAATCAA TGTTATCCAGCATATAAACCGGGCCAAAGACAAGAACCACATGATTATTAGTATCGATGCTGAGAA AGCCTTTGACAAAATACAACAACCCTTCATGCTGAAAACATTGAATAAGCTGGGAATTGATGGCAC CTACTTCAAAATCATCAGAGCCATATATGACAAACCAACAGCAAATATCATTCTGAATGGTCAGAA ATTGGAAGCATTCCCCTTGAAAACCGGCACACGGCAGGGTTGCCCTCTGTCACCACTCCTCTTCAAC ATCGTGTTGGAAGTTCTTGCCCGCGCAATCCGGCAGGAAAAGGAAATCAAGGGCATTCAACTGGGC AAAGAGGAAGTTAAATTGAGCCTGTTTGCAGACGACATGATCGTCTATTTGGAAAACCCCATAGTT AGTGCACAAAATCTGCTGAAGTTGATCAGTAATTTCTCCAAAGTGAGTGGGTACAAAATCAATGTG CAAAAGAGCCAAGCTTTCTTGTACACCAACAACAGGCAAACTGAGTCTCAAATCATGGGCGAACTC CCCTTCGTGATTGCATCCAAGCGGATCAAATACCTGGGGATTCAATTGACTCGTGATGTGAAGGAC CTCTTCAAGGAGAACTACAAACCCCTGCTCAAGGAAATCAAAGAGGACACAAACAAATGGAAGAA CATTCCATGCTCTTGGGTGGGAAGGATCAATATCGTCAAAATGGCCATCCTGCCCAAGGTAATTTA CAGGTTCAATGCTATACCCATCAAGCTCCCCATGACATTCTTCACAGAACTTGAAAAGACGACGCT GAAGTTCATTTGGAACCAGAAACGTGCCAGGATTGCTAAATCTATTCTCTCCCAAAAGAACAAAGC TGGCGGAATCACACTCCCAGACTTCAAACTTTACTACAAGGCGACCGTGACGAAAACGGCTTGGTA CTGGTACCAAAACAGGGATATAGATCAATGGAACCGAACGGAGCCCAGCGAAATTATGCCTCATA

[0687]

TATACAACTATCTGATCTTTGACAAACCGGAGAAGAACAAGCAATGGGGAAAGGATAGTCTGTTTA
ATAAATGGTGCTGGGAAAACTGGCTCGCAATCTGTAGGAAGCTGAAACTGGATCCATTCTTGACGC
CTTATACAAAGATAAATTCCCGATGGATTAAAGATCTCAACGTGAAACCCAAAACAATTAAAACCC
TCGAGGAAAACCTGGGTATTACGATTCAGGACATTGGGGTGGGAAAGGACTTCATGTCCAAAACCC
CAAAAGCGATGGCAACCAAAGACAAAATCGACAAATGGGATCTCATAAAACTTAAGTCATTTTGC
ACAGCTAAAGAAACGACAATTAGGGTGAACCGACAACCGACCACTTGGGAGAAAATCTTCGCAAC
ATACAGTTCTGACAAAGGCCTGATTTCCAGGATCTACAATGAATTGAAACAAATTTACAAGAAGAA
GACGAACAACCCTATAAAGAAATGGGCCAAGGACATGAACAGACACTTCTCTAAGGAAGACATTT
ATGCAGCCAAGAAACACATGAAGAAATGCAGCTCTTCACTGGCAATCAGGGAAATGCAAATCAAA
ACAACAATGAGATATCATCTCACACCCGTCAGAATGGCCATCATTAAGAAGAGCGGAAACAACCG
GTGCTGGCGTGGTTGCGGAGAAATCGGTACTCTCCTTCACTGTTGGTGGGACTGTAAACTCGTTCAA
CCACTGTGGAAGTCTGTGTGGCGGTTCCTCAGAGATCTGGAACTCGAAATCCCATTTGACCCAGCC
ATCCCTCTCCTGGGTATATACCCGAATGAGTATAAATCCTGCTGCTATAAAGACACCTGCACAAGG
ATGTTTATTGCAGCTCTCTTCACAATCGCGAAGACGTGGAACCAACCCAAATGTCCGACTATGATTG
ACTGGATTAAGAAGATGTGGCACATATACACTATGGAATACTATGCTGCGATCAAGAACGATGAGT
TCATATCATTTGTGGGCACATGGATGAAACTCGAAACCATCATACTCTCTAAATTGAGTCAAGAAC
AGAAAACTAAACACCGTATATTTTCCCTGATCGGTGGGAATTAGCTACAAAGACGATGACGACAAG
GACCATGGAGACGGTGAGAGACACAAAAAATTCCAACACACTATTGCAATGAAAATAAATTTCCTT
TATTAGCCAGAAGTCAGATGCTCAAGGGGCTTCATGATGTCCCCATAATTTTTGGCAGAGGGAAAA
AGATCTCAGTGGTATTTGTGAGCCAGGGCATTGGCCTTCTGATAGGCAGCCTGCACCTGAGGAGTG
CGGCCGCTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGC
AGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGG
TTGTCGGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCGAGCTGC
ACGCTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGG
CCATGATATAGACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTC
CTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCG
CGGGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCTTCGGGCATG
GCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTAG
GTCAGGGTGGTCACGAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAG
GGTCAGCTTGCCGTAGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTGGCCGTTTAC
GTCGCCGTCCAGCTCGACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACCAT
GGTGGCGGGATCTGACGGTTCACTAAACCAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTA
CCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTTGGTGCC
AAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATC
CACGCCCATTGATGTACTGCCAAAACCGCATCACCATGGTAATAGCGATGACTAATACGTAGATGT
ACTGCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGT
CATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTT
ACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACG
TCATTATTGACGTCAATGGGCGGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTT
ATGTAACGACGTCTCAGCTGACAATGAGATCACATGGACACAGGAAGGGGAATATCACACTCTGG
GGACTGTGGTGGGGTCGGGGGAGGGGGGAGGGATAGCATTGGGAGATATACCTAATGCTAGATGA
CACATTAGTGGGTGCAGCGCACCAGCATGGCACATGTATACATATGTAACTAACCTGCACAATGTG
CACATGTACCCTAAAACTTAGAGTATAATGGATCCGCAGGCCTCTGCTAGCTTGACTGACTGAGAT
ACAGCGTACCTTCAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGA
ATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAA
GCTGCAATAAACAAGTT**(서열번호 36)**

LINE-1 플라스미드 GFP**(서열번호 37)**
CGGCCGCGGGGGGAGGAGCCAAGATGGCCGAATAGGAACAGCTCCGGTCTACAGCTCCCAGCGTG
AGCGACGCAGAAGACGGTGATTTCTGCATTTCCATCTGAGGTACCGGGTTCATCTCACTAGGGAGT
GCCAGACAGTGGGCGCAGGCCAGTGTGTGTGCGCACCGTGCGCGAGCCGAAGCAGGGCGAGGCAT
TGCCTCACCTGGGAAGCGCAAGGGGTCAGGGAGTTCCCTTTCCGAGTCAAAGAAAGGGGTGACGG
ACGCACCTGGAAAATCGGGTCACTCCCACCCGAATATTGCGCTTTTCAGACCGGCTTAAGAAACGG
CGCACCACGAGACTATATCCCACACCTGGCTCGGAGGGTCCTACGCCCACGGAATCTCGCTGATTG
CTAGCACAGCAGTCTGAGATCAAACTGCAAGGCGGCAACGAGGCTGGGGGAGGGGCGCCCGCCAT
TGCCCAGGCTTGCTTAGGTAAACAAAGCAGCAGGGAAGCTCGAACTGGGTGGAGCCCACCACAGC
TCAAGGAGGCCTGCCTGCCTCTGTAGGCTCCACCTCTGGGGGCAGGGCACAGACAAACAAAAAGA
CAGCAGTAACCTCTGCAGACTTAAGTGTCCCTGTCTGACAGCTTTGAAGAGAGCAGTGGTTCTCCC
AGCACGCAGCTGGAGATCTGAGAACGGGCAGACTGCCTCCTCAAGTGGGTCCCTGACCCCTGACCC
CCGAGCAGCCTAACTGGGAGGCACCCCCCAGCAGGGGCACACTGACACCTCACACGGCAGGGTAT
TCCAACAGACCTGCAGCTGAGGGTCCTGTCTGTTAGAAGGAAAACTAACAACCAGAAAGGACATCT
ACACCGAAAACCCATCTGTACATCACCATCATCAAAGACCAAAAGTAGATAAAACCACAAAGATG
GGGAAAAAACAGAACAGAAAAACTGGAAACTCTAAAACGCAGAGCGCCTCTCCTCCTCCAAAGGA

[0688]

[0689]

ACGCAGTTCCTCACCAGCAACAGAACAAAGCTGGATGGAGAATGATTTTGATGAGCTGAGAGAAG
AAGGCTTCAGACGATCAAATTACTCTGAGCTACGGGAGGACATTCAAACCAAAGGCAAAGAAGTT
GAAAACTTTGAAAAAAATTTAGAAGAATGTATAACTAGAATAACCAATACAGAGAAGTGCTTAAA
GGAGCTGATGGAGCTGAAAACCAAGGCTCGAGAACTACGTGAAGAATGCAGAAGCCTCAGGAGCC
GATGCGATCAACTGGAAGAAAGGGTATCAGCAATGGAAGATGAAATGAATGAAATGAAGCGAGA
AGGGAAGTTTAGAGAAAAAAGAATAAAAAGAAATGAGCAAAGCCTCCAAGAAATATGGGACTAT
GTGAAAAGACCAAATCTACGTCTGATTGGTGTACCTGAAAGTGATGTGGAGAATGGAACCAAGTTG
GAAAACACTCTGCAGGATATTATCCAGGAGAACTTCCCCAATCTAGCAAGGCAGGCCAACGTTCAG
ATTCAGGAAATACAGAGAACGCCACAAAGATACTCCTCGAGAAGAGCAACTCCAAGACACATAAT
TGTCAGATTCACCAAAGTTGAAATGAAGGAAAAAATGTTAAGGGCAGCCAGAGAGAAAGGTCGGG
TTACCCTCAAAGGAAAGCCCATCAGACTAACAGCGGATCTCTCGGCAGAAACCCTACAAGCCAGA
AGAGAGTGGGGGCCAATATTCAACATTCTTAAAGAAAAGAATTTTCAACCCAGAATTTCATATCCA
GCCAAACTAAGCTTCATAAGTGAAGGAGAAATAAAATACTTTATAGACAAGCAAATGTTGAGAGA
TTTTGTCACCACCAGGCCTGCCCTAAAAGAGCTCCTGAAGGAAGCGCTAAACATGGAAAGGAACA
ACCGGTACCAGCCGCTGCAAAATCATGCCAAAATGTAAAGACCATCAAGACTAGGAAGAAACTGC
ATCAACTAATGAGCAAAATCACCAGCTAACATCATAATGACAGGATCAACTTCACACATAACAATA
TTAACTTTAAATATAAATGGACTAAATTCTGCAATTAAAAGACACAGACTGGCAAGTTGGATAAAG
AGTCAAGACCCATCAGTGTGCTGTATTCAGGAAACCCATCTCACGTGCAGAGACACACATAGGCTC
AAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCAA
TCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTAC
ATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTATGCACCCAATACA
GGAGCACCCAGATTCATAAAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTA
ATAATGGGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAA
GGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTCCA
CCCCAAATCAACAGAATATACATTTTTTTCAGCACCACACCACACCTATTCCAAAATTGACCACATA
GTTGGAAGTAAAGCTCTCCTCAGCAAATGTAAAAGAACAGAAATTATAACAAACTATCTCTCAGAC
CACAGTGCAATCAAACTAGAACTCAGGATTAAGAATCTCACTCAAAGCCGCTCAACTACATGGAAA
CTGAACAACCTGCTCCTGAATGACTACTGGGTACATAACGAAATGAAGGCAGAAATAAAGATGTTC
TTTGAAACCAACGAGAACAAAGACACCACATACCAGAATCTCTGGGACGCATTCAAAGCAGTGTGT
AGAGGGAAATTTATAGCACTAAATGCCTACAAGAGAAAGCAGGAAAGATCCAAAATTGACACCCT
AACATCACAATTAAAAGAACTAGAAAAGCAAGAGCAAACACATTCAAAAGCTAGCAGAAGGCAA
GAAATAACTAAAATCAGAGCAGAACTGAAGGAAATAGAGACACAAAAAACCCTTCAAAAAATCAA
TGAATCCAGGAGCTGGTTTTTTGAAAGGATCAACAAAATTGATAGACCGCTAGCAAGACTAATAAA
GAAAAAAAGAGAGAAGAATCAAATAGACACAATAAAAAATGATAAAGGGGATATCACCACCGAT
CCCACAGAAATACAAACTACCATCAGAGAATACTACAAACACCTCTACGCAAATAAACTAGAAAA
TCTAGAAGAAATGGATACATTCCTCGACACATACACTCTCCCAAGACTAAACCAGGAAGAAGTTGA
ATCTCTGAATCGACCAATAACAGGCTCTGAAATTGTGGCAATAATCAATAGTTTACCAACCAAAAA
GAGTCCAGGACCAGATGGATTCACAGCCGAATTCTACCAGAGGTACAAGGAGGAACTGGTACCAT
TCCTTCTGAAACTATTCCAATCAATAGAAAAAGAGGGAATCCTCCCTAACTCATTTTATGAGGCCA
GCATCATTCTGATACCAAAGCCGGGCAGAGACACAACCAAAAAAGAGAATTTTAGACCAATATCCT
TGATGAACATTGATGCAAAAATCCTCAATAAAATACTGGCAAACCGAATCCAGCAGCACATCAAA
AAGCTTATCCACCATGATCAAGTGGGCTTCATCCCTGGGATGCAAGGCTGGTTCAATATACGCAAA
TCAATAAATGTAATCCAGCATATAAACAGAGCCAAAGACAAAAACCACATGATTATCTCAATAGAT
GCAGAAAAAGCCTTTGACAAAATTCAACAACCCTTCATGCTAAAAACTCTCAATAAATTAGGTATT
GATGGGACGTATTTCAAAATAATAAGAGCTATCTATGACAAACCCACAGCCAATATCATACTGAAT
GGGCAAAAACTGGAAGCATTCCCTTTGAAAACCGGCACAAGACAGGGATGCCCTCTCTCACCGCTC
CTATTCAACATAGTGTTGGAAGTTCTGGCCAGGGCAATCAGGCAGGAGAAGGAAATAAAGGGTAT
TCAATTAGGAAAAGAGGAAGTCAAATTGTCCCTGTTTGCAGACGACATGATTGTTTATCTAGAAAA
CCCCATCGTCTCAGCCCAAAATCTCCTTAAGCTGATAAGCAACTTCAGCAAAGTCTCAGGATACAA
AATCAATGTACAAAAATCACAAGCATTCTTATACACCAACAACAGACAAACAGAGAGCCAAATCA
TGGGTGAACTCCCATTCACAATTGCTTCAAAGAGAATAAAATACCTAGGAATCCAACTTACAAGGG
ATGTGAAGGACCTCTTCAAGGAGAACTACAAACCACTGCTCAAGGAAATAAAAGAGGAGACAAAC
AAATGGAAGAACATTCCATGCTCATGGGTAGGAAGAATCAATATCGTGAAAATGGCCATACTGCCC
AAGGTAATTTACAGATTCAATGCCATCCCCATCAAGCTACCAATGACTTTCTTCACAGAATTGGAA
AAAACTACTTTAAAGTTCATATGGAACCAAAAAAGAGCCCGCATTGCCAAGTCAATCCTAAGCCAA
AAGAACAAAGCTGGAGGCATCACACTACCTGACTTCAAACTATACTACAAGGCTACAGTAACCAA
AACAGCATGGTACTGGTACCAAAACAGAGATATAGATCAATGGAACAGAACAGAGCCCTCAGAAA
TAATGCCGCATATCTACAACTATCTGATCTTTGACAAACCTGAGAAAAACAAGCAATGGGGAAAGG
ATTCCCTATTTAATAAATGGTGCTGGGAAAACTGGCTAGCCATATGTAGAAAGCTGAAACTGGATC
CCTTCCTTACACCTTATACAAAAATCAATTCAAGATGGATTAAAGATTTAAACGTTAAACCTAAAA
CCATAAAAACCCTAGAAGAAAACCTAGGCATTACCATTCAGGACATAGGCGTGGGCAAGGACTTC
ATGTCCAAAACACCAAAAGCAATGGCAACAAAAGACAAAATTGACAAATGGGATCTAATTAAACT
AAAGAGCTTCTGCACAGCAAAAGAAACTACCATCAGAGTGAACAGGCAACCTACAACATGGGAGA

[0690]

AAATTTTTGCAACCTACTCATCTGACAAAGGGCTAATATCCAGAATCTACAATGAACTCAAACAAA
TTTACAAGAAAAAAACAAACAACCCCATCAAAAAGTGGGCGAAGGACATGAACAGACACTTCTCA
AAAGAAGACATTTATGCAGCCAAAAAACACATGAAGAAATGCTCATCATCACTGGCCATCAGAGA
AATGCAAATCAAAACCACTATGAGATATCATCTCACACCAGTTAGAATGGCAATCATTAAAAAGTC
AGGAAACAACAGGTGCTGGAGAGGATGCGGAGAAATAGGAACACTTTTACACTGTTGGTGGGACT
GTAAACTAGTTCAACCATTGTGGAAGTCAGTGTGGCGATTCCTCAGGGATCTAGAACTAGAAATAC
CATTTGACCCAGCCATCCCATTACTGGGTATATACCCAAATGAGTATAAATCATGCTGCTATAAAG
ACACATGCACACGTATGTTTATTGCGGCACTATTCACAATAGCAAAGACTTGGAACCAACCCAAAT
GTCCAACAATGATAGACTGGATTAAGAAAATGTGGCACATATACACCATGGAATACTATGCAGCCA
TAAAAAATGATGAGTTCATATCCTTTGTAGGGACATGGATGAAATTGGAAACCATCATTCTCAGTA
AACTATCGCAAGAACAAAAAACCAAACACCGCATATTCTCACTCATAGGTGGGAATTGAACAATG
AGATCACATGGACACAGGAAGGGGAATATCACACTCTGGGGACTGTGGTGGGGTCGGGGGAGGGG
GGAGGGATAGCATTGGGAGATATACCTAATGCTAGATGACACATTAGTGGGTGCAGCGCACCAGC
ATGGCACATGTATACGGATCCGAATTCTCGACGGATCGATCCGAACAAACGACCCAACACCCGTGC
GTTTTATTCTGTCTTTTTATTGCCGATCCCCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGAT
GCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAA
GCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCGGCCGCTTTACTTGTACA
GCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGACCATGTGATCGCGCT
TCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAGCACGG
GGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCCAGGTGAGTCCAGGAGATGTTTCAGC
ACTGTTGCCTTTAGTCTCGAGGCAACTTAGACAACTGAGTATTGATCTGAGCACAGCAGGGTGTGA
GCTGTTTGAAGATACTGGGGTTGGGGGTGAAGAAACTGCAGAGGACTAACTGGGCTGAGACCCAG
TGGCAATGTTTTAGGGCCTAAGGAATGCCTCTGAAAATCTAGATGGACAACTTTGACTTTGAGAAA
AGAGAGGTGGAAATGAGGAAAATGACTTTTCTTTATTAGATTTCGGTAGAAAGAACTTTCATCTTTC
CCCTATTTTTGTTATTCGTTTTAAAACATCTATCTGGAGGCAGGACAAGTATGGTCATTAAAAAGAT
GCAGGCAGAAGGCATATATTGGCTCAGTCAAAGTGGGGAACTTTGGTGGCCAAACATACATTGCTA
AGGCTATTCCTATATCAGCTGGACACATATAAAATGCTGCTAATGCTTCATTACAAACTTATATCCT
TTAATTCCAGATGGGGGCAAAGTATGTCCAGGGGTGAGGAACAATTGAAACATTTGGGCTGGAGTA
GATTTTGAAAGTCAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGAGAGCGTGTGTTTCTTTTAAC
GTTTTCAGCCTACAGCATACAGGGTTCATGGTGGCAAGAAGATAACAAGATTTAAATTATGGCCAG
TGACTAGTGCTGCAAGAAGAACAACTACCTGCATTTAATGGGAAAGCAAAATCTCAGGCTTTGAGG
GAAGTTAACATAGGCTTGATTCTGGGTGGAAGCTGGGTGTGTAGTTATCTGGAGGCCAGGCTGGAG
CTCTCAGCTCACTATGGGTTCATCTTTATTGTCTCCTTTCATCTCAACAGCTGCACGCTGCCGTCCTC
GATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGATATAGACG
TTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGC
CCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCTTGTAGTT
GCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAA
GTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTCAC
GAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGT
AGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTGGCCGTTTACGTCGCCGTCCAGCTC
GACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACCATGGTGGCGAATTCGAA
GCTTGAGCTCGAGATCTGAGTCCGGTAGCGCTAGCGGATCTGACGGTTCACTAAACCAGCTCTGCT
TATATAGACCTCCCACCGTACACGCCTACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACA
TTTTGGAAAGTCCCGTTGATTTTGGTGCCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTT
GGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTGATGTACTGCCAAAACCGCATCACCATG
GTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGG
GCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATAC
ACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTC
CCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGGTCGTTGGGCGGT
CAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCGGAACTCCATATATGGGCTATGAACTAAT
GACCCCGTAATTGATTACTATTAGCCCGGGGGATCCAGACATGATAAGATACATTGATGAGTTTGG
ACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTA
TTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGG
TTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGCTGATT
ATGATCCGGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGA
GACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGG
GTGTTGGCGGGTGTCGGGGCGCAGCCATGAGGTCGATCGACTCTAGAGGATCGATCCCCGCCCCGG
ACGAACTAAACCTGACTACGACATCTCTGCCCCTTCTTCGCGGGGCAGTGCATGTAATCCCTTCAGT
TGGTTGGTACAACTTGCCAACTGGGCCCTGTTCCACATGTGACACGGGGGGGGACCAAACACAAAG
GGGTTCTCTGACTGTAGTTGACATCCTTATAAATGGATGTGCACATTTGCCAACACTGAGTGGCTTT
CATCCTGGAGCAGACTTTGCAGTCTGTGGACTGCAACACAACATTGCCTTTATGTGTAACTCTTGGC
TGAAGCTCTTACACCAATGCTGGGGGACATGTACCTCCCAGGGGCCCAGGAAGACTACGGGAGGC
TACACCAACGTCAATCAGAGGGGCCTGTGTAGCTACCGATAAGCGGACCCTCAAGAGGGCATTAG

[0691]

CAATAGTGTTTATAAGGCCCCCTTGTTAACCCTAAACGGGTAGCATATGCTTCCCGGGTAGTAGTAT
ATACTATCCAGACTAACCCTAATTCAATAGCATATGTTACCCAACGGGAAGCATATGCTATCGAAT
TAGGGTTAGTAAAAGGGTCCTAAGGAACAGCGATATCTCCCACCCCATGAGCTGTCACGGTTTTAT
TTACATGGGGTCAGGATTCCACGAGGGTAGTGAACCATTTTAGTCACAAGGGCAGTGGCTGAAGAT
CAAGGAGCGGGCAGTGAACTCTCCTGAATCTTCGCCTGCTTCTTCATTCTCCTTCGTTTAGCTAATA
GAATAACTGCTGAGTTGTGAACAGTAAGGTGTATGTGAGGTGCTCGAAAACAAGGTTTCAGGTGAC
GCCCCCAGAATAAAATTTGGACGGGGGGTTCAGTGGTGGCATTGTGCTATGACACCAATATAACCC
TCACAAACCCCTTGGGCAATAAATACTAGTGTAGGAATGAAACATTCTGAATATCTTTAACAATAG
AAATCCATGGGGTGGGGACAAGCCGTAAAGACTGGATGTCCATCTCACACGAATTTATGGCTATGG
GCAACACATAATCCTAGTGCAATATGATACTGGGGTTATTAAGATGTGTCCCAGGCAGGGACCAAG
ACAGGTGAACCATGTTGTTACACTCTATTTGTAACAAGGGGAAAGAGAGTGGACGCCGACAGCAG
CGGACTCCACTGGTTGTCTCTAACACCCCCGAAAATTAAACGGGGCTCCACGCCAATGGGGCCCAT
AAACAAAGACAAGTGGCCACTCTTTTTTTTGAAATTGTGGAGTGGGGGCACGCGTCAGCCCCCACA
CGCCGCCCTGCGGTTTTGGACTGTAAAATAAGGGTGTAATAACTTGGCTGATTGTAACCCCGCTAA
CCACTGCGGTCAAACCACTTGCCCACAAAACCACTAATGGCACCCCGGGGAATACCTGCATAAGTA
GGTGGGCGGGCCAAGATAGGGGCGCGATTGCTGCGATCTGGAGGACAAATTACACACACTTGCGC
CTGAGCGCCAAGCACAGGGTTGTTGGTCCTCATATTCACGAGGTCGCTGAGAGCACGGTGGGCTAA
TGTTGCCATGGGTAGCATATACTACCCAAATATCTGGATAGCATATGCTATCCTAATCTATATCTGG
GTAGCATAGGCTATCCTAATCTATATCTGGGTAGCATATGCTATCCTAATCTATATCTGGGTAGTAT
ATGCTATCCTAATTTATATCTGGGTAGCATAGGCTATCCTAATCTATATCTGGGTAGCATATGCTAT
CCTAATCTATATCTGGGTAGTATATGCTATCCTAATCTGTATCCGGGTAGCATATGCTATCCTAATA
GAGATTAGGGTAGTATATGCTATCCTAATTTATATCTGGGTAGCATATACTACCCAAATATCTGGAT
AGCATATGCTATCCTAATCTATATCTGGGTAGCATATGCTATCCTAATCTATATCTGGGTAGCATAG
GCTATCCTAATCTATATCTGGGTAGCATATGCTATCCTAATCTATATCTGGGTAGTATATGCTATCCT
AATTTATATCTGGGTAGCATAGGCTATCCTAATCTATATCTGGGTAGCATATGCTATCCTAATCTAT
ATCTGGGTAGTATATGCTATCCTAATCTGTATCCGGGTAGCATATGCTATCCTCATGCATATACAGT
CAGCATATGATACCCAGTAGTAGAGTGGGAGTGCTATCCTTTGCATATGCCGCCACCTCCCAAGGG
GGCGTGAATTTTCGCTGCTTGTCCTTTTCCTGCATGCTGGTTGCTCCCATTCTTAGGTGAATTTAAGG
AGGCCAGGCTAAAGCCGTCGCATGTCTGATTGCTCACCAGGTAAATGTCGCTAATGTTTTCCAACG
CGAGAAGGTGTTGAGCGCGGAGCTGAGTGACGTGACAACATGGGTATGCCCAATTGCCCCATGTTG
GGAGGACGAAAATGGTGACAAGACAGATGGCCAGAAATACACCAACAGCACGCATGATGTCTACT
GGGGATTTATTCTTTAGTGCGGGGGAATACACGGCTTTTAATACGATTGAGGGCGTCTCCTAACAA
GTTACATCACTCCTGCCCTTCCTCACCCTCATCTCCATCACCTCCTTCATCTCCGTCATCTCCGTCAT
CACCCTCCGCGGCAGCCCCTTCCACCATAGGTGGAAACCAGGGAGGCAAATCTACTCCATCGTCAA
AGCTGCACACAGTCACCCTGATATTGCAGGTAGGAGCGGGCTTTGTCATAACAAGGTCCTTAATCG
CATCCTTCAAAACCTCAGCAAATATATGAGTTTGTAAAAAGACCATGAAATAACAGACAATGGACT
CCCTTAGCGGGCCAGGTTGTGGGCCGGGTCCAGGGGCCATTCCAAAGGGGAGACGACTCAATGGT
GTAAGACGACATTGTGGAATAGCAAGGGCAGTTCCTCGCCTTAGGTTGTAAAGGGAGGTCTTACTA
CCTCCATATACGAACACACCGGCGACCCAAGTTCCTTCGTCGGTAGTCCTTTCTACGTGACTCCTAG
CCAGGAGAGCTCTTAAACCTTCTGCAATGTTCTCAAATTTCGGGTTGGAACCTCCTTGACCACGATG
CTTTCCAAACCACCCTCCTTTTTTGCGCCTGCCTCCATCACCCTGACCCCGGGGTCCAGTGCTTGGG
CCTTCTCCTGGGTCATCTGCGGGGCCCTGCTCTATCGCTCCCGGGGGCACGTCAGGCTCACCATCTG
GGCCACCTTCTTGGTGGTATTCAAAATAATCGGCTTCCCCTACAGGGTGGAAAAATGGCCTTCTACC
TGGAGGGGGCCTGCGCGGTGGAGACCCGGATGATGATGACTGACTACTGGGACTCCTGGGCCTCTT
TTCTCCACGTCCACGACCTCTCCCCCTGGCTCTTTCACGACTTCCCCCCCTGGCTCTTTCACGTCCTC
TACCCCGGCGGCCTCCACTACCTCCTCGACCCCGGCCTCCACTACCTCCTCGACCCCGGCCTCCACT
GCCTCCTCGACCCCGGCCTCCACCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTCCTGCTCCTG
CCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCT
CCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGC
CCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTC
CTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCT
GCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTGC
TCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTG
CCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCT
CCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCC
CTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCT
GCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTGC
TCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTG
CCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCCCCT
CCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGC
CCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCCCCTC
CTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCCGCTCCTGCTCCTGCTCCTGTTCCACCGTGGGTC
CCTTTGCAGCCAATGCAACTTGGACGTTTTTGGGGGTCTCCGGACACCATCTCTATGTCTTGGCCCTG
ATCCTGAGCCGCCCGGGGCTCCTGGTCTTCCGCCTCCTCGTCCTCGTCCTCTTCCCCGTCCTCGTCCA
TGGTTATCACCCCCTCTTCTTTGAGGTCCACTGCCGCCGGAGCCTTCTGGTCCAGATGTGTCTCCCTT
CTCTCCTAGGCCATTTCCAGGTCCTGTACCTGGCCCCTCGTCAGACATGATTCACACTAAAAGAGAT
CAATAGACATCTTTATTAGACGACGCTCAGTGAATACAGGGAGTGCAGACTCCTGCCCCCTCCAAC
AGCCCCCCCACCCTCATCCCCTTCATGGTCGCTGTCAGACAGATCCAGGTCTGAAAATTCCCCATCC

[0692]

TCCGAACCATCCTCGTCCTCATCACCAATTACTCGCAGCCCGGAAAACTCCCGCTGAACATCCTCAA
GATTTGCGTCCTGAGCCTCAAGCCAGGCCTCAAATTCCTCGTCCCCCTTTTTGCTGGACGGTAGGGA
TGGGGATTCTCGGGACCCCTCCTCTTCCTCTTCAAGGTCACCAGACAGAGATGCTACTGGGGCAAC
GGAAGAAAAGCTGGGTGCGGCCTGTGAGGATCAGCTTATCGATGATAAGCTGTCAAACATGAGAA
TTCTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTT
TCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAA
TACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAA
GGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCT
GTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTG
GGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTT
CCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAG
AGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAA
AGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACA
CTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACA
TGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACG
AGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTAC
TTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTC
TGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCG
CGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGG
GAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGC
ATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTT
AAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGT
TCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGT
AATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCT
ACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTG
TAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCC
TGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT
TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGA
ACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGG
GAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTT
CCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGAT
TTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGT
TCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAAC
CGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCA
GTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCA
CACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCTGTGGAAT
GTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCAT
CTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAG
CATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCG
CCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCG
CCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAA
AGCTTGCATGCCTGCAGGTCGGCCGCCACGACCGGTGCCGCCACCATCCCCTGACCCACGCCCCTG
ACCCCTCACAAGGAGACGACCTTCCATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGA
CGACGTCCCCCGGGCCGTACGCACCCTCGCCGCCGCGTTCGCCGACTACCCCGCCACGCGCCACAC
CGTCGACCCGGACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGG
GCTCGACATCGGCAAGGTGTGGGTCGCGGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGG
AGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCC
CGGCTGGCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGCGTG
GTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCT
CCCCGGAGTGGAGGCGGCCGAGCGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCCGCGCCCCGCA
ACCTCCCCTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGC
GCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACGCCCGCCCCACGACCCGCAGCGCCCGACCG
AAAGGAGCGCACGACCCCATGGCTCCGACCGAAGCCGACCCGGGCGGCCCCGCCGACCCCGCACC
CGCCCCCGAGGCCCACCGACTCTAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACT
TGCTTTAAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTT
AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAA
GCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGGAT
CACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGGGGAGATGG
GGGAGGCTAACTGAAACACGGAAGGAGACAATACCGGAAGGAACCCGCGCTATGACGGCAATAA
AAAGACAGAATAAAACGCACGGGTGTTGGGTCGTTTGTTCATAAACGCGGGGTTCGGTCCCAGGGC
TGGCACTCTGTCGATACCCCACCGAGACCCCATTGGGGCCAATACGCCCGCGTTTCTTCCTTTTCCC
CACCCCACCCCCCAAGTTCGGGTGAAGGCCCAGGGCTCGCAGCCAACGTCGGGGCGGCAGGCCCT
GCCATAGCCACTGGCCCCGTGGGTTAGGGACGGGGTCCCCCATGGGGAATGGTTTATGGTTCGTGG

GGGTTATTATTTTGGGCGTTGCGTGGGGTCTGGTCCACGACTGGACTGAGCAGACAGACCCATGGT
TTTTGGATGGCCTGGGCATGGACCGCATGTACTGGCGCGACACGAACACCGGGCGTCTGTGGCTGC
CAAACACCCCCGACCCCCAAAAACCACCGCGCGGATTTCTGGCGTGCCAAGCTAGTCGACCAATTC
TCATGTTTGACAGCTTATCATCGCAGATCCGGGCAACGTTGTTGCATTGCTGCAGGCGCAGAACTG
GTAGGTATGGAAGATCTCTAGAAGCTGGGTACCAGCTGCTAGCAAGCTTGCTAGCGGCCGGCTCGA
GTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGATG
TCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAAC
GTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTATCCCTATCAGTGATAGA
GAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGGTAGGCGTGTACGGTG
GGAGGCCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCG**(서열번호 37)**

LINE1-GFP mRNA**(서열번호 38)**

TAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGG
AATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAG
TCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAA
CTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTG
CATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCA
GGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGC
GCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGA
GGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGC
GTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGA
AAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCG
GTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGA
GAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGA
CGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGA
AGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGA
TTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAAC
TGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGA
TGTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCATCAACTAATGAGCAAAATCACC
AGCTAACATCATAGTATACATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAAC
GGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTT
TGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGA
AAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGAC
GGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCAT
ACAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAA
ACAGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACA
CCATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAG
CGCACTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACA
TTCTTCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGT
CAAAGTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGC
TGAGAATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACG
ATTATTGGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAG
GATACTACCTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCA
ACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGG
AGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAG
TTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAG
CGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGAT
TGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCG
GGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGA
TACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAG
CGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGC
TGAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAA
AAGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCG
ATACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACA
AGATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTA
TACCTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAG
CTAAGGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGC
CATTTATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAAT
TTACGATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGAC
TGGAACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGG
GCTATTCGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTT
GTTCGCCGATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTT
ATTTCTAACTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACA

[0693]

CAAATAATCGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGA
TAAAGTATCTCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTC
TCCTGAAAGAGATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAA
TCAACATAGTGAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACT
GCCTATGACCTTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGC
AAGAATAGCGAAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAA
GTTGTATTATAAAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCA
GTGGAATCGGACCGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGATAAGCC
AGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGG
CCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTGGA
TCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATA
CAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAA
GATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGT
TAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAG
CAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGG
CAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAG
TGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACCTTACCCCA
GTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGC
ACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTC
TGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGA
ATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCT
AAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTGGCATATTTAT
ACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACCTGGATGAAG
CTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTC
ATTGGTGGTAACGACTACAAAGACGATGACGACAAGTAAAGCGCTTCTAGAAGTTGTCTCCTCCTG
CACTGACTGACTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTT
GTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAAT
GCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCT
GTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGAC
GCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCC
TCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTT
GGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTT
GCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTC
CTTCCCGCTGAGAGACACAAAAAATTCCAACACACTATTGCAATGAAAATAAATTTCCTTTATTAG
CCAGAAGTCAGATGCTCAAGGGGCTTCATGATGTCCCCATAATTTTTGGCAGAGGGAAAAAGATCT
CAGTGGTATTTGTGAGCCAGGGCATTGGCCTTCTGATAGGCAGCCTGCACCTGAGGAGTGCGGCCG
CTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGACC
ATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTGTCGG
GCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCGAGCTGCACGCTGC
CGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGAT
ATAGACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAG
TCGATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCT
TGTAGTTGCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACT
TGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTAGGTCAGGG
TGGTCACGAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGC
TTGCCGTAGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTGGCCGTTTACGTCGCCGT
CCAGCTCGACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACCATGGTGGCGG
GATCTGACGGTTCACTAAACCAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCCATT
TGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTTGGTGCCAAAACAAA
CTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCA
TTGATGTACTGCCAAAACCGCATCACCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAA
GTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACG
TCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAA
TACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTG
ACGTCAATGGGCGGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGG
GCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCT
TTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGACAT
GATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTG
TGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTT**(서열번호 38)**

LINE-1-플라스미드_CD5-인트론-FCR-PI3K**(서열번호 39)**

[0694] CGGCCGCGGGGGGAGGAGCCAAGATGGCCGAATAGGAACAGCTCCGGTCTACAGCTCCCAGCGTG
AGCGACGCAGAAGACGGTGATTTCTGCATTTCCATCTGAGGTACCGGGTTCATCTCACTAGGGAGT

[0695]

GCCAGACAGTGGGCGCAGGCCAGTGTGTGTGCGCACCGTGCGCGAGCCGAAGCAGGGCGAGGCAT
TGCCTCACCTGGGAAGCGCAAGGGGTCAGGGAGTTCCCTTTCCGAGTCAAAGAAAGGGGTGACGG
ACGCACCTGGAAAATCGGGTCACTCCCACCCGAATATTGCGCTTTTCAGACCGGCTTAAGAAACGG
CGCACCACGAGACTATATCCCACACCTGGCTCGGAGGGTCCTACGCCCACGGAATCTCGCTGATTG
CTAGCACAGCAGTCTGAGATCAAACTGCAAGGCGGCAACGAGGCTGGGGGAGGGGCGCCCGCCAT
TGCCCAGGCTTGCTTAGGTAAACAAAGCAGCAGGGAAGCTCGAACTGGGTGGAGCCCACCACAGC
TCAAGGAGGCCTGCCTGCCTCTGTAGGCTCCACCTCTGGGGGCAGGGCACAGACAAACAAAAAGA
CAGCAGTAACCTCTGCAGACTTAAGTGTCCCTGTCTGACAGCTTTGAAGAGAGCAGTGGTTCTCCC
AGCACGCAGCTGGAGATCTGAGAACGGGCAGACTGCCTCCTCAAGTGGGTCCCTGACCCCTGACCC
CCGAGCAGCCTAACTGGGAGGCACCCCCCAGCAGGGGCACACTGACACCTCACACGGCAGGGTAT
TCCAACAGACCTGCAGCTGAGGGTCCTGTCTGTTAGAAGGAAAACTAACAACCAGAAAGGACATCT
ACACCGAAAACCCATCTGTACATCACCATCATCAAAGACCAAAAGTAGATAAAACCACAAAGATG
GGGAAAAAACAGAACAGAAAAACTGGAAACTCTAAAACGCAGAGCGCCTCTCCTCCTCCAAAGGA
ACGCAGTTCCTCACCAGCAACAGAACAAAGCTGGATGGAGAATGATTTTGATGAGCTGAGAGAAG
AAGGCTTCAGACGATCAAATTACTCTGAGCTACGGGAGGACATTCAAACCAAAGGCAAAGAAGTT
GAAAACTTTGAAAAAAATTTAGAAGAATGTATAACTAGAATAACCAATACAGAGAAGTGCTTAAA
GGAGCTGATGGAGCTGAAAACCAAGGCTCGAGAACTACGTGAAGAATGCAGAAGCCTCAGGAGCC
GATGCGATCAACTGGAAGAAAGGGTATCAGCAATGGAAGATGAAATGAATGAAATGAAGCGAGA
AGGGAAGTTTAGAGAAAAAAGAATAAAAAGAAATGAGCAAAGCCTCCAAGAAATATGGGACTAT
GTGAAAAGACCAAATCTACGTCTGATTGGTGTACCTGAAAGTGATGTGGAGAATGGAACCAAGTTG
GAAAACACTCTGCAGGATATTATCCAGGAGAACTTCCCCAATCTAGCAAGGCAGGCCAACGTTCAG
ATTCAGGAAATACAGAGAACGCCACAAAGATACTCCTCGAGAAGAGCAACTCCAAGACACATAAT
TGTCAGATTCACCAAAGTTGAAATGAAGGAAAAAATGTTAAGGGCAGCCAGAGAGAAAGGTCGGG
TTACCCTCAAAGGAAAGCCCATCAGACTAACAGCGGATCTCTCGGCAGAAACCCTACAAGCCAGA
AGAGAGTGGGGGCCAATATTCAACATTCTTAAAGAAAAGAATTTTCAACCCAGAATTTCATATCCA
GCCAAACTAAGCTTCATAAGTGAAGGAGAAATAAAATACTTTATAGACAAGCAAATGTTGAGAGA
TTTTGTCACCACCAGGCCTGCCCTAAAAGAGCTCCTGAAGGAAGCGCTAAACATGGAAAGGAACA
ACCGGTACCAGCCGCTGCAAAATCATGCCAAAATGTAAAGACCATCAAGACTAGGAAGAAACTGC
ATCAACTAATGAGCAAAATCACCAGCTAACATCATAATGACAGGATCAACTTCACACATAACAATA
TTAACTTTAAATATAAATGGACTAAATTCTGCAATTAAAAGACACAGACTGGCAAGTTGGATAAAG
AGTCAAGACCCATCAGTGTGCTGTATTCAGGAAACCCATCTCACGTGCAGAGACACACATAGGCTC
AAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCAA
TCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTAC
ATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTATGCACCCAATACA
GGAGCACCCAGATTCATAAAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTA
ATAATGGGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAA
GGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTCCA
CCCCAAATCAACAGAATATACATTTTTTTCAGCACCACACCACACCTATTCCAAAATTGACCACATA
GTTGGAAGTAAAGCTCTCCTCAGCAAATGTAAAAGAACAGAAATTATAACAAACTATCTCTCAGAC
CACAGTGCAATCAAACTAGAACTCAGGATTAAGAATCTCACTCAAAGCCGCTCAACTACATGGAAA
CTGAACAACCTGCTCCTGAATGACTACTGGGTACATAACGAAATGAAGGCAGAAATAAAGATGTTC
TTTGAAACCAACGAGAACAAAGACACCACATACCAGAATCTCTGGGACGCATTCAAAGCAGTGTGT
AGAGGGAAATTTATAGCACTAAATGCCTACAAGAGAAAGCAGGAAAGATCCAAAATTGACACCCT
AACATCACAATTAAAAGAACTAGAAAAGCAAGAGCAAACACATTCAAAAGCTAGCAGAAGGCAA
GAAATAACTAAAATCAGAGCAGAACTGAAGGAAATAGAGACACAAAAAACCCTTCAAAAAATCAA
TGAATCCAGGAGCTGGTTTTTTGAAAGGATCAACAAAATTGATAGACCGCTAGCAAGACTAATAAA
GAAAAAAAGAGAGAAGAATCAAATAGACACAATAAAAAATGATAAAGGGGATATCACCACCGAT
CCCACAGAAATACAAACTACCATCAGAGAATACTACAAACACCTCTACGCAAATAAACTAGAAAA
TCTAGAAGAAATGGATACATTCCTCGACACATACACTCTCCCAAGACTAAACCAGGAAGAAGTTGA
ATCTCTGAATCGACCAATAACAGGCTCTGAAATTGTGGCAATAATCAATAGTTTACCAACCAAAAA
GAGTCCAGGACCAGATGGATTCACAGCCGAATTCTACCAGAGGTACAAGGAGGAACTGGTACCAT
TCCTTCTGAAACTATTCCAATCAATAGAAAAAGAGGGAATCCTCCCTAACTCATTTTATGAGGCCA
GCATCATTCTGATACCAAAGCCGGGCAGAGACACAACCAAAAAAGAGAATTTTAGACCAATATCCT
TGATGAACATTGATGCAAAAATCCTCAATAAAATACTGGCAAACCGAATCCAGCAGCACATCAAA
AAGCTTATCCACCATGATCAAGTGGGCTTCATCCCTGGGATGCAAGGCTGGTTCAATATACGCAAA
TCAATAAATGTAATCCAGCATATAAACAGAGCCAAAGACAAAAACCACATGATTATCTCAATAGAT
GCAGAAAAAGCCTTTGACAAAATTCAACAACCCTTCATGCTAAAAACTCTCAATAAATTAGGTATT
GATGGGACGTATTTCAAAATAATAAGAGCTATCTATGACAAACCCACAGCCAATATCATACTGAAT
GGGCAAAAACTGGAAGCATTCCCTTTGAAAACCGGCACAAGACAGGGATGCCCTCTCTCACCGCTC
CTATTCAACATAGTGTTGGAAGTTCTGGCCAGGGCAATCAGGCAGGAGAAGGAAATAAAGGGTAT
TCAATTAGGAAAAGAGGAAGTCAAATTGTCCCTGTTTGCAGACGACATGATTGTTTATCTAGAAAA
CCCCATCGTCTCAGCCCAAAATCTCCTTAAGCTGATAAGCAACTTCAGCAAAGTCTCAGGATACAA
AATCAATGTACAAAAATCACAAGCATTCTTATACACCAACAACAGACAAACAGAGAGCCAAATCA

[0696]

TGGGTGAACTCCCATTCACAATTGCTTCAAAGAGAATAAAATACCTAGGAATCCAACTTACAAGGG
ATGTGAAGGACCTCTTCAAGGAGAACTACAAACCACTGCTCAAGGAAATAAAAGAGGAGACAAAC
AAATGGAAGAACATTCCATGCTCATGGGTAGGAAGAATCAATATCGTGAAAATGGCCATACTGCCC
AAGGTAATTTACAGATTCAATGCCATCCCCATCAAGCTACCAATGACTTTCTTCACAGAATTGGAA
AAAACTACTTTAAAGTTCATATGGAACCAAAAAAGAGCCCGCATTGCCAAGTCAATCCTAAGCCAA
AAGAACAAAGCTGGAGGCATCACACTACCTGACTTCAAACTATACTACAAGGCTACAGTAACCAA
AACAGCATGGTACTGGTACCAAAACAGAGATATAGATCAATGGAACAGAACAGAGCCCTCAGAAA
TAATGCCGCATATCTACAACTATCTGATCTTTGACAAACCTGAGAAAAACAAGCAATGGGGAAAGG
ATTCCCTATTTAATAAATGGTGCTGGGAAAACTGGCTAGCCATATGTAGAAAGCTGAAACTGGATC
CCTTCCTTACACCTTATACAAAAATCAATTCAAGATGGATTAAAGATTTAAACGTTAAACCTAAAA
CCATAAAAAACCCTAGAAGAAAACCTAGGCATTACCATTCAGGACATAGGCGTGGGCAAGGACTTC
ATGTCCAAAACACCAAAAGCAATGGCAACAAAAGACAAAATTGACAAATGGGATCTAATTAAACT
AAAGAGCTTCTGCACAGCAAAAGAAACTACCATCAGAGTGAACAGGCAACCTACAACATGGGAGA
AAATTTTTGCAACCTACTCATCTGACAAAGGGCTAATATCCAGAATCTACAATGAACTCAAACAAA
TTTACAAGAAAAAAACAAACAACCCCATCAAAAAGTGGGCGAAGGACATGAACAGACACTTCTCA
AAAGAAGACATTTATGCAGCCAAAAAACACATGAAGAAATGCTCATCATCACTGGCCATCAGAGA
AATGCAAATCAAAACCACTATGAGATATCATCTCACACCAGTTAGAATGGCAATCATTAAAAAGTC
AGGAAACAACAGGTGCTGGAGAGGATGCGGAGAAATAGGAACACTTTTACACTGTTGGTGGGACT
GTAAACTAGTTCAACCATTGTGGAAGTCAGTGTGGCGATTCCTCAGGGATCTAGAACTAGAAATAC
CATTTGACCCAGCCATCCCATTACTGGGTATATACCCAAATGAGTATAAATCATGCTGCTATAAAG
ACACATGCACACGTATGTTTATTGCGGCACTATTCACAATAGCAAAGACTTGGAACCAACCCAAAT
GTCCAACAATGATAGACTGGATTAAGAAAATGTGGCACATATACACCATGGAATACTATGCAGCCA
TAAAAAATGATGAGTTCATATCCTTTGTAGGGACATGGATGAAATTGGAAACCATCATTCTCAGTA
AACTATCGCAAGAACAAAAAACCAAACACCGCATATTCTCACTCATAGGTGGGAATTGAACAATG
AGATCACATGGACACAGGAAGGGGAATATCACACTCTGGGGACTGTGGTGGGGTCGGGGGAGGGG
GGAGGGATAGCATTGGGAGATATACCTAATGCTAGATGACACATTAGTGGGTGCAGCGCACCAGC
ATGGCACATGTATACGGATCCGAATTCTCGACGGATCGATCCGAACAAACGACCCAACACCCGTGC
GTTTTATTCTGTCTTTTTATTGCCGATCCCCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGAT
GCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAA
GCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCGGCCGCTCATGTTCTCGT
AGGAGTCGGCGTCCTCTTCGTGGTTAGGTCCAGGTTGGCCTCTGATAGACCGCAGCTGAGGAGCGG
CGTACAGAATGCCTCTCATGTCCTCATAGCTGCCGCTGCCTTGTGGAGGCTTCTCGTGCTTCAGTGT
CTCGTATGTCTCTTGATTCCGGGTGCTCAGGCCGGTGTACACGCCATCAGATTTCTCGTAGCTGGTG
ATGGCGGCCTTCCGCACTTGGATCTTCAGCCGTCTGCAGTACAGGGTGATGACCAGAGACAGCAGC
AGGACACCACATGTGCCAGCCAGAGGGGCCCAAATGTAGATATCCAGGCCTCTGGTATGCACAGCT
CCGCCTGCAGCAGGTCTACAGGCTTCAGGTCTGAGAGACAGAGGCTGGCTGGCGATTGTAGGAGCT
GGTGTAGGTGGTCTAGGAGCGGGTGTTGTTGTAGGCTTGGCGGGCAGAAACACGGGCACGAAGTG
GCTGAAGTACATGATGCTATTGCTCAGGGCTCCGCTTCCTCCGCCGCCTGATTTGATTTCCAGCTTG
GTGCCTCCGCCAAATGTCCAAGGGCTCTCGTCGTACTGCTGGCAGTAGTAGATGCCGAAGTCCTCG
TACTGCAGGCTGCTGATTGTCAGGGTGTAGTCGGTGCCAGAGCCGCTGCCAGAAAATCTGCTTGGC
ACGCCGCTTTCCAGTCTGTTGGCCCGGTAGATCAGTGTCTTAGGGGCCTTGCCAGGCTTCTGCTGGA
ACCAGCTCAGGTAGCTGTTGATGTCCTGGCTGGCTCTACAGGTGATGGTCACTCTATCGCCCACAGA
GGCAGACAGGCTGCTAGGGCTCTGTGTCATCTGGATATCAGAGCCACCACCGCCAGATCCACCGCC
ACCTGATCCTCCGCCTCCGCTAGAAACTGTCACTGTGGTGCCCTGGCCCCACACATCGAAGTACCA
GTCGTAGCCTCTTCTGGTGCAGAAGTACACGGCGGTATCCTCGGCTCTCAGGCTGTTGATCTGCAGG
TAGGCGGTGTTCTTGCTGTCGTCCAGGCTGAAGGTGAATCTGCCCTTAAAGCTATCGGCGTAGGTTG
GCTCGCCGGTGTGGGTATTGATCCAGCCCATCCACTCAAGGCCAGGTGAGTCCAGGAGATGTTTCA
GCACTGTTGCCTTTAGTCTCGAGGCAACTTAGACAACTGAGTATTGATCTGAGCACAGCAGGGTGT
GAGCTGTTTGAAGATACTGGGGTTGGGGGTGAAGAAACTGCAGAGGACTAACTGGGCTGAGACCC
AGTGGCAATGTTTTAGGGCCTAAGGAATGCCTCTGAAAATCTAGATGGACAACTTTGACTTTGAGA
AAAGAGAGGTGGAAATGAGGAAAATGACTTTTCTTTATTAGATTTCGGTAGAAAGAACTTTCATCT
TTCCCCTATTTTTGTTATTCGTTTTAAAACATCTATCTGGAGGCAGGACAAGTATGGTCATTAAAAA
GATGCAGGCAGAAGGCATATATTGGCTCAGTCAAAGTGGGGAACTTTGGTGGCCAAACATACATTG
CTAAGGCTATTCCTATATCAGCTGGACACATATAAAATGCTGCTAATGCTTCATTACAAACTTATAT
CCTTTAATTCCAGATGGGGGCAAAGTATGTCCAGGGGTGAGGAACAATTGAAACATTTGGGCTGGA
GTAGATTTTGAAAGTCAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGAGAGCGTGTGTTTCTTTT
AACGTTTTCAGCCTACAGCATACAGGGTTCATGGTGGCAAGAAGATAACAAGATTTAAATTATGGC
CAGTGACTAGTGCTGCAAGAAGAACAACTACCTGCATTTAATGGGAAAGCAAAATCTCAGGCTTTG
AGGGAAGTTAACATAGGCTTGATTCTGGGTGGAAGCTGGGTGTGTAGTTATCTGGAGGCCAGGCTG
GAGCTCTCAGCTCACTATGGGTTCATCTTTATTGTCTCCTTTTTCCAGGGGCCTGTCGGACCCAGTTC
ATGCCGTAGTTGGTGAAGGTGTAGCCGCTGGCGGCACAGCTGATTCTGACAGATCCGCCAGGTTTC
ACAAGTCCGCCGCCAGACTGAACCAGCTGGATCTCAGAGATGCTACAGGCCACTGTTCCCAGCAGC
AGCAGAGACTGCAGCCACATCTGGTGGCGAATTCGAAGCTTGAGCTCGAGATCTGAGTCCGGTAGC

[0697]

GCTAGCGGATCTGACGGTTCACTAAACCAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACC
GCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTTGGTGCCAA
AACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCA
CGCCCATTGATGTACTGCCAAAACCGCATCACCATGGTAATAGCGATGACTAATACGTAGATGTAC
TGCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCA
TTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTAC
CGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTC
ATTATTGACGTCAATGGGCGGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTAT
GTAACGCGGAACTCCATATATGGGCTATGAACTAATGACCCCGTAATTGATTACTATTAGCCCGGG
GGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAA
AATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACA
AGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAA
AGCAAGTAAAACCTCTACAAATGTGGTATGGCTGATTATGATCCGGCTGCCTCGCGCGTTTCGGTG
ATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATG
CCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATG
AGGTCGATCGACTCTAGAGGATCGATCCCCGCCCCGGACGAACTAAACCTGACTACGACATCTCTG
CCCCTTCTTCGCGGGGCAGTGCATGTAATCCCTTCAGTTGGTTGGTACAACTTGCCAACTGGGCCCT
GTTCCACATGTGACACGGGGGGGGACCAAACACAAAGGGGTTCTCTGACTGTAGTTGACATCCTTA
TAAATGGATGTGCACATTTGCCAACACTGAGTGGCTTTCATCCTGGAGCAGACTTTGCAGTCTGTGG
ACTGCAACACAACATTGCCTTTATGTGTAACTCTTGGCTGAAGCTCTTACACCAATGCTGGGGGACA
TGTACCTCCCAGGGGCCCAGGAAGACTACGGGAGGCTACACCAACGTCAATCAGAGGGGCCTGTG
TAGCTACCGATAAGCGGACCCTCAAGAGGGCATTAGCAATAGTGTTTATAAGGCCCCCTTGTTAAC
CCTAAACGGGTAGCATATGCTTCCCGGGTAGTAGTATATACTATCCAGACTAACCCTAATTCAATA
GCATATGTTACCCAACGGGAAGCATATGCTATCGAATTAGGGTTAGTAAAAGGGTCCTAAGGAACA
GCGATATCTCCCACCCCATGAGCTGTCACGGTTTTATTTACATGGGGTCAGGATTCCACGAGGGTAG
TGAACCATTTTAGTCACAAGGGCAGTGGCTGAAGATCAAGGAGCGGGCAGTGAACTCTCCTGAATC
TTCGCCTGCTTCTTCATTCTCCTTCGTTTAGCTAATAGAATAACTGCTGAGTTGTGAACAGTAAGGT
GTATGTGAGGTGCTCGAAAACAAGGTTTCAGGTGACGCCCCCAGAATAAAATTTGGACGGGGGGTT
CAGTGGTGGCATTGTGCTATGACACCAATATAACCCTCACAAACCCCTTGGGCAATAAATACTAGT
GTAGGAATGAAACATTCTGAATATCTTTAACAATAGAAATCCATGGGGTGGGGACAAGCCGTAAA
GACTGGATGTCCATCTCACACGAATTTATGGCTATGGGCAACACATAATCCTAGTGCAATATGATA
CTGGGGTTATTAAGATGTGTCCCAGGCAGGGACCAAGACAGGTGAACCATGTTGTTACACTCTATT
TGTAACAAGGGGAAAGAGAGTGGACGCCGACAGCAGCGGACTCCACTGGTTGTCTCTAACACCCC
CGAAAATTAAACGGGGCTCCACGCCAATGGGGCCCATAAACAAAGACAAGTGGCCACTCTTTTTTT
TGAAATTGTGGAGTGGGGGCACGCGTCAGCCCCCACACGCCGCCCTGCGGTTTTGGACTGTAAAAT
AAGGGTGTAATAACTTGGCTGATTGTAACCCCGCTAACCACTGCGGTCAAACCACTTGCCCACAAA
ACCACTAATGGCACCCCGGGGAATACCTGCATAAGTAGGTGGGCGGGCCAAGATAGGGGCGCGAT
TGCTGCGATCTGGAGGACAAATTACACACACTTGCGCCTGAGCGCCAAGCACAGGGTTGTTGGTCC
TCATATTCACGAGGTCGCTGAGAGCACGGTGGGCTAATGTTGCCATGGGTAGCATATACTACCCAA
ATATCTGGATAGCATATGCTATCCTAATCTATATCTGGGTAGCATAGGCTATCCTAATCTATATCTG
GGTAGCATATGCTATCCTAATCTATATCTGGGTAGTATATGCTATCCTAATTTATATCTGGGTAGCA
TAGGCTATCCTAATCTATATCTGGGTAGCATATGCTATCCTAATCTATATCTGGGTAGTATATGCTA
TCCTAATCTGTATCCGGGTAGCATATGCTATCCTAATAGAGATTAGGGTAGTATATGCTATCCTAAT
TTATATCTGGGTAGCATATACTACCCAAATATCTGGATAGCATATGCTATCCTAATCTATATCTGGG
TAGCATATGCTATCCTAATCTATATCTGGGTAGCATAGGCTATCCTAATCTATATCTGGGTAGCATA
TGCTATCCTAATCTATATCTGGGTAGTATATGCTATCCTAATTTATATCTGGGTAGCATAGGCTATC
CTAATCTATATCTGGGTAGCATATGCTATCCTAATCTATATCTGGGTAGTATATGCTATCCTAATCT
GTATCCGGGTAGCATATGCTATCCTCATGCATATACAGTCAGCATATGATACCCAGTAGTAGAGTG
GGAGTGCTATCCTTTGCATATGCCGCCACCTCCCAAGGGGGCGTGAATTTTCGCTGCTTGTCCTTTT
CCTGCATGCTGGTTGCTCCCATTCTTAGGTGAATTTAAGGAGGCCAGGCTAAAGCCGTCGCATGTCT
GATTGCTCACCAGGTAAATGTCGCTAATGTTTTCCAACGCGAGAAGGTGTTGAGCGCGGAGCTGAG
TGACGTGACAACATGGGTATGCCCAATTGCCCCATGTTGGGAGGACGAAAATGGTGACAAGACAG
ATGGCCAGAAATACACCAACAGCACGCATGATGTCTACTGGGGATTTATTCTTTAGTGCGGGGGAA
TACACGGCTTTTAATACGATTGAGGGCGTCTCCTAACAAGTTACATCACTCCTGCCCTTCCTCACCC
TCATCTCCATCACCTCCTTCATCTCCGTCATCTCCGTCATCACCCTCCGCGGCAGCCCCTTCCACCAT
AGGTGGAAACCAGGGAGGCAAATCTACTCCATCGTCAAAGCTGCACACAGTCACCCTGATATTGCA
GGTAGGAGCGGGCTTTGTCATAACAAGGTCCTTAATCGCATCCTTCAAAACCTCAGCAAATATATG
AGTTTGTAAAAAGACCATGAAATAACAGACAATGGACTCCCTTAGCGGGCCAGGTTGTGGGCCGG
GTCCAGGGGCCATTCCAAAGGGGAGACGACTCAATGGTGTAAGACGACATTGTGGAATAGCAAGG
GCAGTTCCTCGCCTTAGGTTGTAAAGGGAGGTCTTACTACCTCCATATACGAACACACCGGCGACCC
CAAGTTCCTTCGTCGGTAGTCCTTTCTACGTGACTCCTAGCCAGGAGAGCTCTTAAACCTTCTGCAA
TGTTCTCAAATTTCGGGTTGGAACCTCCTTGACCACGATGCTTTCCAAACCACCCTCCTTTTTTGCGC
CTGCCTCCATCACCCTGACCCCGGGGTCCAGTGCTTGGGCCTTCTCCTGGGTCATCTGCGGGGCCCT

[0698]

GCTCTATCGCTCCCGGGGGCACGTCAGGCTCACCATCTGGGCCACCTTCTTGGTGGTATTCAAAATA
ATCGGCTTCCCCTACAGGGTGGAAAAATGGCCTTCTACCTGGAGGGGGCCTGCGCGGTGGAGACCC
GGATGATGATGACTGACTACTGGGACTCCTGGGCCTCTTTTCTCCACGTCCACGACCTCTCCCCCTG
GCTCTTTCACGACTTCCCCCCCTGGCTCTTTCACGTCCTCTACCCCGGCGGCCTCCACTACCTCCTCG
ACCCCGGCCTCCACTACCTCCTCGACCCCGGCCTCCACTGCCTCCTCGACCCCGGCCTCCACCTCCT
GCTCCTGCCCCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCC
TCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCC
CTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCT
CCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCT
CCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCT
TCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTC
CTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCT
CCTCCTGCTCCTGCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGC
CCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTCCTGCTCC
TGCCCCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCC
CTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCC
CCTCCCGCTCCTGCTCCTGCTCCTGTTCCACCGTGGGTCCCTTTGCAGCCAATGCAACTTGGACGTTT
TTGGGGTCTCCGGACACCATCTCTATGTCTTGGCCCTGATCCTGAGCCGCCCGGGGCTCCTGGTCTT
CCGCCTCCTCGTCCTCGTCCTCTTCCCCGTCCTCGTCCATGGTTATCACCCCCTCTTCTTTGAGGTCC
ACTGCCGCCGGAGCCTTCTGGTCCAGATGTGTCTCCCTTCTCTCCTAGGCCATTTCCAGGTCCTGTA
CCTGGCCCCTCGTCAGACATGATTCACACTAAAAGAGATCAATAGACATCTTTATTAGACGACGCT
CAGTGAATACAGGGAGTGCAGACTCCTGCCCCCTCCAACAGCCCCCCCACCCTCATCCCCTTCATG
GTCGCTGTCAGACAGATCCAGGTCTGAAAATTCCCCATCCTCCGAACCATCCTCGTCCTCATCACCA
ATTACTCGCAGCCCGGAAAACTCCCGCTGAACATCCTCAAGATTTGCGTCCTGAGCCTCAAGCCAG
GCCTCAAATTCCTCGTCCCCCTTTTTGCTGGACGGTAGGGATGGGGATTCTCGGGACCCCTCCTCTT
CCTCTTCAAGGTCACCAGACAGAGATGCTACTGGGGCAACGGAAGAAAAGCTGGGTGCGGCCTGT
GAGGATCAGCTTATCGATGATAAGCTGTCAAACATGAGAATTCTTGAAGACGAAAGGGCCTCGTGA
TACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCG
GGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATG
AGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTC
CGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGT
GAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACA
GCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCT
GCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTA
TTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGT
AAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAAC
GATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGA
TCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGCAG
CAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAAT
TAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCT
GGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGC
CAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAAC
GAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTT
ACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCT
TTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTA
GAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAA
AACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAC
TGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTC
AAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTG
GCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGG
GCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATAC
CTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGT
AAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT
ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCG
GAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCT
CACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA
TACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCC
TGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTAC
AATCTGCTCTGATGCCGCATAGTTAAGCCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTC
CCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGG
AAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAT
AGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCAT
GGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGT
AGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTGCATGCCTGCAGGTCGGCCGCC

ACGACCGGTGCCGCCACCATCCCCTGACCCACGCCCCTGACCCCTCACAAGGAGACGACCTTCCAT
GACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCACCCT
CGCCGCCGCGTTCGCCGACTACCCCGCCACGCGCCACACCGTCGACCCGGACCGCCACATCGAGCG
GGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGC
GGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGTTCG
CCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAA
GGCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCGCCC
GACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGAGCGCGC
CGGGGTGCCCGCCTTCCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTCTACGAGCGGCTCGGCTTC
ACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGG
TGCCTGACGCCCGCCCCACGACCCGCAGCGCCCGACCGAAAGGAGCGCACGACCCCATGGCTCCG
ACCGAAGCCGACCCGGGCGGCCCGCCGACCCCGCACCCGCCCCCGAGGCCCACCGACTCTAGAG
GATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCC
CTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTT
ACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGG
TTTGTCCAAACTCATCAATGTATCTTATCATGTCTGGATCACTCGCCGATAGTGGAAACCGACGCCC
CAGCACTCGTCCGAGGGCAAAGGAATAGGGGAGATGGGGGAGGCTAACTGAAACACGGAAGGAG
ACAATACCGGAAGGAACCCGCGCTATGACGGCAATAAAAAGACAGAATAAAACGCACGGGTGTTG
GGTCGTTTGTTCATAAACGCGGGGTTCGGTCCCAGGGCTGGCACTCTGTCGATACCCCACCGAGAC
CCCATTGGGGCCAATACGCCCGCGTTTCTTCCTTTTCCCCACCCCACCCCCCAAGTTCGGGTGAAGG
CCCAGGGCTCGCAGCCAACGTCGGGGCGGCAGGCCCTGCCATAGCCACTGGCCCCGTGGGTTAGG
GACGGGGTCCCCCATGGGGAATGGTTTATGGTTCGTGGGGGTTATTATTTTGGGCGTTGCGTGGGGT
CTGGTCCACGACTGGACTGAGCAGACAGACCCATGGTTTTTGGATGGCCTGGGCATGGACCGCATG
TACTGGCGCGACACGAACACCGGGCGTCTGTGGCTGCCAAACACCCCCGACCCCCAAAAACCACC
GCGCGGATTTCTGGCGTGCCAAGCTAGTCGACCAATTCTCATGTTTGACAGCTTATCATCGCAGATC
CGGGCAACGTTGTTGCATTGCTGCAGGCGCAGAACTGGTAGGTATGGAAGATCTCTAGAAGCTGGG
TACCAGCTGCTAGCAAGCTTGCTAGCGGCCGGCTCGAGTTTACTCCCTATCAGTGATAGAGAACGT
ATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGATGTCGAGTTTACTCCCTATCAGTGATAGAGA
ACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGAT
AGAGAACGTATGTCGAGTTTATCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAG
TGATAGAGAACGTATGTCGAGGTAGGCGTGTACGGTGGGAGGCCTATATAAGCAGAGCTCGTTTAG
TGAACCGTCAGATCGCCG**(서열번호 39)**

LINE-1 플라스미드-CD5_FCR-PI3K_T2A-GFP 인트론**(서열번호 40)**
CGGCCGCGGGGGGAGGAGCCAAGATGGCCGAATAGGAACAGCTCCGGTCTACAGCTCCCAGCGTG
AGCGACGCAGAAGACGGTGATTTCTGCATTTCCATCTGAGGTACCGGGTTCATCTCACTAGGGAGT
GCCAGACAGTGGGCGCAGGCCAGTGTGTGTGCGCACCGTGCGCGAGCCGAAGCAGGGCGAGGCAT
TGCCTCACCTGGGAAGCGCAAGGGGTCAGGGAGTTCCCTTTCCGAGTCAAAGAAAGGGGTGACGG
ACGCACCTGGAAAATCGGGTCACTCCCACCCGAATATTGCGCTTTTCAGACCGGCTTAAGAAACGG
CGCACCACGAGACTATATCCCACACCTGGCTCGGAGGGTCCTACGCCCACGGAATCTCGCTGATTG
CTAGCACAGCAGTCTGAGATCAAACTGCAAGGCGGCAACGAGGCTGGGGGAGGGGCGCCCGCCAT
TGCCCAGGCTTGCTTAGGTAAACAAAGCAGCAGGGAAGCTCGAACTGGGTGGAGCCCACCACAGC
TCAAGGAGGCCTGCCTGCCTCTGTAGGCTCCACCTCTGGGGGCAGGGCACAGACAAACAAAAAGA
CAGCAGTAACCTCTGCAGACTTAAGTGTCCCTGTCTGACAGCTTTGAAGAGAGCAGTGGTTCTCCC
AGCACGCAGCTGGAGATCTGAGAACGGGCAGACTGCCTCCTCAAGTGGGTCCCTGACCCCTGACCC
CCGAGCAGCCTAACTGGGAGGCACCCCCCAGCAGGGGCACACTGACACCTCACACGGCAGGGTAT
TCCAACAGACCTGCAGCTGAGGGTCCTGTCTGTTAGAAGGAAAACTAACAACCAGAAAGGACATCT
ACACCGAAAACCCATCTGTACATCACCATCATCAAAGACCAAAAGTAGATAAAACCACAAAGATG
GGGAAAAAACAGAACAGAAAAACTGGAAACTCTAAAACGCAGAGCGCCTCTCCTCCTCCAAAGGA
ACGCAGTTCCTCACCAGCAACAGAACAAAGCTGGATGGAGAATGATTTTGATGAGCTGAGAGAAG
AAGGCTTCAGACGATCAAATTACTCTGAGCTACGGGAGGACATTCAAACCAAAGGCAAAGAAGTT
GAAAACTTTGAAAAAAATTTAGAAGAATGTATAACTAGAATAACCAATACAGAGAAGTGCTTAAA
GGAGCTGATGGAGCTGAAAACCAAGGCTCGAGAACTACGTGAAGAATGCAGAAGCCTCAGGAGCC
GATGCGATCAACTGGAAGAAAGGGTATCAGCAATGGAAGATGAAATGAATGAAATGAAGCGAGA
AGGGAAGTTTAGAGAAAAAAGAATAAAAAGAAATGAGCAAAGCCTCCAAGAAATATGGGACTAT
GTGAAAAGACCAAATCTACGTCTGATTGGTGTACCTGAAAGTGATGTGGAGAATGGAACCAAGTTG
GAAAACACTCTGCAGGATATTATCCAGGAGAACTTCCCCAATCTAGCAAGGCAGGCCAACGTTCAG
ATTCAGGAAATACAGAGAACGCCACAAAGATACTCCTCGAGAAGAGCAACTCCAAGACACATAAT
TGTCAGATTCACCAAAGTTGAAATGAAGGAAAAAATGTTAAGGGCAGCCAGAGAGAAAGGTCGGG
TTACCCTCAAAGGAAAGCCCATCAGACTAACAGCGGATCTCTCGGCAGAAACCCTACAAGCCAGA
AGAGAGTGGGGGCCAATATTCAACATTCTTAAAGAAAAGAATTTTCAACCCAGAATTTCATATCCA
GCCAAACTAAGCTTCATAAGTGAAGGAGAAATAAAATACTTTATAGACAAGCAAATGTTGAGAGA
TTTTGTCACCACCAGGCCTGCCCTAAAAGAGCTCCTGAAGGAAGCGCTAAACATGGAAAGGAACA

[0699]

[0700]

ACCGGTACCAGCCGCTGCAAAATCATGCCAAAATGTAAAGACCATCAAGACTAGGAAGAAACTGC
ATCAACTAATGAGCAAAATCACCAGCTAACATCATAATGACAGGATCAACTTCACACATAACAATA
TTAACTTTAAATATAAATGGACTAAATTCTGCAATTAAAAGACACAGACTGGCAAGTTGGATAAAG
AGTCAAGACCCATCAGTGTGCTGTATTCAGGAAACCCATCTCACGTGCAGAGACACACATAGGCTC
AAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCAA
TCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTAC
ATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTATGCACCCAATACA
GGAGCACCCAGATTCATAAAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTA
ATAATGGGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAA
GGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTCCA
CCCCAAATCAACAGAATATACATTTTTTTCAGCACCACACCACACCTATTCCAAAATTGACCACATA
GTTGGAAGTAAAGCTCTCCTCAGCAAATGTAAAAGAACAGAAATTATAACAAACTATCTCTCAGAC
CACAGTGCAATCAAACTAGAACTCAGGATTAAGAATCTCACTCAAAGCCGCTCAACTACATGGAAA
CTGAACAACCTGCTCCTGAATGACTACTGGGTACATAACGAAATGAAGGCAGAAATAAAGATGTTC
TTTGAAACCAACGAGAACAAAGACACCACATACCAGAATCTCTGGGACGCATTCAAAGCAGTGTGT
AGAGGGAAATTTATAGCACTAAATGCCTACAAGAGAAAGCAGGAAAGATCCAAAATTGACACCCT
AACATCACAATTAAAAGAACTAGAAAAGCAAGAGCAAACACATTCAAAAGCTAGCAGAAGGCAA
GAAATAACTAAAATCAGAGCAGAACTGAAGGAAATAGAGACACAAAAAACCCTTCAAAAAATCAA
TGAATCCAGGAGCTGGTTTTTTGAAAGGATCAACAAAATTGATAGACCGCTAGCAAGACTAATAAA
GAAAAAAAGAGAGAAGAATCAAATAGACACAATAAAAAATGATAAAGGGGATATCACCACCGAT
CCCACAGAAATACAAACTACCATCAGAGAATACTACAAACACCTCTACGCAAATAAACTAGAAAA
TCTAGAAGAAATGGATACATTCCTCGACACATACACTCTCCCAAGACTAAACCAGGAAGAAGTTGA
ATCTCTGAATCGACCAATAACAGGCTCTGAAATTGTGGCAATAATCAATAGTTTACCAACCAAAAA
GAGTCCAGGACCAGATGGATTCACAGCCGAATTCTACCAGAGGTACAAGGAGGAACTGGTACCAT
TCCTTCTGAAACTATTCCAATCAATAGAAAAAGAGGGAATCCTCCCTAACTCATTTTATGAGGCCA
GCATCATTCTGATACCAAAGCCGGGCAGAGACACAACCAAAAAAGAGAATTTTAGACCAATATCCT
TGATGAACATTGATGCAAAAATCCTCAATAAAATACTGGCAAACCGAATCCAGCAGCACATCAAA
AAGCTTATCCACCATGATCAAGTGGGCTTCATCCCTGGGATGCAAGGCTGGTTCAATATACGCAAA
TCAATAAATGTAATCCAGCATATAAACAGAGCCAAAGACAAAAACCACATGATTATCTCAATAGAT
GCAGAAAAAGCCTTTGACAAAATTCAACAACCCTTCATGCTAAAAACTCTCAATAAATTAGGTATT
GATGGGACGTATTTCAAAATAATAAGAGCTATCTATGACAAACCCACAGCCAATATCATACTGAAT
GGGCAAAAACTGGAAGCATTCCCTTTGAAAACCGGCACAAGACAGGGATGCCCTCTCTCACCGCTC
CTATTCAACATAGTGTTGGAAGTTCTGGCCAGGGCAATCAGGCAGGAGAAGGAAATAAAGGGTAT
TCAATTAGGAAAAGAGGAAGTCAAATTGTCCCTGTTTGCAGACGACATGATTGTTTATCTAGAAAA
CCCCATCGTCTCAGCCCAAAATCTCCTTAAGCTGATAAGCAACTTCAGCAAAGTCTCAGGATACAA
AATCAATGTACAAAAATCACAAGCATTCTTATACACCAACAACAGACAAACAGAGAGCCAAATCA
TGGGTGAACTCCCATTCACAATTGCTTCAAAGAGAATAAAATACCTAGGAATCCAACTTACAAGGG
ATGTGAAGGACCTCTTCAAGGAGAACTACAAACCACTGCTCAAGGAAATAAAAGAGGAGACAAAC
AAATGGAAGAACATTCCATGCTCATGGGTAGGAAGAATCAATATCGTGAAAATGGCCATACTGCCC
AAGGTAATTTACAGATTCAATGCCATCCCCATCAAGCTACCAATGACTTTCTTCACAGAATTGGAA
AAAACTACTTTAAAGTTCATATGGAACCAAAAAAGAGCCCGCATTGCCAAGTCAATCCTAAGCCAA
AAGAACAAAGCTGGAGGCATCACACTACCTGACTTCAAACTATACTACAAGGCTACAGTAACCAA
AACAGCATGGTACTGGTACCAAAACAGAGATATAGATCAATGGAACAGAACAGAGCCCTCAGAAA
TAATGCCGCATATCTACAACTATCTGATCTTTGACAAACCTGAGAAAAACAAGCAATGGGGAAAGG
ATTCCCTATTTAATAAATGGTGCTGGGAAAACTGGCTAGCCATATGTAGAAAGCTGAAACTGGATC
CCTTCCTTACACCTTATACAAAAATCAATTCAAGATGGATTAAAGATTTAAACGTTAAACCTAAAA
CCATAAAAACCCTAGAAGAAAACCTAGGCATTACCATTCAGGACATAGGCGTGGGCAAGGACTTC
ATGTCCAAAACACCAAAAGCAATGGCAACAAAAGACAAAATTGACAAATGGGATCTAATTAAACT
AAAGAGCTTCTGCACAGCAAAAGAAACTACCATCAGAGTGAACAGGCAACCTACAACATGGGAGA
AAATTTTTGCAACCTACTCATCTGACAAAGGGCTAATATCCAGAATCTACAATGAACTCAAACAAA
TTTACAAGAAAAAAACAAACAACCCCATCAAAAAGTGGGCGAAGGACATGAACAGACACTTCTCA
AAAGAAGACATTTATGCAGCCAAAAAACACATGAAGAAATGCTCATCATCACTGGCCATCAGAGA
AATGCAAATCAAAACCACTATGAGATATCATCTCACACCAGTTAGAATGGCAATCATTAAAAAGTC
AGGAAACAACAGGTGCTGGAGAGGATGCGGAGAAATAGGAACACTTTTACACTGTTGGTGGGACT
GTAAACTAGTTCAACCATTGTGGAAGTCAGTGTGGCGATTCCTCAGGGATCTAGAACTAGAAATAC
CATTTGACCCAGCCATCCCATTACTGGGTATATACCCAAATGAGTATAAATCATGCTGCTATAAAG
ACACATGCACACGTATGTTTATTGCGGCACTATTCACAATAGCAAAGACTTGGAACCAACCCAAAT
GTCCAACAATGATAGACTGGATTAAGAAAATGTGGCACATATACACCATGGAATACTATGCAGCCA
TAAAAAATGATGAGTTCATATCCTTTGTAGGGACATGGATGAAATTGGAAACCATCATTCTCAGTA
AACTATCGCAAGAACAAAAAACCAAACACCGCATATTCTCACTCATAGGTGGGAATTGAACAATG
AGATCACATGGACACAGGAAGGGGAATATCACACTCTGGGGACTGTGGTGGGGTCGGGGGAGGGG
GGAGGGATAGCATTGGGAGATATACCTAATGCTAGATGACACATTAGTGGGTGCAGCGCACCAGC
ATGGCACATGTATACGGATCCGAATTCTCGACGGATCGATCCGAACAAACGACCCAACACCCGTGC

[0701]

GTTTTATTCTGTCTTTTTATTGCCGATCCCCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGAT
GCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAA
GCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCGGCCGCTTTACTTGTACA
GCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGACCATGTGATCGCGCT
TCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAGCACGG
GGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCCAGGTGAGTCCAGGAGATGTTTCAGC
ACTGTTGCCTTTAGTCTCGAGGCAACTTAGACAACTGAGTATTGATCTGAGCACAGCAGGGTGTGA
GCTGTTTGAAGATACTGGGGTTGGGGGTGAAGAAACTGCAGAGGACTAACTGGGCTGAGACCCAG
TGGCAATGTTTTAGGGCCTAAGGAATGCCTCTGAAAATCTAGATGGACAACTTTGACTTTGAGAAA
AGAGAGGTGGAAATGAGGAAAATGACTTTTCTTTATTAGATTTCGGTAGAAAGAACTTTCATCTTTC
CCCTATTTTTGTTATTCGTTTTAAAACATCTATCTGGAGGCAGGACAAGTATGGTCATTAAAAAGAT
GCAGGCAGAAGGCATATATTGGCTCAGTCAAAGTGGGGAACTTTGGTGGCCAAACATACATTGCTA
AGGCTATTCCTATATCAGCTGGACACATATAAAATGCTGCTAATGCTTCATTACAAACTTATATCCT
TTAATTCCAGATGGGGGCAAAGTATGTCCAGGGGTGAGGAACAATTGAAACATTTGGGCTGGAGTA
GATTTTGAAAGTCAGCTCTGTGTGTGTGTGTGTGTGTGTGTGAGAGCGTGTGTTTCTTTTAAC
GTTTTCAGCCTACAGCATACAGGGTTCATGGTGGCAAGAAGATAACAAGATTTAAATTATGGCCAG
TGACTAGTGCTGCAAGAAGAACAACTACCTGCATTTAATGGGAAAGCAAAATCTCAGGCTTTGAGG
GAAGTTAACATAGGCTTGATTCTGGGTGGAAGCTGGGTGTGTAGTTATCTGGAGGCCAGGCTGGAG
CTCTCAGCTCACTATGGGTTCATCTTTATTGTCTCCTTTCATCTCAACAGCTGCACGCTGCCGTCCTC
GATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGATATAGACG
TTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGC
CCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCTTGTAGTT
GCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAA
GTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTCAC
GAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGT
AGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTGGCCGTTTACGTCGCCGTCCAGCTC
GACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACCATAGGGCCGGGATTCTC
CTCCACGTCACCGCATGTTAGAAGACTTCCTCTGCCCTCCATGTTCTCGTAGGAGTCGGCGTCCTCT
TCGTGGTTAGGTCCAGGTTGGCCTCTGATAGACCGCAGCTGAGGAGCGGCGTACAGAATGCCTCTC
ATGTCCTCATAGCTGCCGCTGCCTTGTGGAGGCTTCTCGTGCTTCAGTGTCTCGTATGTCTCTTGATT
CCGGGTGCTCAGGCCGGTGTACACGCCATCAGATTTCTCGTAGCTGGTGATGGCGGCCTTCCGCAC
TTGGATCTTCAGCCGTCTGCAGTACAGGGTGATGACCAGAGACAGCAGCAGGACACCACATGTGCC
AGCCAGAGGGGCCCAAATGTAGATATCCAGGCCTCTGGTATGCACAGCTCCGCCTGCAGCAGGTCT
ACAGGCTTCAGGTCTGAGAGACAGAGGCTGGCTGGCGATTGTAGGAGCTGGTGTAGGTGGTCTAGG
AGCGGGTGTTGTTGTAGGCTTGGCGGGCAGAAACACGGGCACGAAGTGGCTGAAGTACATGATGC
TATTGCTCAGGGCTCCGCTTCCTCCGCCGCCTGATTTGATTTCCAGCTTGGTGCCTCCGCCAAATGT
CCAAGGGCTCTCGTCGTACTGCTGGCAGTAGTAGATGCCGAAGTCCTCGTACTGCAGGCTGCTGAT
TGTCAGGGTGTAGTCGGTGCCAGAGCCGCTGCCAGAAAATCTGCTTGGCACGCCGCTTTCCAGTCT
GTTGGCCCGGTAGATCAGTGTCTTAGGGGCCTTGCCAGGCTTCTGCTGGAACCAGCTCAGGTAGCT
GTTGATGTCCTGGCTGGCTCTACAGGTGATGGTCACTCTATCGCCCACAGAGGCAGACAGGCTGCT
AGGGCTCTGTGTCATCTGGATATCAGAGCCACCACCGCCAGATCCACCGCCACCTGATCCTCCGCC
TCCGCTAGAAACTGTCACTGTGGTGCCCTGGCCCCACACATCGAAGTACCAGTCGTAGCCTCTTCTG
GTGCAGAAGTACACGGCGGTATCCTCGGCTCTCAGGCTGTTGATCTGCAGGTAGGCGGTGTTCTTG
CTGTCGTCCAGGCTGAAGGTGAATCTGCCCTTAAAGCTATCGGCGTAGGTTGGCTCGCCGGTGTGG
GTATTGATCCAGCCCATCCACTCAAGGCCTTTTCCAGGGGCCTGTCGGACCCAGTTCATGCCGTAGT
TGGTGAAGGTGTAGCCGCTGGCGGCACAGCTGATTCTGACAGATCCGCCAGGTTTCACAAGTCCGC
CGCCAGACTGAACCAGCTGGATCTCAGAGATGCTACAGGCCACTGTTCCCAGCAGCAGCAGAGACT
GCAGCCACATTCGAAGCTTGAGCTCGAGATCTGAGTCCGGTAGCGCTAGCGGATCTGACGGTTCAC
TAAACCAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCCATTTGCGTCAATGGGGCG
GAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTTGGTGCCAAAACAAACTCCCATTGACGTCAA
TGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTGATGTACTGCCAAA
ACCGCATCACCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATA
AGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTAC
TTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACG
TCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGG
GTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCGGAACTCCATATATGG
GCTATGAACTAATGACCCCGTAATTGATTACTATTAGCCCGGGGGATCCAGACATGATAAGATACA
TTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTG
ATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCA
TTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGT
GGTATGGCTGATTATGATCCGGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACA
TGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAG
GGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGAGGTCGATCGACTCTAGAGGATC

[0702]

GATCCCCGCCCCGGACGAACTAAACCTGACTACGACATCTCTGCCCCTTCTTCGCGGGGCAGTGCA
TGTAATCCCTTCAGTTGGTTGGTACAACTTGCCAACTGGGCCCTGTTCCACATGTGACACGGGGGGG
GACCAAACACAAAGGGGTTCTCTGACTGTAGTTGACATCCTTATAAATGGATGTGCACATTTGCCA
ACACTGAGTGGCTTTCATCCTGGAGCAGACTTTGCAGTCTGTGGACTGCAACACAACATTGCCTTTA
TGTGTAACTCTTGGCTGAAGCTCTTACACCAATGCTGGGGGACATGTACCTCCCAGGGGCCCAGGA
AGACTACGGGAGGCTACACCAACGTCAATCAGAGGGGCCTGTGTAGCTACCGATAAGCGGACCCT
CAAGAGGGCATTAGCAATAGTGTTTATAAGGCCCCCTTGTTAACCCTAAACGGGTAGCATATGCTT
CCCGGGTAGTAGTATATACTATCCAGACTAACCCTAATTCAATAGCATATGTTACCCAACGGGAAG
CATATGCTATCGAATTAGGGTTAGTAAAAGGGTCCTAAGGAACAGCGATATCTCCCACCCCATGAG
CTGTCACGGTTTTATTTACATGGGGTCAGGATTCCACGAGGGTAGTGAACCATTTTAGTCACAAGG
GCAGTGGCTGAAGATCAAGGAGCGGGCAGTGAACTCTCCTGAATCTTCGCCTGCTTCTTCATTCTCC
TTCGTTTAGCTAATAGAATAACTGCTGAGTTGTGAACAGTAAGGTGTATGTGAGGTGCTCGAAAAC
AAGGTTTCAGGTGACGCCCCCAGAATAAAATTTGGACGGGGGGTTCAGTGGTGGCATTGTGCTATG
ACACCAATATAACCCTCACAAACCCCTTGGGCAATAAATACTAGTGTAGGAATGAAACATTCTGAA
TATCTTTAACAATAGAAATCCATGGGGTGGGGACAAGCCGTAAAGACTGGATGTCCATCTCACACG
AATTTATGGCTATGGGCAACACATAATCCTAGTGCAATATGATACTGGGGTTATTAAGATGTGTCCC
AGGCAGGGACCAAGACAGGTGAACCATGTTGTTACACTCTATTTGTAACAAGGGGAAAGAGAGTG
GACGCCGACAGCAGCGGACTCCACTGGTTGTCTCTAACACCCCCGAAAATTAAACGGGGCTCCACG
CCAATGGGGCCCATAAACAAAGACAAGTGGCCACTCTTTTTTTTGAAATTGTGGAGTGGGGGCACG
CGTCAGCCCCCACACGCCGCCCTGCGGTTTTGGACTGTAAAATAAGGGTGTAATAACTTGGCTGAT
TGTAACCCCGCTAACCACTGCGGTCAAACCACTTGCCCACAAAACCACTAATGGCACCCCGGGGAA
TACCTGCATAAGTAGGTGGGCGGGCCAAGATAGGGGCGCGATTGCTGCGATCTGGAGGACAAATT
ACACACACTTGCGCCTGAGCGCCAAGCACAGGGTTGTTGGTCCTCATATTCACGAGGTCGCTGAGA
GCACGGTGGGCTAATGTTGCCATGGGTAGCATATACTACCCAAATATCTGGATAGCATATGCTATC
CTAATCTATATCTGGGTAGCATAGGCTATCCTAATCTATATCTGGGTAGCATATGCTATCCTAATCT
ATATCTGGGTAGTATATGCTATCCTAATTTATATCTGGGTAGCATAGGCTATCCTAATCTATATCTG
GGTAGCATATGCTATCCTAATCTATATCTGGGTAGTATATGCTATCCTAATCTGTATCCGGGTAGCA
TATGCTATCCTAATAGAGATTAGGGTAGTATATGCTATCCTAATTTATATCTGGGTAGCATATACTA
CCCAAATATCTGGATAGCATATGCTATCCTAATCTATATCTGGGTAGCATATGCTATCCTAATCTAT
ATCTGGGTAGCATAGGCTATCCTAATCTATATCTGGGTAGCATATGCTATCCTAATCTATATCTGGG
TAGTATATGCTATCCTAATTTATATCTGGGTAGCATAGGCTATCCTAATCTATATCTGGGTAGCATA
TGCTATCCTAATCTATATCTGGGTAGTATATGCTATCCTAATCTGTATCCGGGTAGCATATGCTATC
CTCATGCATATACAGTCAGCATATGATACCCAGTAGTAGAGTGGGAGTGCTATCCTTTGCATATGCC
GCCACCTCCCAAGGGGGCGTGAATTTTCGCTGCTTGTCCTTTTCCTGCATGCTGGTTGCTCCCATTCT
TAGGTGAATTTAAGGAGGCCAGGCTAAAGCCGTCGCATGTCTGATTGCTCACCAGGTAAATGTCGC
TAATGTTTTCCAACGCGAGAAGGTGTTGAGCGCGGAGCTGAGTGACGTGACAACATGGGTATGCCC
AATTGCCCCATGTTGGGAGGACGAAAATGGTGACAAGACAGATGGCCAGAAATACACCAACAGCA
CGCATGATGTCTACTGGGGATTTATTCTTTAGTGCGGGGGAATACACGGCTTTTAATACGATTGAGG
GCGTCTCCTAACAAGTTACATCACTCCTGCCCTTCCTCACCCTCATCTCCATCACCTCCTTCATCTCC
GTCATCTCCGTCATCACCCTCCGCGGCAGCCCCTTCCACCATAGGTGGAAACCAGGGAGGCAAATC
TACTCCATCGTCAAAGCTGCACACAGTCACCCTGATATTGCAGGTAGGAGCGGGCTTTGTCATAAC
AAGGTCCTTAATCGCATCCTTCAAAACCTCAGCAAATATATGAGTTTGTAAAAAGACCATGAAATA
ACAGACAATGGACTCCCTTAGCGGGCCAGGTTGTGGGCCGGGTCCAGGGGCCATTCCAAAGGGGA
GACGACTCAATGGTGTAAGACGACATTGTGGAATAGCAAGGGCAGTTCCTCGCCTTAGGTTGTAAA
GGGAGGTCTTACTACCTCCATATACGAACACACCGGCGACCCAAGTTCCTTCGTCGGTAGTCCTTTC
TACGTGACTCCTAGCCAGGAGAGCTCTTAAACCTTCTGCAATGTTCTCAAATTTCGGGTTGGAACCT
CCTTGACCACGATGCTTTCCAAACCACCCTCCTTTTTTGCGCCTGCCTCCATCACCCTGACCCCGGG
GTCCAGTGCTTGGGCCTTCTCCTGGGTCATCTGCGGGGCCCTGCTCTATCGCTCCCGGGGGCACGTC
AGGCTCACCATCTGGGCCACCTTCTTGGTGGTATTCAAAATAATCGGCTTCCCCTACAGGGTGGAA
AAATGGCCTTCTACCTGGAGGGGGCCTGCGCGGTGGAGACCCGGATGATGATGACTGACTACTGGG
ACTCCTGGGCCTCTTTTCTCCACGTCCACGACCTCTCCCCCTGGCTCTTTCACGACTTCCCCCCCTGG
CTCTTTCACGTCCTCTACCCCGGCGGCCTCCACTACCTCCTCGACCCCGGCCTCCACTACCTCCTCG
ACCCCGGCCTCCACTGCCTCCTCGACCCCGGCCTCCACCTCCTGCTCCTGCCCCTCCTGCTCCTGCCC
CTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCT
CCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGC
TCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTC
CTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCT
CCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCC
TGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCC
CTCCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCTCCTG
CCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCT
CCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCT
GCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCCGCTCCTGCTCCTGCTCCTGT

[0703]

TCCACCGTGGGTCCCTTTGCAGCCAATGCAACTTGGACGTTTTTGGGGTCTCCGGACACCATCTCTA
TGTCTTGGCCCTGATCCTGAGCCGCCCGGGGCTCCTGGTCTTCCGCCTCCTCGTCCTCGTCCTCTTCC
CCGTCCTCGTCCATGGTTATCACCCCCTCTTCTTTGAGGTCCACTGCCGCCGGAGCCTTCTGGTCCA
GATGTGTCTCCCTTCTCTCCTAGGCCATTTCCAGGTCCTGTACCTGGCCCCTCGTCAGACATGATTC
ACACTAAAAGAGATCAATAGACATCTTTATTAGACGACGCTCAGTGAATACAGGGAGTGCAGACTC
CTGCCCCCTCCAACAGCCCCCCCACCCTCATCCCCTTCATGGTCGCTGTCAGACAGATCCAGGTCTG
AAAATTCCCCATCCTCCGAACCATCCTCGTCCTCATCACCAATTACTCGCAGCCCGGAAAACTCCCG
CTGAACATCCTCAAGATTTGCGTCCTGAGCCTCAAGCCAGGCCTCAAATTCCTCGTCCCCCTTTTTG
CTGGACGGTAGGGATGGGGATTCTCGGGACCCCTCCTCTTCCTCTTCAAGGTCACCAGACAGAGAT
GCTACTGGGGCAACGGAAGAAAAGCTGGGTGCGGCCTGTGAGGATCAGCTTATCGATGATAAGCT
GTCAAACATGAGAATTCTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTC
ATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTT
GTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCA
ATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCG
GCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGT
TGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCC
CCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGT
TGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTC
ACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAAC
CATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCG
CTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAG
CCATACCAAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTA
TTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAA
GTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCG
GTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAG
TTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGT
GCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAA
AACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCC
TTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGA
TCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGT
TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCA
AATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACAT
ACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTT
GGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACAC
AGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGC
GCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAG
AGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACC
TCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCA
ACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCC
CCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACG
ACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACG
CATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGT
TAAGCCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAA
GTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAG
GCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCA
TCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTAT
GCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGC
CTAGGCTTTTGCAAAAAGCTTGCATGCCTGCAGGTCGGCCGCCACGACCGGTGCCGCCACCATCCC
CTGACCCACGCCCCTGACCCCTCACAAGGAGACGACCTTCCATGACCGAGTACAAGCCCACGGTGC
GCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCACCCTCGCCGCCGCGTTCGCCGACTACC
CCGCCACGCGCCACACCGTCGACCCGGACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCT
TCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGCGGACGACGGCGCCGCGGTGGCG
GTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCCCGCGCATGGC
CGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGCCGCACCGGC
CCAAGGAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGG
GCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGAGCGCGCCGGGGTGCCCGCCTTCCTGGAG
ACCTCCGCGCCCCGCAACCTCCCCTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTCGAG
GTGCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACGCCCGCCCCACGAC
CCGCAGCGCCCGACCGAAAGGAGCGCACGACCCCATGGCTCCGACCGAAGCCGACCCGGGCGGCC
CCGCCGACCCCGCACCCGCCCCCGAGGCCCACCGACTCTAGAGGATCATAATCAGCCATACCACAT
TTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAA
TGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACA
AATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATC

TTATCATGTCTGGATCACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGG
AATAGGGGAGATGGGGGAGGCTAACTGAAACACGGAAGGAGACAATACCGGAAGGAACCCGCGC
TATGACGGCAATAAAAAGACAGAATAAAACGCACGGGTGTTGGGTCGTTTGTTCATAAACGCGGG
GTTCGGTCCCAGGGCTGGCACTCTGTCGATACCCCACCGAGACCCCATTGGGGCCAATACGCCCGC
GTTTCTTCCTTTTCCCCACCCCACCCCCCAAGTTCGGGTGAAGGCCCAGGGCTCGCAGCCAACGTCG
GGGCGGCAGGCCCTGCCATAGCCACTGGCCCCGTGGGTTAGGGACGGGGTCCCCCATGGGGAATG
GTTTATGGTTCGTGGGGGTTATTATTTTGGGCGTTGCGTGGGGTCTGGTCCACGACTGGACTGAGCA
GACAGACCCATGGTTTTTGGATGGCCTGGGCATGGACCGCATGTACTGGCGCGACACGAACACCGG
GCGTCTGTGGCTGCCAAACACCCCCGACCCCCAAAAACCACCGCGCGGATTTCTGGCGTGCCAAGC
TAGTCGACCAATTCTCATGTTTGACAGCTTATCATCGCAGATCCGGGCAACGTTGTTGCATTGCTGC
AGGCGCAGAACTGGTAGGTATGGAAGATCTCTAGAAGCTGGGTACCAGCTGCTAGCAAGCTTGCTA
GCGGCCGGCTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTG
ATAGAGAACGATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATC
AGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTATCC
CTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGGT
AGGCGTGTACGGTGGGAGGCCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCG(서열번호 **40)**

LINE-1 플라스미드 Her2-Cd3z-T2A GFP 인트론(서열번호 **41)**

CGGCCGCGGGGGGAGGAGCCAAGATGGCCGAATAGGAACAGCTCCGGTCTACAGCTCCCAGCGTG
AGCGACGCAGAAGACGGTGATTTCTGCATTTCCATCTGAGGTACCGGGTTCATCTCACTAGGGAGT
GCCAGACAGTGGGCGCAGGCCAGTGTGTGTGCGCACCGTGCGCGAGCCGAAGCAGGGCGAGGCAT
TGCCTCACCTGGGAAGCGCAAGGGGTCAGGGAGTTCCCTTTCCGAGTCAAAGAAAGGGGTGACGG
ACGCACCTGGAAAATCGGGTCACTCCCACCCGAATATTGCGCTTTTCAGACCGGCTTAAGAAACGG
CGCACCACGAGACTATATCCCACACCTGGCTCGGAGGGTCCTACGCCCACGGAATCTCGCTGATTG
CTAGCACAGCAGTCTGAGATCAAACTGCAAGGCGGCAACGAGGCTGGGGGAGGGGCGCCCGCCAT
TGCCCAGGCTTGCTTAGGTAAACAAAGCAGCAGGGAAGCTCGAACTGGGTGGAGCCCACCACAGC
TCAAGGAGGCCTGCCTGCCTCTGTAGGCTCCACCTCTGGGGGCAGGGCACAGACAAACAAAAAGA
CAGCAGTAACCTCTGCAGACTTAAGTGTCCCTGTCTGACAGCTTTGAAGAGAGCAGTGGTTCTCCC
AGCACGCAGCTGGAGATCTGAGAACGGGCAGACTGCCTCCTCAAGTGGGTCCCTGACCCCTGACCC
CCGAGCAGCCTAACTGGGAGGCACCCCCCAGCAGGGGCACACTGACACCTCACACGGCAGGGTAT
TCCAACAGACCTGCAGCTGAGGGTCCTGTCTGTTAGAAGGAAAACTAACAACCAGAAAGGACATCT
ACACCGAAAACCCATCTGTACATCACCATCATCAAAGACCAAAAGTAGATAAAACCACAAAGATG
GGGAAAAAACAGAACAGAAAAACTGGAAACTCTAAAACGCAGAGCGCCTCTCCTCCTCCAAAGGA
ACGCAGTTCCTCACCAGCAACAGAACAAAGCTGGATGGAGAATGATTTTGATGAGCTGAGAGAAG
AAGGCTTCAGACGATCAAATTACTCTGAGCTACGGGAGGACATTCAAACCAAAGGCAAAGAAGTT
GAAAACTTTGAAAAAAATTTAGAAGAATGTATAACTAGAATAACCAATACAGAGAAGTGCTTAAA
GGAGCTGATGGAGCTGAAAACCAAGGCTCGAGAACTACGTGAAGAATGCAGAAGCCTCAGGAGCC
GATGCGATCAACTGGAAGAAAGGGTATCAGCAATGGAAGATGAAATGAATGAAATGAAGCGAGA
AGGGAAGTTTAGAGAAAAAAGAATAAAAAGAAATGAGCAAAGCCTCCAAGAAATATGGGACTAT
GTGAAAAGACCAAATCTACGTCTGATTGGTGTACCTGAAAGTGATGTGGAGAATGGAACCAAGTTG
GAAAACACTCTGCAGGATATTATCCAGGAGAACTTCCCCAATCTAGCAAGGCAGGCCAACGTTCAG
ATTCAGGAAATACAGAGAACGCCACAAAGATACTCCTCGAGAAGAGCAACTCCAAGACACATAAT
TGTCAGATTCACCAAAGTTGAAATGAAGGAAAAAATGTTAAGGGCAGCCAGAGAGAAAGGTCGGG
TTACCCTCAAAGGAAAGCCCATCAGACTAACAGCGGATCTCTCGGCAGAAACCCTACAAGCCAGA
AGAGAGTGGGGGCCAATATTCAACATTCTTAAAGAAAAGAATTTTCAACCCAGAATTTCATATCCA
GCCAAACTAAGCTTCATAAGTGAAGGAGAAATAAAATACTTTATAGACAAGCAAATGTTGAGAGA
TTTTGTCACCACCAGGCCTGCCCTAAAAGAGCTCCTGAAGGAAGCGCTAAACATGGAAAGGAACA
ACCGGTACCAGCCGCTGCAAAATCATGCCAAAATGTAAAGACCATCAAGACTAGGAAGAAACTGC
ATCAACTAATGAGCAAAATCACCAGCTAACATCATAATGACAGGATCAACTTCACACATAACAATA
TTAACTTTAAATATAAATGGACTAAATTCTGCAATTAAAAGACACAGACTGGCAAGTTGGATAAAG
AGTCAAGACCCATCAGTGTGCTGTATTCAGGAAACCCATCTCACGTGCAGAGACACACATAGGCTC
AAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCAA
TCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTAC
ATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTATGCACCCAATACA
GGAGCACCCAGATTCATAAAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTA
ATAATGGGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAA
GGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTCCA
CCCCAAATCAACAGAATATACATTTTTTTCAGCACCACACCACACCTATTCCAAAATTGACCACATA
GTTGGAAGTAAAGCTCTCCTCAGCAAATGTAAAAGAACAGAAATTATAACAAACTATCTCTCAGAC
CACAGTGCAATCAAACTAGAACTCAGGATTAAGAATCTCACTCAAAGCCGCTCAACTACATGGAAA
CTGAACAACCTGCTCCTGAATGACTACTGGGTACATAACGAAATGAAGGCAGAAATAAAGATGTTC
TTTGAAACCAACGAGAACAAAGACACCACATACCAGAATCTCTGGGACGCATTCAAAGCAGTGTGT

[0704]

[0705]

AGAGGGAAATTTATAGCACTAAATGCCTACAAGAGAAAGCAGGAAAGATCCAAAATTGACACCCT
AACATCACAATTAAAAGAACTAGAAAAGCAAGAGCAAACACATTCAAAAGCTAGCAGAAGGCAA
GAAATAACTAAAATCAGAGCAGAACTGAAGGAAATAGAGACACAAAAAACCCTTCAAAAAATCAA
TGAATCCAGGAGCTGGTTTTTTGAAAGGATCAACAAAATTGATAGACCGCTAGCAAGACTAATAAA
GAAAAAAAGAGAGAAGAATCAAATAGACACAATAAAAAATGATAAAGGGGATATCACCACCGAT
CCCACAGAAATACAAACTACCATCAGAGAATACTACAAACACCTCTACGCAAATAAACTAGAAAA
TCTAGAAGAAATGGATACATTCCTCGACACATACACTCTCCCAAGACTAAACCAGGAAGAAGTTGA
ATCTCTGAATCGACCAATAACAGGCTCTGAAATTGTGGCAATAATCAATAGTTTACCAACCAAAAA
GAGTCCAGGACCAGATGGATTCACAGCCGAATTCTACCAGAGGTACAAGGAGGAACTGGTACCAT
TCCTTCTGAAACTATTCCAATCAATAGAAAAAGAGGGAATCCTCCCTAACTCATTTTATGAGGCCA
GCATCATTCTGATACCAAAGCCGGGCAGAGACACAACCAAAAAAGAGAATTTTAGACCAATATCCT
TGATGAACATTGATGCAAAAATCCTCAATAAAATACTGGCAAACCGAATCCAGCAGCACATCAAA
AAGCTTATCCACCATGATCAAGTGGGCTTCATCCCTGGGATGCAAGGCTGGTTCAATATACGCAAA
TCAATAAATGTAATCCAGCATATAAACAGAGCCAAAGACAAAAACCACATGATTATCTCAATAGAT
GCAGAAAAAGCCTTTGACAAAATTCAACAACCCTTCATGCTAAAAACTCTCAATAAATTAGGTATT
GATGGGACGTATTTCAAAATAATAAGAGCTATCTATGACAAACCCACAGCCAATATCATACTGAAT
GGGCAAAAACTGGAAGCATTCCCTTTGAAAACCGGCACAAGACAGGGATGCCCTCTCTCACCGCTC
CTATTCAACATAGTGTTGGAAGTTCTGGCCAGGGCAATCAGGCAGGAGAAGGAAATAAAGGGTAT
TCAATTAGGAAAAGAGGAAGTCAAATTGTCCCTGTTTGCAGACGACATGATTGTTTATCTAGAAAA
CCCCATCGTCTCAGCCCAAAATCTCCTTAAGCTGATAAGCAACTTCAGCAAAGTCTCAGGATACAA
AATCAATGTACAAAAATCACAAGCATTCTTATACACCAACAACAGACAAACAGAGAGCCAAATCA
TGGGTGAACTCCCATTCACAATTGCTTCAAAGAGAATAAAATACCTAGGAATCCAACTTACAAGGG
ATGTGAAGGACCTCTTCAAGGAGAACTACAAACCACTGCTCAAGGAAATAAAAGAGGAGACAAAC
AAATGGAAGAACATTCCATGCTCATGGGTAGGAAGAATCAATATCGTGAAAATGGCCATACTGCCC
AAGGTAATTTACAGATTCAATGCCATCCCCATCAAGCTACCAATGACTTTCTTCACAGAATTGGAA
AAAACTACTTTAAAGTTCATATGGAACCAAAAAAGAGCCCGCATTGCCAAGTCAATCCTAAGCCAA
AAGAACAAAGCTGGAGGCATCACACTACCTGACTTCAAACTATACTACAAGGCTACAGTAACCAA
AACAGCATGGTACTGGTACCAAAACAGAGATATAGATCAATGGAACAGAACAGAGCCCTCAGAAA
TAATGCCGCATATCTACAACTATCTGATCTTTGACAAACCTGAGAAAAACAAGCAATGGGGAAAGG
ATTCCCTATTTAATAAATGGTGCTGGGAAAACTGGCTAGCCATATGTAGAAAGCTGAAACTGGATC
CCTTCCTTACACCTTATACAAAAATCAATTCAAGATGGATTAAAGATTTAAACGTTAAACCTAAAA
CCATAAAAACCCTAGAAGAAAACCTAGGCATTACCATTCAGGACATAGGCGTGGGCAAGGACTTC
ATGTCCAAAACACCAAAAGCAATGGCAACAAAAGACAAAATTGACAAATGGGATCTAATTAAACT
AAAGAGCTTCTGCACAGCAAAAGAAACTACCATCAGAGTGAACAGGCAACCTACAACATGGGAGA
AAATTTTTGCAACCTACTCATCTGACAAAGGGCTAATATCCAGAATCTACAATGAACTCAAACAAA
TTTACAAGAAAAAAACAAACAACCCCATCAAAAAGTGGGCGAAGGACATGAACAGACACTTCTCA
AAAGAAGACATTTATGCAGCCAAAAAACACATGAAGAAATGCTCATCATCACTGGCCATCAGAGA
AATGCAAATCAAAACCACTATGAGATATCATCTCACACCAGTTAGAATGGCAATCATTAAAAAGTC
AGGAAACAACAGGTGCTGGAGAGGATGCGGAGAAATAGGAACACTTTTACACTGTTGGTGGGACT
GTAAACTAGTTCAACCATTGTGGAAGTCAGTGTGGCGATTCCTCAGGGATCTAGAACTAGAAATAC
CATTTGACCCAGCCATCCCATTACTGGGTATATACCCAAATGAGTATAAATCATGCTGCTATAAAG
ACACATGCACACGTATGTTTATTGCGGCACTATTCACAATAGCAAAGACTTGGAACCAACCCAAAT
GTCCAACAATGATAGACTGGATTAAGAAAATGTGGCACATATACACCATGGAATACTATGCAGCCA
TAAAAAATGATGAGTTCATATCCTTTGTAGGGACATGGATGAAATTGGAAACCATCATTCTCAGTA
AACTATCGCAAGAACAAAAAACCAAACACCGCATATTCTCACTCATAGGTGGGAATTGAACAATG
AGATCACATGGACACAGGAAGGGGAATATCACACTCTGGGGACTGTGGTGGGGTCGGGGGAGGGG
GGAGGGATAGCATTGGGAGATATACCTAATGCTAGATGACACATTAGTGGGTGCAGCGCACCAGC
ATGGCACATGTATACGGATCCGAATTCTCGACGGATCGATCCGAACAAACGACCCAACACCCGTGC
GTTTTATTCTGTCTTTTTATTGCCGATCCCCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGAT
GCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAA
GCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCGGCCGCTTTACTTGTACA
GCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGACCATGTGATCGCGCT
TCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAGCACGG
GGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCCAGGTGAGTCCAGGAGATGTTTCAGC
ACTGTTGCCTTTAGTCTCGAGGCAACTTAGACAACTGAGTATTGATCTGAGCACAGCAGGGTGTGA
GCTGTTTGAAGATACTGGGGTTGGGGGTGAAGAAACTGCAGAGGACTAACTGGGCTGAGACCCAG
TGGCAATGTTTTAGGGCCTAAGGAATGCCTCTGAAAATCTAGATGGACAACTTTGACTTTGAGAAA
AGAGAGGTGGAAATGAGGAAAATGACTTTTCTTTATTAGATTTCGGTAGAAAGAACTTTCATCTTTC
CCCTATTTTTGTTATTCGTTTTAAAACATCTATCTGGAGGCAGGACAAGTATGGTCATTAAAAAGAT
GCAGGCAGAAGGCATATATTGGCTCAGTCAAAGTGGGGAACTTTGGTGGCCAAACATACATTGCTA
AGGCTATTCCTATATCAGCTGGACACATATAAAATGCTGCTAATGCTTCATTACAAACTTATATCCT
TTAATTCCAGATGGGGGCAAAGTATGTCCAGGGGTGAGGAACAATTGAAACATTTGGGCTGGAGTA
GATTTTGAAAGTCAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGAGAGCGTGTGTTTCTTTTAAC

[0706]

GTTTTCAGCCTACAGCATACAGGGTTCATGGTGGCAAGAAGATAACAAGATTTAAATTATGGCCAG
TGACTAGTGCTGCAAGAAGAACAACTACCTGCATTTAATGGGAAAGCAAAATCTCAGGCTTTGAGG
GAAGTTAACATAGGCTTGATTCTGGGTGGAAGCTGGGTGTGTAGTTATCTGGAGGCCAGGCTGGAG
CTCTCAGCTCACTATGGGTTCATCTTTATTGTCTCCTTTCATCTCAACAGCTGCACGCTGCCGTCCTC
GATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGATATAGACG
TTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGC
CCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCTTGTAGTT
GCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAA
GTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTCAC
GAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGT
AGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTGGCCGTTTACGTCGCCGTCCAGCTC
GACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACCATAGGGCCGGGATTCTC
CTCCACGTCACCGCATGTTAGAAGACTTCCTCTGCCCTCTCTTGGAGGCAGGGCCTGCATGTGCAGG
GCATCGTAGGTATCCTTGGTGGCTGTGCTCAGTCCCTGGTACAGTCCATCGTGGCCCTTGCCTCTTC
TTCTCTCGCCCTTCATGCCGATCTCGCTGTAGGCCTCGGCCATCTTGTCTTTCTGCAGCTCATTATAC
AGGCCCTCTTGAGGATTCTTTCTCCGCTGGGGCTTGCCGCCCATCTCAGGATCTCTGCCTCTCCGCTT
ATCCAGCACGTCGTACTCTTCTCTTCTCCCCAGGTTCAGCTCGTTGTACAGCTGATTCTGGCCCTGCT
GGTAAGCAGGAGCGTCGGCGGATCTGCTGAACTTCACTCTGCAGTACAGGGTGATGACCAGAGAG
AGCAGCAGAACGCCACATGTGCCAGCCAGAGGGGCCCAAATGTAGATATCCAGGCCTCTGGTATG
CACAGCTCCGCCAGCTGCAGGTCTACAGGCTTCAGGTCTGAGAGACAGAGGCTGGCTGGCGATTGT
AGGAGCTGGTGTAGGTGGTCTAGGAGCGGGTGTTGTTGTAGGCTTGGCGGGCAGAAACACGGGCA
CGAAGTGGCTGAAGTACATGATGCTATTGCTCAGGGCTCCGCTTCCTCCGCCTCCGCTAGAAGAAA
CTGTGACCAGGGTGCCCTGTCCCCAAACATCCATGGCGTAGAAGCCGTCGCCTCCCCATCTAGAAC
AGTAGTACACGGCGGTGTCCTCGGCTCTCAGGCTGTTCATCTGCAGGTAGGCGGTGTTCTTGCTGGT
GTCGGCGCTGATGGTGAATCTGCCCTTCACGCTATCGGCGTATCTGGTGTAGCCGTTGGTGGGGTAG
ATTCTGGCGACCCATTCAAGTCCCTTTCCAGGGGCCTGTCGGACCCAGTGGATGTAGGTGTCCTTGA
TGTTGAAGCCGCTGGCGGCACAAGACAGTCTCAGAGAGCCGCCAGGCTGAACAAGTCCTCCGCCA
GATTCAACCAGCTGCACCTCAGATCCTTCGCCAGATCCAGGCTTTCCAGAGCCGCTGGTGCTGCCTG
TTCTCTTGATTTCCACCTTGGTGCCCTGGCCAAAGGTTGGAGGTGTGGTGTAGTGCTGCTGGCAGTA
GTAGGTGGCGAAGTCCTCAGGCTGCAGGCTAGAGATGGTCAGGGTGAAGTCGGTGCCAGATCTGCT
GCCGCTGAATCTGCTTGGCACGCCGCTGTACAGAAAGCTGGCGCTGTAGATCAGCAGCTTAGGGGC
TTTTCCAGGCTTCTGCTGATACCAGGCCACGGCGGTATTCACATCCTGGCTGGCTCTACAGGTGATG
GTCACTCTATCGCCCACAGAGGCAGACAGGCTGCTAGGGCTCTGTGTCATCTGGATGTCGCTGATG
CTGCAGGCCACTGTTCCCAGCAGCAGCAGAGACTGCAGCCACATTCGAAGCTTGAGCTCGAGATCT
GAGTCCGGTAGCGCTAGCGGATCTGACGGTTCACTAAACCAGCTCTGCTTATATAGACCTCCCACC
GTACACGCCTACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTG
ATTTTGGTGCCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTC
AAACCGCTATCCACGCCCATTGATGTACTGCCAAAACCGCATCACCATGGTAATAGCGATGACTAA
TACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGG
CCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAG
TGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGG
GAACATACGTCATTATTGACGTCAATGGGCGGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTA
CCGTAAGTTATGTAACGCGGAACTCCATATATGGGCTATGAACTAATGACCCCGTAATTGATTACT
ATTAGCCCGGGGGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATG
CAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCT
GCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGG
AGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGCTGATTATGATCCGGCTGCCTCGCG
CGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTG
TAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGG
CGCAGCCATGAGGTCGATCGACTCTAGAGGATCGATCCCCGCCCCGGACGAACTAAACCTGACTAC
GACATCTCTGCCCCTTCTTCGCGGGGCAGTGCATGTAATCCCTTCAGTTGGTTGGTACAACTTGCCA
ACTGGGCCCTGTTCCACATGTGACACGGGGGGGGACCAAACACAAAGGGGTTCTCTGACTGTAGTT
GACATCCTTATAAATGGATGTGCACATTTGCCAACACTGAGTGGCTTTCATCCTGGAGCAGACTTTG
CAGTCTGTGGACTGCAACACAACATTGCCTTTATGTGTAACTCTTGGCTGAAGCTCTTACACCAATG
CTGGGGGACATGTACCTCCCAGGGGCCCAGGAAGACTACGGGAGGCTACACCAACGTCAATCAGA
GGGGCCTGTGTAGCTACCGATAAGCGGACCCTCAAGAGGGCATTAGCAATAGTGTTTATAAGGCCC
CCTTGTTAACCCTAAACGGGTAGCATATGCTTCCCGGGTAGTAGTATATACTATCCAGACTAACCCT
AATTCAATAGCATATGTTACCCAACGGGAAGCATATGCTATCGAATTAGGGTTAGTAAAAGGGTCC
TAAGGAACAGCGATATCTCCCACCCCATGAGCTGTCACGGTTTTATTTACATGGGGTCAGGATTCCA
CGAGGGTAGTGAACCATTTTAGTCACAAGGGCAGTGGCTGAAGATCAAGGAGCGGGCAGTGAACT
CTCCTGAATCTTCGCCTGCTTCTTCATTCTCCTTCGTTTAGCTAATAGAATAACTGCTGAGTTGTGAA
CAGTAAGGTGTATGTGAGGTGCTCGAAAACAAGGTTTCAGGTGACGCCCCCAGAATAAAATTTGGA
CGGGGGGTTCAGTGGTGGCATTGTGCTATGACACCAATATAACCCTCACAAACCCCTTGGGCAATA

[0707]

AATACTAGTGTAGGAATGAAACATTCTGAATATCTTTAACAATAGAAATCCATGGGGTGGGGACAA
GCCGTAAAGACTGGATGTCCATCTCACACGAATTTATGGCTATGGGCAACACATAATCCTAGTGCA
ATATGATACTGGGGTTATTAAGATGTGTCCCAGGCAGGGACCAAGACAGGTGAACCATGTTGTTAC
ACTCTATTTGTAACAAGGGGAAAGAGAGTGGACGCCGACAGCAGCGGACTCCACTGGTTGTCTCTA
ACACCCCCGAAAATTAAACGGGGCTCCACGCCAATGGGGCCCATAAACAAAGACAAGTGGCCACT
CTTTTTTTTGAAATTGTGGAGTGGGGGCACGCGTCAGCCCCCACACGCCGCCCTGCGGTTTTGGACT
GTAAAATAAGGGTGTAATAACTTGGCTGATTGTAACCCCGCTAACCACTGCGGTCAAACCACTTGC
CCACAAAACCACTAATGGCACCCCGGGGAATACCTGCATAAGTAGGTGGGCGGGCCAAGATAGGG
GCGCGATTGCTGCGATCTGGAGGACAAATTACACACACTTGCGCCTGAGCGCCAAGCACAGGGTTG
TTGGTCCTCATATTCACGAGGTCGCTGAGAGCACGGTGGGCTAATGTTGCCATGGGTAGCATATAC
TACCCAAATATCTGGATAGCATATGCTATCCTAATCTATATCTGGGTAGCATAGGCTATCCTAATCT
ATATCTGGGTAGCATATGCTATCCTAATCTATATCTGGGTAGTATATGCTATCCTAATTTATATCTG
GGTAGCATAGGCTATCCTAATCTATATCTGGGTAGCATATGCTATCCTAATCTATATCTGGGTAGTA
TATGCTATCCTAATCTGTATCCGGGTAGCATATGCTATCCTAATAGAGATTAGGGTAGTATATGCTA
TCCTAATTTATATCTGGGTAGCATATACTACCCAAATATCTGGATAGCATATGCTATCCTAATCTAT
ATCTGGGTAGCATATGCTATCCTAATCTATATCTGGGTAGCATAGGCTATCCTAATCTATATCTGGG
TAGCATATGCTATCCTAATCTATATCTGGGTAGTATATGCTATCCTAATTTATATCTGGGTAGCATA
GGCTATCCTAATCTATATCTGGGTAGCATATGCTATCCTAATCTATATCTGGGTAGTATATGCTATC
CTAATCTGTATCCGGGTAGCATATGCTATCCTCATGCATATACAGTCAGCATATGATACCCAGTAGT
AGAGTGGGAGTGCTATCCTTTGCATATGCCGCCACCTCCCAAGGGGGCGTGAATTTTCGCTGCTTGT
CCTTTTCCTGCATGCTGGTTGCTCCCATTCTTAGGTGAATTTAAGGAGGCCAGGCTAAAGCCGTCGC
ATGTCTGATTGCTCACCAGGTAAATGTCGCTAATGTTTTCCAACGCGAGAAGGTGTTGAGCGCGGA
GCTGAGTGACGTGACAACATGGGTATGCCCAATTGCCCCATGTTGGGAGGACGAAAATGGTGACA
AGACAGATGGCCAGAAATACACCAACAGCACGCATGATGTCTACTGGGGATTTATTCTTTAGTGCG
GGGGAATACACGGCTTTTAATACGATTGAGGGCGTCTCCTAACAAGTTACATCACTCCTGCCCTTCC
TCACCCTCATCTCCATCACCTCCTTCATCTCCGTCATCTCCGTCATCACCCTCCGCGGCAGCCCCTTC
CACCATAGGTGGAAACCAGGGAGGCAAATCTACTCCATCGTCAAAGCTGCACACAGTCACCCTGAT
ATTGCAGGTAGGAGCGGGCTTTGTCATAACAAGGTCCTTAATCGCATCCTTCAAAACCTCAGCAAA
TATATGAGTTTGTAAAAAGACCATGAAATAACAGACAATGGACTCCCTTAGCGGGCCAGGTTGTGG
GCCGGGTCCAGGGGCCATTCCAAAGGGGAGACGACTCAATGGTGTAAGACGACATTGTGGAATAG
CAAGGGCAGTTCCTCGCCTTAGGTTGTAAAGGGAGGTCTTACTACCTCCATATACGAACACACCGG
CGACCCAAGTTCCTTCGTCGGTAGTCCTTTCTACGTGACTCCTAGCCAGGAGAGCTCTTAAACCTTC
TGCAATGTTCTCAAATTTCGGGTTGGAACCTCCTTGACCACGATGCTTTCCAAACCACCCTCCTTTTT
TGCGCCTGCCTCCATCACCCTGACCCCGGGGTCCAGTGCTTGGGCCTTCTCCTGGGTCATCTGCGGG
GCCCTGCTCTATCGCTCCCGGGGGCACGTCAGGCTCACCATCTGGGCCACCTTCTTGGTGGTATTCA
AAATAATCGGCTTCCCCTACAGGGTGGAAAAATGGCCTTCTACCTGGAGGGGGCCTGCGCGGTGGA
GACCCGGATGATGATGACTGACTACTGGGACTCCTGGGCCTCTTTTCTCCACGTCCACGACCTCTCC
CCCTGGCTCTTTCACGACTTCCCCCCCTGGCTCTTTCACGTCCTCTACCCCGGCGGCCTCCACTACCT
CCTCGACCCCGGCCTCCACTACCTCCTCGACCCCGGCCTCCACTGCCTCCTCGACCCCGGCCTCCAC
CTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCT
GCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTCC
TGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTG
CCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCT
GCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGC
CCCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTC
CTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTCCT
GCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCC
TGCCCCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTC
CTGCCCCTCCCGCTCCTGCTCCTGCTCCTGTTCCACCGTGGGTCCCTTTGCAGCCAATGCAACTTGG
ACGTTTTTGGGGTCTCCGGACACCATCTCTATGTCTTGGCCCTGATCCTGAGCCGCCCGGGGCTCCT
GGTCTTCCGCCTCCTCGTCCTCGTCCTCTTCCCCGTCCTCGTCCATGGTTATCACCCCCTCTTCTTTG
AGGTCCACTGCCGCCGGAGCCTTCTGGTCCAGATGTGTCTCCCTTCTCTCCTAGGCCATTTCCAGGT
CCTGTACCTGGCCCCTCGTCAGACATGATTCACACTAAAAGAGATCAATAGACATCTTTATTAGAC
GACGCTCAGTGAATACAGGGAGTGCAGACTCCTGCCCCCTCCAACAGCCCCCCCACCCTCATCCCC
TTCATGGTCGCTGTCAGACAGATCCAGGTCTGAAAATTCCCCATCCTCCGAACCATCCTCGTCCTCA
TCACCAATTACTCGCAGCCCGGAAAACTCCCGCTGAACATCCTCAAGATTTGCGTCCTGAGCCTCA
AGCCAGGCCTCAAATTCCTCGTCCCCCTTTTTGCTGGACGGTAGGGATGGGGATTCTCGGGACCCCT
CCTCTTCCTCTTCAAGGTCACCAGACAGAGATGCTACTGGGGCAACGGAAGAAAAGCTGGGTGCGG
CCTGTGAGGATCAGCTTATCGATGATAAGCTGTCAAACATGAGAATTCTTGAAGACGAAAGGGCCT
CGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACT
TTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGC
TCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAAC

[0708]

ATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACG
CTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTC
AACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAA
GTTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATA
CACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATG
ACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTG
ACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGC
CTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCC
TGCAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCA
ACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGC
TGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTG
GGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGAT
GAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAA
GTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGA
TCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCC
CGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACA
AAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAG
GTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCAC
CACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTG
CCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGC
GGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTG
AGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTA
TCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGT
ATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGG
GGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTT
TTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGA
GCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGA
GCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTC
AGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGA
AAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAG
GTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGC
AACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCG
CCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCC
AGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTGCATGCCTGCAGGTCG
GCCGCCACGACCGGTGCCGCCACCATCCCCTGACCCACGCCCCTGACCCCTCACAAGGAGACGACC
TTCCATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACG
CACCCTCGCCGCCGCGTTCGCCGACTACCCCGCCACGCGCCACACCGTCGACCCGGACCGCCACAT
CGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTG
GGTCGCGGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCGG
TGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGA
TGGAAGGCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCT
CGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGAG
CGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTCTACGAGCGGCTC
GGCTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAG
CCCGGTGCCTGACGCCCGCCCCACGACCCGCAGCGCCCGACCGAAAGGAGCGCACGACCCCATGG
CTCCGACCGAAGCCGACCCGGGCGGCCCCGCCGACCCCGCACCCGCCCCCGAGGCCCACCGACTCT
AGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCT
CCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAAT
GGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTT
GTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGGATCACTCGCCGATAGTGGAAACCGAC
GCCCCAGCACTCGTCCGAGGGCAAAGGAATAGGGGAGATGGGGGAGGCTAACTGAAACACGGAA
GGAGACAATACCGGAAGGAACCCGCGCTATGACGGCAATAAAAAGACAGAATAAAACGCACGGG
TGTTGGGTCGTTTGTTCATAAACGCGGGGTTCGGTCCCAGGGCTGGCACTCTGTCGATACCCCACCG
AGACCCCATTGGGGCCAATACGCCCGCGTTTCTTCCTTTTCCCCACCCCACCCCCCAAGTTCGGGTG
AAGGCCCAGGGCTCGCAGCCAACGTCGGGGCGGCAGGCCCTGCCATAGCCACTGGCCCCGTGGGT
TAGGGACGGGGTCCCCCATGGGGAATGGTTTATGGTTCGTGGGGGTTATTATTTTGGGCGTTGCGTG
GGGTCTGGTCCACGACTGGACTGAGCAGACAGACCCATGGTTTTTGGATGGCCTGGGCATGGACCG
CATGTACTGGCGCGACACGAACACCGGGCGTCTGTGGCTGCCAAACACCCCCGACCCCCAAAAACC
ACCGCGCGGATTTCTGGCGTGCCAAGCTAGTCGACCAATTCTCATGTTTGACAGCTTATCATCGCAG
ATCCGGGCAACGTTGTTGCATTGCTGCAGGCGCAGAACTGGTAGGTATGGAAGATCTCTAGAAGCT
GGGTACCAGCTGCTAGCAAGCTTGCTAGCGGCCGGCTCGAGTTTACTCCCTATCAGTGATAGAGAA
CGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGATGTCGAGTTTACTCCCTATCAGTGATAG
AGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGT

GATAGAGAACGTATGTCGAGTTTATCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTAT
CAGTGATAGAGAACGTATGTCGAGGTAGGCGTGTACGGTGGGAGGCCTATATAAGCAGAGCTCGTT
TAGTGAACCGTCAGATCGCCG(서열번호 **41**)

LINE-1 ORF2-NLS mRNA(서열번호 **42**)

TAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGG
AATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAG
TCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAA
CTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTG
CATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCA
GGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGC
GCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGA
GGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGC
GTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGA
AAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCG
GTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGA
GAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGA
CGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGA
AGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGA
TTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAAC
TGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGA
TGTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCATCAACTAATGAGCAAAATCACC
AGCTAACATCATAGTATACATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAAC
GGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTT
TGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGA
AAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGAC
GGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCAT
ACAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAA
ACAGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACA
CCATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAG
CGCACTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACA
TTCTTCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGT
CAAAGTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGC
TGAGAATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACG
ATTATTGGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAG
GATACTACCTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCA
ACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGG
AGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAG
TTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAG
CGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGAT
TGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCG
GGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGA
TACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAG
CGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGC
TGAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAA
AAGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCG
ATACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACA
AGATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTA
TACCTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAG
CTAAGGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGC
CATTTATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAAT
TTACGATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGAC
TGGAACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGG
GCTATTCGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTT
GTTCGCCGATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTT
ATTTCTAACTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACA
CAAATAATCGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGA
TAAAGTATCTCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTC
TCCTGAAAGAGATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAA
TCAACATAGTGAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACT
GCCTATGACCTTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGC
AAGAATAGCGAAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAA
GTTGTATTATAAAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCA

[0709]

GTGGAATCGGACCGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGATAAGCC
AGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGG
CCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTGGA
TCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATA
CAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAA
GATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGT
TAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAG
CAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGG
CAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAG
TGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACCTTACCCCA
GTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGC
ACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTC
TGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGA
ATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCT
AAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTGGCATATTTAT
ACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACCTGGATGAAG
CTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTC
ATTGGTGGTAACGACTACAAAGACGATGACGACAAGCCCGCCGCCAAGAGGGTGAAGCTGGACTA
AAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGATACAATCGATTTCTGGATCCGCAGG
CCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTT
TTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATT
TTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACG
TGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAG
CTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGC
CCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGAC
GTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTC
CCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCTGAGAGACACAAAAAATTCCAACACACTAT
TGCAATGAAAATAAATTTCCTTTATTAGCCAGAAGTCAGATGCTCAAGGGGCTTCATGATGTCCCC
ATAATTTTTGGCAGAGGGAAAAAGATCTCAGTGGTATTTGTGAGCCAGGGCATTGGCCTTCTGATA
GGCAGCCTGCACCTGAGGAGTGCGGCCGCTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCCC
GGCGGCGGTCACGAACTCCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGC
GGACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCTG
CTGGTAGTGGTCGGCGAGCTGCACGCTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTG
ATGCCGTTCTTCTGCTTGTCGGCCATGATATAGACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTG
CCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCG
CCCTCGAACTTCACCTCGGCGCGGGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTGCGCTCCT
GGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGC
TGAAGCACTGCACGCCGTAGGTCAGGGTGGTCACGAGGGTGGGCCAGGGCACGGGCAGCTTGCCG
GTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGTAGGTGGCATCGCCCTCGCCCTCGCCGGACACG
CTGAACTTGTGGCCGTTTACGTCGCCGTCCAGCTCGACCAGGATGGGCACCACCCCGGTGAACAGC
TCCTCGCCCTTGCTCACCATGGTGGCGGGATCTGACGGTTCACTAAACCAGCTCTGCTTATATAGAC
CTCCCACCGTACACGCCTACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAA
GTCCCGTTGATTTTGGTGCCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCC
CCGTGAGTCAAACCGCTATCCACGCCCATTGATGTACTGCCAAAACCGCATCACCATGGTAATAGC
GATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATG
CCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATG
TACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGGAAAGTCCCTATTGG
CGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGGTCGTTGGGCGGTCAGCCAGG
CGGGCCATTTACCGTAAGTTATGTAACGGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGC
CTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGA
GATACAGCGTACCTTCAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACT
AGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTA
TAAGCTGCAATAAACAAGTT(서열번호 **42**)

LINE-1 alu mRNA GFP(서열번호 **43**)
TAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGG
AATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAG
TCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAA
CTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTG
CATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCA
GGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGC
GCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGA
GGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGC

[0710]

[0711]

GTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGA
AAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCG
GTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGA
GAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGA
CGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGA
AGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGA
TTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAAC
TGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGA
TGTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCATCAACTAATGAGCAAAATCACC
AGCTAACATCATAGTATACATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAAC
GGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTT
TGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGA
AAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGAC
GGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCAT
ACAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAA
ACAGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACA
CCATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAG
CGCACTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACA
TTCTTCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGT
CAAAGTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGC
TGAGAATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACG
ATTATTGGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAG
GATACTACCTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCA
ACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGG
AGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAG
TTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAG
CGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGAT
TGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCG
GGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGA
TACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAG
CGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGC
TGAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAA
AAGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCG
ATACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACA
AGATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTA
TACCTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAG
CTAAGGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGC
CATTTATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAAT
TTACGATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGAC
TGGAACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGG
GCTATTCGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTT
GTTCGCCGATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTT
ATTTCTAACTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACA
CAAATAATCGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGA
TAAAGTATCTCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTC
TCCTGAAAGAGATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAA
TCAACATAGTGAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACT
GCCTATGACCTTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGC
AAGAATAGCGAAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAA
GTTGTATTATAAAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCA
GTGGAATCGGACCGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGATAAGCC
AGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGG
CCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTGGA
TCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATA
CAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAA
GATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGT
TAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAG
CAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGG
CAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAG
TGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACCTTACCCCA
GTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGC
ACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTC
TGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGA

ATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCT
AAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTGGCATATTTAT
ACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACCTGGATGAAG
CTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTC
ATTGGTGGTAACGACTACAAAGACGATGACGACAAGTAAAGCGGCCGGGCGCGGTGGCTCACGCC
TGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGAGGATCGCAGTTCGAGACCAGCGCGAGACCCC
GTCTCTACAAAAATACAAAAATTAGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGATACAA
TCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTAT
TCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTG
CTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTG
TGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGG
GGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGG
AACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCG
TGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCG
CGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCTGAGAGACA
CAAAAAATTCCAACACACTATTGCAATGAAAATAAATTTCCTTTATTAGCCAGAAGTCAGATGCTC
AAGGGGCTTCATGATGTCCCCATAATTTTTGGCAGAGGGAAAAAGATCTCAGTGGTATTTGTGAGC
CAGGGCATTGGCCTTCTGATAGGCAGCCTGCACCTGAGGAGTGCGGCCGCTTTACTTGTACAGCTC
GTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGACCATGTGATCGCGCTTCTC
GTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAGCACGGGGCC
GTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCGAGCTGCACGCTGCCGTCCTCGATGTTGTG
GCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGATATAGACGTTGTGGCTG
TTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCT
CGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCTTGTAGTTGCCGTCGTC
CTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGTGCTG
CTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTCACGAGGGTGGG
CCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGTAGGTGGCATC
GCCCTCGCCCTCGCCGGACACGCTGAACTTGTGGCCGTTTACGTCGCCGTCCAGCTCGACCAGGAT
GGGCACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACCATGGTGGCGGGATCTGACGGTTCACT
AAACCAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCCATTTGCGTCAATGGGGCG
GAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTTGGTGCCAAAACAAACTCCCATTGACGTCAA
TGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTGATGTACTGCCAAA
ACCGCATCACCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATA
AGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTAC
TTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACG
TCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGG
GTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGGGCCTGCTGCCGGCTCT
GCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCG
CCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGACATGATAAGATACATTGA
TGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGC
TATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTT**(서열번호 43)**

LINE-1 플라스미드 CVBE IRES GFP**(서열번호 44)**

TAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGG
AATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAG
TCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAA
CTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTG
CATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCA
GGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGC
GCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGA
GGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGC
GTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGA
AAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCG
GTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGA
GAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGA
CGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGA
AGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGA
TTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAAC
TGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGA
TGTGAGACAGTTAAAACAGCCTGTGGGTTGATCCCACCCACAGGCCCATTGGGCGCTAGCACTCTG
GTATCACGGTACCTTTGTGCGCCTGTTTTATACCCCCTCCCCCAACTGTAACTTAGAAGTAACACAC
ACCGATCAACAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTG
AGTATCAATAGACTGCTCACGCGGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTCGAAA

[0712]

[0713]

AACCTAGTAACACCGTGGAAGTTGCAGAGTGTTTCGCTCAGCACTACCCCAGTGTAGATCAGGTCG
ATGAGTCACCGCATTCCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCCATGGGGAA
ACCCATGGGACGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGG
CCCCTGAATGCGGCTAATCCTAACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGTGTGTCGTAA
CGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCCTATACTGGCTG
CTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGACTAATAGA
GCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAACATTACAATTCA
TTGTTAAGTTGAATACAGCAAATACATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAAC
ATTAACGGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCA
AGCGTTTGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGA
TGGCGAAAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGA
CAAGACGGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAG
GCAGCATACAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCT
TTATCAAACAGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTT
CAATACACCATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGC
TCAATAGCGCACTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGA
GTACACATTCTTCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGC
TCTGCTGTCAAAGTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAA
GCTCGAGCTGAGAATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCT
GCTCAACGATTATTGGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGA
GAATAAGGATACTACCTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCAT
CGCCCTCAACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAA
GGAGTTGGAGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTC
GCGCCGAGTTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGT
TCTTCGAGCGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAG
AACCAGATTGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGAC
CACTATTCGGGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACAC
TTTTCTGGATACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATT
ACAGGAAGCGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGG
TTTACAGCTGAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGT
CTATAGAAAAGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACC
AGGACGCGATACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAAT
ATTGAACAAGATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGT
GGGGTTTATACCTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACAT
TAATAGAGCTAAGGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGAT
TCAGCAGCCATTTATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATT
CGCGCAATTTACGATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCG
CTCAAGACTGGAACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGC
TGGCTAGGGCTATTCGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAG
CTTTCCTTGTTCGCCGATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCT
TAAACTTATTTCTAACTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTT
CTGTACACAAATAATCGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGC
AAAAGGATAAAGTATCTCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTAC
AAGCCTCTCCTGAAAGAGATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTT
GGCAGAATCAACATAGTGAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCA
ATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAA
AAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCC
TGATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGA
CATCGACCAGTGGAATCGGACCGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATT
CGATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGA
ATTGGCTGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACA
GTAGGTGGATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGG
ATCACAATACAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACT
AAGGATAAGATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACG
ATCAGAGTTAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGG
TTGATAAGCAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAA
GAAGTGGGCAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATA
TGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACC
TTACCCCAGTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCG
AGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCT
GGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTA
TCCTAACGAATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTT
ACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTG

[0714]

GCATATTTATACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACC
TGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAAT
CTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACGACAAGTAAAGCGCTTCTAGAAGTTG
TCTCCTCCTGCACTGACTGACTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGAT
TACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACG
CTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAAT
CCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGT
GTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTC
GCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGG
CTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCT
CGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCA
GCGGACCTTCCTTCCCGCGAACAAACGACCCAACACCCGTGCGTTTTATTCTGTCTTTTTATTGCCG
ATCCCCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGA
TACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAG
CCAACGCTATGTCCTGATAGCGGTCGGCCGCTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCC
CGGCGGCGGTCACGAACTCCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGG
CGGACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCT
GCTGGTAGTGGTCGGCCAGGTGAGTCCAGGAGATGTTTCAGCACTGTTGCCTTTAGTCTCGAGGCA
ACTTAGACAACTGAGTATTGATCTGAGCACAGCAGGGTGTGAGCTGTTTGAAGATACTGGGGTTGG
GGGTGAAGAAACTGCAGAGGACTAACTGGGCTGAGACCCAGTGGCAATGTTTTAGGGCCTAAGGA
ATGCCTCTGAAAATCTAGATGGACAACTTTGACTTTGAGAAAAGAGAGGTGGAAATGAGGAAAAT
GACTTTTCTTTATTAGATTTCGGTAGAAAGAACTTTCATCTTTCCCCTATTTTTGTTATTCGTTTTAAA
ACATCTATCTGGAGGCAGGACAAGTATGGTCATTAAAAAGATGCAGGCAGAAGGCATATATTGGCT
CAGTCAAAGTGGGGAACTTTGGTGGCCAAACATACATTGCTAAGGCTATTCCTATATCAGCTGGAC
ACATATAAAATGCTGCTAATGCTTCATTACAAACTTATATCCTTTAATTCCAGATGGGGGCAAAGTA
TGTCCAGGGGTGAGGAACAATTGAAACATTTGGGCTGGAGTAGATTTTGAAAGTCAGCTCTGTGTG
TGTGTGTGTGTGTGTGTGTGAGAGCGTGTGTTTCTTTTAACGTTTTCAGCCTACAGCATACAGGG
TTCATGGTGGCAAGAAGATAACAAGATTTAAATTATGGCCAGTGACTAGTGCTGCAAGAAGAACA
ACTACCTGCATTTAATGGGAAAGCAAAATCTCAGGCTTTGAGGGAAGTTAACATAGGCTTGATTCT
GGGTGGAAGCTGGGTGTGTAGTTATCTGGAGGCCAGGCTGGAGCTCTCAGCTCACTATGGGTTCAT
CTTTATTGTCTCCTTTCATCTCAACAGCTGCACGCTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGT
TCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGATATAGACGTTGTGGCTGTTGTAGTTGTACTCC
AGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTCACCA
GGGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGT
GCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGG
GTAGCGGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTCACGAGGGTGGGCCAGGGCACGGGCA
GCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGTAGGTGGCATCGCCCTCGCCCTCGC
CGGACACGCTGAACTTGTGGCCGTTTACGTCGCCGTCCAGCTCGACCAGGATGGGCACCACCCCGG
TGAACAGCTCCTCGCCCTTGCTCACCATGGTGGCGAATTCGAAGCTTGAGCACGAGATCTGAGTCC
GGTAGGCCTAGCGGATCTGACGGTTCACTAAACCAGCTCTGCTTATATAGACCTCCCACCGTACAC
GCCTACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTTG
GTGCCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCG
CTATCCACGCCCATTGATGTACTGCCAAAACCGCATCACCATGGTAATAGCGATGACTAATACGTA
GATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTT
ACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGC
AGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAAC
ATACGTCATTATTGACGTCAATGGGCGGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGT
AAGTTATGTAACGGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACG
AGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTT
CAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAA
AAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAA
CAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTT
AAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCCCTT
GGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGCATT
GGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCA
ACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAG
GTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCTGG
ACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTTGG
GCTATCGAAACTTAATTAAGCGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTC
CGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTTTTT
ATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTT
GGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTTC
TCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGA

[0715]

TGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGT
GCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGC
GCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGG
CAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGC
GGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAG
CACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCG
CGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCC
ACGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGG
CCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCT
TGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATC
GCCTTCTATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATC
TGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAAT
AAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGC
AGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATG
GTTAATTAACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAA
TTGCGTCAGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAA
TCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC
ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCG
ACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCC
TGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACG
CTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT
CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA
TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGA
GTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTG
AAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGC
GGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTG
ATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGA
TTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTA
TATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG
TCTATTTCGTTCATCCATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAATCTTC
AAACCTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGC
CTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGA
TCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAATA
TCGAAATCGGGGCGCGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATAT
CGTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTG
TACTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCA
ACGTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGC
ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCT
CACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACAT
CGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGAT
GAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACT
CGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATCTT
ACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCC
AACTTACTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGAT
CATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGAC
ACCACGATGCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTA
GCTTCCCGGCAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCG
GCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCA
TTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAAC
CGATTCTAGGTGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACA
TATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCA
GAAATTTATCCTTAAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTA
CGCGAACAGCCGTGGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCT
TTTTGGCAGCGATCGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAG
TGGTCATAACAGCAGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCC
CAAGGGGTTATGCTATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATG
GATAGCGATTTTATTATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCAT
TAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC
GCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAC
TTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTAT
CATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAG
TACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGC

TGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTC
TCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCG
TAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG
AGCTGGTTTAGTGAACCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAG
CGGCCGC(서열번호 **44**)

LINE-1 플라스미드 EV71 IRES(서열번호 **45**)
TAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGG
AATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAG
TCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAA
CTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTG
CATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCA
GGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGC
GCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGA
GGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGC
GTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGA
AAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCG
GTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGA
GAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGA
CGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGA
AGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGA
TTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAAC
TGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGA
TGTGAGACAGTTAAAACAGCTGTGGGTTGTCACCCACCCACAGGGTCCACTGGGCGCTAGTACACT
GGTATCTCGGTACCTTTGTACGCCTGTTTTATACCCCCTCCCTGATTTGCAACTTAGAAGCAACGCA
AACCAGATCAATAGTAGGTGTGACATACCAGTCGCATCTTGATCAAGCACTTCTGTATCCCCGGAC
CGAGTATCAATAGACTGTGCACACGGTTGAAGGAGAAAACGTCCGTTACCCGGCTAACTACTTCGA
GAAGCCTAGTAACGCCATTGAAGTTGCAGAGTGTTTCGCTCAGCACTCCCCCCGTGTAGATCAGGT
CGATGAGTCACCGCATTCCCCACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCTATGGGG
TAACCCATAGGACGCTCTAATACGGACATGGCGTGAAGAGTCTATTGAGCTAGTTAGTAGTCCTCC
GGCCCCTGAATGCGGCTAATCCTAACTGCGGAGCACATACCCTTAATCCAAAGGGCAGTGTGTCGT
AACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCTTTTTATTCTTGTATTGGC
TGCTTATGGTGACAATTAAAGAATTGTTACCATATAGCTATTGGATTGGCCATCCAGTGTCAAACAG
AGCTATTGTATATCTCTTTGTTGGATTCACACCTCTCACTCTTGAAACGTTACACACCCTCAATTACA
TTATACTGCTGAACACGAAGCGTACATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAAC
ATTAACGGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCA
AGCGTTTGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGA
TGGCGAAAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGA
CAAGACGGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAG
GCAGCATACAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCT
TTATCAAACAGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTT
CAATACACCATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGC
TCAATAGCGCACTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGA
GTACACATTCTTCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGC
TCTGCTGTCAAAGTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAA
GCTCGAGCTGAGAATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCT
GCTCAACGATTATTGGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGA
GAATAAGGATACTACCTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCAT
CGCCCTCAACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAA
GGAGTTGGAGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTC
GCGCCGAGTTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGT
TCTTCGAGCGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAG
AACCAGATTGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGAC
CACTATTCGGGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACAC
TTTTCTGGATACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATT
ACAGGAAGCGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGG
TTTACAGCTGAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGT
CTATAGAAAAGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACC
AGGACGCGATACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAAT
ATTGAACAAGATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGT
GGGGTTTATACCTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACAT
TAATAGAGCTAAGGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGAT
TCAGCAGCCATTTATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATT

[0716]

[0717]

CGCGCAATTTACGATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCG
CTCAAGACTGGAACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGC
TGGCTAGGGCTATTCGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAG
CTTTCCTTGTTCGCCGATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCT
TAAACTTATTTCTAACTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTT
CTGTACACAAATAATCGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGC
AAAAGGATAAAGTATCTCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTAC
AAGCCTCTCCTGAAAGAGATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTT
GGCAGAATCAACATAGTGAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCA
ATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAA
AAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCC
TGATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGA
CATCGACCAGTGGAATCGGACCGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATT
CGATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGA
ATTGGCTGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACA
GTAGGTGGATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGG
ATCACAATACAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACT
AAGGATAAGATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACG
ATCAGAGTTAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGG
TTGATAAGCAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAA
GAAGTGGGCAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATA
TGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACC
TTACCCCAGTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCG
AGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCT
GGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTA
TCCTAACGAATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTT
ACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTG
GCATATTTATACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACC
TGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAAT
CTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACGACAAGTAAAGCGCTTCTAGAAGTTG
TCTCCTCCTGCACTGACTGACTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGAT
TACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACG
CTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAAT
CCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGT
GTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTC
GCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGG
CTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCT
CGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCA
GCGGACCTTCCTTCCCGCGAACAAACGACCCAACACCCGTGCGTTTTATTCTGTCTTTTTATTGCCG
ATCCCCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGA
TACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAG
CCAACGCTATGTCCTGATAGCGGTCGGCCGCTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCC
CGGCGGCGGTCACGAACTCCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGG
CGGACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCT
GCTGGTAGTGGTCGGCCAGGTGAGTCCAGGAGATGTTTCAGCACTGTTGCCTTTAGTCTCGAGGCA
ACTTAGACAACTGAGTATTGATCTGAGCACAGCAGGGTGTGAGCTGTTTGAAGATACTGGGGTTGG
GGGTGAAGAAACTGCAGAGGACTAACTGGGCTGAGACCCAGTGGCAATGTTTTAGGGCCTAAGGA
ATGCCTCTGAAAATCTAGATGGACAACTTTGACTTTGAGAAAAGAGAGGTGGAAATGAGGAAAAT
GACTTTTCTTTATTAGATTTCGGTAGAAAGAACTTTCATCTTTCCCCTATTTTTGTTATTCGTTTTAAA
ACATCTATCTGGAGGCAGGACAAGTATGGTCATTAAAAAGATGCAGGCAGAAGGCATATATTGGCT
CAGTCAAAGTGGGGAACTTTGGTGGCCAAACATACATTGCTAAGGCTATTCCTATATCAGCTGGAC
ACATATAAAATGCTGCTAATGCTTCATTACAAACTTATATCCTTTAATTCCAGATGGGGGCAAAGTA
TGTCCAGGGGTGAGGAACAATTGAAACATTTGGGCTGGAGTAGATTTTGAAAGTCAGCTCTGTGTG
TGTGTGTGTGTGTGTGTGTGAGAGCGTGTGTTTCTTTTAACGTTTTCAGCCTACAGCATACAGGG
TTCATGGTGGCAAGAAGATAACAAGATTTAAATTATGGCCAGTGACTAGTGCTGCAAGAAGAACA
ACTACCTGCATTTAATGGGAAAGCAAAATCTCAGGCTTTGAGGGAAGTTAACATAGGCTTGATTCT
GGGTGGAAGCTGGGTGTGTAGTTATCTGGAGGCCAGGCTGGAGCTCTCAGCTCACTATGGGTTCAT
CTTTATTGTCTCCTTTCATCTCAACAGCTGCACGCTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGT
TCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGATATAGACGTTGTGGCTGTTGTAGTTGTACTCC
AGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTCACCA
GGGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGT
GCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGG
GTAGCGGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTCACGAGGGTGGGCCAGGGCACGGGCA

[0718]

GCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGTAGGTGGCATCGCCCTCGCCCTCGC
CGGACACGCTGAACTTGTGGCCGTTTACGTCGCCGTCCAGCTCGACCAGGATGGGCACCACCCCGG
TGAACAGCTCCTCGCCCTTGCTCACCATGGTGGCGAATTCGAAGCTTGAGCACGAGATCTGAGTCC
GGTAGGCCTAGCGGATCTGACGGTTCACTAAACCAGCTCTGCTTATATAGACCTCCCACCGTACAC
GCCTACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTTG
GTGCCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCG
CTATCCACGCCCATTGATGTACTGCCAAAACCGCATCACCATGGTAATAGCGATGACTAATACGTA
GATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTT
ACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGC
AGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAAC
ATACGTCATTATTGACGTCAATGGGCGGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGT
AAGTTATGTAACGGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACG
AGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTT
CAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAA
AAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAA
CAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTT
AAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCCCTT
GGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGCATT
GGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCA
ACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAG
GTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCTGG
ACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTTGG
GCTATCGAAACTTAATTAAGCGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTC
CGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTTTTT
ATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTT
GGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTTC
TCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGA
TGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGT
GCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGC
GCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGG
CAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGC
GGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAG
CACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCG
CGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCC
ACGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGG
CCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCT
TGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATC
GCCTTCTATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATC
TGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAAT
AAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGC
AGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATG
GTTAATTAACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAA
TTGCGTCAGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAA
TCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC
ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCG
ACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCC
TGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACG
CTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT
CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA
TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGA
GTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTG
AAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGC
GGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTG
ATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGA
TTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTA
TATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG
TCTATTTCGTTCATCCATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAATCTTC
AAACCTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGC
CTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGA
TCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAATA
TCGAAATCGGGGCGCGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATAT
CGTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTG
TACTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCA

ACGTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGC
ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCT
CACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACAT
CGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGAT
GAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACT
CGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATCTT
ACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCC
AACTTACTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGAT
CATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGAC
ACCACGATGCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTA
GCTTCCCGGCAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCG
GCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCA
TTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAAC
CGATTCTAGGTGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACA
TATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCA
GAAATTTATCCTTAAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTA
CGCGAACAGCCGTGGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCT
TTTTGGCAGCGATCGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAG
TGGTCATAACAGCAGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCC
CAAGGGGTTATGCTATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATG
GATAGCGATTTTATTATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCAT
TAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC
GCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAC
TTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTAT
CATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAG
TACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGC
TGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTC
TCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCG
TAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG
AGCTGGTTTAGTGAACCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAG
CGGCCGC**(서열번호 45)**

LINE-1 플라스미드 ORF1-E2A-ORF2 GFP**(서열번호 46)**
TAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGGG
AATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAG
TCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAA
CTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTG
CATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCA
GGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGC
GCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGA
GGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGC
GTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGA
AAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCG
GTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGA
GAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGA
CGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGA
AGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGA
TTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAAC
TGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGA
TGGGAAGCGGACAGTGTACTAATTATGCTCTCTTGAAATTGGCTGGAGATGTTGAGAGCAACCCTG
GACCTATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAACGGCCTCAACTCAGC
TATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTTTGTTGCATCCAAGA
GACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGAAAGATTTATCAGGC
GAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGACGGATTTCAAGCCCA
CCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATACAGCAGGAAGAA
CTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAACAGGTCCTGTCCG
ATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACACCATTGAGCACCCT
GGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCACTGCATCAGG
CAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCTTCAGCGCCCC
ACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAAGTGCAAGCG
CACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGAATCAAGAA
CCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATTGGGTCCA

[0719]

[0720]

CAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTACCTATC
AAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTATAAAA
GAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAAACAGGAA
CAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGAT
CGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAA
GATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGATTGATACCATCA
AGAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCGGGAGTATTAT
AAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGATACCTATACT
CTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCGAGATTGTG
GCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGAGTTTTATC
AACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAGGAAGGCA
TCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCGATACCACAAA
GAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAAGATTCTCGC
CAACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATACCTGGCAT
GCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTAAGGATAA
GAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTTATGCT
CAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTACGATAA
GCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCCG
CCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCGTC
AAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGAT
GATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACTT
TAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAATCG
ACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCT
CGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGA
GATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGT
GAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACC
TTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCG
AAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATA
AAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGG
ACCGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAAT
AAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCG
GAAACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTT
GAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATACAAGATATAG
GCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAAGATTGATAAGT
GGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGTTAATAGGCAGC
CCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAGCAGAATATATA
ACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGGCAAAAGATATG
AACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAGTGTAGTTCAAG
CTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACCTTACCCCAGTGAGGATGGC
AATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGCACCCTGCTCCA
TTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTCTGAGGGACCTC
GAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGAATACAAGAGCT
GTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCTAAGACGTGGA
ATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTGGCATATTTATACCATGGAGT
ATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACCTGGATGAAGCTGGAGACTA
TTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTCATTGGTGGTA
ACGACTACAAAGACGATGACGACAAGTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGA
CTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATT
GACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATC
ATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATG
AGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCA
CTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCC
ACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGAC
AATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGA
TTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGA
ACAAACGACCCAACACCCGTGCGTTTTATTCTGTCTTTTTATTGCCGATCCCCTCAGAAGAACTCGT
CAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAA
GCGGTCAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATA
GCGGTCGGCCGCTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACT
CCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGT
AGTGGTTGTCGGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCCA
GGTGAGTCCAGGAGATGTTTCAGCACTGTTGCCTTTAGTCTCGAGGCAACTTAGACAACTGAGTATT
GATCTGAGCACAGCAGGGTGTGAGCTGTTTGAAGATACTGGGGTTGGGGGTGAAGAAACTGCAGA

[0721]

GGACTAACTGGGCTGAGACCCAGTGGCAATGTTTTAGGGCCTAAGGAATGCCTCTGAAAATCTAGA
TGGACAACTTTGACTTTGAGAAAAGAGAGGTGGAAATGAGGAAAATGACTTTTCTTTATTAGATTT
CGGTAGAAAGAACTTTCATCTTTCCCCTATTTTTGTTATTCGTTTTAAAACATCTATCTGGAGGCAG
GACAAGTATGGTCATTAAAAAGATGCAGGCAGAAGGCATATATTGGCTCAGTCAAAGTGGGGAAC
TTTGGTGGCCAAACATACATTGCTAAGGCTATTCCTATATCAGCTGGACACATATAAAATGCTGCTA
ATGCTTCATTACAAACTTATATCCTTTAATTCCAGATGGGGGCAAAGTATGTCCAGGGGTGAGGAA
CAATTGAAACATTTGGGCTGGAGTAGATTTTGAAAGTCAGCTCTGTGTGTGTGTGTGTGTGTGTG
TGTGAGAGCGTGTGTTTCTTTTAACGTTTTCAGCCTACAGCATACAGGGTTCATGGTGGCAAGAAGA
TAACAAGATTTAAATTATGGCCAGTGACTAGTGCTGCAAGAAGAACAACTACCTGCATTTAATGGG
AAAGCAAAATCTCAGGCTTTGAGGGAAGTTAACATAGGCTTGATTCTGGGTGGAAGCTGGGTGTGT
AGTTATCTGGAGGCCAGGCTGGAGCTCTCAGCTCACTATGGGTTCATCTTTATTGTCTCCTTTCATCT
CAACAGCTGCACGCTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTC
TGCTTGTCGGCCATGATATAGACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGT
TGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTT
CACCTCGGCGCGGGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCT
TCGGGCATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGC
ACGCCGTAGGTCAGGGTGGTCACGAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGAT
GAACTTCAGGGTCAGCTTGCCGTAGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTG
GCCGTTTACGTCGCCGTCCAGCTCGACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTT
GCTCACCATGGTGGCGAATTCGAAGCTTGAGCACGAGATCTGAGTCCGGTAGGCCTAGCGGATCTG
ACGGTTCACTAAACCAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCCATTTGCGTC
AATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTTGGTGCCAAAACAAACTCCCA
TTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTGATG
TACTGCCAAAACCGCATCACCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGA
AAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATA
GGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCC
ACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCA
ATGGGCGGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGGGCCTGC
TGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGC
CGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGACATGATAA
GATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAA
TTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTG
CATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTAC
AAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTT
GAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGCATTGGCGCAGAAAAAAATGCCT
GATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAA
CTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGGTCGTCAGCTATCCTGCAG
GCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCTGGACACCACTAGGGGTCAGAA
GTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTTGGGCTATCGAAACTTAATTAAG
CGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACT
CCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGG
CCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAA
AGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTTCTCCCGCCGCTTGGGTGGAGA
GGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGT
CAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAGG
ACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTG
TCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTC
ACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATC
CGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAA
GCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTC
GCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACGGCGATGCCTGCTTG
CCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCG
GACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT
GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCT
TGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCC
CCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGC
ATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGA
GGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAATTAACCAGTCAA
GTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCGTCAGTCAAGTTC
GATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCATGTGAGCAAAAGG
CCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC
TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGAT
ACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA

CCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTT
CGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCG
CCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAG
CCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGC
CTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG
AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGC
AAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCT
GACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTC
ACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGT
CTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCAT
AGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAATCTTCAAACCTTTCGTCCGATC
CATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGCCTTGCGCCTTGGCTATT
GCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGATCGGGATCACCCGAGA
GAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAATATCGAAATCGGGGCGCG
CCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATATCGTTGCTTGGCAGTCAG
CCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTGTACTCAAGCCTGGTCAC
GGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCAACGTATAATCGAGTCCT
AGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGCATGAGTATTCAACATTT
CCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGG
TGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACATCGAACTGGATCTCAACA
GCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGATGAGCACTTTTAAAGTTCT
GCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTA
TTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGT
AAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAAC
GATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGA
TCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAG
CAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAGTT
GATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTG
GTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCC
AGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACG
AAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCGATTCTAGGTGCATT
GGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCA
ACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAG
ATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAACAGCCGTGGC
GCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGGCAGCGATCGC
GGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCATAACAGCAGC
TTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTAT
CAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGATAGCGATTTTATTAT
CTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATAT
ATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGC
CCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGC
CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGG
ACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATGCGGTTTTGGCA
GTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGT
CAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCC
ATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAA
CCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCGC**(서열번호 46)**

LINE-1 플라스미드 ORF1-P2A-ORF2 GFP**(서열번호 47)**

TAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGG
AATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAG
TCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAA
CTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTG
CATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCA
GGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGC
GCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGA
GGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGC
GTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGA
AAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCG
GTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGA
GAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGA

[0722]

[0723]

CGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGA
AGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGA
TTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAAC
TGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGA
TGGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTGGA
CCTATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAACGGCCTCAACTCAGCTA
TCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTTTGTTGCATCCAAGAGA
CCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGAAAGATTTATCAGGCGA
ACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGACGGATTTCAAGCCCACC
AAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATACAGCAGGAAGAACT
TACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAACAGGTCCTGTCCGA
TCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACACCATTGAGCACCCTG
GATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCACTGCATCAGGC
AGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCTTCAGCGCCCCA
CATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAAGTGCAAGCGC
ACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGAATCAAGAAC
CTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATTGGGTCCAC
AATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTACCTATCA
AAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTATAAAAG
AAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGGAAC
AGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGATC
GAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAG
ATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGATTGATACCATCAA
GAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCGGGAGTATTATA
AGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGATACCTATACTCT
GCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCGAGATTGTGG
CCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGAGTTTTATCA
ACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAGGAAGGCAT
CTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCGATACCACAAAG
AAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAAGATTCTCGCC
AACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATACCTGGCATG
CAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTAAGGATAAG
AATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTTATGCTC
AAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTACGATAAG
CCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCCGC
CAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCGTC
AAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGAT
GATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACTT
TAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAATCG
ACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCT
CGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGA
GATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGT
GAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACC
TTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCG
AAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATA
AAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGG
ACCGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAAT
AAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCG
GAAACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTT
GAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATACAAGATATAG
GCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAAGATTGATAAGT
GGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGTTAATAGGCAGC
CCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAGCAGAATATATA
ACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGGCAAAAGATATG
AACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAGTGTAGTTCAAG
CTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACCTTACCCCAGTGAGGATGGC
AATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGCACCCTGCTCCA
TTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTCTGAGGGACCTC
GAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGAATACAAGAGCT
GTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCTAAGACGTGGA
ATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTGGCATATTTATACCATGGAGT
ATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACCTGGATGAAGCTGGAGACTA
TTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTCATTGGTGGTA

[0724]

ACGACTACAAAGACGATGACGACAAGTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGA
CTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATT
GACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATC
ATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATG
AGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCA
CTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCC
ACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGAC
AATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGA
TTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGA
ACAAACGACCCAACACCCGTGCGTTTTATTCTGTCTTTTTATTGCCGATCCCCTCAGAAGAACTCGT
CAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAA
GCGGTCAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATA
GCGGTCGGCCGCTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACT
CCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGT
AGTGGTTGTCGGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCCA
GGTGAGTCCAGGAGATGTTTCAGCACTGTTGCCTTTAGTCTCGAGGCAACTTAGACAACTGAGTATT
GATCTGAGCACAGCAGGGTGTGAGCTGTTTGAAGATACTGGGGTTGGGGGTGAAGAAACTGCAGA
GGACTAACTGGGCTGAGACCCAGTGGCAATGTTTTAGGGCCTAAGGAATGCCTCTGAAAATCTAGA
TGGACAACTTTGACTTTGAGAAAAGAGAGGTGGAAATGAGGAAAATGACTTTTCTTTATTAGATTT
CGGTAGAAAGAACTTTCATCTTTCCCCTATTTTTGTTATTCGTTTTAAAACATCTATCTGGAGGCAG
GACAAGTATGGTCATTAAAAAGATGCAGGCAGAAGGCATATATTGGCTCAGTCAAAGTGGGGAAC
TTTGGTGGCCAAACATACATTGCTAAGGCTATTCCTATATCAGCTGGACACATATAAAATGCTGCTA
ATGCTTCATTACAAACTTATATCCTTTAATTCCAGATGGGGGCAAAGTATGTCCAGGGGTGAGGAA
CAATTGAAACATTTGGGCTGGAGTAGATTTTGAAAGTCAGCTCTGTGTGTGTGTGTGTGTGTGTG
TGTGAGAGCGTGTGTTTCTTTTAACGTTTTCAGCCTACAGCATACAGGGTTCATGGTGGCAAGAAGA
TAACAAGATTTAAATTATGGCCAGTGACTAGTGCTGCAAGAAGAACAACTACCTGCATTTAATGGG
AAAGCAAAATCTCAGGCTTTGAGGGAAGTTAACATAGGCTTGATTCTGGGTGGAAGCTGGGTGTGT
AGTTATCTGGAGGCCAGGCTGGAGCTCTCAGCTCACTATGGGTTCATCTTTATTGTCTCCTTTCATCT
CAACAGCTGCACGCTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTC
TGCTTGTCGGCCATGATATAGACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGT
TGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTT
CACCTCGGCGCGGGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCT
TCGGGCATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGC
ACGCCGTAGGTCAGGGTGGTCACGAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGAT
GAACTTCAGGGTCAGCTTGCCGTAGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTG
GCCGTTTACGTCGCCGTCCAGCTCGACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTT
GCTCACCATGGTGGCGAATTCGAAGCTTGAGCACGAGATCTGAGTCCGGTAGGCCTAGCGGATCTG
ACGGTTCACTAAACCAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCCATTTGCGTC
AATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTTGGTGCCAAAACAAACTCCCA
TTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTGATG
TACTGCCAAAACCGCATCACCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGA
AAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATA
GGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCC
ACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCA
ATGGGCGGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGGGCCTGC
TGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGC
CGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGACATGATAA
GATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAA
TTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTG
CATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTAC
AAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTT
GAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGCATTGGCGCAGAAAAAAATGCCT
GATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAA
CTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGGTCGTCAGCTATCCTGCAG
GCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCTGGACACCACTAGGGGTCAGAA
GTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTTGGGCTATCGAAACTTAATTAAG
CGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACT
CCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGG
CCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAA
AGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTTCTCCCGCCGCTTGGGTGGAGA
GGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGT
CAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAGG
ACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTG

[0725]

TCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTC
ACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATC
CGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAA
GCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTC
GCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACGGCGATGCCTGCTTG
CCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCG
GACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT
GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCT
TGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCC
CCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGC
ATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGA
GGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAATTAACCAGTCAA
GTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCGTCAGTCAAGTTC
GATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCATGTGAGCAAAAGG
CCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC
TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGAT
ACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA
CCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTT
CGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCG
CCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAG
CCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGC
CTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG
AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGC
AAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCT
GACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTC
ACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGT
CTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCAT
AGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAATCTTCAAACCTTTCGTCCGATC
CATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGCCTTGCGCCTTGGCTATT
GCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGATCGGGATCACCCGAGA
GAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAATATCGAAATCGGGGCGCG
CCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATATCGTTGCTTGGCAGTCAG
CCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTGTACTCAAGCCTGGTCAC
GGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCAACGTATAATCGAGTCCT
AGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGCATGAGTATTCAACATTT
CCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGG
TGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACATCGAACTGGATCTCAACA
GCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGATGAGCACTTTTAAAGTTCT
GCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTA
TTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGT
AAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAAC
GATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGA
TCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAG
CAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAGTT
GATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTG
GTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCC
AGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACG
AAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCGATTCTAGGTGCATT
GGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCA
ACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAG
ATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAACAGCCGTGGC
GCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGGCAGCGATCGC
GGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCATAACAGCAGC
TTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTAT
CAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGATAGCGATTTTATTAT
CTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATAT
ATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGC
CCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGC
CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGG
ACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATGCGGTTTTGGCA
GTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGT
CAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCC

ATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAA
CCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCGC(서열번호 47)

LINE-1 플라스미드 ORF1-T2A ORF2 GFP(서열번호 48)

TAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGG
AATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAG
TCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAA
CTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTG
CATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCA
GGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGC
GCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGA
GGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGC
GTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGA
AAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCG
GTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGA
GAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGA
CGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGA
AGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGA
TTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAAC
TGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGA
TGGGAAGCGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTGACGTCGAGGAGAATCCTGGACCT
ATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAACGGCCTCAACTCAGCTATCA
AGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTTTGTTGCATCCAAGAGACCC
ACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGAAAGATTTATCAGGCGAACG
GTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGACGGATTTCAAGCCCACCAAA
ATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATACAGCAGGAAGAACTTAC
CATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAACAGGTCCTGTCCGATCT
GCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACACCATTGAGCACCCTGGAT
CGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCACTGCATCAGGCAGA
TCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCTTCAGCGCCCCACAT
CATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAAGTGCAAGCGCACA
GAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGAATCAAGAACCTG
ACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATTGGGTCCACAAT
GAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTACCTATCAAAA
CCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTATAAAAGAAA
ACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGGAACAGA
CCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGATCGAA
ACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAGATA
GACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGATTGATACCATCAAGAA
CGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCGGGAGTATTATAAGCA
TTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGATACCTATACTCTGCCA
CGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCGAGATTGTGGCCATA
ATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGAGTTTTATCAACGGT
ATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAGGAAGGCATCTTGCC
CAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCGATACCACAAAGAAGGA
AAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAAGATTCTCGCCAACAG
AATCCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATACCTGGCATGCAGGG
CTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTAAGGATAAGAATCA
TATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTTATGCTCAAGAC
TCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTACGATAAGCCGAC
TGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCCGCCAAGG
CTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCGTCAAGAG
AAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGATGATATG
ATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACTTTAGCAA
GGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAATCGACAGAC
CGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCTCGGAAT
CCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGAGATTAA
GGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGTGAAGAT
GGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACCTTCTTT
ACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCGAAGTC
CATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATAAAGCC
ACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGGACCGA

[0726]

[0727]

ACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAATAAACA
GTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAAACT
CAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTTGAATGT
CAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATACAAGATATAGGCGTCG
GCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAAGATTGATAAGTGGGAC
CTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGTTAATAGGCAGCCCACT
ACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAGCAGAATATATAACGAG
CTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGGCAAAAGATATGAACAG
GCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAGTGTAGTTCAAGCTTGGC
CATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACCTTACCCCAGTGAGGATGGCAATTAT
CAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGCACCCTGCTCCATTGCTG
GTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTCTGAGGGACCTCGAGCT
TGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGAATACAAGAGCTGTTGT
TACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCTAAGACGTGGAATCAG
CCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTGGCATATTTATACCATGGAGTATTAC
GCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACCTGGATGAAGCTGGAGACTATTATTC
TGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTCATTGGTGGTAACGACT
ACAAAGACGATGACGACAAGTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGATA
CAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTG
GTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCT
ATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGA
GTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT
TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGG
CGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATT
CCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCT
GCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGAACAA
ACGACCCAACACCCGTGCGTTTTATTCTGTCTTTTTATTGCCGATCCCCTCAGAAGAACTCGTCAAG
AAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGT
CAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGT
CGGCCGCTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGC
AGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGG
TTGTCGGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCCAGGTGA
GTCCAGGAGATGTTTCAGCACTGTTGCCTTTAGTCTCGAGGCAACTTAGACAACTGAGTATTGATCT
GAGCACAGCAGGGTGTGAGCTGTTTGAAGATACTGGGGTTGGGGGTGAAGAAACTGCAGAGGACT
AACTGGGCTGAGACCCAGTGGCAATGTTTTAGGGCCTAAGGAATGCCTCTGAAAATCTAGATGGAC
AACTTTGACTTTGAGAAAAGAGAGGTGGAAATGAGGAAAATGACTTTTCTTTATTAGATTTCGGTA
GAAAGAACTTTCATCTTTCCCCTATTTTTGTTATTCGTTTTAAAACATCTATCTGGAGGCAGGACAA
GTATGGTCATTAAAAAGATGCAGGCAGAAGGCATATATTGGCTCAGTCAAAGTGGGGAACTTTGGT
GGCCAAACATACATTGCTAAGGCTATTCCTATATCAGCTGGACACATATAAAATGCTGCTAATGCTT
CATTACAAACTTATATCCTTTAATTCCAGATGGGGGCAAAGTATGTCCAGGGGTGAGGAACAATTG
AAACATTTGGGCTGGAGTAGATTTTGAAAGTCAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGAG
AGCGTGTGTTTCTTTTAACGTTTTCAGCCTACAGCATACAGGGTTCATGGTGGCAAGAAGATAACA
AGATTTAAATTATGGCCAGTGACTAGTGCTGCAAGAAGAACAACTACCTGCATTTAATGGGAAAGC
AAAATCTCAGGCTTTGAGGGAAGTTAACATAGGCTTGATTCTGGGTGGAAGCTGGGTGTGTAGTTA
TCTGGAGGCCAGGCTGGAGCTCTCAGCTCACTATGGGTTCATCTTTATTGTCTCCTTTCATCTCAAC
AGCTGCACGCTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCT
TGTCGGCCATGATATAGACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCC
GTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACC
TCGGCGCGGGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCTTCGG
GCATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGC
CGTAGGTCAGGGTGGTCACGAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAAC
TTCAGGGTCAGCTTGCCGTAGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTGGCCG
TTTACGTCGCCGTCCAGCTCGACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTTGCTC
ACCATGGTGGCGAATTCGAAGCTTGAGCACGAGATCTGAGTCCGGTAGGCCTAGCGGATCTGACGG
TTCACTAAACCAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCCATTTGCGTCAATG
GGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTTGGTGCCAAAACAAACTCCCATTGA
CGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTGATGTACT
GCCAAAACCGCATCACCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGT
CCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGG
GCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCA
TTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGG
GCGGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGGGCCTGCTGCC
GGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCC

[0728]

TCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGACATGATAAGATA
CATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTG
TGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATT
CATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAA
TGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAG
GGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGCATTGGCGCAGAAAAAAATGCCTGAT
GCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTT
GCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGGTCGTCAGCTATCCTGCAGGCG
ATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCTGGACACCACTAGGGGTCAGAAGTA
GTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTTGGGCTATCGAAACTTAATTAAGCGA
TCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCG
CCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCG
CCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAG
GAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTTCTCCCGCCGCTTGGGTGGAGAGG
CTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCA
GCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAGGAC
GAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTC
ACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCAC
CTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCG
GCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCC
GGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCC
AGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACGGCGATGCCTGCTTGCCG
AATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGAC
CGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGAC
CGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGA
CGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCG
TGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATC
GCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGG
ATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAATTAACCAGTCAAGT
CAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCGTCAGTCAAGTTCGA
TCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCATGTGAGCAAAAGGCC
AGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTG
ACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAC
CAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACC
TGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTC
GGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGC
CTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGC
CACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCC
TAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG
AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGC
AAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCT
GACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTC
ACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGT
CTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCAT
AGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAATCTTCAAACCTTTCGTCCGATC
CATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGCCTTGCGCCTTGGCTATT
GCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGATCGGGATCACCCGAGA
GAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAATATCGAAATCGGGGCGCG
CCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATATCGTTGCTTGGCAGTCAG
CCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTGTACTCAAGCCTGGTCAC
GGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCAACGTATAATCGAGTCCT
AGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGCATGAGTATTCAACATTT
CCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGG
TGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACATCGAACTGGATCTCAACA
GCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGATGAGCACTTTTAAAGTTCT
GCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTA
TTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGT
AAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAAC
GATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGA
TCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAG
CAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAGTT
GATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTG
GTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCC

AGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACG
AAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCGATTCTAGGTGCATT
GGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCA
ACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAG
ATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAACAGCCGTGGC
GCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGGCAGCGATCGC
GGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCATAACAGCAGC
TTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTAT
CAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGATAGCGATTTTATTAT
CTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATAT
ATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGC
CCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGC
CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGG
ACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATGCGGTTTTGGCA
GTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGT
CAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCC
ATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAA
CCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCGC**(서열번호 48)**

LINE-1_ORF2-MCP_MS2_mRNA**(서열번호 49)**
TAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGG
AATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAG
TCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAA
CTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTG
CATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCA
GGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGC
GCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGA
GGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGC
GTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGA
AAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCG
GTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGA
GAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGA
CGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGA
AGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGA
TTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAAC
TGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGA
TGTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCATCAACTAATGAGCAAAATCACC
AGCTAACATCATAGTATACATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAAC
GGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTT
TGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGA
AAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGAC
GGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCAT
ACAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAA
ACAGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACA
CCATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAG
CGCACTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACA
TTCTTCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGT
CAAAGTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGC
TGAGAATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACG
ATTATTGGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAG
GATACTACCTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCA
ACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGG
AGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAG
TTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAG
CGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGAT
TGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCG
GGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGA
TACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAG
CGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGC
TGAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAA
AAGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCG

[0729]

[0730]

ATACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACA
AGATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTA
TACCTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAG
CTAAGGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGC
CATTTATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAAT
TTACGATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGAC
TGGAACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGG
GCTATTCGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTT
GTTCGCCGATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTT
ATTTCTAACTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACA
CAAATAATCGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGA
TAAAGTATCTCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTC
TCCTGAAAGAGATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAA
TCAACATAGTGAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACT
GCCTATGACCTTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGC
AAGAATAGCGAAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAA
GTTGTATTATAAAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCA
GTGGAATCGGACCGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGATAAGCC
AGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGG
CCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTGGA
TCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATA
CAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAA
GATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGT
TAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAG
CAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGG
CAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAG
TGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACCTTACCCCA
GTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGC
ACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTC
TGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGA
ATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCT
AAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTGGCATATTTAT
ACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACCTGGATGAAG
CTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTC
ATTGGTGGTAACGCTTCTAACTTTACTCAGTTCGTTCTCGTCGACAATGGCGGAACTGGCGACGTGA
CTGTCGCCCCAAGCAACTTCGCTAACGGGATCGCTGAATGGATCAGCTCTAACTCGCGTTCACAGG
CTTACAAAGTAACCTGTAGCGTTCGTCAGAGCTCTGCGCAGAATCGCAAATACACCATCAAAGTCG
AGGTGCCTAAAGGCGCCTGGCGTTCGTACTTAAATATGGAACTAACCATTCCAATTTTCGCCACGA
ATTCCGACTGCGAGCTTATTGTTAAGGCAATGCAAGGTCTCCTAAAAGATGGAAACCCGATTCCCT
CAGCAATCGCAGCAAACTCCGGCATCTACGCCATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGG
ACAACGGCGGCACCGGCGACGTGACCGTGGCCCCCAGCAACTTCGCCAACGGCATCGCCGAGTGG
ATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGACCTGCAGCGTGAGACAGAGCAGCGCCCA
GAACAGAAAGTACACCATCAAGGTGGAGGTGCCCAAGGGCGCCTGGAGAAGCTACCTGAACATGG
AGCTGACCATCCCCATCTTCGCCACCAACAGCGACTGCGAGCTGATCGTGAAGGCCATGCAGGGCC
TGCTGAAGGACGGCAACCCCATCCCCAGCGCCATCGCCGCCAACAGCGGCATCTACGACTACAAA
GACGATGACGACAAGTAAAGCAACCTACAAACGGGTGGAGGATCACCCCACCCGACACTTCACAA
TCAAGGGGTACAATACACAAGGGTGGAGGAACACCCCACCCTCCAGACACATTACACAGAAATCC
AATCAAACAGAAGCACCATCAGGGCTTCTGCTACCAAATTTATCTCAAAAAACTACAACAAGGAAT
CACCATCAGGGATTCCCTGTGCAATATACGTCAAACGAGGGCCACGACGGGAGGACGATCACGCC
TCCCGAATATCGGCATGTCTGGCTTTCGAATTCAGTGCGTGGAGCATCAGCCCACGCAGCCAATCA
GAGTCGAATACAAGTCGACTTTCGCGAAGAGCATCAGCCTTCGCGCCATTCTTACACAAACCACAC
TCTCCCCTACAGGAACAGCATCAGCGTTCCTGCCCAGTACCCAACTCAAGAAAATTTATGTCCCCAT
GCAGCATCAGCGCATGGGCCCCAAGAATACATCCCCAACAAAATCACATCCGAGCACCAACAGGG
CTCGGAGTGTTGTTTCTTGTCCAACTGGACAAACCCTCCATGGACCATCAGGCCATGGACTCTCACC
AACAAGACAAAAACTACTCTTCTCGAAGCAGCATCAGCGCTTCGAAACACTCGAGCATACATTGTG
CCTATTTCTTGGGTGGACGATCACGCCACCCATGCTCTCACGAATTTCAAAACACGGACAAGGACG
AGCACCACCAGGGCTCGTCGTTCCACGTCCAATACGATTACTTACCTTTCGGGATCACGATCACGG
ATCCCGCAGCTACATCACTTCCACTCAGGACATTCAAGCATGCACGATCACGGCATGCTCCACAAG
TCTCAACCACAGAAACTACCAAATGGGTTCAGCACCAGCGAACCCACTCCTACCTCAAACCTCTTC
CCACAAAACTGGCAAGCAGGATCACCGCTTGCCCATTCCAACATACCAAATCAAAAACAATTACTG
GTACAGCATCAGCGTACCAGCCCACATCTCTCACTACTATCAAAAACCAAACCGTTCAGCAACAGC
GAACGGTACACACGGAAAAATCAACTGGTTTACAAATACGAAAGACGATCACGCTTTCGTCCAGCG
CAAACTATTACGAAAAACATCCGACGGGAAGAGCAACAGCCTTCCCGCGGCGGAAAACCTCACAA

AAACACGACAAACGGATGCACGAACACGGCATCCGCCGACAACCCACAAACTTACAACCAGGCAA
ACGGTGCAGGATCACCGCACCGTACATCAAACACCTCAGATCTCATGCTTCTAGAAGTTGTCTCCTC
CTGCACTGACTGACTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAA
ATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTT
AATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTT
GCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCT
GACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCC
CCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT
GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGT
GTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACC
TTCCTTCCCGCTGAGAGACACAAAAAATTCCAACACACTATTGCAATGAAAATAAATTTCCTTTATT
AGCCAGAAGTCAGATGCTCAAGGGGCTTCATGATGTCCCCATAATTTTTGGCAGAGGGAAAAAGAT
CTCAGTGGTATTTGTGAGCCAGGGCATTGGCCTTCTGATAGGCAGCCTGCACCTGAGGAGTGCGGC
CGCTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGA
CCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTGTC
GGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCGAGCTGCACGCT
GCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATG
ATATAGACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGA
AGTCGATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGCGGG
TCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGG
ACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTAGGTCA
GGGTGGTCACGAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTC
AGCTTGCCGTAGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTGGCCGTTTACGTCGC
CGTCCAGCTCGACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACCATGGTGG
CGGGATCTGACGGTTCACTAAACCAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCC
ATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTTGGTGCCAAAACA
AACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCC
CATTGATGTACTGCCAAAACCGCATCACCATGGTAATAGCGATGACTAATACGTAGATGTACTGCC
AAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGA
CGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTA
AATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTAT
TGACGTCAATGGGCGGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAAC
GGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCC
CTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGAC
ATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATT
TGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTT**(서열번호 49)**

LINE1 ORF2-핑크 mRNA GFP**(서열번호 50)**

TAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGG
AATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAG
TCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAA
CTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTG
CATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCA
GGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGC
GCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGA
GGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGC
GTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGA
AAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCG
GTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGA
GAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGA
CGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGA
AGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGA
TTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAAC
TGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGA
TGTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCATCAACTAATGAGCAAAATCACC
AGCTAACATCATAGTATACATGGTCATAGGAACTTACATTTCGATTATTACCTTAAACGTGAATGGG
TTAAATGCCCCAACCAAGAGACATCGGCTGGCTGAATGGATTCAGAAACAGGACCCCTATATTTGC
TGTCTGCAGGAGACCCACTTCCGTCCTCGCGACACATACAGACTGAAAGTGAGGGGCTGGAAAAA
GATCTTCCATGCCAATGGAAATCAAAAGAAAGCTGGAGTGGCTATTCTCATCTCAGATAAAATTGA
CTTCAAAATAAAGAATGTTACTCGAGATAAGGAGGGACACTACATAATGATCCAGGGGTCCATCCA
AGAAGAGGATATAACTATTATTAATATTTATGCACCCAACATTGGCGCCCCTCAGTACATCAGGCA
GCTGCTTACAGCTATCAAGGAGGAAATCGACAGTAACACGATTATCGTGGGGGACTTTAACACCAG

[0731]

[0732]

CCTTACTCCGATGGATAGATCATCCAAAATGAAAATAAATAAGGAAACAGAGGCTCTTAATGACAC
CATTGACCAGATAGATCTGATTGATATATATAGGACATTCCATCCAAAAACTGCCGATTACACTTTC
TTCAGCAGTGCGCATGGAACCTTCTCCAGGATAGATCACATCTTGGGTCACAAAAGTAGCCTCAGT
AAGTTTAAGAAAATTGAAATCATTAGCAGCATCTTTTCTGACCATAACGCTATGCGCCTGGAGATG
AATCACAGGGAGAAGAACGTAAAGAAGACAAACACCTGGAGGCTGAACAATACGCTGCTAAATAA
CCAAGAGATCACTGAGGAAATCAAACAGGAAATAAAAAAATACTTGGAGACAAATGACAATGAAA
ACACGACCACCCAGAACTTGTGGGATGCAGCTAAAGCGGTTCTGAGAGGGAAGTTTATAGCTATTC
AAGCCTACCTTAAGAAACAGGAAAAATCTCAAGTGAACAATTTGACCTTACACCTAAAGAAACTGG
AGAAGGAGGAGCAGACCAAACCCAAAGTGAGCAGGAGGAAAGAAATCATCAAGATCAGAGCCGA
AATCAATGAAATAGAAACTAAGAAGACAATTGCCAAGATCAATAAAACTAAATCCTGGTTCTTTGA
GAAGATCAACAAAATTGATAAGCCATTAGCCAGACTCATCAAGAAAAAGAGGGAGAGGACTCAGA
TCAATAAGATCAGAAATGAGAAAGGGGAAGTTACAACCGACACCGCGGAGATTCAGAACATCCTG
AGAGACTACTACAAGCAACTTTATGCCAATAAAATGGACAACCTGGAAGAAATGGACAAATTCCT
GGAAAGGTATAACCTTCCCCGGCTGAACCAGGAGGAGACTGAAAATATCAACCGCCCAATCACAA
GTAATGAGATTGAGACTGTGATTAAGAATCTTCCAACTAACAAAAGTCCCGGCCCCGATGGCTTCA
CAGGTGAATTCTATCAGACCTTTCGGGAGGAGTTGACACCCATCCTTCTCAAGCTCTTCCAAAAAAT
TGCAGAGGAGGGCACACTCCCGAACTCATTCTATGAGGCCACCATCACCCTGATCCCAAAGCCCGA
CAAGGACACTACAAAGAAAGAAAATTACCGACCAATTTCCCTGATGAATATCGATGCCAAGATCCT
CAACAAAATCTTGGCAAACAGAATCCAGCAGCACATTAAGAGGATCATACACCACGATCAGGTGG
GCTTTATCCCGGGGATGCAAGGATTCTTCAATATCCGCAAATCAATCAATGTGATCCACCATATTAA
CAAGTTGAAGAAGAAGAACCATATGATCATCTCCATCGATGCAGAGAAAGCTTTTGACAAAATTCA
ACACCCATTTATGATCAAAACTCTCCAGAAGGTGGGCATCGAGGGGACCTACCTCAACATAATTAA
GGCCATCTATGATAAGCCCACAGCCAACATCATTCTCAATGGTGAAAAGCTGAAGGCATTTCCTCT
GCGGTCCGGAACGAGACAGGGATGTCCTCTCTCCTCTTCTGTTCAACATCGTTCTGGAAGTCCTA
GCCACCGCTATCCGCGAGGAAAAGGAAATTAAAGGCATACAGATTGGAAAGGAAGAGGTAAAACT
GTCTCTGTTTGCGGATGATATGATACTGTACATAGAGAATCCTAAAACTGCCACCCGGAAGCTGTT
GGAGCTAATTAATGAGTATGGTAAGGTCGCCGGTTACAAGATTAATGCTCAGAAGTCTCTTGCTTTC
CTGTACACTAATGATGAAAAGTCTGAACGGGAAATTATGGAGACACTCCCCTTTACCATTGCAACC
AAACGTATTAAATACCTTGGCATTAACCTGCCTAAGGAGACAAAAGACCTGTATGCTGAAAACTAT
AAGACACTGATGAAAGAGATTAAAGATGATACCAACCGGTGGCGGGATATCCCATGTTCTTGGATT
GGCAGAATCAACATTGTGAAGATGAGCATCCTGCCCAAGGCCATCTACAGATTCAATGCCATCCCT
ATCAAATTACCTATGGCATTTTTTACGGAGCTGGAACAGATCATCTTAAAATTTGTGTGGCGCCACA
AGCGGCCCCGAATCGCCAAAGCGGTCTTGAGGCAGAAGAATGGCGCTGGGGGAATCCGACTCCCT
GACTTCAGATTGTACTACAAAGCTACCGTCATCAAGACAATCTGGTACTGGCACAAGAACAGAAAC
ATCGATCAGTGGAACAAGATCGAAAGCCCTGAGATTAACCCCCGCACCTATGGTCAACTGATCTAT
GACAAAGGGGGCAAGGATATACAATGGCGCAAGGACAGCCTCTTCAATAAGTGGTGCTGGGAAAA
CTGGACAGCCACCTGCAAGCGTATGAAGCTGGAGTACTCCCTGACACCATACACAAAAATAAACTC
AAAGTGGATTCGAGACCTCAATATTCGGCTGGACACTATAAAACTCCTGGAGGAGAACATTGGGCG
TACACTCTTTGACATTAATCATAGCAAGATCTTTTTCGATCCCCCTCCTCGTGTAATGGAAATAAAA
ACAAAAATAAACAAGTGGGATCTGATGAAACTTCAGAGCTTTTGCACCGCAAAGGAGACCATAAA
CAAGACGAAGCGCCAACCCTCAGAATGGGAGAAAATATTTGCGAATGAGTCTACGGACAAAGGCT
TAATCTCCAAAATATATAAGCAGCTCATTCAGCTCAATATCAAGGAAACAAACACCCCGATCCAAA
AGTGGGCAGAGGACCTAAATCGGCATTTCTCCAAGGAAGACATCCAGACGGCCACGAAGCACATG
AAGCGATGCTCAACTTCCCTGATTATTCGCGAAATGCAGATCAAGACTACTATGCGCTATCACCTCA
CTCCTGTTCGGATGGGCATCATCCGGAAATCTACAAACAACAAGTGCTGGAGAGGGTGTGGCGAAA
AGGGAACCCTCTTGCATTGTTGGTGGGAGTGTAAGTTGATCCAGCCACTATGGCGGACCATATGGA
GGTTCCTTAAAAAACTGAAGATTGAGCTGCCATATGACCCAGCAATCCCACTGCTGGGCATATACC
CGGAGAAAACCGTGATTCAGAAAGACACTTGCACCCGAATGTTCATTGCAGCATTGTTTACAATAG
CCAGGTCATGGAAGCAGCCTAAGTGCCCCTCGACAGACGAGTGGATCAAGAAGATGTGGTACATTT
ATACTATGGAATATTACAGCGCCATCAAACGCAACGAAATTGGGTCTTTTCTGGAGACGTGGATGG
ATCTAGAGACTGTCATCCAGAGTGAGGTAAGTCAGAAAGAGAAGAACAAATATCGTATTTTAACGC
ATATTTGTGGAACCTGGAAGAATGGTACAGATGAGCCGGTCTGCCGAACCGAGATTGAGACCCAG
ATGGACTACAAAGACGATGACGACAAGTGAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTG
ACTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGA
TTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTA
TCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTA
TGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCC
CACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTG
CCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTG
ACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTG
GATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGC
TGAGAGACACAAAAAATTCCAACACACTATTGCAATGAAAATAAATTTCCTTTATTAGCCAGAAGT
CAGATGCTCAAGGGGCTTCATGATGTCCCCATAATTTTTGGCAGAGGGAAAAAGATCTCAGTGGTA

TTTGTGAGCCAGGGCATTGGCCTTCTGATAGGCAGCCTGCACCTGAGGAGTGCGGCCGCTTTACTTG
TACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGACCATGTGATCG
CGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAGC
ACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCGAGCTGCACGCTGCCGTCCTCG
ATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGATATAGACGT
TGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCC
CTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCTTGTAGTTG
CCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAG
TCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTCACG
AGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGTA
GGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTGGCCGTTTACGTCGCCGTCCAGCTCG
ACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACCATGGTGGCGGGATCTGAC
GGTTCACTAAACCAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCCATTTGCGTCAA
TGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTTGGTGCCAAAACAAACTCCCATTG
ACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTGATGTAC
TGCCAAAACCGCATCACCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAG
TCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGG
GGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCC
ATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATG
GGCGGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGGGCCTGCTGC
CGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGC
CTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGACATGATAAGAT
ACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTT
GTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTT**(서열번호 50)**

[0733]

[0734] **실시예 19. 안정적으로 역전위된 세포의 농후화**

[0735] 안정적으로 통합된 핵산 서열을 가진 세포 수율을 증가시키기 위한 노력으로, 본 실시예에 기재된 바와 같은 분류 및 배양 방법을 시도하였다. 293T 세포를 IVT에 의해 생성된 LINE1-GFP mRNA로 전기천공하고 적어도 3일 동안 시험관내에서 배양하였다. 도 40에 나타낸 바와 같이, 유세포분석을 이용하여 GFP의 발현을 주기적으로 측정하였다. 정량적 PCR을 이용하여 게놈당 게놈 통합을 평가하였다. GFP 및 하우스킵핑 유전자(FAU)에 대한 표준 곡선을 사용하여 게놈당 게놈에서 GFP를 코딩하는 핵산의 보간을 평가하였다. GFP 양성 세포의 분류 및 농후화 배양에서, 도 40에 나타낸 바와 같이, 다회 세포 계대배양(EP 후 적어도 18일) 동안 통합이 안정적이며 상당한 농후화가 가능하다는 것은 분명하였다. GFP 발현은 EP 후 5일째 날에 293T 세포의 약 1%에서 검출될 수 있었다. GFP+ 세포는 1차 분류 후 약 28%까지 농후화되었고, 2차 분류 후 약 74%의 세포까지 더 농후화되었다(도 40 및 도 41c).

[0736] 게놈 통합의 표준 곡선 및 예시적인 정량은 각각 도 41a 및 41b에 제시되어 있다. 도 41c는 10^3 유닛의 GFP 형광 강도 및 10^4 유닛의 GFP 형광 강도에서 게이팅될 때 게놈당 평균 GFP 통합 수를 보여준다.

[0737] **실시예 20. 증가된 트랜스포존 매개 통합을 위한 mRNA 농도의 적정**

[0738] 다양한 세포 유형, 즉 293T 세포, K562 및 THP-1 세포에서 세포당 최적 게놈 통합을 위해 전기천공에 사용된 LINE1-GFP mRNA의 농도를 적정하였다(도 42 내지 46). 100, 500, 1000, 1500 및 2000 ng/㎕의 mRNA를 GFP 발현 및 세포당 통합 수에 대해 시험하였다. 1000 ng/㎕보다 더 높은 농도는 세포 사멸을 야기한다. 도 42, 43 및 44에 나타낸 결과는 1000 ng/㎕가 역전위된 통합된 핵산에 의해 코딩된 GFP의 더 높은 장기간 발현을 야기함을 보여준다. 통합된 DNA 코딩된 단백질 발현은 3일째 날에 검출될 수 있기 시작하고 6일 내지 7일쯤에 피크에 도달한다(도 45). 그러나, K562 및 THP-1에서 LINE-1 GFP mRNA의 게놈 통합 및 발현은 매우 낮았고; 통합은 K562 세포에서 세포당 약 0.067 내지 0.155로 검출되었다(도 46)(THP-1 데이터는 표시되지 않음). 더 높은 LINE1-GFP mRNA 농도(1500 및 2000 ng/㎕)는 이 세포들에서 세포 사멸을 야기하였다. PD-0015 단핵구에서의 GFP mRNA 발현은 전기천공 후 3일째 날에 검출되었으며, 세포당 검출 가능한 통합이 있었다(도 47). 보다 더 광범위한 DNase 1 처리를 위한 단계를 취해야 했으며, 전기천공 전에 시험 mRNA 배치(batch)를 잔류 플라스미드에 대해 평가해야 했었다. 게놈에서 통합 수준의 측정의 정확도는 먼저 통합된 DNA 서열을 PCR로 농후화한 후, 페어링 말단 시퀀싱을 수행하여 게놈 내의 통합 부위를 맵핑함으로써 개선될 수 있었다. gDNA 추출 → 초음파처리에 의한 전단 → DNA 말단에 라이게이션된 DNA 링커 → 네스티드 PCR(1: 링커를 위한 하나의 프라이머, 통합된 DNA에 대한 두 번째 프라이머, 2: 추가된 일루미나(Illumina) 시퀀싱 어댑터) → 페어링 말단 시퀀싱을 포함하는 차세대 시퀀싱은 이러한 측면에서 최고의 표준으로 간주된다.

[0739] **실시예 21. 트랜스포존 매개 통합을 방해하는 후보의 녹다운에 의한 통합 효율의 향상**

[0740] 본 실시예에서, 녹다운이 GFP를 코딩하는 시험 핵산의 통합을 높일 수 있는지를 확인하기 위해 siRNA를 사용하여 다수의 내생성 후보들을 녹다운시켰다. 후보는 LINE1 역전위의 억제제인 ADAR1, ADAR2(ADAR1B), APOBEC3C, BRCA1, let-7 miRNA, RNase L, TASOR(HUSH 복합체), RAD51을 포함하였다. siRNA(표적 후보당 3개)를 제조하고 LINE1-GFP mRNA와 함께 시험 세포 내로 전기천공하고 유세포분석으로 LINE-1 GFP 발현의 변경에 대해 시험하고

qPCR로 그의 게놈 통합에 대해 시험하였고, 추가 적정을 위해 LINE-1 GFP 통합 및 발현을 증가시키는 데 도움이 되는 siRNA의 칵테일을 선택하였다. 시험된 다양한 siRNA들의 결과는 도 48 내지 51에 제시되어 있다. 개별적으로 시험된 ADAR1, BRCA 및 RNASEL의 녹다운은 LINE1-GFP 통합의 약 2배 증가를 유도하였다. ADAR2, RAD51 및 APOEBEC3C는 각각 1.5배 미만의 증가를 이끌어내었고, let7 miRNA 및 TASOR는 각각 증가를 이끌어내지 않았다. 도 48에 나타낸 연구에서, LINE-1 GFP(2000 ng/㎕)를 293 세포에서 100, 200 및 300 ng/㎕의 siBRCA와 함께 전기천공하였고, 전기천공한 지 4일 후에 얻은 데이터가 표시되어 있다. 100 ng/㎕를 사용한 경우, 통합률은 세포당 약 0.06 GFP 카피이었고, siBRCA1_s459(100 ng/㎕)는 통합을 약 2배 증가시킨다. 도 49에 나타낸 데이터는 전기천공 후 6일째 날에, siRNASEL 및 siADAR1 siRNA 각각이 개별적으로 통합을 약 2배 증가시켰음을 입증한다. 다른 한편으로, siAPOBEC3C_s2617은 전기천공 후 6일째 날에 GFP 통합을 1.5배 미만으로 증가시킨다(도 50).

표 11

게놈 통합률에 대한 특정 녹다운의 효과

표적	293T 세포에서 GFP 통합 배수 변화
ADAR1	약 2 배 증가
ADAR2	<1.5 배 증가
APOEBEC3C	<1.5 배 증가
BRCA	약 2 배 증가
Let7 miRNA	증가 없음
RNASEL	약 2 배 증가
TASOR(허쉬 복합체)	증가 없음
RAD51	<1.5 배 증가

[0741]

[0742] ADAR, APOEBEC3C, BRCA 및 RNASEL에 대한 siRNA를 siRNA 칵테일용으로 선택하였다. 두 세트의 실험에서 1000 ng/㎕ 및 1500 ng/㎕ LINE1-GFP mRNA를 사용하여 전기천공을 위한 siRNA의 농도를 적정하였다. 1500 ng/㎕의 LINE1-GFP mRNA는 약간 독성이 있는 것으로 관찰되었다(도 51). 1000 ng/㎕를 사용한 경우, 75 ng/㎕의 siRNA 각각은 293T 세포에서 GFP의 통합을 약 5배 개선하였다. 이 결과는 매우 고무적이며 추가 개발을 뒷받침한다. K562 세포에서의 유사한 실험으로부터 얻은 결과는 도 52에 표시되어 있다.

[0743] **실시예 20. 역전위 효율의 개선**

[0744] 역전위의 효율은 다수의 변수, 예컨대, 레트로트랜스포존 유전자 통합의 수준 및 통합된 유전자의 후속 전사와 번역 효율에 의해 영향을 받을 수 있다. 이 실시예에서, LINE-1 돌연변이, mRNA 서열 변경/추가, 및 mRNA 화학 반응의 변경과 같은 변수가 세포에서의 구축물의 역전위 효율 및 발현의 증가에 미치는 영향을 시험한다. 카고 유전자 통합을 개선하기 위해, 하기 예시적인 실험을 수행하여 세포에서의 구축물의 역전위 효율 및 발현의 증가에 미치는 이들의 영향을 시험할 것이다: (1) mRNA 변형, 예컨대, mRNA 안정성 및 단백질 발현을 증가시키는 mRNA 변형; (2) 서열 변형, 예컨대, 국소화 및 통합을 향상시키기 위한 핵 국소화 서열(NLS)의 추가 및 LINE-1 단백질 내로의 돌연변이의 도입; (3) 카고 발현을 개선하는 방법; 및 (4) 다양한 유기체들로부터의 다양한 역요소들에 대한 생물정보학적 분석. 이러한 실험으로부터 얻은 지식을, 골수 세포, T 세포, 간세포, 심근세포, 뉴런 및 망막 색소 상피 세포를 비롯한 다수의 인간 세포 유형에 맞게 적용하여, 다양한 세포 유형에 대한 이러한 접근법의 특이성과 다양성을 확인한다. 레트로트랜스포존의 생체내 전달도 마우스에서 수행될 것이다.

[0745] 이 실험에 포함된 예시적인 세포는 간세포, 심근세포, 망막 색소 상피 세포 및 뉴런을 포함한다. 1차 세포와 세포주가 사용될 것이고 각각의 세포 유형/세포주에 최적인 조건에서 배양될 수 있다. 시험 구축물은 본 명세서의 다른 부분에 설명된 바와 같이 레트로트랜스포존 복합체의 ORF1/2 유전자에 대해 역방향으로 삽입되는 프로모터 및 폴리 A 서열과 함께 GFP를 코딩하는 서열을 포함하는 플라스미드 구축물 및 mRNA 구축물을 포함한다. 플라스미드 구축물의 경우, 시험 구축물은 센스 방향의 인트론에 의해 불연속된 안티센스 방향의 GFP 유전자를 함유하

는 구축물을 포함한다. mRNA 구축물의 경우, 시험 구축물은 인트론이 없는 구축물을 포함한다. 역전위 조건은 전기천공 및/또는 형질감염 시약을 통해 시험관내에서 세포로 전달되는 GFP-레트로트랜스포존과 함께, 통합을 향상시키는 시약을 포함할 수 있다. 각각의 특정 세포 유형에 최적화된 전기천공 조건이 사용된다. 예를 들어, SSC 및 FSC를 기반으로 설정된 게이트를 사용하여 유세포분석으로 통합의 효율을 측정한다. 유세포분석을 위한 게이트를 설정하기 위해 음성 대조군 및 양성 대조군에서 GFP를 측정한다. 사용된 예시적인 음성 대조군은 <0.1% GFP+로 설정될 것이다. 사용된 예시적인 양성 대조군은 >90% GFP+로 설정될 것이다. 유세포분석을 위한 게이트를 설정하기 위해 음성 대조군 및 양성 대조군에서 GFP를 측정한다. 사용된 예시적인 음성 대조군은 <0.1% GFP+로 설정될 것이다. 사용된 예시적인 양성 대조군은 >90% GFP+로 설정될 것이다. 성공적인 통합을 표시하는 예시적인 측정은 유세포분석에 의해 측정될 때 역전위 후 10일째 날까지 세포의 >2%에서 GFP 발현을 측정하는 것이다. 성공적인 통합을 표시하는 또 다른 예시적인 측정은 유세포분석에 의해 측정될 때 역전위 후 10일째 날까지 세포의 >10%에서 GFP 발현을 측정하는 것이다.

[0746] 예를 들어, 통합을 확인하기 위해 PCR 및 차세대 시퀀싱(NGS)과 같은 추가 어세이를 수행한다. 형질감염/전기천공된 세포에 대해 차세대 시퀀싱을 수행할 수 있다.

[0747] 안정성 및 단백질 발현을 증가시키는 변형이 시험될, mRNA 안정성 및 번역 효율에 직접적으로 영향을 미칠 수 있는 예시적인 mRNA 구축물은 5' CAP, 폴리(A) 꼬리 및 비번역 영역(UTR)을 포함한다. 예를 들어, 5' CAP의 효소 통합 및 7-메틸구아노신과 같은 CAP 유사체의 사용이 시험될 것이다. 5' CAP의 효소 통합은 CAP 유사체를 사용하는 것보다 효율적일 수 있다. 예를 들어, 폴리(A) 꼬리 길이가 번역 효율에 영향을 미치는지를 확인하기 위해 폴리(A) 꼬리 길이의 변경도 시험될 것이다. 예를 들어, GC 함량이 번역 효율에 영향을 미치는지를 확인하기 위해 mRNA의 퍼센트 GC 함량의 변경이 시험될 것이다. 높은 GC 함량은 mRNA 2차 구조 안정성을 증가시킬 수 있다. 낮은 GC 함량은 mRNA의 번역성을 감소시킬 수 있다. 서열 특이적 mRNA 변경은 최적 GC 함량에 도달하기 위해 시험될 수 있다. 또 다른 예시적인 방법에서, 특정 코돈을 희귀 코돈으로 치환하고 변형 뉴클레오타이드를 삽입하는 것이 번역 효율에 영향을 미칠 수 있는지를 확인하기 위해 수행될 것이다. 안정성 및 단백질 발현을 증가시키기 위해 시험될 다른 예시적인 뉴클레오타이드 변형은 5-메틸사이티딘(m5C) 및 슈도우리딘(Ψ) 뉴클레오타이드의 사용을 포함한다. 성공적인 통합을 표시하는 예시적인 측정은 유세포분석에 의해 측정될 때 역전위 후 10일째 날까지 세포의 >2%에서 GFP 발현을 측정하는 것이다. 성공적인 통합을 표시하는 또 다른 예시적인 측정은 유세포분석에 의해 측정될 때 역전위 후 10일째 날까지 세포의 >10%에서 GFP 발현을 측정하는 것이다.

[0748] 전체 길이 mRNA의 증가된 전사에 대해 다양한 mRNA 생성 과정도 시험될 것이다.

[0749] 예컨대, 향상된 국소화 및 통합에 대해 시험될 예시적인 NLS 서열 변형은 다양한 NLS 서열, LINE-1 요소의 상류 또는 하류에의 NLS 서열의 배치, 및 사용된 NLS 서열의 수를 포함한다. 전술된 바와 같이, 유세포분석을 위한 게이트를 설정하기 위해 음성 대조군 및 양성 대조군에서 GFP를 측정한다. 전술된 바와 같이, 사용된 예시적인 음성 대조군은 <0.1% GFP+로 설정될 것이다. 전술된 바와 같이, 사용된 예시적인 양성 대조군은 >90% GFP+로 설정될 것이다. 성공적인 카고 유전자 발현을 표시하는 예시적인 측정은 유세포분석에 의해 측정될 때 역전위 후 10일째 날까지 세포의 >10%에서 증가된 GFP 발현을 측정하는 것이다.

[0750] 예컨대, 향상된 국소화 및 통합에 대해 시험될 LINE-1 단백질의 돌연변이를 코딩하는 예시적인 서열 변형은 ORF1 단백질의 K3R 및 ORF2의 Y1180A를 포함한다. 전술된 바와 같이, 유세포분석을 위한 게이트를 설정하기 위해 음성 대조군과 양성 대조군에서 GFP를 측정한다. 전술된 바와 같이, 사용된 예시적인 음성 대조군은 <0.1% GFP+로 설정될 것이다. 전술된 바와 같이, 사용된 예시적인 양성 대조군은 >90% GFP+로 설정될 것이다. 성공적인 카고 유전자 발현을 표시하는 예시적인 측정은 유세포분석에 의해 측정될 때 역전위 후 10일째 날까지 세포의 >10%에서 증가된 GFP 발현을 측정하는 것이다.

[0751] **실시예 21. 인간 LINE1 구축물에 의한 역전위에서 핵 국소화 신호 도입의 효과**

[0752] 이 실시예에서, 핵 국소화 신호가 이하 기재된 바와 같이 다양한 위치에 삽입되어 있는 여러 구축물들을 생성하고 시험하였고, 역전위 활성을 HEK293T 세포에서 시험하였다. 하나 이상의 다양한 NLS 및 본원에 기재된 레트로트랜스포존 서열을 포함하는 상이한 수의 전체 NLS를 함유하는 mRNA 구축물을 293T 세포 내로 전기천공하고, 세포의 게놈 내로의 시험 유전자의 삽입 수를 조사하였다. NLS를 도 53에 요약된 바와 같이 LINE1-GFP mRNA의 하기 위치에서 개별적으로 시험하였다: (i) NLS가 LINE1-GFP 구축물의 ORF1을 코딩하는 서열의 N-말단에서 삽입되어 있는 ORF1-N(도 54a); (ii) NLS가 LINE1-GFP 구축물의 ORF1을 코딩하는 서열의 C-말단에서 삽입되어 있는 ORF1-C(도 55a), (iii) NLS가 LINE1-GFP 구축물의 ORF2를 코딩하는 서열의 N-말단에서 삽입되어 있는 ORF2-N(도 56a); (iv) SV40 NLS 및 링커가 LINE1-GFP 구축물의 ORF2를 코딩하는 서열의 N-말단에서 삽입되어 있는

ORF2-N(도 57a); 및 (v) NLS가 LINE1-GFP 구축물의 ORF2를 코딩하는 서열의 C-말단에서 삽입되어 있는 ORF2-C(도 58a). 각각의 경우, LINE GFP NLS 구축물을 사용한 역전위를 세포(게놈)당 GFP 삽입의 수로 측정하였다.

[0753] 표 12A는 도 54a 내지 55c에 대해 설명된 시험의 구축물과 관련된 서열을 상세히 설명한다.

[0754] [표 12A]

NLS	**아미노산 서열 및 핵산 서열**
(#3) SV40 NLS	PKKKRKV(**서열번호** 80) ccaaagaagaagcggaaggtc(**서열번호** 81)
(#4) SV40 NLS + **링커(밑줄로 표시된 링커 서열)**	PKKKRKVGGGS(**서열번호** 82) ccaaagaagaagcggaaggtcggcggcggcagc(**서열번호** 83)
(#5) **뉴클레오플라스민** NLS	KRPAATKKAGQAKKKK(**서열번호** 84) aaaaggccggcggccacgaaaaaggccggccaggcaaaaaagaaaaag(**서열번호** 85)
(#6) **뉴클레오플라스민** NLS + **링커(밑줄로 표시된 링커 서열)**	KRPAATKKAGQAKKKKGGGS(**서열번호** 86) aaaaggccggcggccacgaaaaaggccggccaggcaaaaaagaaaaagggcggcggcagc(**서열번호** 87)

[0755]

[0756] 도 54b 및 54c, 및 도 55b 및 55c에 나타낸 바와 같이, ORF1 서열의 N-말단 및 C-말단에서의 NLS 삽입은 LINE-1 구축물의 역전위 활성을 제거한다. 다른 한편으로, 도 56a 내지 58c에 나타낸 구축물들 각각은 ORF2 N-말단 또는 C-말단에서의 NLS의 삽입에 의한 역전위의 일부 개선을 보여주었다.

[0757] 표 12B는 도 56 내지 56c에 대해 설명된 시험의 구축물과 관련된 서열을 상세히 설명한다.

[0758] [표 12B]

NLS	**아미노산 서열 및 핵산 서열**
(#3) SV40 NLS	PKKKRKV(**서열번호** 80) ccaaagaagaagcggaaggtc(**서열번호** 81)
(#4) **뉴클레오플라스민** NLS	KRPAATKKAGQAKKKK(**서열번호** 84) aaaaggccggcggccacgaaaaaggccggccaggcaaaaaagaaaaag(**서열번호** 85)
(#5) **링커** + **뉴클레오플라스민** NLS **(밑줄로 표시된 링커 서열)**	GGGSKRPAATKKAGQAKKKK(서열번호 88) ggcggcggcagcaaaaggccggcggccacgaaaaaggccggccaggcaaaaaagaaaaag(**서열번호** 89)

[0759]

[0760] ORF2 서열의 N-말단에서의 NLS의 삽입은 세포당 삽입 수를 약 2배 증가시킨다(도 56b 및 56c). 그러나, 표 12C에 표시된 링커 서열이 추가된 SV40 NLS의 사용은 GFP 통합의 수를 약 3.5배 증가시킨다(도 57b 및 57c).

[0761] 표 12C는 도 57 내지 57c에 대해 설명된 시험의 구축물과 관련된 서열을 상세히 설명한다.

[0762] [표 12C]

NLS	**아미노산 서열 및 핵산 서열**
(#4a) SV40 NLS + **링커** **(밑줄로 표시된 링커 서열)**	PKKKRKVGGGS(**서열번호** 82) ccaaagaagaagcggaaggtcggcggcggcagc(**서열번호** 83)

[0763]

[0764] 그러나, 놀랍게도, ORF2 C-말단에서의 NLS의 삽입은 역전위를 약 5배 증가시킨다(도 58b 및 58c).

[0765] 표 12D는 도 58a 내지 58c에 대해 설명된 시험의 구축물과 관련된 서열을 상세히 설명한다.

[0766] [표 12D]

NLS	**아미노산 서열 및 핵산 서열**
(#3) SV40 NLS	PKKKRKV(**서열번호** 80) ccaaagaagaagcggaaggtc(**서열번호** 81)
(#4) **링커** + SV40 NLS	GGGSPKKKRKV(**서열번호** 90) ggcggcggcagcccaaagaagaagcggaaggtc(**서열번호** 91)
(#5) **뉴클레오플라스민** NLS	KRPAATKKAGQAKKKK(**서열번호** 84) aaaaggccggcggccacgaaaaaggccggccaggcaaaaaagaaaaag(**서열번호** 85)
(#6) **링커** + **뉴클레오플라스민** NLS	GGGSKRPAATKKAGQAKKKK(**서열번호** 88) ggcggcggcagcaaaaggccggcggccacgaaaaaggccggccaggcaaaaaagaaaaag(**서열번호** 89)

[0767]

[0768] SV40 NLS 및 뉴클레오플라스민 NLS 둘 다를 ORF2의 C-말단에 삽입하면 NLS 서열이 없는 LINE1-GFP mRNA 구축물로 전기천공한 세포의 2% 미만에 비해 >5% GFP 양성 세포가 확인된다. 이 결과는 NLS를 사용한 ORF2 변형이 역전위 효율을 증가시킨다는 것을 시사한다. 가장 유의미한 점은 C-말단에서의 삽입이 LINE1 요소의 역전위의 더 높은 증가를 제공한다는 것이다.

[0769] 이 실험에 사용된 예시적인 플라스미드는 표 13에 제시되어 있다.

표 13

예시적인 플라스미드 서열

명칭	서열
>RET-002 LINE1-GFP ORF1-CSV40-NLS	CTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGCCTTG CGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGATCGG GATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAATATCGAA ATCGGGGCGCGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATATCGTTGC TTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTGTACTCAA GCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCAACGTATA ATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGCATGAGTAT TCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCA GAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACATCGAACTGG ATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGATGAGCACTTT TAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGC ATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATCTTACGGATGGCA TGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCT GACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGC CTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTG TAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACA GTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGC TGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGC CAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACG AAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCGATTCTAGGTGCATTG GCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAAC GGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGAT CGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAACAGCCGTGGCGC TCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGGCAGCGATCGCGG CTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCATAACAGCAGCTT CAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTATCAA TCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGATAGCGATTTTATTATCTA ACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATG GAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCAT TGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGT GGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCT ATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTC CTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATGCGGTTTTGGCAGTACAT CAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGG AGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGC AAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGAT CAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCGCTAATACGACTCACTATA GGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGGAATTCCAAGACACAATCCG CTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAGTCCTGGATGGAAAACGACTT CGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAACTCCGCGAAGATATCCAGACA AAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTGCATCACCCGTATCACAAACACTG AGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCAGGGAGCTTCGAGAGGAGTGTCGGAG TCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGCGCGATGGAAGACGAGATGAACGAGATG AAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGAGGAACGAACAGAGTCTGCAGGAGATTTGGG ATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGCGTCCCCGAGAGCGACGTAGAAAACGGGACTAA ACTGGAGAATACACTTCAAGACATCATTCAAGAAAATTTTCCAAACCTGGCTCGGCAAGCTAATGTG CAAATCCAAGAGATCCAACGCACACCCCAGCGGTATAGCTCTCGGCGTGCCACCCCTAGGCATATTA TCGTGCGCTTTACTAAGGTGGAGATGAAAGAGAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGT GACTTTGAAGGGCAAACCTATTCGGCTGACGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGG GAATGGGGCCCCATCTTTAATATCCTGAAGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAA AGTTGAGTTTTATCTCCGAGGGTGAGATTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGT GACAACTCGCCCAGCTCTCAAGGAACTGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATAT CAACCCTTGCAGAACCACGCAAAGATGccaaagaagaagcggaaggtcTGAGACAGCCGTCAGACCA TCAAGACTAGGAAGAAACTGCATCAACTAATGAGCAAAATCACCAGCTAACATCATAGTATACATGA CCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAACGGCCTCAACTCAGCTATCAAGCG CCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTTTGTTGCATCCAAGAGACCCACCTG ACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGAAAGATTTATCAGGCGAACGGTAAGC AGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGACGGATTTCAAGCCCACCAAAATTAAGCG TGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATACAGCAGGAAGAACTTACCATATTGAAC ATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAACAGGTCCTGTCCGATCTGCAGCGAGATC

[0770]

[0771]	TGGATTCTCATACGTTGATTATGGGTGATTTCAATACACCATTGAGCACCCTGGATCGCAGCACCAG GCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCACTGCATCAGGCAGATCTCATTGATATT TATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCTTCAGCGCCCCACATCATACATACTCAA AGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAAGTGCAAGCGCACAGAGATAATTACAAA TTACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGAATCAAGAACCTGACCCAGAGCCGGAGT ACCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATTGGGTCCACAATGAGATGAAGGCAGAGA TTAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTACCTATCAAAACCTTTGGGATGCCTTTAA GGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTATAAAAGAAAACAAGAGAGATCTAAGATC GATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGGAACAGACCCACTCCAAGGCGTCAAGAC GGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAAT TAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAGATAGACAGACCTCTGGCACGACTGATT AAGAAGAAGCGCGAAAAGAACCAGATTGATACCATCAAGAACGACAAGGGCGACATCACTACTGACC CGACCGAGATCCAGACCACTATTCGGGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCT GGAAGAGATGGACACTTTTCTGGATACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCC CTCAACCGCCCAATTACAGGAAGCGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTC CTGGTCCGGACGGGTTTACAGCTGAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCT CAAGCTCTTTCAGTCTATAGAAAAGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATA CTTATTCCCAAACCAGGACGCGATACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATA TCGACGCTAAAATATTGAACAAGATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACA TCACGACCAGGTGGGGTTTATACCTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTC ATTCAACACATTAATAGAGCTAAGGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCAT TCGATAAGATTCAGCAGCCATTTATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTT TAAGATTATTCGCGCAATTTACGATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAG GCCTTTCCGCTCAAGACTGGAACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTAC TCGAGGTGCTGGCTAGGGCTATTCGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGA GGTCAAGCTTTCCTTGTTCGCCGATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAG AACCTTCTTAAACTTATTTCTAACTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTC AGGCCTTTCTGTACACAAATAATCGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCAT AGCCAGCAAAAGGATAAAGTATCTCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAA AATTACAAGCCTCTCCTGAAAGAGATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCAT GGGTTGGCAGAATCAACATAGTGAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCAT CCCAATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAAT CAAAAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGC CTGATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGA CATCGACCAGTGGAATCGGACCGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTC GATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATT GGCTGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAG GTGGATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACA ATACAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATA AGATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGT TAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAGC AGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGGCAA AAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAGTGTAG TTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACCTTACCCCAGTGAGG ATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGCACCCTGC TCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTCTGAGGGA CCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGAATACAAG AGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCTAAGACGT GGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTGGCATATTTATACCATGGA GTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACCTGGATGAAGCTGGAGACT ATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTCATTGGTGGTA ACGACTACAAAGACGATGACGACAAGTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGAC TGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATTG ACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATC ATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTA TGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCC ACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTG CCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGA CAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGG ATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCt gagagacacaaaaaattccaacacactattgcaatgaaaataaatttcctttattagccagaagtca gatgctcaaggggcttcatgatgtccccataatttttggcagagggaaaaagatctcagtggtattt gtgagccagggcattggccttctgataggcagcctgcacctgaggagtgcggccgctttacttgtac agctcgtccatgccgagagtgatcccggcggcggtcacgaactccagcaggaccatgtgatcgcgct

	tctcgttggggtctttgctcagggcggactgggtgctcaggtagtggttgtcgggcagcagcacggg gccgtcgccgatgggggtgttctgctggtagtggtcggcgagctgcacgctgccgtcctcgatgttg tggcggatcttgaagttcaccttgatgccgttcttctgcttgtcggccatgatatagacgttgtggc tgttgtagttgtactccagcttgtgccccaggatgttgccgtcctccttgaagtcgatgcccttcag ctcgatgcggttcaccagggtgtcgccctcgaacttcacctcggcgcgggtcttgtagttgccgtcg tccttgaagaagatggtgcgctcctggacgtagccttcgggcatggcggacttgaagaagtcgtgct gcttcatgtggtcggggtagcggctgaagcactgcacgccgtaggtcagggtggtcacgagggtggg ccagggcacgggcagcttgccggtggtgcagatgaacttcagggtcagcttgccgtaggtggcatcg ccctcgccctcgccggacacgctgaacttgtggccgtttacgtcgccgtccagctcgaccaggatgg gcaccaccccggtgaacagctcctcgcccttgctcaccatggtggcgggatctgacggttcactaaa ccagctctgcttatatagacctcccaccgtacacgcctaccgcccatttgcgtcaatggggcggagt tgttacgacattttggaaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggg gtggagacttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaaccgc atcaccatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcccataaggtca tgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaataggggggcgtacttggcat atgatacacttgatgtactgccaagtgggcagtttaccgtaaatactccacccattgacgtcaatgg aaagtccctattggcgttactatgggaacatacgtcattattgacgtcaatgggcgggggtcgttgg gcggtcagccaggcgggccatttaccgtaagttatgtaacgGGCCTGCTGCCGGCTCTGCGGCCTCT TCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAG CTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGACATGATAAGATACATTGATGAGTTTGGA CAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTAT TTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGT TCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCCCATCT CTATCGGTATCGTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTGCCCCTC GGGCCGGATTGCTATCTACCGGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTAT ACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCC TCCTTACCAGAAATTTATCCTTAAGGTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGATCAAG ACATTCCTTTAATGGTCTTTTCTGGACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTCTTCCC TCCCTAATCTCATTGGTTACCTTGGGCTATCGAAACTTAATTAAGCGATCTGCATCTCAATTAGTCA GCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTC CGCCCCATCGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATT CCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATGATTGA ACAAGATGGATTGCACGCAGGTTCTCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCA CAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTT TTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCT GGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTG CTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCA TCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGC GAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGAC GAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCG AGGATCTCGTCGTGACCCACGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTC TGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGT GATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTC CCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCTGTGCC TTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACT CCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC TGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGA TGCGGTGGGCTCTATGGTTAATTAACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCTGGGTT TCGACTACGCTCAGAATTGCGTCAGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCTCCGAT TACGAGTTTCATTTAAATCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCG TTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGA GGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTC TCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTT TCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGC ACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGT AAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGC GGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCT GCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC CGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAA GATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGG TCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCA

[0772]

	GCGATCTGTCTATTTCGTTCATCCATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGG AAAATCTTCAAAC(서열번호 92)
>ret-003-line1-gfp-orf1-c-링커_sv40-nls	CCTCTCAGTGCGAGTCTCGACGATCCATATCGTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGG GACCCAGGAAGTCCAATCGTCAGATATTGTACTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTA AACCTAGATATTGATAGTCTGATCGGTCAACGTATAATCGAGTCCTAGCTTTTGCAAACATCTATCA AGAGACAGGATCAGCAGGAGGCTTTCGCATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTT TTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAG ATCAGTTGGGTGCGCGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTT TCGCCCCGAAGAACGCTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCC CGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGT ATTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCAT AACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATTGGAGGACCGAAGGAGCTAACC GCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAG CCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACCTTGCGTAAACTATT AACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAGTTGATAGACTGGATGGAGGCGGATAAAGTT GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTG AGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTAT CTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCA CTGATTAAGCATTGGTAACCGATTCTAGGTGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGC GCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCA TACGTGTCCTCCTTACCAGAAATTTATCCTTAAGATCGTTTAAACTCGACTCTGGCTCTATCGAATC TCCGTCGTTTCGAGCTTACGCGAACAGCCGTGGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCA GCTATCGTCAGCTTACCTTTTTGGCAGCGATCGCGGCTCCCGACATCTTGGACCATTAGCTCCACAG GTATCTTCTTCCCTCTAGTGGTCATAACAGCAGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAA GACCCGTTTAGAGGCCCCAAGGGGTTATGCTATCAATCGTTGCGTTACACACACAAAAAACCAACAC ACATCCATCTTCGATGGATAGCGATTTTATTATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAA TCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATG GCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGT AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCA GTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCT GGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT CGCTATTACCATGCTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGG GGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACT TTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGT CTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGA CACACCCGCCAGCGGCCGCTAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGC AAAATCGCAAGACGGGGAATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTC CCCTGCTACTGAGCAGTCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGA TCCAACTATTCAGAACTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGA ACCTCGAGGAGTGCATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAA GACAAAAGCCAGGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAG CGCGTGAGCGCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCA TTAAGAGGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGAT CGGCGTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAA GAAAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGC GGTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGA GAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGACG GTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGAAGG AGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGATTAA GTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAACTGCTC AAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGATGggcg gcggcagcccaaagaagaagcggaaggtcTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACT GCATCAACTAATGAGCAAAATCACCAGCTAACATCATAGTATACATGACCGGCTCTAACTCACATAT CACCATCCTTACACTTAACATTAACGGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGG ATCAAATCACAGGATCCAAGCGTTTGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACC GCCTCAAGATCAAGGGATGGCGAAAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGC AATTCTGGTCTCAGACAAGACGGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTAT ATTATGGTGAAAGGCAGCATACAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCG GCGCACCTCGCTTTATCAAACAGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGAT TATGGGTGATTTCAATACACCATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGAC ACGCAAGAGCTCAATAGCGCACTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTA AGAGTACCGAGTACACATTCTTCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGG CTCAAAGGCTCTGCTGTCAAAGTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGC

[0773]

[0774]

	GCGATCAAGCTCGAGCTGAGAATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATA ACCTGCTGCTCAACGATTATTGGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAAC AAATGAGAATAAGGATACTACCTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAG TTCATCGCCCTCAACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGC TGAAGGAGTTGGAGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGAT TCGCGCCGAGTTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGG TTCTTCGAGCGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGA ACCAGATTGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCAC TATTCGGGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTT CTGGATACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAG GAAGCGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTAC AGCTGAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATA GAAAAGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGAC GCGATACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAA CAAGATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTT ATACCTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAG CTAAGGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCC ATTTATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATT TACGATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTG GAACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGC TATTCGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTC GCCGATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTT CTAACTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAA TAATCGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAG TATCTCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGA AAGAGATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACAT AGTGAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATG ACCTTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAG CGAAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTA TAAAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGG ACCGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAATA AACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAA ACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTTGAAT GTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATACAAGATATAGGCGTCG GCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAAGATTGATAAGTGGGACCT TATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGTTAATAGGCAGCCCACTACA TGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAGCAGAATATATAACGAGCTGA AGCAGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGGCAAAAGATATGAACAGGCATTT TAGCAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGT GAGATGCAGATTAAGACGACCATGCGATACCACCTTACCCCAGTGAGGATGGCAATTATCAAGAAAT CTGGCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGCACCCTGCTCCATTGCTGGTGGGATTG CAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCC TTCGATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGAATACAAGAGCTGTTGTTACAAGGATA CGTGTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCC CACAATGATCGATTGGATCAAGAAAATGTGGCATATTTATACCATGGAGTATTACGCAGCAATTAAG AATGACGAATTTATTTCCTTCGTTGGGACCTGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGT CTCAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGA CGACAAGTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGATACAATCGATTTCTGG ATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATG TTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTAT GGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTT GTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCA CCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGC CGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCG GGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCT TCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCtgagagacacaaaaaattcc aacacactattgcaatgaaaataaatttcctttattagccagaagtcagatgctcaaggggcttcat gatgtccccataatttttggcagagggaaaaagatctcagtggtatttgtgagccagggcattggcc ttctgataggcagcctgcacctgaggagtgcggccgctttacttgtacagctcgtccatgccgagag tgatcccggcggcggtcacgaactccagcaggaccatgtgatcgcgcttctcgttggggtctttgct cagggcggactgggtgctcaggtagtggttgtcgggcagcagcacggggccgtcgccgatgggggtg ttctgctggtagtggtcggcgagctgcacgctgccgtcctcgatgttgtggcggatcttgaagttca ccttgatgccgttcttctgcttgtcggccatgatatagacgttgtggctgttgtagttgtactccag cttgtgccccaggatgttgccgtcctccttgaagtcgatgcccttcagctcgatgcggttcaccagg

	gtgtcgccctcgaacttcacctcggcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgc gctcctggacgtagccttcgggcatggcggacttgaagaagtcgtgctgcttcatgtggtcggggta gcggctgaagcactgcacgccgtaggtcagggtggtcacgagggtgggccagggcacgggcagcttg ccggtggtgcagatgaacttcagggtcagcttgccgtaggtggcatcgccctcgccctcgccggaca cgctgaacttgtggccgtttacgtcgccgtccagctcgaccaggatgggcaccaccccggtgaacag ctcctcgcccttgctcaccatggtggcgggatctgacggttcactaaaccagctctgcttatataga cctcccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacattttggaaa gtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggagacttggaaatccc cgtgagtcaaaccgctatccacgcccattgatgtactgccaaaaccgcatcaccatggtaatagcga tgactaatacgtagatgtactgccaagtaggaaagtcccataaggtcatgtactgggcataatgcca ggcgggccatttaccgtcattgacgtcaataggggcgtacttggcatatgatacacttgatgtact gccaagtgggcagtttaccgtaaatactccacccattgacgtcaatggaaagtccctattggcgtta ctatgggaacatacgtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggcc atttaccgtaagttatgtaacgGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGC CCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACA GCGTACCTTCAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGC AGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTG CAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAG GTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATA ACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTAC CGGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGAT TCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATC CTTAAGGTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTT TTCTGGACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTA CCTTGGGCTATCGAAACTTAATTAAGCGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCC TAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAAT TTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGC TTTTTTGGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCA GGTTCTCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCT CTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTC CGGTGCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCT TGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGG GGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCG GCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGA GCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCG CGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCA CGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGC CGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTG GCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGC CTTCTATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCT GTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAAT AAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCA GGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTT AATTAACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTG CGTCAGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATC ATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATA GGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCG CTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTA GGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCC CGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCA CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGA AGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTA CGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAG GATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAA ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTT CATCCATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAATCTTCAAACCTTTCG TCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGCCTTGCGCCTT GGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGATCGGGATCAC CCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAATATCGAAATCGGG GCGCGCCTGGTGTACCGAGAACGAT(서열번호 93)

[0775]

>RET-004 LINE1-GFP ORF1-C 뉴클레오플라스민-NLS	TCGAAATCGGGGCGCGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATATC GTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTGTA CTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCAAC GTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGCATG AGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTC ACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACATCGA ACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGATGAGC ACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTC GCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATCTTACGGA TGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTA CTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAA CTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGAT GCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGG CAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGG CTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACT GGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGAT GAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCGATTCTAGGTG CATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCA GCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTT AAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAACAGCCGT GGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGGCAGCGAT CGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCATAACAGC AGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCT ATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGATAGCGATTTTATT ATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTAGTTCATAGCCCAT ATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCG CCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAA TGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGC CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGA CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATGCGGTTTTGGCAG TACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCA ATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATT GACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGT CAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCGCTAATACGACTCA CTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGGAATTCCAAGACACA ATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAGTCCTGGATGGAAAAC GACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAACTCCGCGAAGATATCC AGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTGCATCACCCGTATCACAAA CACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCAGGGAGCTTCGAGAGGAGTGT CGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGCGCGATGGAAGACGAGATGAACG AGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGAGGAACGAACAGAGTCTGCAGGAGAT TTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGCGTCCCCGAGAGCGACGTAGAAAACGGG ACTAAACTGGAGAATACACTTCAAGACATCATTCAAGAAAATTTTCCAAACCTGGCTCGGCAAGCTA ATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCGGTATAGCTCTCGGCGTGCCACCCCTAGGCA TATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGAGAAGATGCTGCGAGCCGCTCGGGAAAAGGGA AGGGTGACTTTGAAGGGCAAACCTATTCGGCTGACGGTTGACCTTAGCGCCGAGACACTCCAGGCAC GCCGGGAATGGGGCCCCATCTTTAATATCCTGAAGGAGAAGAACTTCCAGCCACGAATCTCTTACCC TGCAAAGTTGAGTTTTATCTCCGAGGGTGAGATTAAGTATTTCATCGATAAACAGATGCTGCGAGAC TTCGTGACAACTCGCCCAGCTCTCAAGGAACTGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATA GATATCAACCCTTGCAGAACCACGCAAAGATGaaaaggccggcggccacgaaaaaggccggccaggc aaaaaagaaaaagTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCATCAACTAATGAGC AAAATCACCAGCTAACATCATAGTATACATGACCGGCTCTAACTCACATATCACCATCCTTACACTT AACATTAACGGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATC CAAGCGTTTGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGG ATGGCGAAAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGAC AAGACGGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCA GCATACAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTAT CAAACAGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAAT ACACCATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATA GCGCACTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACAC ATTCTTCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTG TCAAAGTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGC TGAGAATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACGA TTATTGGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAGGAT

[0776]

	ACTACCTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACG CCTATAAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAA ACAGGAACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAA GAGATCGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTA ATAAGATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGATTGATACCAT CAAGAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCGGGAGTATTAT AAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGATACCTATACTC TGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCGAGATTGTGGC CATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGAGTTTTATCAA CGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAGGAAGGCATCT TGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCGATACCACAAAGAA GGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAAGATTCTCGCCAAC AGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATACCTGGCATGCAGG GCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTAAGGATAAGAATCA TATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTTATGCTCAAGACT CTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTACGATAAGCCGACTG CTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCCGCCAAGGCTG TCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCGTCAAGAGAAA GAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGATGATATGATTG TGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACTTTAGCAAGGT CAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAATCGACAGACCGAA TCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCTCGGAATCCAGC TGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGAGATTAAGGAAGA TACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGTGAAGATGGCAATA CTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACCTTCTTTACGGAGC TCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCGAAGTCCATCTTGAG CCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATAAAGCCACAGTAACT AAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGGACCGAACCATCAGAGA TAATGCCCCACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGA CAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAAACTCAAGCTCGACCCC TTTCTTACACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTTGAATGTCAAGCCAAAGACTA TAAAGACACTGGAAGAGAATCTTGGGATCACAATACAAGATATAGGCGTCGGCAAAGATTTTATGTC AAAGACGCCCAAGGCCATGGCCACTAAGGATAAGATTGATAAGTGGGACCTTATTAAGCTCAAAAGC TTCTGTACTGCCAAGGAGACCACGATCAGAGTTAATAGGCAGCCCACTACATGGGAAAAGATTTTCG CCACTTATTCATCAGATAAGGGGTTGATAAGCAGAATATATAACGAGCTGAAGCAGATCTACAAGAA GAAAACGAATAATCCCATCAAGAAGTGGGCAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATC TACGCCGCGAAGAAGCATATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGA CGACCATGCGATACCACCTTACCCCAGTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATG TTGGCGGGGCTGTGGCGAGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCG CTTTGGAAATCAGTCTGGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTC CCTTGCTCGGAATCTATCCTAACGAATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTT CATCGCGGCCTTGTTTACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGG ATCAAGAAAATGTGGCATATTTATACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTT CCTTCGTTGGGACCTGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGAC AAAGCATAGAATCTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACGACAAGTAAAGCGCT TCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGATACAATCGATTTCTGGATCCGCAGGCCTAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCT ATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCC TCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCG TGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCT TTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGC TGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCT TTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTC GGCCCTCAATCCAGCGGACCTTCCTTCCCGCtgagagacacaaaaaattccaacacactattgcaat gaaaataaatttcctttattagccagaagtcagatgctcaaggggcttcatgatgtccccataattt ttggcagagggaaaaagatctcagtggtatttgtgagccagggcattggccttctgataggcagcct gcacctgaggagtgcggccgctttacttgtacagctcgtccatgccgagagtgatcccggcggcggt cacgaactccagcaggaccatgtgatcgcgcttctcgttggggtctttgctcagggcggactgggtg ctcaggtagtggttgtcgggcagcagcacggggccgtcgccgatgggggtgttctgctggtagtggt cggcgagctgcacgctgccgtcctcgatgttgtggcggatcttgaagttcaccttgatgccgttctt ctgcttgtcggccatgatatagacgttgtggctgttgtagttgtactccagcttgtgccccaggatg ttgccgtcctccttgaagtcgatgcccttcagctcgatgcggttcaccagggtgtcgccctcgaact tcacctcggcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgcgctcctggacgtagcc ttcgggcatggcggacttgaagaagtcgtgctgcttcatgtggtcggggtagcggctgaagcactgc

[0777]

	acgccgtaggtcagggtggtcacgagggtgggccagggcacgggcagcttgccggtggtgcagatga acttcagggtcagcttgccgtaggtggcatcgccctcgccctcgccggacacgctgaacttgtggcc gtttacgtcgccgtccagctcgaccaggatgggcaccaccccggtgaacagctcctcgcccttgctc accatggtggcgggatctgacggttcactaaaccagctctgcttatatagacctcccaccgtacacg cctaccgcccatttgcgtcaatggggcggagttgttacgacattttggaaagtcccgttgattttgg tgccaaaacaaactcccattgacgtcaatggggtggagacttggaaatccccgtgagtcaaaccgct atccacgcccattgatgtactgccaaaaccgcatcaccatggtaatagcgatgactaatacgtagat gtactgccaagtaggaaagtcccataaggtcatgtactgggcataatgccaggcgggccatttaccg tcattgacgtcaataggggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagttt accgtaaatactccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatacgt cattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttaccgtaagttat gtaacgGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGA TCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCA CAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTT TATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAC AACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGT AAAACCTCTACAAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCCCTTGGGGCCTCT AAACGGGTCTTGAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGCATTGGCGCAGAA AAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGG CTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGGTCGTCAGCT ATCCTGCAGGCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCTGGACACCACTAG GGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTTGGGCTATCGAAA CTTAATTAAGCGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCC GCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTTTTTATTTATGCA GAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAG GCTTTTGCAAAGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTTCTCCCGCCGCTT GGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTT CCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAA CTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCG ACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTC ATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTT GATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGG AAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTT CGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACGGCGATGCCTGCTTG CCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGG ACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGA CCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTT GACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCC CCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGC ATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAG GATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAATTAACCAGTCAAGT CAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCGTCAGTCAAGTTCGA TCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCATGTGAGCAAAAGGCC AGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGA CGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGT CCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGT GTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTA TCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTG GTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA CGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGC AGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCA GTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATC CTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTT ACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCATT TAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAATCTTCAAACCTTTCGTCCGATCCATCTTGCA GGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGCCTTGCGCCTTGGCTATTGCTTGGCAG CGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGATCGGGATCACCCGAGAGAAGTTCAAC CTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAATA(서열번호 94)
RET-005 LINE1-	TATCGAAATCGGGGCGCGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATA TCGTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTG TACTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCA

[0778]

[0779]

GFP ORF1-C 링커_뉴클레오플라스민-NLS	ACGTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGCA TGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGC TCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACATC GAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGATGA GCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGG TCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATCTTACG GATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACT TACTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGT AACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACG ATGCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCC GGCAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCC GGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCA CTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGG ATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCGATTCTAGG TGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATT CAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCC TTAAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAACAGCC GTGGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGGCAGCG ATCGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCATAACA GCAGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATG CTATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGATAGCGATTTTA TTATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTAGTTCATAGCCC ATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCC CGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTC AATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTAC GCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGG GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATGCGGTTTTGGC AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGT CAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCA TTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACC GTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCGCTAATACGACT CACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGGAATTCCAAGACA CAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAGTCCTGGATGGAAA ACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAACTCCGCGAAGATAT CCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTGCATCACCCGTATCACA AACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCAGGGAGCTTCGAGAGGAGT GTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGCGCGATGGAAGACGAGATGAA CGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGAGGAACGAACAGAGTCTGCAGGAG ATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGCGTCCCCGAGAGCGACGTAGAAAACG GGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGAAAATTTTCCAAACCTGGCTCGGCAAGC TAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCGGTATAGCTCTCGGCGTGCCACCCCTAGG CATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGAGAAGATGCTGCGAGCCGCTCGGGAAAAGG GAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGACGGTTGACCTTAGCGCCGAGACACTCCAGGC ACGCCGGGAATGGGGCCCCATCTTTAATATCCTGAAGGAGAAGAACTTCCAGCCACGAATCTCTTAC CCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGATTAAGTATTTCATCGATAAACAGATGCTGCGAG ACTTCGTGACAACTCGCCCAGCTCTCAAGGAACTGCTCAAAGAGGCTCTTAATATGGAGCGCAATAA TAGATATCAACCCTTGCAGAACCACGCAAAGATGggcggcggcagcaaaaggccggcggccacgaaa aaggccggccaggcaaaaaagaaaaagTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGC ATCAACTAATGAGCAAAATCACCAGCTAACATCATAGTATACATGACCGGCTCTAACTCACATATCA CCATCCTTACACTTAACATTAACGGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGAT CAAATCACAGGATCCAAGCGTTTGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGC CTCAAGATCAAGGGATGGCGAAAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAA TTCTGGTCTCAGACAAGACGGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATAT TATGGTGAAAGGCAGCATACAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGC GCACCTCGCTTTATCAAACAGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTA TGGGTGATTTCAATACACCATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACAC GCAAGAGCTCAATAGCGCACTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAG AGTACCGAGTACACATTCTTCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCT CAAAGGCTCTGCTGTCAAAGTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGC GATCAAGCTCGAGCTGAGAATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAAC CTGCTGCTCAACGATTATTGGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAA ATGAGAATAAGGATACTACCTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTT CATCGCCCTCAACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTG AAGGAGTTGGAGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTC

	GCGCCGAGTTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTT CTTCGAGCGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAAC CAGATTGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTA TTCGGGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCT GGATACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGA AGCGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAG CTGAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGA AAAGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGC GATACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACA AGATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTAT ACCTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCT AAGGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCAT TTATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTA CGATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGA ACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTA TTCGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGC CGATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCT AACTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATA ATCGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTA TCTCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAA GAGATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAG TGAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGAC CTTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCG AAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATA AAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGGAC CGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAATAAA CAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAAAC TCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTTGAATGT CAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATACAAGATATAGGCGTCGGC AAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAAGATTGATAAGTGGGACCTTA TTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGTTAATAGGCAGCCCACTACATG GGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAGCAGAATATATAACGAGCTGAAG CAGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGGCAAAAGATATGAACAGGCATTTTA GCAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGA GATGCAGATTAAGACGACCATGCGATACCACCTTACCCCAGTGAGGATGGCAATTATCAAGAAATCT GGCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCA AGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTT CGATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGAATACAAGAGCTGTTGTTACAAGGATACG TGTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCA CAATGATCGATTGGATCAAGAAAATGTGGCATATTTATACCATGGAGTATTACGCAGCAATTAAGAA TGACGAATTTATTTCCTTCGTTGGGACCTGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCT CAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACG ACAAGTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGATACAATCGATTTCTGGAT CCGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTT GCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGG CTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGT CAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACC ACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCG CCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGG GAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTC TGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCtgagagacacaaaaaattccaa cacactattgcaatgaaaataaatttcctttattagccagaagtcagatgctcaaggggcttcatga tgtccccataatttttggcagagggaaaaagatctcagtggtatttgtgagccagggcattggcctt ctgataggcagcctgcacctgaggagtgcggccgctttacttgtacagctcgtccatgccgagagtg atcccggcggcggtcacgaactccagcaggaccatgtgatcgcgcttctcgttggggtctttgctca gggcggactgggtgctcaggtagtggttgtcgggcagcagcacgggggccgtcgccgatgggggtgtt ctgctggtagtggtcggcgagctgcacgctgccgtcctcgatgttgtggcggatcttgaagttcacc ttgatgccgttcttctgcttgtcggccatgatatagacgttgtggctgttgtagttgtactccagct tgtgccccaggatgttgccgtcctccttgaagtcgatgcccttcagctcgatgcggttcaccagggt gtcgccctcgaacttcacctcggcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgcgc tcctggacgtagccttcgggcatggcggacttgaagaagtcgtgctgcttcatgtggtcggggtagc ggctgaagcactgcacgccgtaggtcagggtggtcacgagggtgggccagggcacgggcagcttgcc ggtggtgcagatgaacttcagggtcagcttgccgtaggtggcatcgccctcgccctcgccggacacg ctgaacttgtggccgtttacgtcgccgtccagctcgaccaggatgggcaccaccccggtgaacagct

[0780]

[0781]

	cctcgcccttgctcaccatggtggcgggatctgacggttcactaaaccagctctgcttatatagacc tcccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacattttggaaagt cccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggagacttggaaatccccg tgagtcaaaccgctatccacgcccattgatgtactgccaaaaccgcatcaccatggtaatagcgatg actaatacgtagatgtactgccaagtaggaaagtcccataaggtcatgtactgggcataatgccagg cgggccatttaccgtcattgacgtcaataggggggcgtacttggcatatgatacacttgatgtactgc caagtgggcagtttaccgtaaatactccacccattgacgtcaatggaaagtccctattggcgttact atgggaacatacgtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccat ttaccgtaagttatgtaacgGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCC TCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGC GTACCTTCAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAG TGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCA ATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGT TTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCG GCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTC AGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCT TAAGGTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTT CTGGACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACC TTGGGCTATCGAAACTTAATTAAGCGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTA ACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTT TTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTT TTTTGGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGG TTCTCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCT GATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCG GTGCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTG CGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGG CAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGC GGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGC ACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCG CCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACG GCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCG GCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGC GGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCT TCTATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGT TGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAA AATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGG ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAA TTAACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCG TCAGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCAT GTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGG CTCCGCCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC TATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCT TACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGG TATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCG ACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAG TGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTA CCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG GGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGA TCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAAC TTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCA TCCATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAATCTTCAAACCTTTCGTC CGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGCCTTGCGCCTTGG CTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGATCGGGATCACCC GAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAA(서열번호 95)
RET-006 LINE1-GFP ORF1-N	GTCGGAAAATCTTCAAACCTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAA GTCTCTTGGCCGGCCTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATC CCGAATATCCGAGATCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGA GATCCGAGGAATATCGAAATCGGGGCGCGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTC GACGATCCATATCGTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATC GTCAGATATTGTACTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGT

SV40 NLS	CTGATCGGTCAACGTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGG AGGCTTTCGCATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTT CCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAG TGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTT TCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAA GAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAA AGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACAC TGCGGCCAACTTACTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATG GGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGC GTGACACCACGATGCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTAC TCTAGCTTCCCGGCAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGC TCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTA TCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCA GGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAA CCGATTCTAGGTGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACA TATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCA GAAATTTATCCTTAAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTA CGCGAACAGCCGTGGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCT TTTTGGCAGCGATCGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAG TGGTCATAACAGCAGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCC AAGGGGTTATGCTATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGA TAGCGATTTTATTATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTA GTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGC CCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTT CCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCAT ATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACA TGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGAT GCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCAC CCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACA ACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGG TTTAGTGAACCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCG CTAATACGACTCACTATAGGGAGAAGTACTGCCACCATGccaaagaagaagcggaaggtcGGCAAGA AGCAAAATCGCAAGACGGGGAATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAG CTCCCCTGCTACTGAGCAGTCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGG CGATCCAACTATTCAGAACTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGA AGAACCTCGAGGAGTGCATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACT TAAGACAAAAGCCAGGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAG GAGCGCGTGAGCGCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGC GCATTAAGAGGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTT GATCGGCGTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATT CAAGAAAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCC AGCGGTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAA AGAGAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTG ACGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGA AGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGAT TAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAACTG CTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGATGT GAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCATCAACTAATGAGCAAAATCACCAGCTA ACATCATAGTATACATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAACGGCCT CAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTTTGTTGC ATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGAAAGATTT ATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGACGGATTTCAA GCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATACAGCAGGAA GAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAACAGGTCCTGT CCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACACCATTGAGCAC CCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCACTGCATCAG GCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCTTCAGCGCCC CACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAAGTGCAAGCG CACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGAATCAAGAAC CTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATTGGGTCCACA ATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTACCTATCAAAA CCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTATAAAAGAAAA CAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGGAACAGACCC ACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGATCGAAACCCA

[0782]

[0783]

	AAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAGATAGACAGA CCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGATTGATACCATCAAGAACGACAAGG GCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCGGGAGTATTATAAGCATTTGTATGC TAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGATACCTATACTCTGCCACGGCTTAAT CAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCGAGATTGTGGCCATAATTAACTCCC TGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGAGTTTTATCAACGGTATATGGAAGA GCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAGGAAGGCATCTTGCCCAATTCCTTC TACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCGATACCACAAAGAAGGAAAACTTCCGGC CCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAAGATTCTCGCCAACAGAATCCAACAACA TATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATACCTGGCATGCAGGGCTGGTTTAACATC CGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTAAGGATAAGAATCATATGATCATCTCTA TAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTTATGCTCAAGACTCTGAACAAACTCGG CATCGACGGAACATATTTTAAGATTATTCGCGCAATTTACGATAAGCCGACTGCTAACATTATCCTT AACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCCGCCAAGGCTGTCCCCTCTCCCCGC TTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCGTCAAGAGAAAGAGATTAAAGGGAT ACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGATGATATGATTGTGTACCTGGAGAAT CCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACTTTAGCAAGGTCAGCGGCTATAAGA TTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAATCGACAGACCGAATCCCAGATAATGGG TGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCTCGGAATCCAGCTGACACGAGACGTT AAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGAGATTAAGGAAGATACTAATAAGTGGA AGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGTGAAGATGGCAATACTTCCTAAAGTGAT ATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAGAAAACAACC CTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGAAGAATAAGG CCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGACAGCCTGGTA TTGGTATCAGAATAGAGACATCGACCAGTGGAATCGGACCGAACCATCAGAGATAATGCCCCACATC TATAATTACCTTATATTCGATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTTCAACA AGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTACACCCTA CACTAAAATCAACAGTAGGTGGATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACACTGGAA GAGAATCTTGGGATCACAATACAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGG CCATGGCCACTAAGGATAAGATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAA GGAGACCACGATCAGAGTTAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTATTCATCA GATAAGGGGTTGATAAGCAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATC CCATCAAGAAGTGGGCAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAA GCATATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATAC CACCTTACCCCAGTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTG GCGAGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGT CTGGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATC TATCCTAACGAATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGT TTACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTG GCATATTTATACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACC TGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCT TCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACGACAAGTAAAGCGCTTCTAGAAGTTGTCT CCTCCTGCACTGACTGACTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTAC AAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTG CTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATC CTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTG TTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCG CTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGC TCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTC GCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAG CGGACCTTCCTTCCCGCtgagagacacaaaaaattccaacacactattgcaatgaaaataaatttcc tttattagccagaagtcagatgctcaaggggcttcatgatgtcccataatttttggcagagggaaa aagatctcagtggtatttgtgagccagggcattggccttctgataggcagcctgcacctgaggagtg cggccgctttacttgtacagctcgtccatgccgagagtgatcccggcggcggtcacgaactccagca ggaccatgtgatcgcgcttctcgttggggtctttgctcagggcggactgggtgctcaggtagtggtt gtcgggcagcagcacgggggccgtcgccgatgggggtgttctgctggtagtggtcggcgagctgcacg ctgccgtcctcgatgttgtggcggatcttgaagttcaccttgatgccgttcttctgcttgtcggcca tgatatagacgttgtggctgttgtagttgtactccagcttgtgccccaggatgttgccgtcctcctt gaagtcgatgcccttcagctcgatgcggttcaccagggtgtcgccctcgaacttcacctcggcgcgg gtcttgtagttgccgtcgtccttgaagaagatggtgcgctcctggacgtagccttcgggcatggcgg acttgaagaagtcgtgctgcttcatgtggtcggggtagcggctgaagcactgcacgccgtaggtcag ggtggtcacgagggtgggccagggcacgggcagcttgccggtggtgcagatgaacttcagggtcagc ttgccgtaggtggcatcgccctcgcctcgccggacacgctgaacttgtggccgtttacgtcgccgt ccagctcgaccaggatgggcaccaccccggtgaacagctcctcgcccttgctcaccatggtggcggg

[0784]

	atctgacggttcactaaaccagctctgcttatatagacctcccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacattttggaaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggagacttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaaccgcatcaccatggtaatagcgatgactaatacgtagatgtactgccaagtagaaaagtcccataaggtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaataggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgtaaatactccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatacgtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttaccgtaagttatgtaacgGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGGTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCTGGACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTTGGGCTATCGAAACTTAATTAAGCGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTTCTCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAATTAACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCGTCAGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTC(서열번호 96)
RET-007 LINE1-GFP ORF1-NSV40 NLS 링커	GGAAAATCTTCAAACCTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGCCTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGATCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAATATCGAAATCGGGGCGCGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATATCGTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTGTACTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCAACGTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGCATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGG

	GTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCC AATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAG CAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGC ATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGC GGCCAACTTACTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGG GATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTG ACACCACGATGCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCT AGCTTCCCGGCAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCG GCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCA TTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGC AACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCG ATTCTAGGTGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATAT GCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAA ATTTATCCTTAAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTACGC GAACAGCCGTGGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCTTTT TGGCAGCGATCGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGG TCATAACAGCAGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAG GGGTTATGCTATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGATAG CGATTTTATTATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTAGTT CATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCA ACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCA TTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATG CCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGA CCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATGCG GTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCC ATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACT CCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTT AGTGAACCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCGCTA ATACGACTCACTATAGGGAGAAGTACTGCCACCATGccaaagaagaagcggaaggtcggcggcggca gcGGCAAGAAGCAAAATCGCAAGACGGGGAATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGA GCGTTCTAGCTCCCCTGCTACTGAGCAGTCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAG GGATTTAGGCGATCCAACTATTCAGAACTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGA ATTTCGAGAAGAACCTCGAGGAGTGCATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACT CATGGAACTTAAGACAAAAGCCAGGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGAC CAGCTCGAGGAGCGCGTGAGCGCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCA GGGAGAAGCGCATTAAGAGGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAA CCTGCGGTTGATCGGCGTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAA GACATCATTCAAGAAAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAAC GCACACCCCAGCGGTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGT GGAGATGAAAGAGAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCT ATTCGGCTGACGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTA ATATCCTGAAGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGA GGGTGAGATTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTC AAGGAACTGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACG CAAAGATGTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCATCAACTAATGAGCAAAAT CACCAGCTAACATCATAGTATACATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAACAT TAACGGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGC GTTTGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGC GAAAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGAC GGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATA CAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAAC AGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACACC ATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCA CTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCT TCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAA GTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGA ATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATT GGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTAC CTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTAT AAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGG AACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGAT CGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAG ATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGATTGATACCATCAAGA ACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCGGGAGTATTATAAGCA

[0785]

	TTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGATACCTATACTCTGCCA CGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCGAGATTGTGGCCATAA TTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGAGTTTTATCAACGGTA TATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAGGAAGGCATCTTGCCC AATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCGATACCACAAAGAAGGAAA ACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAAGATTCTCGCCAACAGAAT CCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATACCTGGCATGCAGGGCTGG TTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTAAGGATAAGAATCATATGA TCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTTATGCTCAAGACTCTGAA CAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTACGATAAGCCGACTGCTAAC ATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCCGCCAAGGCTGTCCCC TCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCGTCAAGAGAAAGAGAT TAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGATGATATGATTGTGTAC CTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACTTTAGCAAGGTCAGCG GCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAATCGACAGACCGAATCCCA GATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCTCGGAATCCAGCTGACA CGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGAGATTAAGGAAGATACTA ATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGTGAAGATGGCAATACTTCC TAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAG AAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGA AGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGAC AGCCTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGGACCGAACCATCAGAGATAATG CCCCACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCC TCTTCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCT TACACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAG ACACTGGAAGAGAATCTTGGGATCACAATACAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGA CGCCCAAGGCCATGGCCACTAAGGATAAGATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTG TACTGCCAAGGAGACCACGATCAGAGTTAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACT TATTCATCAGATAAGGGGTTGATAAGCAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAA CGAATAATCCCATCAAGAAGTGGGCAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGC CGCGAAGAAGCATATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACC ATGCGATACCACCTTACCCCAGTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGC GGGGCTGTGGCGAGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTG GAAATCAGTCTGGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTG CTCGGAATCTATCCTAACGAATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCG CGGCCTTGTTTACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAA GAAAATGTGGCATATTTATACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTC GTTGGGACCTGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGC ATAGAATCTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACGACAAGTAAAGCGCTTCTAG AAGTTGTCTCCTCCTGCACTGACTGACTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCT CTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTG GATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTT GTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTG TGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCG GGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTG GACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCA TGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCC TCAATCCAGCGGACCTTCCTTCCCGCtgagagacacaaaaaattccaacacactattgcaatgaaaa taaatttccttttattagccagaagtcagatgctcaaggggcttcatgatgtccccataatttttggc agagggaaaaagatctcagtggtatttgtgagccagggcattggccttctgataggcagcctgcacc tgaggagtgcggccgctttacttgtacagctcgtccatgccgagagtgatcccggcggcggtcacga actccagcaggaccatgtgatcgcgcttctcgttggggtctttgctcagggcggactgggtgctcag gtagtggttgtcgggcagcagcacggggccgtcgccgatgggggtgttctgctggtagtggtcggcg agctgcacgctgccgtcctcgatgttgtggcggatcttgaagttcaccttgatgccgttcttctgct tgtcggccatgatatagacgttgtggctgttgtagttgtactccagcttgtgccccaggatgttgcc gtcctccttgaagtcgatgcccttcagctcgatgcggttcaccagggtgtcgccctcgaacttcacc tcggcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgcgctcctggacgtagccttcgg gcatggcggacttgaagaagtcgtgctgcttcatgtggtcggggtagcggctgaagcactgcacgcc gtaggtcagggtggtcacgagggtgggccagggcacgggcagcttgccggtggtgcagatgaacttc agggtcagcttgccgtaggtggcatcgccctcgccctcgccggacacgctgaacttgtggccgttta cgtcgccgtccagctcgaccaggatgggcaccaccccggtgaacagctcctcgcccttgctcaccat ggtggcgggatctgacggttcactaaaccagctctgcttatatagacctcccaccgtacacgcctac cgcccatttgcgtcaatggggcggagttgttacgacattttggaaagtcccgttgattttggtgcca aaacaaactcccattgacgtcaatggggtggagacttggaaatccccgtgagtcaaaccgctatcca

[0786]

	cgcccattgatgtactgccaaaaccgcatcaccatggtaatagcgatgactaatacgtagatgtact gccaagtaggaaagtcccataaggtcatgtactgggcataatgccaggcgggccatttaccgtcatt gacgtcaataggggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgt aaatactccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatacgtcatta ttgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttaccgtaagttatgtaac gGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCC CTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGAC ATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTT GTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAA CAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAAC CTCTACAAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCCCTTGGGGCCTCTAAACG GGTCTTGAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGCATTGGCGCAGAAAAAAA TGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCC CCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGGTCGTCAGCTATCCT GCAGGCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCTGGACACCACTAGGGGTC AGAAGTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTTGGGCTATCGAAACTTAA TTAAGCGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCC TAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTTTTTATTTATGCAGAGGC CGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTT TGCAAAGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTTCTCCCGCCGCTTGGGTG GAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGC TGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCA GGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTT GTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTC ACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCC GGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCC GGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCA GGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACGGCGATGCCTGCTTGCCGAA TATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGC TATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCT TCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGA GTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTG CCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGC ATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTG GGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAATTAACCAGTCAAGTCAGCT ACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCGTCAGTCAAGTTCGATCTGG TCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCATGTGAGCAAAAGGCCAGCAA AAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGC ATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTT TCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCC TTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGG TCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGG TAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCT ACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATT ACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGA ACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTT AAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAA TGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCATTTAAAT TTCCGAACTCTCCAAGGCCCTCGTC(서열번호 97)
RET-008 LINE1-GFP ORF1-N 뉴클레오플라스민 NLS	GGAAAATCTTCAAACCTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTC TCTTGGCCGGCCTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCG AATATCCGAGATCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGAT CCGAGGAATATCGAAATCGGGGCGCGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGAC GATCCATATCGTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTC AGATATTGTACTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTG ATCGGTCAACGTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGG CTTTCGCATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCT GTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGG GTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCC AATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAG CAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGC

[0787]

	ATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGC GGCCAACTTACTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGG GATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTG ACACCACGATGCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCT AGCTTCCCGGCAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCG GCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCA TTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGC AACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCG ATTCTAGGTGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATAT GCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAA ATTTATCCTTAAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTACGC GAACAGCCGTGGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCTTTT TGGCAGCGATCGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGG TCATAACAGCAGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAG GGGTTATGCTATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGATAG CGATTTTATTATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTAGTT CATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCA ACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCA TTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATG CCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGA CCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATGCG GTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCC ATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACT CCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTT AGTGAACCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCGCTA ATACGACTCACTATAGGGAGAAGTACTGCCACCATGaaaaggccggcggccacgaaaaaggccggcc aggcaaaaaagaaaaagGGCAAGAAGCAAAATCGCAAGACGGGGAATTCCAAGACACAATCCGCTAG CCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAGTCCTGGATGGAAAACGACTTCGAT GAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAACTCCGCGAAGATATCCAGACAAAGG GGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTGCATCACCCGTATCACAAACACTGAGAA ATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCAGGGAGCTTCGAGAGGAGTGTCGGAGTCTG AGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGCGCGATGGAAGACGAGATGAACGAGATGAAAA GAGAGGGCAAATTCAGGGAGAAGCGCATTAAGAGGAACGAACAGAGTCTGCAGGAGATTTGGGATTA CGTCAAGAGGCCTAACCTGCGGTTGATCGGCGTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTG GAGAATACACTTCAAGACATCATTCAAGAAAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAA TCCAAGAGATCCAACGCACACCCCAGCGGTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGT GCGCTTTACTAAGGTGGAGATGAAAGAGAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACT TTGAAGGGCAAACCTATTCGGCTGACGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAAT GGGGCCCCATCTTTAATATCCTGAAGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTT GAGTTTTATCTCCGAGGGTGAGATTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACA ACTCGCCCAGCTCTCAAGGAACTGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAAC CCTTGCAGAACCACGCAAAGATGTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCATCA ACTAATGAGCAAAATCACCAGCTAACATCATAGTATACATGACCGGCTCTAACTCACATATCACCAT CCTTACACTTAACATTAACGGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAA TCACAGGATCCAAGCGTTTGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCA AGATCAAGGGATGGCGAAAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCT GGTCTCAGACAAGACGGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATG GTGAAAGGCAGCATACAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCAC CTCGCTTTATCAAACAGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGG TGATTTCAATACACCATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAA GAGCTCAATAGCGCACTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTA CCGAGTACACATTCTTCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAA GGCTCTGCTGTCAAAGTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATC AAGCTCGAGCTGAGAATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGC TGCTCAACGATTATTGGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGA GAATAAGGATACTACCTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATC GCCCTCAACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGG AGTTGGAGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGC CGAGTTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTC GAGCGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGA TTGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCG GGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGAT ACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCG AGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGA

[0788]

[0789]

	GTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAG GAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCGATA CCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAAGAT TCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATACCT GGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTAAGG ATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTTAT GCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTACGAT AAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCC GCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCG TCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGAT GATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACT TTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAATCG ACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCTC GGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGAGA TTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGTGAA GATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACCTTC TTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCGAAGT CCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATAAAGC CACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGGACCGAA CCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAATAAACAGT GGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAAACTCAA GCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTTGAATGTCAAG CCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATACAAGATATAGGCGTCGGCAAAG ATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAAGATTGATAAGTGGGACCTTATTAA GCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGTTAATAGGCAGCCCACTACATGGGAA AAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAGCAGAATATATAACGAGCTGAAGCAGA TCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGGCAAAAGATATGAACAGGCATTTTAGCAA AGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGATG CAGATTAAGACGACCATGCGATACCACCTTACCCCAGTGAGGATGGCAATTATCAAGAAATCTGGCA ATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAAGCT GGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCGAT CCCGCAATTCCCTTGCTCGGAATCTATCCTAACGAATACAAGAGCTGTTGTTACAAGGATACGTGTA CCCGGATGTTCATCGCGGCCTTGTTTACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCACAAT GATCGATTGGATCAAGAAAATGTGGCATATTTATACCATGGAGTATTACGCAGCAATTAAGAATGAC GAATTTATTTCCTTCGTTGGGACCTGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTCAGG AGCAAAAGACAAAGCATAGAATCTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACGACAA GTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGATACAATCGATTTCTGGATCCGC AGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTC CTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTT CATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGG CAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCT GTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTG CCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAG CTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCT ACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCtgagagacacaaaaaattccaacaca ctattgcaatgaaaataaatttcctttattagccagaagtcagatgctcaaggggcttcatgatgtc cccataatttttggcagagggaaaaagatctcagtggtatttgtgagccagggcattggccttctga taggcagcctgcacctgaggagtgcggccgctttacttgtacagctcgtccatgccgagagtgatcc cggcggcggtcacgaactccagcaggaccatgtgatcgcgcttctcgttggggtctttgctcagggc ggactgggtgctcaggtagtggttgtcgggcagcagcacggggccgtcgccgatgggggtgttctgc tggtagtggtcggcgagctgcacgctgccgtcctcgatgttgtggcggatcttgaagttcaccttga tgccgttcttctgcttgtcggccatgatatagacgttgtggctgttgtagttgtactccagcttgtg ccccaggatgttgccgtcctccttgaagtcgatgcccttcagctcgatgcggttcaccagggtgtcg ccctcgaacttcacctcggcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgcgctcct ggacgtagccttcgggcatggcggacttgaagaagtcgtgctgcttcatgtggtcggggtagcggct gaagcactgcacgccgtaggtcagggtggtcacgagggtgggccagggcacgggcagcttgccggtg gtgcagatgaacttcagggtcagcttgccgtaggtggcatcgccctcgccctcgccggacacgctga acttgtggccgtttacgtcgccgtccagctcgaccaggatgggcaccaccccggtgaacagctcctc gcccttgctcaccatggtggcgggatctgacggttcactaaaccagctctgcttatatagacctccc accgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacattttggaaagtcccg ttgattttggtgccaaaacaaactcccattgacgtcaatggggtggagacttggaaatccccgtgag tcaaaccgctatccacgcccattgatgtactgccaaaaccgcatcaccatggtaatagcgatgacta atacgtagatgtactgccaagtaggaaagtcccataaggtcatgtactgggcataatgccaggcggg ccatttaccgtcattgacgtcaataggggggcgtacttggcatatgatacacttgatgtactgccaag

	tgggcagtttaccgtaaatactccacccattgacgtcaatggaaagtccctattggcgttactatgg gaacatacgtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttac cgtaagttatgtaacgGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAG ACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTAC CTTCAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAA AAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAA ACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTT TAAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCCCT TGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGCAT TGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCA ACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAG GTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCTGG ACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTTGG GCTATCGAAACTTAATTAAGCGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTC CGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTTTT TATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTT GGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTTCT CCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATG CCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGC CCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCA GCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGG ATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCT GCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGT ACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAG CCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACGGCGA TGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTG GGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCG AATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTA TCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTT TGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATG AGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAG CAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAATTAA CCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCGTCAG TCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCATGTGA GCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCC GCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA AAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACC GGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATC TCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCG CTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCA GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGT GGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTT TGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGT CTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTT CACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGG TCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCA TAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTC(서열번호 98)
RET-009 LINE1-GFP ORF1-N 뉴클레오플라스민 NLS 링커	CAAACCTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGG CCTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAG ATCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAATA TCGAAATCGGGGCGCGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATATC GTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTGTA CTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCAAC GTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGCATG AGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTC ACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACATCGA ACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGATGAGC ACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTC GCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATCTTACGGA TGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTA CTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAA CTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGAT

[0790]

	GCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGG CAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGG CTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACT GGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGAT GAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCGATTCTAGGTG CATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCA GCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTT AAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAACAGCCGT GGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGGCAGCGAT CGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCATAACAGC AGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCT ATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGATAGCGATTTTATT ATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTAGTTCATAGCCCAT ATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCG CCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAA TGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGC CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGA CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATGCGGTTTTGGCAG TACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCA ATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATT GACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGT CAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCGCTAATACGACTCA CTATAGGGAGAAGTACTGCCACCATGaaaaggccggcggccacgaaaaaggccggccaggcaaaaaa gaaaaagggcggcggcagcGGCAAGAAGCAAAATCGCAAGACGGGGAATTCCAAGACACAATCCGCT AGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAGTCCTGGATGGAAAACGACTTCG ATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAACTCCGCGAAGATATCCAGACAAA GGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTGCATCACCCGTATCACAAACACTGAG AAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCAGGGAGCTTCGAGAGGAGTGTCGGAGTC TGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGCGCGATGGAAGACGAGATGAACGAGATGAA AAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGAGGAACGAACAGAGTCTGCAGGAGATTTGGGAT TACGTCAAGAGGCCTAACCTGCGGTTGATCGGCGTCCCCGAGAGCGACGTAGAAAACGGGACTAAAC TGGAGAATACACTTCAAGACATCATTCAAGAAAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCA AATCCAAGAGATCCAACGCACACCCCAGCGGTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATC GTGCGCTTTACTAAGGTGGAGATGAAAGAGAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGA CTTTGAAGGGCAAACCTATTCGGCTGACGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGA ATGGGGCCCCATCTTTAATATCCTGAAGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAG TTGAGTTTTATCTCCGAGGGTGAGATTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGA CAACTCGCCCAGCTCTCAAGGAACTGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCA ACCCTTGCAGAACCACGCAAAGATGTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCAT CAACTAATGAGCAAAATCACCAGCTAACATCATAGTATACATGACCGGCTCTAACTCACATATCACC ATCCTTACACTTAACATTAACGGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCA AATCACAGGATCCAAGCGTTTGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCT CAAGATCAAGGGATGGCGAAAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATT CTGGTCTCAGACAAGACGGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTA TGGTGAAAGGCAGCATACAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGC ACCTCGCTTTATCAAACAGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATG GGTGATTTCAATACACCATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGC AAGAGCTCAATAGCGCACTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAG TACCGAGTACACATTCTTCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCA AAGGCTCTGCTGTCAAAGTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGA TCAAGCTCGAGCTGAGAATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCT GCTGCTCAACGATTATTGGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAAT GAGAATAAGGATACTACCTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCA TCGCCCTCAACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAA GGAGTTGGAGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGC GCCGAGTTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCT TCGAGCGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCA GATTGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATT CGGGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGG ATACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAG CGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCT GAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAA AGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCGA TACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAAG

[0791]

[0792]

	ATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATAC CTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTAA GGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTT ATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTACG ATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAAC CCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATT CGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCG ATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAA CTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAAT CGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATC TCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGA GATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGTG AAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACCT TCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCGAA GTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATAAA GCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGGACCG AACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAATAAACA GTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAAACTC AAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTTGAATGTCA AGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATACAAGATATAGGCGTCGGCAA AGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAAGATTGATAAGTGGGACCTTATT AAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGTTAATAGGCAGCCCACTACATGGG AAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAGCAGAATATATAACGAGCTGAAGCA GATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGGCAAAAGATATGAACAGGCATTTTAGC AAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGA TGCAGATTAAGACGACCATGCGATACCACCTTACCCCAGTGAGGATGGCAATTATCAAGAAATCTGG CAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAAG CTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCG ATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGAATACAAGAGCTGTTGTTACAAGGATACGTG TACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCACA ATGATCGATTGGATCAAGAAAATGTGGCATATTTATACCATGGAGTATTACGCAGCAATTAAGAATG ACGAATTTATTTCCTTCGTTGGGACCTGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTCA GGAGCAAAAGACAAAGCATAGAATCTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACGAC AAGTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGATACAATCGATTTCTGGATCC GCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGC TCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCT TTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCA GGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCAC CTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCC TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGA AGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTG CTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCtgagagacacaaaaaattccaaca cactattgcaatgaaaataaatttcctttattagccagaagtcagatgctcaaggggcttcatgatg tccccataatttttggcagagggaaaaagatctcagtggtatttgtgagccagggcattggccttct gataggcagcctgcacctgaggagtgcggccgctttacttgtacagctcgtccatgccgagagtgat cccggcggcggtcacgaactccagcaggaccatgtgatcgcgcttctcgttggggtctttgctcagg gcggactgggtgctcaggtagtggttgtcgggcagcagcacggggccgtcgccgatgggggtgttct gctggtagtggtcggcgagctgcacgctgccgtcctcgatgttgtggcggatcttgaagttcacctt gatgccgttcttctgcttgtcggccatgatatagacgttgtggctgttgtagttgtactccagcttg tgccccaggatgttgccgtcctccttgaagtcgatgcccttcagctcgatgcggttcaccagggtgt cgccctcgaacttcacctcggcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgcgctc ctggacgtagccttcgggcatggcggacttgaagaagtcgtgctgcttcatgtggtcggggtagcgg ctgaagcactgcacgccgtaggtcagggtggtcacgagggtgggccagggcacgggcagcttgccgg tggtgcagatgaacttcagggtcagcttgccgtaggtggcatcgccctcgccctcgccggacacgct gaacttgtggccgtttacgtcgccgtccagctcgaccaggatgggcaccaccccggtgaacagctcc tcgcccttgctcaccatggtggcgggatctgacggttcactaaaccagctctgcttatatagacctc ccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacattttggaaagtcc cgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggagacttggaaatccccgtg agtcaaaccgctatccacgcccattgatgtactgccaaaaccgcatcaccatggtaatagcgatgac taatacgtagatgtactgccaagtaggaaagtcccataaggtcatgtactgggcataatgccaggcg ggccatttaccgtcattgacgtcaataggggggcgtacttggcatatgatacacttgatgtactgcca agtgggcagtttaccgtaaatactccacccattgacgtcaatggaaagtccctattggcgttactat gggaacatacgtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccattt accgtaagttatgtaacgGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTC

	AGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGT ACCTTCAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTG AAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAAT AAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTT TTTAAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCC CTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGC ATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAG CAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTA AGGTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCT GGACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTT GGGCTATCGAAACTTAATTAAGCGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAAC TCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTT TTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTT TTGGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTT CTCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGA TGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGT GCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCG CAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCA GGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGG CTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCAC GTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCC AGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACGGC GATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGC TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGG CGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTC TATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTTG TTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAA TGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGAC AGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAATT AACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCGTC AGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCATGT GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCT CCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTA TAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTA CCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTA TCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGAC CGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGG CAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG GTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC TTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTG TTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGG GTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATC TTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTT GGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATC CATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAATCTT(서열번호 99)
RET-010 LINE1-GFP ORF2-NSV40-NLS	GGGGCGCGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATATCGTTGCTTG GCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTGTACTCAAGCC TGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCAACGTATAATC GAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGCATGAGTATTCA ACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAA ACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACATCGAACTGGATC TCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGATGAGCACTTTTAA AGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATA CACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGA CAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGAC AACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTT GATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAG CAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAGTT GATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGG TTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAG ATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAA TAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCGATTCTAGGTGCATTGGCG CAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGA

[0793]

	TACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGATCGT TTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAACAGCCGTGGCGCTCA TTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGGCAGCGATCGCGGCTC CCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCATAACAGCAGCTTCAG CTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTATCAATCG TTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGATAGCGATTTTATTATCTAACT GCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAG TTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGA CGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGA GTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATT GACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTA CTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATGCGGTTTTGGCAGTACATCAA TGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGT TTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAA TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCAG ATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCGCTAATACGACTCACTATAGGG AGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGGAATTCCAAGACACAATCCGCTA GCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAGTCCTGGATGGAAAACGACTTCGA TGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAACTCCGCGAAGATATCCAGACAAAG GGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTGCATCACCCGTATCACAAACACTGAGA AATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCAGGGAGCTTCGAGAGGAGTGTCGGAGTCT GAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGCGCGATGGAAGACGAGATGAACGAGATGAAA AGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGAGGAACGAACAGAGTCTGCAGGAGATTTGGGATT ACGTCAAGAGGCCTAACCTGCGGTTGATCGGCGTCCCCGAGAGCGACGTAGAAAACGGGACTAAACT GGAGAATACACTTCAAGACATCATTCAAGAAAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAA ATCCAAGAGATCCAACGCACACCCCAGCGGTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCG TGCGCTTTACTAAGGTGGAGATGAAAGAGAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGAC TTTGAAGGGCAAACCTATTCGGCTGACGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAA TGGGGCCCCATCTTTAATATCCTGAAGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGT TGAGTTTTATCTCCGAGGGTGAGATTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGAC AACTCGCCCAGCTCTCAAGGAACTGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAA CCCTTGCAGAACCACGCAAAGATGTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCATC AACTAATGAGCAAAATCACCAGCTAACATCATAGTATACATGccaaagaagaagcggaaggtcACCG GCTCTAACTCACATATCACCATCCTTACACTTAACATTAACGGCCTCAACTCAGCTATCAAGCGCCA TCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTTTGTTGCATCCAAGAGACCCACCTGACC TGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGAAAGATTTATCAGGCGAACGGTAAGCAGA AGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGACGGATTTCAAGCCCACCAAAATTAAGCGTGA TAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATACAGCAGGAAGAACTTACCATATTGAACATC TACGCGCCAAACACCGGCGCACCTCGCTTTATCAAACAGGTCCTGTCCGATCTGCAGCGAGATCTGG ATTCTCATACGTTGATTATGGGTGATTTCAATACACCATTGAGCACCCTGGATCGCAGCACCAGGCA AAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCACTGCATCAGGCAGATCTCATTGATATTTAT CGCACTCTTCATCCTAAGAGTACCGAGTACACATTCTTCAGCGCCCCACATCATACATACTCAAAGA TCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAAGTGCAAGCGCACAGAGATAATTACAAATTA CCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGAATCAAGAACCTGACCCAGAGCCGGAGTACC ACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATTGGGTCCACAATGAGATGAAGGCAGAGATTA AAATGTTCTTCGAAACAAATGAGAATAAGGATACTACCTATCAAAACCTTTGGGATGCCTTTAAGGC CGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCGAT ACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACGGC AGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATTAA CGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTAAG AAGAAGCGCGAAAAGAACCAGATTGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCCGA CCGAGATCCAGACCACTATTCGGGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTGGA AGAGATGGACACTTTTCTGGATACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTC AACCGCCCAATTACAGGAAGCGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTG GTCCGGACGGGTTTACAGCTGAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAA GCTCTTTCAGTCTATAGAAAAGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTT ATTCCCAAACCAGGACGCGATACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCG ACGCTAAAATATTGAACAAGATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACATCA CGACCAGGTGGGGTTTATACCTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATT CAACACATTAATAGAGCTAAGGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCG ATAAGATTCAGCAGCCATTTATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAA GATTATTCGCGCAATTTACGATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCC TTTCCGCTCAAGACTGGAACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCG AGGTGCTGGCTAGGGCTATTCGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGT

[0794]

	CAAGCTTTCCTTGTTCGCCGATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAAC CTTCTTAAACTTATTTCTAACTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGG CCTTTCTGTACACAAATAATCGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGC CAGCAAAAGGATAAAGTATCTCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAAT TACAAGCCTCTCCTGAAAGAGATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGG TTGGCAGAATCAACATAGTGAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCC AATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAA AAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTG ATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACAT CGACCAGTGGAATCGGACCGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGAT AAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGC TGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTG GATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATA CAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAAGA TTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGTTAA TAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAGCAGA ATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGGCAAAAG ATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAGTGTAGTTC AAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACCTTACCCCAGTGAGGATG GCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGCACCCTGCTCC ATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTCTGAGGGACCT CGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGAATACAAGAGC TGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCTAAGACGTGGA ATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTGGCATATTTATACCATGGAGTA TTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACCTGGATGAAGCTGGAGACTATT ATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTCATTGGTGGTAACG ACTACAAAGACGATGACGACAAGTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGA TACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACT GGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATG CTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGA GGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACT GGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCA CGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAA TTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATT CTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCtgag agacacaaaaaattccaacacactattgcaatgaaaataaatttcctttattagccagaagtcagat gctcaaggggcttcatgatgtccccataatttttggcagagggaaaaagatctcagtggtatttgtg agccagggcattggccttctgataggcagcctgcacctgaggagtgcggccgctttacttgtacagc tcgtccatgccgagagtgatcccggcggcggtcacgaactccagcaggaccatgtgatcgcgcttct cgttggggtctttgctcagggcggactgggtgctcaggtagtggttgtcgggcagcagcacggggcc gtcgccgatgggggtgttctgctggtagtggtcggcgagctgcacgctgccgtcctcgatgttgtgg cggatcttgaagttcaccttgatgccgttcttctgcttgtcggccatgatatagacgttgtggctgt tgtagttgtactccagcttgtgccccaggatgttgccgtcctccttgaagtcgatgcccttcagctc gatgcggttcaccagggtgtcgccctcgaacttcacctcggcgcgggtcttgtagttgccgtcgtcc ttgaagaagatggtgcgctcctggacgtagccttcgggcatggcggacttgaagaagtcgtgctgct tcatgtggtcggggtagcggctgaagcactgcacgccgtaggtcagggtggtcacgagggtgggcca gggcacgggcagcttgccggtggtgcagatgaacttcagggtcagcttgccgtaggtggcatcgccc tcgccctcgccggacacgctgaacttgtggccgtttacgtcgccgtccagctcgaccaggatgggca ccaccccggtgaacagctcctcgcccttgctcaccatggtggcgggatctgacggttcactaaacca gctctgcttatatagacctcccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgt tacgacattttggaaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtg gagacttggaaatcccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaaccgcatc accatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcccataaggtcatgt actgggcataatgccaggcgggccatttaccgtcattgacgtcaataggggggcgtacttggcatatg atacacttgatgtactgccaagtgggcagtttaccgtaaatactccacccattgacgtcaatggaaa gtccctattggcgttactatgggaacatacgtcattattgacgtcaatgggcgggggtcgttgggcg gtcagccaggcgggccatttaccgtaagttatgtaacgGGCCTGCTGCCGGCTCTGCGGCCTCTTCC GCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTT GACTGACTGAGATACAGCGTACCTTCAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAA ACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG TAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCA GGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCCCATCTCTA TCGGTATCGTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTGCCCCTCGGG CCGGATTGCTATCTACCGGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACT

[0795]

	CCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCC TTACCAGAAATTTATCCTTAAGGTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGATCAAGACA TTCCTTTAATGGTCTTTTCTGGACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCC CTAATCTCATTGGTTACCTTGGGCTATCGAAACTTAATTAAGCGATCTGCATCTCAATTAGTCAGCA ACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGC CCCATCGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCA GAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATGATTGAACA AGATGGATTGCACGCAGGTTCTCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAA CAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTG TCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGC CACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTA TTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCA TGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAA ACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAA GAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGG ATCTCGTCGTGACCCACGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGG ATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGAT ATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCG ATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTC TAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCC ACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGG GGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC GGTGGGCTCTATGGTTAATTAACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCG ACTACGCTCAGAATTGCGTCAGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTAC GAGTTTCATTTAAATCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTG CTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGT GGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT CATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACG AACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAG ACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCG CTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGC TGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGAT CCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCA TGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTA AAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCG ATCTGTCTATTTCGTTCATCCATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAA ATCTTCAAACCTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTG GCCGGCCTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATAT CCGAGATCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAG GAATATCGAAATC(서열번호 100)
ret-011-line1-gfp-orf2-n-sv40-nls-링커	TGGCCGGCCTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAAT ATCCGAGATCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCG AGGAATATCGAAATCGGGGCGCGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGAT CCATATCGTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGA TATTGTACTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATC GGTCAACGTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTT TCGCATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTT TTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTT ACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAAT GATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAA CTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATC TTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGC CAACTTACTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGAT CATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACA CCACGATGCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTAGC TTCCCGGCAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCC CTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTG CAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAAC TATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCGATT CTAGGTGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCC AGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATT

[0796]

	TATCCTTAAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAA CAGCCGTGGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGG CAGCGATCGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCA TAACAGCAGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGG TTATGCTATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGATAGCGA TTTTATTATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTAGTTCAT AGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACG ACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTG ACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCA AGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCT TATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATGCGGTT TTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCG CCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGT GAACCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCGCTAATA CGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGGAATTCCA AGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAGTCCTGGAT GGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAACTCCGCGAA GATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTGCATCACCCGTA TCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCAGGGAGCTTCGAGA GGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGCGCGATGGAAGACGAG ATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGAGGAACGAACAGAGTCTGC AGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGCGTCCCCGAGAGCGACGTAGA AAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGAAAATTTTCCAAACCTGGCTCGG CAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCGGTATAGCTCTCGGCGTGCCACCC CTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGAGAAGATGCTGCGAGCCGCTCGGGA AAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGACGGTTGACCTTAGCGCCGAGACACTC CAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGAAGGAGAAGAACTTCCAGCCACGAATCT CTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGATTAAGTATTTCATCGATAAACAGATGCT GCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAACTGCTCAAAGAGGCTCTTAATATGGAGCGC AATAATAGATATCAACCCTTGCAGAACCACGCAAAGATGTGAGACAGCCGTCAGACCATCAAGACTA GGAAGAAACTGCATCAACTAATGAGCAAAATCACCAGCTAACATCATAGTATACATGccaaagaaga agcggaaggtcggcggcggcagcACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAA CGGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTT TGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGAA AGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGACGGA TTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATACAG CAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAACAGG TCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACACCATT GAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCACTG CATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCTTCA GCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAAGTG CAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGAATC AAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATTGGG TCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTACCTA TCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTATAAA AGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGGAAC AGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGATCGA AACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAGATA GACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGATTGATACCATCAAGAACG ACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCGGGAGTATTATAAGCATTT GTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGATACCTATACTCTGCCACGG CTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCGAGATTGTGGCCATAATTA ACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGAGTTTTATCAACGGTATAT GGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAGGAAGGCATCTTGCCCAAT TCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCGATACCACAAAGAAGGAAAACT TCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAAGATTCTCGCCAACAGAATCCA ACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATACCTGGCATGCAGGGCTGGTTT AACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTAAGGATAAGAATCATATGATCA TCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTTATGCTCAAGACTCTGAACAA ACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTACGATAAGCCGACTGCTAACATT ATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCCGCCAAGGCTGTCCCCTCT CCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCGTCAAGAGAAAGAGATTAA AGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGATGATATGATTGTGTACCTG

[0797]

[0798]

	GAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACTTTAGCAAGGTCAGCGGCT ATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAATCGACAGACCGAATCCCAGAT AATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCTCGGAATCCAGCTGACACGA GACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGAGATTAAGGAAGATACTAATA AGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGTGAAGATGGCAATACTTCCTAA AGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAGAAA ACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGAAGA ATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGACAGC CTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGGACCGAACCATCAGAGATAATGCCC CACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCT TCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTAC ACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACA CTGGAAGAGAATCTTGGGATCACAATACAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGC CCAAGGCCATGGCCACTAAGGATAAGATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTAC TGCCAAGGAGACCACGATCAGAGTTAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTAT TCATCAGATAAGGGGTTGATAAGCAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGA ATAATCCCATCAAGAAGTGGGCAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGC GAAGAAGCATATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATG CGATACCACCTTACCCCAGTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGG GCTGTGGCGAGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAA ATCAGTCTGGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTC GGAATCTATCCTAACGAATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGG CCTTGTTTACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAA AATGTGGCATATTTATACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTT GGGACCTGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATA GAATCTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACGACAAGTAAAGCGCTTCTAGAAG TTGTCTCCTCCTGCACTGACTGACTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTG GATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGAT ACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTA TAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGC ACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGA CTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGAC AGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGG CTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCA ATCCAGCGGACCTTCCTTCCCGCtgagagacacaaaaaattccaacacactattgcaatgaaaataa atttcctttattagccagaagtcagatgctcaaggggcttcatgatgtccccataatttttggcaga gggaaaaagatctcagtggtatttgtgagccagggcattggccttctgataggcagcctgcacctga ggagtgcggccgctttacttgtacagctcgtccatgccgagagtgatcccggcggcggtcacgaact ccagcaggaccatgtgatcgcgcttctcgttggggtctttgctcagggcggactgggtgctcaggta gtggttgtcgggcagcagcacggggccgtcgccgatgggggtgttctgctggtagtggtcggcgagc tgcacgctgccgtcctcgatgttgtggcggatcttgaagttcaccttgatgccgttcttctgcttgt cggccatgatatagacgttgtggctgttgtagttgtactccagcttgtgccccaggatgttgccgtc ctccttgaagtcgatgcccttcagctcgatgcggttcaccagggtgtcgccctcgaacttcacctcg gcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgcgctcctggacgtagccttcgggca tggcggacttgaagaagtcgtgctgcttcatgtggtcggggtagcggctgaagcactgcacgccgta ggtcagggtggtcacgagggtgggccagggcacgggcagcttgccggtggtgcagatgaacttcagg gtcagcttgccgtaggtggcatcgccctcgccctcgccggacacgctgaacttgtggccgtttacgt cgccgtccagctcgaccaggatgggcaccaccccggtgaacagctcctcgcccttgctcaccatggt ggcgggatctgacggttcactaaaccagctctgcttatatagacctcccaccgtacacgcctaccgc ccatttgcgtcaatggggcggagttgttacgacattttggaaagtcccgttgattttggtgccaaaa caaactcccattgacgtcaatggggtggagacttggaaatccccgtgagtcaaaccgctatccacgc ccattgatgtactgccaaaaccgcatcaccatggtaatagcgatgactaatacgtagatgtactgcc aagtaggaaagtcccataaggtcatgtactgggcataatgccaggcgggccatttaccgtcattgac gtcaatagggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgtaaa tactccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatacgtcattattg acgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttaccgtaagttatgtaacgGG CCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTT TGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGACATG ATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTG AAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAA TTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTC TACAAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCCCTTGGGGCCTCTAAACGGGT CTTGAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGCATTGGCGCAGAAAAAAATGC CTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCA

[0799]

	ACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGGTCGTCAGCTATCCTGCA GGCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCTGGACACCACTAGGGGTCAGA AGTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTTGGGCTATCGAAACTTAATTA AGCGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAA CTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTTTTTATTTATGCAGAGGCCGA GGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGC AAAGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTTCTCCCGCCGCGTTGGGTGGAG AGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGT CAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAGGA CGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTC ACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACC TTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGC TACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGT CTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGC TCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACGGCGATGCCTGCTTGCCGAATAT CATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTAT CAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCC TCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTT CTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCT TCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATT GTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGA AGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAATTAACCAGTCAAGTCAGCTACT TGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCGTCAGTCAAGTTCGATCTGGTCC TTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCATGTGAGCAAAAGGCCAGCAAAAG GCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATC ACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCC CCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTT CTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCG TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAA CTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACA CTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG CTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACG CGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACG AAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAA TTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGC TTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCATTTAAATTTC CGAACTCTCCAAGGCCCTCGTCGGAAAATCTTCAAACCTTTCGTCCGATCCATCTTGCAGGCTACCT CTCGAACGAACTATCGCAAGTCTCT(서열번호 101)
RET-012 LINE1-GFP ORF2-N 뉴클레오플라스민-NLS	GACGATCCATATCGTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATC GTCAGATATTGTACTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGT CTGATCGGTCAACGTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGG AGGCTTTCGCATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTT CCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAG TGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTT TCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAA GAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAA AGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACAC TGCGGCCAACTTACTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATG GGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGC GTGACACCACGATGCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTAC TCTAGCTTCCCGGCAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGC TCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTA TCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCA GGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAA CCGATTCTAGGTGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACA TATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCA GAAATTTATCCTTAAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTA CGCGAACAGCCGTGGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCT TTTTGGCAGCGATCGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAG TGGTCATAACAGCAGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCC AAGGGGTTATGCTATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGA TAGCGATTTTATTATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTA

	GTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGC CCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTT CCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCAT ATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACA TGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGAT GCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCAC CCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACA ACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGG TTTAGTGAACCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCG CTAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGGA ATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAGTC CTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAACTC CGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTGCATCA CCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCAGGGAGCT TCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGCGCGATGGAA GACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGAGGAACGAACAGA GTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGCGTCCCCGAGAGCGA CGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGAAAATTTTCCAAACCTG GCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCGGTATAGCTCTCGGCGTG CCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGAGAAGATGCTGCGAGCCGC TCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGACGGTTGACCTTAGCGCCGAG ACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGAAGGAGAAGAACTTCCAGCCAC GAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGATTAAGTATTTCATCGATAAACA GATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAACTGCTCAAAGAGGCTCTTAATATG GAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGATGTGAGACAGCCGTCAGACCATCA AGACTAGGAAGAAACTGCATCAACTAATGAGCAAAATCACCAGCTAACATCATAGTATACATGaaaa ggccggcggccacgaaaaaggccggccaggcaaaaaagaaaaagACCGGCTCTAACTCACATATCAC CATCCTTACACTTAACATTAACGGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATC AAATCACAGGATCCAAGCGTTTGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCC TCAAGATCAAGGGATGGCGAAAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAAT TCTGGTCTCAGACAAGACGGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATT ATGGTGAAAGGCAGCATACAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCG CACCTCGCTTTATCAAACAGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTAT GGGTGATTTCAATACACCATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACG CAAGAGCTCAATAGCGCACTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGA GTACCGAGTACACATTCTTCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTC AAAGGCTCTGCTGTCAAAGTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCG ATCAAGCTCGAGCTGAGAATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACC TGCTGCTCAACGATTATTGGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAA TGAGAATAAGGATACTACCTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTC ATCGCCCTCAACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGA AGGAGTTGGAGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCG CGCCGAGTTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTC TTCGAGCGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACC AGATTGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTAT TCGGGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTG GATACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAA GCGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGC TGAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAA AAGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCG ATACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAA GATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATA CCTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTA AGGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATT TATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTAC GATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAA CCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTAT TCGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCC GATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTA ACTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAA TCGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTAT CTCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAG AGATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGT GAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACC

[0800]

[0801]

	TTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCGA AGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATAA AGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGGACC GAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAATAAAC AGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAAACT CAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTTGAATGTC AAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATACAAGATATAGGCGTCGGCA AAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAAGATTGATAAGTGGGACCTTAT TAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGTTAATAGGCAGCCCACTACATGG GAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAGCAGAATATATAACGAGCTGAAGC AGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGGCAAAAGATATGAACAGGCATTTTAG CAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAG ATGCAGATTAAGACGACCATGCGATACCACCTTACCCCAGTGAGGATGGCAATTATCAAGAAATCTG GCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAA GCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTC GATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGAATACAAGAGCTGTTGTTACAAGGATACGT GTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCAC AATGATCGATTGGATCAAGAAAATGTGGCATATTTATACCATGGAGTATTACGCAGCAATTAAGAAT GACGAATTTATTTCCTTCGTTGGGACCTGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTC AGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACGA CAAGTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGATACAATCGATTTCTGGATC CGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTG CTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTC AGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCA CCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGC CTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGG AAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCT GCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCtgagagacacaaaaaattccaac acactattgcaatgaaaataaatttccttttattagccagaagtcagatgctcaaggggcttcatgat gtcccataatttttggcagagggaaaaagatctcagtggtatttgtgagccagggcattggccttc tgataggcagcctgcacctgaggagtgcggccgctttacttgtacagctcgtccatgccgagagtga tcccggcggcggtcacgaactccagcaggaccatgtgatcgcgcttctcgttggggtctttgctcag ggcggactgggtgctcaggtagtggttgtcgggcagcagcacggggccgtcgccgatgggggtgttc tgctggtagtggtcggcgagctgcacgctgccgtcctcgatgtttgtggcggatcttgaagttcacct tgatgccgttcttctgcttgtcggccatgatatagacgttgtggctgttgtagttgtactccagctt gtgccccaggatgttgccgtcctccttgaagtcgatgcccttcagctcgatgcggttcaccagggtg tcgccctcgaacttcacctcggcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgcgct cctggacgtagccttcgggcatggcggacttgaagaagtcgtgctgcttcatgtggtcggggtagcg gctgaagcactgcacgccgtaggtcagggtggtcacgagggtgggccagggcacgggcagcttgccg gtggtgcagatgaacttcagggtcagcttgccgtaggtggcatcgccctcgcccctcgccggacacgc tgaacttgtggccgtttacgtcgccgtccagctcgaccaggatgggcaccaccccggtgaacagctc ctcgcccttgctcaccatggtggcgggatctgacggttcactaaaccagctctgcttatatagacct cccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacattttggaaagtc ccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggagacttggaaatccccgt gagtcaaaccgctatccacgcccattgatgtactgccaaaaccgcatcaccatggtaatagcgatga ctaatacgtagatgtactgccaagtaggaaagtcccataaggtcatgtactgggcataatgccaggc gggccatttaccgtcattgacgtcaataggggggcgtacttggcatatgatacacttgatgtactgcc aagtgggcagtttaccgtaaatactccacccattgacgtcaatggaaagtccctattggcgttacta tgggaacatacgtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatt taccgtaagttatgtaacgGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCT CAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCG TACCTTCAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGT GAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAA TAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTT TTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACC CCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGG CATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCA GCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTT AAGGTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTC TGGACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCT TGGGCTATCGAAACTTAATTAAGCGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAA CTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTT TTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTT

[0802]

	TTTGGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGT TCTCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTG ATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGG TGCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGC GCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGC AGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCG GCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCA CGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGC CAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACGG CGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGG CTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCG GCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTT CTATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTT GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAA ATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGA CAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAAT TAACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCGT CAGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCATG TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGC TCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT ATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTT ACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGT ATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGA CCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT GGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTT GTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGG GGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGAT CTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACT TGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCAT CCATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAATCTTCAAACCTTTCGTCC GATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGCCTTGCGCCTTGGC TATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGATCGGGATCACCCG AGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAATATCGAAATCGGGGCG CGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTC(서열번호 102)
RET-013 LINE1-GFP ORF2-N 뉴클레오플라스민 NLS 링커	AGATCCGAGGAATATCGAAATCGGGGCGCGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCT CGACGATCCATATCGTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAAT CGTCAGATATTGTACTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAG TCTGATCGGTCAACGTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAG GAGGCTTTCGCATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCT TCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGA GTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCT TTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCA AGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAA AAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACA CTGCGGCCAACTTACTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACAT GGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAG CGTGACACCACGATGCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTA CTCTAGCTTCCCGGCAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCG CTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGT ATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTC AGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTA ACCGATTCTAGGTGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCAC ATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACC AGAAATTTATCCTTAAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTT ACGCGAACAGCCGTGGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACC TTTTTGGCAGCGATCGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTA GTGGTCATAACAGCAGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCC CAAGGGGTTATGCTATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGG ATAGCGATTTTATTATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATT AGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCG CCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTT

[0803]

	TCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCA TATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTAC ATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGA TGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCA CCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAAC AACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTG GTTTAGTGAACCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCC GCTAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGG AATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAGT CCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAACT CCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTGCATC ACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCAGGGAGC TTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGCGCGATGGA AGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGAGGAACGAACAG AGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGCGTCCCCGAGAGCG ACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGAAAATTTTCCAAACCT GGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCGGTATAGCTCTCGGCGT GCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGAGAAGATGCTGCGAGCCG CTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGACGGTTGACCTTAGCGCCGA GACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGAAGGAGAAGAACTTCCAGCCA CGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGATTAAGTATTTCATCGATAAAC AGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAACTGCTCAAAGAGGCTCTTAATAT GGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGATGTGAGACAGCCGTCAGACCATC AAGACTAGGAAGAAACTGCATCAACTAATGAGCAAAATCACCAGCTAACATCATAGTATACATGaaa aggccggcggccacgaaaaaggccggccaggcaaaaaagaaaaagggcggcggcagcACCGGCTCTA ACTCACATATCACCATCCTTACACTTAACATTAACGGCCTCAACTCAGCTATCAAGCGCCATCGGCT GGCCAGCTGGATCAAATCACAGGATCCAAGCGTTTGTTGCATCCAAGAGACCCACCTGACCTGTAGA GATACTCACCGCCTCAAGATCAAGGGATGGCGAAAGATTTATCAGGCGAACGGTAAGCAGAAGAAAG CCGGAGTCGCAATTCTGGTCTCAGACAAGACGGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGA AGGTCACTATATTATGGTGAAAGGCAGCATACAGCAGGAAGAACTTACCATATTGAACATCTACGCG CCAAACACCGGCGCACCTCGCTTTATCAAACAGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTC ATACGTTGATTATGGGTGATTTCAATACACCATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGT AAATAAAGACACGCAAGAGCTCAATAGCGCACTGCATCAGGCAGATCTCATTGATATTTATCGCACT CTTCATCCTAAGAGTACCGAGTACACATTCTTCAGCGCCCCACATCATACATACTCAAAGATCGATC ATATCGTCGGCTCAAAGGCTCTGCTGTCAAAGTGCAAGCGCACAGAGATAATTACAAATTACCTGTC AGATCATAGCGCGATCAAGCTCGAGCTGAGAATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGG AAGCTTAATAACCTGCTGCTCAACGATTATTGGGTCCACAATGAGATGAAGGCAGAGATTAAAATGT TCTTCGAAACAAATGAGAATAAGGATACTACCTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTG CAGAGGCAAGTTCATCGCCCTCAACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCGATACTCTC ACCTCTCAGCTGAAGGAGTTGGAGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGA TCACAAAGATTCGCGCCGAGTTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTC TCGTAGTTGGTTCTTCGAGCGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTAAGAAGAAG CGCGAAAAGAACCAGATTGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCCGACCGAGA TCCAGACCACTATTCGGGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGAT GGACACTTTTCTGGATACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGC CCAATTACAGGAAGCGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGG ACGGGTTTACAGCTGAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTT TCAGTCTATAGAAAAGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATACTTATTCCC AAACCAGGACGCGATACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATATCGACGCTA AAATATTGAACAAGATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACATCACGACCA GGTGGGGTTTATACCTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCATTCAACAC ATTAATAGAGCTAAGGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATTCGATAAGA TTCAGCAGCCATTTATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTTAAGATTAT TCGCGCAATTTACGATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCG CTCAAGACTGGAACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGC TGGCTAGGGCTATTCGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCT TTCCTTGTTCGCCGATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTT AAACTTATTTCTAACTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTC TGTACACAAATAATCGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAA AAGGATAAAGTATCTCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAG CCTCTCCTGAAAGAGATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCA GAATCAACATAGTGAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATCCCAATTAA ACTGCCTATGACCTTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATCAAAAGAGA GCAAGAATAGCGAAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTA

[0804]

AGTTGTATTATAAAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGACATCGACCA
GTGGAATCGGACCGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCGATAAGCCA
GAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTGGCTGGCCA
TATGCCGGAAACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGGTGGATCAA
GGACTTGAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAATACAAGAT
ATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAAGATTGATA
AGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGTTAATAGGCA
GCCCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAGCAGAATATAT
AACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGGCAAAAGATATGA
ACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAGTGTAGTTCAAGCTT
GGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACCTTACCCCAGTGAGGATGGCAATT
ATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGCACCCTGCTCCATTGCT
GGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTCTGAGGGACCTCGAGCT
TGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGAATACAAGAGCTGTTGT
TACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCTAAGACGTGGAATCAGC
CTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTGGCATATTTATACCATGGAGTATTACGC
AGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACCTGGATGAAGCTGGAGACTATTATTCTG
AGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTCATTGGTGGTAACGACTACA
AAGACGATGACGACAAGTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGATACAAT
CGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATT
CTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTG
CTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTT
GTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGG
GGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGG
AACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGT
GGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGC
GGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCtgagagacac
aaaaaattccaacacactattgcaatgaaaataaatttcctttattagccagaagtcagatgctcaa
ggggcttcatgatgtccccataatttttggcagagggaaaaagatctcagtggtatttgtgagccag
ggcattggccttctgataggcagcctgcacctgaggagtgcggccgctttacttgtacagctcgtcc
atgccgagagtgatcccggcggcggtcacgaactccagcaggaccatgtgatcgcgcttctcgttgg
ggtctttgctcagggcggactgggtgctcaggtagtggttgtcgggcagcagcacggggccgtcgcc
gatgggggtgttctgctggtagtggtcggcgagctgcacgctgccgtcctcgatgttgtggcggatc
ttgaagttcaccttgatgccgttcttctgcttgtcggccatgatatagacgttgtggctgttgtagt
tgtactccagcttgtgccccaggatgttgccgtcctccttgaagtcgatgcccttcagctcgatgcg
gttcaccagggtgtcgccctcgaacttcacctcggcgcgggtcttgtagttgccgtcgtccttgaag
aagatggtgcgctcctggacgtagccttcgggcatggcggacttgaagaagtcgtgctgcttcatgt
ggtcggggtagcggctgaagcactgcacgccgtaggtcagggtggtcacgagggtgggccagggcac
gggcagcttgccggtggtgcagatgaacttcagggtcagcttgccgtaggtggcatcgccctcgccc
tcgccggacacgctgaacttgtggccgtttacgtcgccgtccagctcgaccaggatgggcaccaccc
cggtgaacagctcctcgcccttgctcaccatggtggcgggatctgacggttcactaaaccagctctg
cttatatagacctcccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgac
attttggaaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggagact
tggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaaccgcatcaccatg
gtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcccataaggtcatgtactggg
cataatgccaggcgggccatttaccgtcattgacgtcaataggggggcgtacttggcatatgatacac
ttgatgtactgccaagtgggcagtttaccgtaaatactccacccattgacgtcaatggaaagtccct
attggcgttactatgggaacatacgtcattattgacgtcaatgggcgggggtcgttgggcggtcagc
caggcgggccatttaccgtaagttatgtaacgGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCT
TCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGA
CTGAGATACAGCGTACCTTCAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAAACCACA
ACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCA
TTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGA
GGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCCCATCTCTATCGGTA
TCGTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTGCCCCTCGGGCCGGAT
TGCTATCTACCGGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACA
TATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCA
GAAATTTATCCTTAAGGTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGATCAAGACATTCCTT
TAATGGTCTTTTCTGGACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCCCTAATC
TCATTGGTTACCTTGGGCTATCGAAACTTAATTAAGCGATCTGCATCTCAATTAGTCAGCAACCATA
GTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATC
GCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTA
GTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATGATTGAACAAGATGG
ATTGCACGCAGGTTCTCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACA

[0805]

	ATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGA CCGACCTGTCCGGTGCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGAC GGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGC GAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTG ATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCG CATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCAT CAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCG TCGTGACCCACGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCAT CGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCT GAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGC AGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTG CCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTC CTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTG GGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGG CTCTATGGTTAATTAACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACG CTCAGAATTGCGTCAGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTT CATTTAAATCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCG TTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCC GACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCC CCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGA CTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACA GAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGC TGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAG CGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTG ATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGAT TATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTAT ATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT CTATTTCGTTCATCCATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAATCTTC AAACCTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGC CTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGA TCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCG(서열번호 103)
RET-014 LINE1-GFP ORF2-CSV40-NLS	AGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATATCGTTGCTTGGCAGTCAGCCAGTCGGAATC CAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTGTACTCAAGCCTGGTCACGGCAGCGTACCGA TCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCAACGTATAATCGAGTCCTAGCTTTTGCAAAC ATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGCATGAGTATTCAACATTTCCGTGTCGCCCTTA TTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGA TGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTT GAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGG TATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTT GGTTGAGTATTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGT GCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATTGGAGGACCGAAGG AGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCT GAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACCTTGCGT AAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAGTTGATAGACTGGATGGAGGCGG ATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGG AGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATC GTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAG GTGCCTCACTGATTAAGCATTGGTAACCGATTCTAGGTGCATTGGCGCAGAAAAAAATGCCTGATGC GACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCC CACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGATCGTTTAAACTCGACTCTGGCTCT ATCGAATCTCCGTCGTTTCGAGCTTACGCGAACAGCCGTGGCGCTCATTTGCTCGTCGGGCATCGAA TCTCGTCAGCTATCGTCAGCTTACCTTTTTGGCAGCGATCGCGGCTCCCGACATCTTGGACCATTAG CTCCACAGGTATCTTCTTCCCTCTAGTGGTCATAACAGCAGCTTCAGCTACCTCTCAATTCAAAAAA CCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTATCAATCGTTGCGTTACACACACAAAAA ACCAACACACATCCATCTTCGATGGATAGCGATTTTATTATCTAACTGCTGATCGAGTGTAGCCAGA TCTAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTAC GGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTT CCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCC ACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATG GCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTA TTAGTCATCGCTATTACCATGCTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTG

[0806]

	ACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCA ACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGG TGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCAGATCTTTGTCGATCCTACCAT CCACTCGACACACCCGCCAGCGGCCGCTAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGG CAAGAAGCAAAATCGCAAGACGGGGAATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGT TCTAGCTCCCCTGCTACTGAGCAGTCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGAT TTAGGCGATCCAACTATTCAGAACTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTT CGAGAAGAACCTCGAGGAGTGCATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATG GAACTTAAGACAAAAGCCAGGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGC TCGAGGAGCGCGTGAGCGCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGA GAAGCGCATTAAGAGGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTG CGGTTGATCGGCGTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACA TCATTCAAGAAAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCAC ACCCCAGCGGTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAG ATGAAAGAGAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTC GGCTGACGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATAT CCTGAAGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGT GAGATTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGG AACTGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAA GATGTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCATCAACTAATGAGCAAAATCACC AGCTAACATCATAGTATACATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAAC GGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTTT GTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGAAA GATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGACGGAT TTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATACAGC AGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAACAGGT CCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACACCATTG AGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCACTGC ATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCTTCAG CGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAAGTGC AAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGAATCA AGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATTGGGT CCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTACCTAT CAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTATAAAA GAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGGAACA GACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGATCGAA ACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAGATAG ACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGATTGATACCATCAAGAACGA CAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCGGGAGTATTATAAGCATTTG TATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGATACCTATACTCTGCCACGGC TTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCGAGATTGTGGCCATAATTAA CTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGAGTTTTATCAACGGTATATG GAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAGGAAGGCATCTTGCCCAATT CCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCGATACCACAAAGAAGGAAAACTT CCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAAGATTCTCGCCAACAGAATCCAA CAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATACCTGGCATGCAGGGCTGGTTTA ACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTAAGGATAAGAATCATATGATCAT CTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTTATGCTCAAGACTCTGAACAAA CTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTACGATAAGCCGACTGCTAACATTA TCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCCGCCAAGGCTGTCCCCTCTC CCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCGTCAAGAGAAAGAGATTAAA GGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGATGATATGATTGTGTACCTGG AGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACTTTAGCAAGGTCAGCGGCTA TAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAATCGACAGACCGAATCCCAGATA ATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCTCGGAATCCAGCTGACACGAG ACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGAGATTAAGGAAGATACTAATAA GTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGTGAAGATGGCAATACTTCCTAAA GTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAGAAAA CAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGAAGAA TAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGACAGCC TGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGGACCGAACCATCAGAGATAATGCCCC ACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTT CAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTACA CCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACAC

	TGGAAGAGAATCTTGGGATCACAATACAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCC CAAGGCCATGGCCACTAAGGATAAGATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACT GCCAAGGAGACCACGATCAGAGTTAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTATT CATCAGATAAGGGGTTGATAAGCAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGAA TAATCCCATCAAGAAGTGGGCAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGCG AAGAAGCATATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGC GATACCACCTTACCCCAGTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGG CTGTGGCGAGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAA TCAGTCTGGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCG GAATCTATCCTAACGAATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGC CTTGTTTACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAA ATGTGGCATATTTATACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTG GGACCTGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAG AATCTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACGACAAGccaaagaagaagcggaag gtcTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGATACAATCGATTTCTGGATCC GCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGC TCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCT TTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCA GGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCAC CTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCC TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGA AGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTG CTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCtgagagacacaaaaaattccaaca cactattgcaatgaaaataaatttcctttattagccagaagtcagatgctcaaggggcttcatgatg tccccataatttttggcagagggaaaaagatctcagtggtatttgtgagccagggcattggccttct gataggcagcctgcacctgaggagtgcggccgctttacttgtacagctcgtccatgccgagagtgat cccggcggcggtcacgaactccagcaggaccatgtgatcgcgcttctcgttggggtctttgctcagg gcggactgggtgctcaggtagtggttgtcgggcagcagcacggggccgtcgccgatgggggtgttct gctggtagtggtcggcgagctgcacgctgccgtcctcgatgttgtggcggatcttgaagttcacctt gatgccgttcttctgcttgtcggccatgatatagacgttgtggctgttgtagttgtactccagcttg tgccccaggatgttgccgtcctccttgaagtcgatgcccttcagctcgatgcggttcaccagggtgt cgccctcgaacttcacctcggcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgcgctc ctggacgtagccttcgggcatggcggacttgaagaagtcgtgctgcttcatgtggtcggggtagcgg ctgaagcactgcacgccgtaggtcagggtggtcacgagggtgggccagggcacgggcagcttgccgg tggtgcagatgaacttcagggtcagcttgccgtaggtggcatcgccctcgccctcgccggacacgct gaacttgtggccgtttacgtcgccgtccagctcgaccaggatgggcaccaccccggtgaacagctcc tcgcccttgctcaccatggtggcgggatctgacggttcactaaaccagctctgcttatatagacctc ccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacattttggaaagtcc cgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggagacttggaaatccccgtg agtcaaaccgctatccacgcccattgatgtactgccaaaaccgcatcaccatggtaatagcgatgac taatacgtagatgtactgccaagtaggaaagtcccataaggtcatgtactgggcataatgccaggcg ggccatttaccgtcattgacgtcaataggggcgtacttggcatatgatacacttgatgtactgcca agtgggcagtttaccgtaaatactccacccattgacgtcaatggaaagtccctattggcgttactat gggaacatacgtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccattt accgtaagttatgtaacgGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTC AGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGT ACCTTCAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTG AAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAAT AAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTT TTTAAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCC CTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGC ATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAG CAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTA AGGTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCT GGACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTT GGGCTATCGAAACTTAATTAAGCGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAAC TCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTT TTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTT TTGGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTT CTCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGA TGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGT GCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCG CAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCA GGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGG

[0807]

	CTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCAC GTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCC AGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACGGC GATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGC TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGG CGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTC TATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTTG TTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAA TGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGAC AGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAATT AACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCGTC AGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCATGT GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCT CCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTA TAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTA CCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTA TCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGAC CGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGG CAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG GTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC TTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTG TTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGG GTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATC TTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTT GGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATC CATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAATCTTCAAACCTTTCGTCCG ATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGCCTTGCGCCTTGGCT ATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGATCGGGATCACCCGA GAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAATATCGAAATCGGGGCGC GCCTGGTGTACCG(서열번호 104)
RET-015 LINE1-GFP ORF2-C_링커_SV40-NLS	TCGACGATCCATATCGTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAA TCGTCAGATATTGTACTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATA GTCTGATCGGTCAACGTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCA GGAGGCTTTCGCATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCC TTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCG AGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGC TTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGC AAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGA AAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAC ACTGCGGCCAACTTACTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACA TGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGA GCGTGACACCACGATGCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTT ACTCTAGCTTCCCGGCAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGC GCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGG TATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGT CAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGT AACCGATTCTAGGTGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCA CATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTAC CAGAAATTTATCCTTAAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCT TACGCGAACAGCCGTGGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTAC CTTTTTGGCAGCGATCGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCT AGTGGTCATAACAGCAGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCC CCAAGGGGTTATGCTATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATG GATAGCGATTTTATTATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCAT TAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC GCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACT TTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATC ATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTA CATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTG ATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAA CAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT GGTTTAGTGAACCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGC

[0808]

	CGCTAATACGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGG GAATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAG TCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAAC TCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTGCAT CACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCAGGGAG CTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGCGCGATGG AAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGAGGAACGAACA GAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGCGTCCCCGAGAGC GACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGAAAATTTTCCAAACC TGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCGGTATAGCTCTCGGCG TGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGAGAAGATGCTGCGAGCC GCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGACGGTTGACCTTAGCGCCG AGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGAAGGAGAAGAACTTCCAGCC ACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGATTAAGTATTTCATCGATAAA CAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAACTGCTCAAAGAGGCTCTTAATA TGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGCAAAGATGTGAGACAGCCGTCAGACCAT CAAGACTAGGAAGAAACTGCATCAACTAATGAGCAAAATCACCAGCTAACATCATAGTATACATGAC CGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAACGGCCTCAACTCAGCTATCAAGCGC CATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTTTGTTGCATCCAAGAGACCCACCTGA CCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGAAAGATTTATCAGGCGAACGGTAAGCA GAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGACGGATTTCAAGCCCACCAAAATTAAGCGT GATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATACAGCAGGAAGAACTTACCATATTGAACA TCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAACAGGTCCTGTCCGATCTGCAGCGAGATCT GGATTCTCATACGTTGATTATGGGTGATTTCAATACACCATTGAGCACCCTGGATCGCAGCACCAGG CAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCACTGCATCAGGCAGATCTCATTGATATTT ATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCTTCAGCGCCCCACATCATACATACTCAAA GATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAAGTGCAAGCGCACAGAGATAATTACAAAT TACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGAATCAAGAACCTGACCCAGAGCCGGAGTA CCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATTGGGTCCACAATGAGATGAAGGCAGAGAT TAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTACCTATCAAAACCTTTGGGATGCCTTTAAG GCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTATAAAAGAAAACAAGAGAGATCTAAGATCG ATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGGAACAGACCCACTCCAAGGCGTCAAGACG GCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGATCGAAACCCAAAAGACTCTTCAGAAAATT AACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAGATAGACAGACCTCTGGCACGACTGATTA AGAAGAAGCGCGAAAAGAACCAGATTGATACCATCAAGAACGACAAGGGCGACATCACTACTGACCC GACCGAGATCCAGACCACTATTCGGGAGTATTATAAGCATTTGTATGCTAACAAGCTTGAGAACCTG GAAGAGATGGACACTTTTCTGGATACCTATACTCTGCCACGGCTTAATCAAGAGGAAGTCGAGTCCC TCAACCGCCCAATTACAGGAAGCGAGATTGTGGCCATAATTAACTCCCTGCCGACAAAGAAATCTCC TGGTCCGGACGGGTTTACAGCTGAGTTTTATCAACGGTATATGGAAGAGCTTGTACCGTTTCTGCTC AAGCTCTTTCAGTCTATAGAAAAGGAAGGCATCTTGCCCAATTCCTTCTACGAAGCTTCTATAATAC TTATTCCCAAACCAGGACGCGATACCACAAAGAAGGAAAACTTCCGGCCCATTAGTCTCATGAATAT CGACGCTAAAATATTGAACAAGATTCTCGCCAACAGAATCCAACAACATATTAAGAAATTGATACAT CACGACCAGGTGGGGTTTATACCTGGCATGCAGGGCTGGTTTAACATCCGGAAGAGTATTAACGTCA TTCAACACATTAATAGAGCTAAGGATAAGAATCATATGATCATCTCTATAGACGCGGAAAAGGCATT CGATAAGATTCAGCAGCCATTTATGCTCAAGACTCTGAACAAACTCGGCATCGACGGAACATATTTT AAGATTATTCGCGCAATTTACGATAAGCCGACTGCTAACATTATCCTTAACGGCCAAAAGCTCGAGG CCTTTCCGCTCAAGACTGGAACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTTTAATATTGTACT CGAGGTGCTGGCTAGGGCTATTCGTCAAGAGAAAGAGATTAAAGGGATACAGCTCGGGAAGGAAGAG GTCAAGCTTTCCTTGTTCGCCGATGATATGATTGTGTACCTGGAGAATCCTATTGTGTCTGCTCAGA ACCTTCTTAAACTTATTTCTAACTTTAGCAAGGTCAGCGGCTATAAGATTAACGTCCAGAAATCTCA GGCCTTTCTGTACACAAATAATCGACAGACCGAATCCCAGATAATGGGTGAGCTTCCGTTTGTCATA GCCAGCAAAAGGATAAAGTATCTCGGAATCCAGCTGACACGAGACGTTAAAGATTTGTTTAAGGAAA ATTACAAGCCTCTCCTGAAAGAGATTAAGGAAGATACTAATAAGTGGAAGAATATCCCCTGTTCATG GGTTGGCAGAATCAACATAGTGAAGATGGCAATACTTCCTAAAGTGATATATCGCTTTAACGCCATC CCAATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAGAAAACAACCCTTAAATTTATATGGAATC AAAAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGGGATTACTTTGCC TGATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGACAGCCTGGTATTGGTATCAGAATAGAGAC ATCGACCAGTGGAATCGGACCGAACCATCAGAGATAATGCCCCACATCTATAATTACCTTATATTCG ATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTGTTGGGAGAATTG GCTGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAAATCAACAGTAGG TGGATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATCTTGGGATCACAA TACAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGCCACTAAGGATAA GATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACCACGATCAGAGTT AATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGGGGTTGATAAGCA

[0809]

[0810]

	GAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAAGAAGTGGGCAAA AGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATATGAAGAAGTGTAGT TCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACCTTACCCCAGTGAGGA TGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCGAGATTGGCACCCTGCT CCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGCTTTCTGAGGGAC CTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTATCCTAACGAATACAAGA GCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTTACGATAGCTAAGACGTG GAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTGGCATATTTATACCATGGAG TATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACCTGGATGAAGCTGGAGACTA TTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCTCATTGGTGGTAA CGACTACAAAGACGATGACGACAAGggcggcggcagcccaaagaagaagcggaaggtcTAAAGCGCT TCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGATACAATCGATTTCTGGATCCGCAGGCCTAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCT ATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCC TCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCG TGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCT TTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGC TGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCT TTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTC GGCCCTCAATCCAGCGGACCTTCCTTCCCGCtgagagacacaaaaaattccaacacactattgcaat gaaaataaatttccttttattagccagaagtcagatgctcaaggggcttcatgatgtccccataattt ttggcagagggaaaaagatctcagtggtatttgtgagccagggcattggccttctgataggcagcct gcacctgaggagtgcggccgctttacttgtacagctcgtccatgccgagagtgatcccggcggcggt cacgaactccagcaggaccatgtgatcgcgcttctcgttggggtctttgctcagggcggactgggtg ctcaggtagtggttgtcgggcagcagcacggggccgtcgccgatgggggtgttctgctggtagtggt cggcgagctgcacgctgccgtcctcgatgttgtggcggatcttgaagttcaccttgatgccgttctt ctgcttgtcggccatgatatagacgttgtggctgttgtagttgtactccagcttgtgccccaggatg ttgccgtcctccttgaagtcgatgcccttcagctcgatgcggttcaccagggtgtcgccctcgaact tcacctcggcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgcgctcctggacgtagcc ttcgggcatggcggacttgaagaagtcgtgctgcttcatgtggtcggggtagcggctgaagcactgc acgccgtaggtcagggtggtcacgagggtgggccagggcacgggcagcttgccggtggtgcagatga acttcagggtcagcttgccgtaggtggcatcgccctcgccctcgccggacacgctgaacttgtggcc gtttacgtcgccgtccagctcgaccaggatgggcaccaccccggtgaacagctcctcgcccttgctc accatggtggcgggatctgacggttcactaaaccagctctgcttatatagacctcccaccgtacacg cctaccgcccatttgcgtcaatgggggcggagttgttacgacattttggaaagtcccgttgattttgg tgccaaaacaaactcccattgacgtcaatggggtggagacttggaaatccccgtgagtcaaaccgct atccacgcccattgatgtactgccaaaaccgcatcaccatggtaatagcgatgactaatacgtagat gtactgccaagtaggaaagtcccataaggtcatgtactgggcataatgccaggcgggccatttaccg tcattgacgtcaataggggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagttt accgtaaatactccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatacgt cattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttaccgtaagttat gtaacgGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGA TCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCA CAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTT TATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAC AACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGT AAAACCTCTACAAATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCCCTTGGGGCCTCT AAACGGGTCTTGAGGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGCATTGGCGCAGAA AAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGG CTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGGTCGTCAGCT ATCCTGCAGGCGATCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCTGGACACCACTAG GGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTTGGGCTATCGAAA CTTAATTAAGCGATCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCC GCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTTTTTATTTATGCA GAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAG GCTTTTGCAAAGGAGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTTCTCCCGCCGCTT GGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTT CCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAA CTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCG ACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTC ATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTT GATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGG AAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTT CGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACGGCGATGCCTGCTTG

	CCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGG ACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGA CCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTT GACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCC CCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGC ATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAG GATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAATTAACCAGTCAAGT CAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCGTCAGTCAAGTTCGA TCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCATGTGAGCAAAAGGCC AGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGA CGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGT CCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGT GTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTA TCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTG GTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA CGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGC AGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCA GTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATC CTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTT ACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCATT TAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAATCTTCAAACCTTTCGTCCGATCCATCTTGCA GGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGCCGGCCTTGCGCCTTGGCTATTGCTTGGCAG CGCCTATCGCCAGGTATTACTCCAATCCCGAATATCCGAGATCGGGATCACCCGAGAGAAGTTCAAC CTACATCCTCAATCCCGATCTATCCGAGATCCGAGGAATATCGAAATCGGGGCGCGCCTGGTGTACC GAGAACGATCCTCTCAGTGCGAGTC(서열번호 105)
ret-016-line1-gfp-orf2-c_뉴클레오플라스민-nls	ACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATATCGTTGCTTGGCAGTCAGCCAGTCGG AATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTGTACTCAAGCCTGGTCACGGCAGCGTA CCGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCAACGTATAATCGAGTCCTAGCTTTTGC AAACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGCATGAGTATTCAACATTTCCGTGTCGCC CTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAA AAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGAT CCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATG ACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATG CAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATTGGAGGACCG AAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGG AGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACCTT GCGTAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAGTTGATAGACTGGATGGAG GCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAAT CTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCG TATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAG ATAGGTGCCTCACTGATTAAGCATTGGTAACCGATTCTAGGTGCATTGGCGCAGAAAAAAATGCCTG ATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACT TGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGATCGTTTAAACTCGACTCTGG CTCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAACAGCCGTGGCGCTCATTTGCTCGTCGGGCAT CGAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGGCAGCGATCGCGGCTCCCGACATCTTGGACCA TTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCATAACAGCAGCTTCAGCTACCTCTCAATTCAA AAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTATCAATCGTTGCGTTACACACACA AAAAACCAACACACATCCATCTTCGATGGATAGCGATTTTATTATCTAACTGCTGATCGAGTGTAGC CAGATCTAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAAC TTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTA TGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTA AATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA CGTATTAGTCATCGCTATTACCATGCTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGG TTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAA ATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCAGATCTTTGTCGATCCTA CCATCCACTCGACACACCCGCCAGCGGCCGCTAATACGACTCACTATAGGGAGAAGTACTGCCACCA TGGGCAAGAAGCAAAATCGCAAGACGGGGAATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGA GCGTTCTAGCTCCCCTGCTACTGAGCAGTCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAG

[0811]

	GGATTTAGGCGATCCAACTATTCAGAACTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGA ATTTCGAGAAGAACCTCGAGGAGTGCATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACT CATGGAACTTAAGACAAAAGCCAGGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGAC CAGCTCGAGGAGCGCGTGAGCGCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCA GGGAGAAGCGCATTAAGAGGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAA CCTGCGGTTGATCGGCGTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAA GACATCATTCAAGAAAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAAC GCACACCCCAGCGGTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGT GGAGATGAAAGAGAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCT ATTCGGCTGACGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTA ATATCCTGAAGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGA GGGTGAGATTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTC AAGGAACTGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACG CAAAGATGTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCATCAACTAATGAGCAAAAT CACCAGCTAACATCATAGTATACATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAACAT TAACGGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGC GTTTGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGC GAAAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGAC GGATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATA CAGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAAC AGGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACACC ATTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCA CTGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCT TCAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAA GTGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGA ATCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATT GGGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTAC CTATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTAT AAAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGG AACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGAT CGAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAG ATAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGATTGATACCATCAAGA ACGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCGGGAGTATTATAAGCA TTTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGATACCTATACTCTGCCA CGGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCGAGATTGTGGCCATAA TTAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGAGTTTTATCAACGGTA TATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAGGAAGGCATCTTGCCC AATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCGATACCACAAAGAAGGAAA ACTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAAGATTCTCGCCAACAGAAT CCAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATACCTGGCATGCAGGGCTGG TTTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTAAGGATAAGAATCATATGA TCATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTTATGCTCAAGACTCTGAA CAAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTACGATAAGCCGACTGCTAAC ATTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCCGCCAAGGCTGTCCCC TCTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCGTCAAGAGAAAGAGAT TAAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGATGATATGATTGTGTAC CTGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACTTTAGCAAGGTCAGCG GCTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAATCGACAGACCGAATCCCA GATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCTCGGAATCCAGCTGACA CGAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGAGATTAAGGAAGATACTA ATAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGTGAAGATGGCAATACTTCC TAAAGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAG AAAACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGA AGAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGAC AGCCTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGGACCGAACCATCAGAGATAATG CCCCACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCC TCTTCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCT TACACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAG ACACTGGAAGAGAATCTTGGGATCACAATACAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGA CGCCCAAGGCCATGGCCACTAAGGATAAGATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTG TACTGCCAAGGAGACCACGATCAGAGTTAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACT TATTCATCAGATAAGGGGTTGATAAGCAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAA CGAATAATCCCATCAAGAAGTGGGCAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGC CGCGAAGAAGCATATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACC

[0812]

[0813]

	ATGCGATACCACCTTACCCCAGTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGC GGGGCTGTGGCGAGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTG GAAATCAGTCTGGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTG CTCGGAATCTATCCTAACGAATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCG CGGCCTTGTTTACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAA GAAAATGTGGCATATTTATACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTC GTTGGGACCTGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGC ATAGAATCTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACGACAAGaaaaggccggcggc cacgaaaaaggccggccaggcaaaaaagaaaaagTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCAC TGACTGACTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTG AAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCC TTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTG TCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACG CAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCT CCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTG GGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTG CCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCC TTCCCGCtgagagacacaaaaaattccaacacactattgcaatgaaaataaatttccttttattagcc agaagtcagatgctcaaggggcttcatgatgtcccctataatttttggcagagggaaaaagatctcag tggtatttgtgagccagggcattggccttctgataggcagcctgcacctgaggagtgcggccgcttt acttgtacagctcgtccatgccgagagtgatcccggcggcggtcacgaactccagcaggaccatgtg atcgcgcttctcgttggggtctttgctcagggcggactgggtgctcaggtagtggttgtcgggcagc agcacggggccgtcgccgatgggggtgttctgctggtagtggtcggcgagctgcacgctgccgtcct cgatgttgtggcggatcttgaagttcaccttgatgccgttcttctgcttgtcggccatgatatagac gttgtggctgttgtagttgtactccagcttgtgccccaggatgttgccgtcctccttgaagtcgatg cccttcagctcgatgcggttcaccagggtgtcgccctcgaacttcacctcggcgcgggtcttgtagt tgccgtcgtccttgaagaagatggtgcgctcctggacgtagccttcgggcatggcggacttgaagaa gtcgtgctgcttcatgtggtcggggtagcggctgaagcactgcacgccgtaggtcagggtggtcacg agggtgggccagggcacgggcagcttgccggtggtgcagatgaacttcagggtcagcttgccgtagg tggcatcgccctcgccctcgccggacacgctgaacttgtggccgtttacgtcgccgtccagctcgac caggatgggcaccaccccggtgaacagctcctcgcccttgctcaccatggtggcgggatctgacggt tcactaaaccagctctgcttatatagacctcccaccgtacacgcctaccgcccatttgcgtcaatgg ggcggagttgttacgacattttggaaagtcccgttgattttggtgccaaaacaaactcccattgacg tcaatggggtggagacttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgcc aaaaccgcatcaccatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtccca taaggtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaataggggggcgta cttggcatatgatacacttgatgtactgccaagtgggcagtttaccgtaaatactccacccattgac gtcaatggaaagtccctattggcgttactatgggaacatacgtcattattgacgtcaatgggcgggg gtcgttgggcggtcagccaggcgggccatttaccgtaagttatgtaacgGGCCTGCTGCCGGCTCTG CGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGC CTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGACATGATAAGATACATTGATG AGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTAT TGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATG TTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATTG GCCCATCTCTATCGGTATCGTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG TGCCCCTCGGGCCGGATTGCTATCTACCGGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCG CGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCAT ACGTGTCCTCCTTACCAGAAATTTATCCTTAAGGTCGTCAGCTATCCTGCAGGCGATCTCTCGATTT CGATCAAGACATTCCTTTAATGGTCTTTTCTGGACACCACTAGGGGTCAGAAGTAGTTCATCAAACT TTCTTCCCTCCCTAATCTCATTGGTTACCTTGGGCTATCGAAACTTAATTAAGCGATCTGCATCTCA ATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGC CCATTCTCCGCCCCATCGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCT GAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAAC ATGATTGAACAAGATGGATTGCACGCAGGTTCTCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATG ACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCC GGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTA TCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGG ACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAA AGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGAC CACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATG ATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCC CGACGGCGAGGATCTCGTCGTGACCCACGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGC CGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGG CTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTAT

	CGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGC CCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAG GTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCA TTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCAT GCTGGGGATGCGGTGGGCTCTATGGTTAATTAACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTG TCTGGGTTTCGACTACGCTCAGAATTGCGTCAGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGT TCTCCGATTACGAGTTTCATTTAAATCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA AGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTC AAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTC GTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCG TGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGG CTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCC AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATT TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA CAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGG GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTT AAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCC CTCGTCGGAAAATCTTCAAACCTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCG CAAGTCTCTTGGCCGGCCTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCA ATCCCGAATATCCGAGATCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATC CGAGATCCGAGGAATATCGAAATCGGGGCGCGCCTGGTGT(서열번호 106)
RET-017 LINE1-GFP ORF2-C_링커_뉴클레오플라스민-NLS	CCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATATCGTTGCTTGGCAGTCAGCCAGTCGGA ATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTGTACTCAAGCCTGGTCACGGCAGCGTAC CGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCAACGTATAATCGAGTCCTAGCTTTTGCA AACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGCATGAGTATTCAACATTTCCGTGTCGCCC TTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAA AGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATC CTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCG CGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGA CTTGGTTGAGTATTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGC AGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATTGGAGGACCGA AGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGA GCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACCTTG CGTAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAGTTGATAGACTGGATGGAGG CGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATC TGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGT ATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGA TAGGTGCCTCACTGATTAAGCATTGGTAACCGATTCTAGGTGCATTGGCGCAGAAAAAAATGCCTGA TGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTT GCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGATCGTTTAAACTCGACTCTGGC TCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAACAGCCGTGGCGCTCATTTGCTCGTCGGGCATC GAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGGCAGCGATCGCGGCTCCCGACATCTTGGACCAT TAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCATAACAGCAGCTTCAGCTACCTCTCAATTCAAA AAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTATCAATCGTTGCGTTACACACACAA AAAACCAACACACATCCATCTTCGATGGATAGCGATTTTATTATCTAACTGCTGATCGAGTGTAGCC AGATCTAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACT TACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTAT GTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTG CCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAA ATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTAC GTATTAGTCATCGCTATTACCATGCTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGT TTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAA TCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTA CGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCAGATCTTTGTCGATCCTAC CATCCACTCGACACACCCGCCAGCGGCCGCTAATACGACTCACTATAGGGAGAAGTACTGCCACCAT GGGCAAGAAGCAAAATCGCAAGACGGGGAATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGAG CGTTCTAGCTCCCCTGCTACTGAGCAGTCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAGG GATTTAGGCGATCCAACTATTCAGAACTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGAA TTTCGAGAAGAACCTCGAGGAGTGCATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACTC ATGGAACTTAAGACAAAAGCCAGGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGACC

[0814]

	AGCTCGAGGAGCGCGTGAGCGCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCAG GGAGAAGCGCATTAAGAGGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAAC CTGCGGTTGATCGGCGTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAAG ACATCATTCAAGAAAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAACG CACACCCCAGCGGTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGTG GAGATGAAAGAGAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCTA TTCGGCTGACGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTAA TATCCTGAAGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGAG GGTGAGATTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTCA AGGAACTGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACGC AAAGATGTGAGACAGCCGTCAGACCATCAAGACTAGGAAGAAACTGCATCAACTAATGAGCAAAATC ACCAGCTAACATCATAGTATACATGACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATT AACGGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCG TTTGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCG AAAGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGACG GATTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATAC AGCAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAACA GGTCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACACCA TTGAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCAC TGCATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCTT CAGCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAAG TGCAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGAA TCAAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATTG GGTCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTACC TATCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTATA AAAGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGGA ACAGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGATC GAAACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAGA TAGACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGATTGATACCATCAAGAA CGACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCGGGAGTATTATAAGCAT TTGTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGATACCTATACTCTGCCAC GGCTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCGAGATTGTGGCCATAAT TAACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGAGTTTTATCAACGGTAT ATGGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAGGAAGGCATCTTGCCCA ATTCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCGATACCACAAAGAAGGAAAA CTTCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAAGATTCTCGCCAACAGAATC CAACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATACCTGGCATGCAGGGCTGGT TTAACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTAAGGATAAGAATCATATGAT CATCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTTATGCTCAAGACTCTGAAC AAACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTACGATAAGCCGACTGCTAACA TTATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCCGCCAAGGCTGTCCCCT CTCCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCGTCAAGAGAAAGAGATT AAAGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGATGATATGATTGTGTACC TGGAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACTTTAGCAAGGTCAGCGG CTATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAATCGACAGACCGAATCCCAG ATAATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCTCGGAATCCAGCTGACAC GAGACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGAGATTAAGGAAGATACTAA TAAGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGTGAAGATGGCAATACTTCCT AAAGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAGA AAACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGAA GAATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGACA GCCTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGGACCGAACCATCAGAGATAATGC CCCACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCCT CTTCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCTT ACACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGA CACTGGAAGAGAATCTTGGGATCACAATACAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGAC GCCCAAGGCCATGGCCACTAAGGATAAGATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGT ACTGCCAAGGAGACCACGATCAGAGTTAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTT ATTCATCAGATAAGGGGTTGATAAGCAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAAC GAATAATCCCATCAAGAAGTGGGCAAAAGATATGAACAGGCATTTTAGCAAAAGAGGATATCTACGCC GCGAAGAAGCATATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCA TGCGATACCACCTTACCCCAGTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCG GGGCTGTGGCGAGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGG AAATCAGTCTGGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGC

[0815]

	TCGGAATCTATCCTAACGAATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGC GGCCTTGTTTACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAG AAAATGTGGCATATTTATACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCG TTGGGACCTGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCA TAGAATCTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACGACAAGggcggcggcagcaaa aggccggcggccacgaaaaaggccggccaggcaaaaaagaaaaagTAAAGCGCTTCTAGAAGTTGTC TCCTCCTGCACTGACTGACTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTA CAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCT GCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAAT CCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGT GTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTC GCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGG CTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCT CGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCA GCGGACCTTCCTTCCCGCtgagagacacaaaaaattccaacacactattgcaatgaaaataaatttc ctttattagccagaagtcagatgctcaaggggcttcatgatgtccccataatttttggcagagggaa aaagatctcagtggtatttgtgagccagggcattggccttctgataggcagcctgcacctgaggagt gcggccgctttacttgtacagctcgtccatgccgagagtgatcccggcggcggtcacgaactccagc aggaccatgtgatcgcgcttctcgttggggtctttgctcagggcggactgggtgctcaggtagtggt tgtcgggcagcagcacggggccgtcgccgatgggggtgttctgctggtagtggtcggcgagctgcac gctgccgtcctcgatgttgtggcggatcttgaagttcaccttgatgccgttcttctgcttgtcggcc atgatatagacgttgtggctgttgtagttgtactccagcttgtgccccaggatgttgccgtcctcct tgaagtcgatgcccttcagctcgatgcggttcaccagggtgtcgccctcgaacttcacctcggcgcg ggtcttgtagttgccgtcgtccttgaagaagatggtgcgctcctggacgtagccttcgggcatggcg gacttgaagaagtcgtgctgcttcatgtggtcggggtagcggctgaagcactgcacgccgtaggtca gggtggtcacgagggtgggccagggcacgggcagcttgccggtggtgcagatgaacttcagggtcag cttgccgtaggtggcatcgccctcgccctcgccggacacgctgaacttgtggccgtttacgtcgccg tccagctcgaccaggatgggcaccaccccggtgaacagctcctcgcccttgctcaccatggtggcgg gatctgacggttcactaaaccagctctgcttatatagacctcccaccgtacacgcctaccgcccatt tgcgtcaatggggcggagttgttacgacattttggaaagtcccgttgattttggtgccaaaacaaac tcccattgacgtcaatggggtggagacttggaaatccccgtgagtcaaaccgctatccacgcccatt gatgtactgccaaaaccgcatcaccatggtaatagcgatgactaatacgtagatgtactgccaagta ggaaagtcccataaggtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaa taggggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgtaaatactc cacccattgacgtcaatggaaagtccctattggcgttactatgggaacatacgtcattattgacgtc aatgggcgggggtcgttgggcggtcagccaggcgggccatttaccgtaagttatgtaacgGGCCTGC TGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGC CGCCTCCCCGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGACATGATAAG ATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATT TGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCA TTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAA ATGTGGTATTGGCCCATCTCTATCGGTATCGTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGA GGGGTTTTTTGTGCCCCTCGGGCCGGATTGCTATCTACCGGCATTGGCGCAGAAAAAAATGCCTGAT GCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTG CCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGGTCGTCAGCTATCCTGCAGGCGA TCTCTCGATTTCGATCAAGACATTCCTTTAATGGTCTTTTCTGGACACCACTAGGGGTCAGAAGTAG TTCATCAAACTTTCTTCCCTCCCTAATCTCATTGGTTACCTTGGGCTATCGAAACTTAATTAAGCGA TCTGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCG CCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCG CCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAGG AGGTAGCCAACATGATTGAACAAGATGGATTGCACGCAGGTTCTCCCGCCGCTTGGGTGGAGAGGCT ATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCG CAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAGGACGAGG CAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGA AGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCT CCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCT GCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGT CGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAG GCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCACGGCGATGCCTGCTTGCCGAATATCATGG TGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGA CATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTG CTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCT AGTATGTAAGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTT GACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTG

[0816]

	AGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACA ATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAATTAACCAGTCAAGTCAGCTACTTGGCG AGATCGACTTGTCTGGGTTTCGACTACGCTCAGAATTGCGTCAGTCAAGTTCGATCTGGTCCTTGCT ATTGCACCCGTTCTCCGATTACGAGTTTCATTTAAATCATGTGAGCAAAAGGCCAGCAAAAGGCCAG GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAA AATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCC TTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGC TCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAG CAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGA AGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTT GATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAG AAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAAC TCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAA AATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAAT CAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCATTTAAATTTCCGAAC TCTCCAAGGCCCTCGTCGGAAAATCTTCAAACCTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGA ACGAACTATCGCAAGTCTCTTGGCCGGCCTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAG GTATTACTCCAATCCCGAATATCCGAGATCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAAT CCCGATCTATCCGAGATCCGAGGAATATCGAAATCGGGGCGCGCCTGGTGTA(서열번호 107)
RET-018 LINE1-GFP ORF2-N_SV40_NLS_링커 ORF2-C_뉴클레오플라스민_NLS	TGGCCGGCCTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAAT ATCCGAGATCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCG AGGAATATCGAAATCGGGGCGCGCCTGGTGTACCGAGAACGATCCCTCTCAGTGCGAGTCTCGACGAT CCATATCGTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGA TATTGTACTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATC GGTCAACGTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTT TCGCATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTT TTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTT ACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAAT GATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAA CTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATC TTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGC CAACTTACTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGAT CATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACA CCACGATGCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTAGC TTCCCGGCAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCC CTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTG CAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAAC TATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCGATT CTAGGTGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCC AGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATT TATCCTTAAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAA CAGCCGTGGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGG CAGCGATCGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCA TAACAGCAGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGG TTATGCTATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGATAGCGA TTTTATTATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTAGTTCAT AGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACG ACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTG ACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCA AGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCT TATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATGCGGTT TTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCG CCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGT GAACCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCGCTAATA CGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGGAATTCCA AGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAGTCCTGGAT GGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAACTCCGCGAA GATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTGCATCACCCGTA TCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCAGGGAGCTTCGAGA GGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGCGCGATGGAAGACGAG

[0817]

	ATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGAGGAACGAACAGAGTCTGC AGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGCGTCCCCGAGAGCGACGTAGA AAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGAAAATTTTCCAAACCTGGCTCGG CAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCGGTATAGCTCTCGGCGTGCCACCC CTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGAGAAGATGCTGCGAGCCGCTCGGGA AAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGACGGTTGACCTTAGCGCCGAGACACTC CAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGAAGGAGAAGAACTTCCAGCCACGAATCT CTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGATTAAGTATTTCATCGATAAACAGATGCT GCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAACTGCTCAAAGAGGCTCTTAATATGGAGCGC AATAATAGATATCAACCCTTGCAGAACCACGCAAAGATGTGAGACAGCCGTCAGACCATCAAGACTA GGAAGAAACTGCATCAACTAATGAGCAAAATCACCAGCTAACATCATAGTATACATGccaaagaaga agcggaaggtcggcggcggcagcACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAA CGGCCTCAACTCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTT TGTTGCATCCAAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGAA AGATTTATCAGGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGACGGA TTTCAAGCCCACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATACAG CAGGAAGAACTTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAACAGG TCCTGTCCGATCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACACCATT GAGCACCCTGGATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCACTG CATCAGGCAGATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCTTCA GCGCCCCACATCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAAGTG CAAGCGCACAGAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGAATC AAGAACCTGACCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATTGGG TCCACAATGAGATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTACCTA TCAAAACCTTTGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTATAAA AGAAAACAAGAGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGGAAC AGACCCACTCCAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGATCGA AACCCAAAAGACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAGATA GACAGACCTCTGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGATTGATACCATCAAGAACG ACAAGGGCGACATCACTACTGACCCGACCGAGATCCAGACCACTATTCGGGAGTATTATAAGCATTT GTATGCTAACAAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGATACCTATACTCTGCCACGG CTTAATCAAGAGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCGAGATTGTGGCCATAATTA ACTCCCTGCCGACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGAGTTTTATCAACGGTATAT GGAAGAGCTTGTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAGGAAGGCATCTTGCCCAAT TCCTTCTACGAAGCTTCTATAATACTTATTCCCAAACCAGGACGCGATACCACAAAGAAGGAAAACT TCCGGCCCATTAGTCTCATGAATATCGACGCTAAAATATTGAACAAGATTCTCGCCAACAGAATCCA ACAACATATTAAGAAATTGATACATCACGACCAGGTGGGGTTTATACCTGGCATGCAGGGCTGGTTT AACATCCGGAAGAGTATTAACGTCATTCAACACATTAATAGAGCTAAGGATAAGAATCATATGATCA TCTCTATAGACGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTTATGCTCAAGACTCTGAACAA ACTCGGCATCGACGGAACATATTTTAAGATTATTCGCGCAATTTACGATAAGCCGACTGCTAACATT ATCCTTAACGGCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCCGCCAAGGCTGTCCCCTCT CCCCGCTTTTGTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCGTCAAGAGAAAGAGATTAA AGGGATACAGCTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGATGATATGATTGTGTACCTG GAGAATCCTATTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACTTTAGCAAGGTCAGCGGCT ATAAGATTAACGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAATCGACAGACCGAATCCCAGAT AATGGGTGAGCTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCTCGGAATCCAGCTGACACGA GACGTTAAAGATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGAGATTAAGGAAGATACTAATA AGTGGAAGAATATCCCCTGTTCATGGGTTGGCAGAATCAACATAGTGAAGATGGCAATACTTCCTAA AGTGATATATCGCTTTAACGCCATCCCAATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAGAAA ACAACCCTTAAATTTATATGGAATCAAAAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGAAGA ATAAGGCCGGTGGGATTACTTTGCCTGATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGACAGC CTGGTATTGGTATCAGAATAGAGACATCGACCAGTGGAATCGGACCGAACCATCAGAGATAATGCCC CACATCTATAATTACCTTATATTCGATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCT TCAACAAGTGGTGTTGGGAGAATTGGCTGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTAC ACCCTACACTAAAATCAACAGTAGGTGGATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACA CTGGAAGAGAATCTTGGGATCACAATACAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGC CCAAGGCCATGGCCACTAAGGATAAGATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTAC TGCCAAGGAGACCACGATCAGAGTTAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTAT TCATCAGATAAGGGGTTGATAAGCAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGA ATAATCCCATCAAGAAGTGGGCAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGC GAAGAAGCATATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATG CGATACCACCTTACCCCAGTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGG GCTGTGGCGAGATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAA ATCAGTCTGGCGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTC

[0818]

	GGAATCTATCCTAACGAATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGG CCTTGTTTACGATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAA AATGTGGCATATTTATACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTT GGGACCTGGATGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATA GAATCTTCTCTCTCATTGGTGGTAACGACTACAAAGACGATGACGACAAGaaaaggccggcggccac gaaaaaggccggccaggcaaaaaagaaaaagTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGA CTGACTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAA GATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTT GTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCT CTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAA CCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCC TATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGC ACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCA CCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTC CCGCtgagagacacaaaaaattccaacacactattgcaatgaaaataaatttcctttattagccaga agtcagatgctcaaggggcttcatgatgtccccataatttttggcagagggaaaaagatctcagtgg tatttgtgagccagggcattggccttctgataggcagcctgcacctgaggagtgcggccgctttact tgtacagctcgtccatgccgagagtgatcccggcggcggtcacgaactccagcaggaccatgtgatc gcgcttctcgttggggtctttgctcagggcggactgggtgctcaggtagtggttgtcgggcagcagc acggggccgtcgccgatgggggtgttctgctggtagtggtcggcgagctgcacgctgccgtcctcga tgttgtggcggatcttgaagttcaccttgatgccgttcttctgcttgtcggccatgatatagacgtt gtggctgttgtagttgtactccagcttgtgccccaggatgttgccgtcctccttgaagtcgatgccc ttcagctcgatgcggttcaccagggtgtcgccctcgaacttcacctcggcgcgggtcttgtagttgc cgtcgtccttgaagaagatggtgcgctcctggacgtagccttcgggcatggcggacttgaagaagtc gtgctgcttcatgtggtcggggtagcggctgaagcactgcacgccgtaggtcagggtggtcacgagg gtgggccagggcacgggcagcttgccggtggtgcagatgaacttcagggtcagcttgccgtaggtgg catcgccctcgccctcgccggacacgctgaacttgtggccgtttacgtcgccgtccagctcgaccag gatgggcaccaccccggtgaacagctcctcgcccttgctcaccatggtggcgggatctgacggttca ctaaaccagctctgcttatatagacctcccaccgtacacgcctaccgcccatttgcgtcaatggggc ggagttgttacgacattttggaaagtcccgttgattttggtgccaaaacaaactcccattgacgtca atggggtggagacttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaa accgcatcaccatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcccataa ggtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaataggggggcgtactt ggcatatgatacacttgatgtactgccaagtgggcagtttaccgtaaatactccacccattgacgtc aatggaaagtccctattggcgttactatgggaacatacgtcattattgacgtcaatgggcgggggtc gttgggcggtcagccaggcgggccatttaccgtaagttatgtaacgGGCCTGCTGCCGGCTCTGCGG CCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTG TCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGACATGATAAGATACATTGATGAGT TTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGC TTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTT CAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCC CATCTCTATCGGTATCGTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTGC CCCTCGGGCCGGATTGCTATCTACCGGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGT CTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACG TGTCCTCCTTACCAGAAATTTATCCTTAAGGTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGA TCAAGACATTCCTTTAATGGTCTTTTCTGGACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTC TTCCCTCCCTAATCTCATTGGTTACCTTGGGCTATCGAAACTTAATTAAGCGATCTGCATCTCAATT AGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCA TTCTCCGCCCCATCGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAG CTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATG ATTGAACAAGATGGATTGCACGCAGGTTCTCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACT GGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGT TCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCG TGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACT GGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGT ATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCAC CAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATC TGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGA CGGCGAGGATCTCGTCGTGACCCACGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGC TTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTA CCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGC CGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCT GTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTG CCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTC

[0819]

	TATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCT GGGGATGCGGTGGGCTCTATGGTTAATTAACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCT GGGTTTCGACTACGCTCAGAATTGCGTCAGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCT CCGATTACGAGTTTCATTTAAATCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG TCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG CGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTG TGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAC CCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATG TAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGG TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAA ACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTC AAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGAT TTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAA TCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTA TCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTC GTCGGAAAATCTTCAAACCTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAA GTCTCT(서열번호 108)
ret-028-line1-orf1-t2a-orf2-n_sv40_nls_링커-gfp	TGGCCGGCCTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAAT ATCCGAGATCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCG AGGAATATCGAAATCGGGGCGCGCCTGGTGTACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGAT CCATATCGTTGCTTGGCAGTCAGCCAGTCGGAATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGA TATTGTACTCAAGCCTGGTCACGGCAGCGTACCGATCTGTTTAAACCTAGATATTGATAGTCTGATC GGTCAACGTATAATCGAGTCCTAGCTTTTGCAAACATCTATCAAGAGACAGGATCAGCAGGAGGCTT TCGCATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTT TTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTT ACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAAT GATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAA CTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATC TTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGC CAACTTACTTCTGACAACGATTGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGAT CATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACA CCACGATGCCTGTAGCAATGGCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTAGC TTCCCGGCAACAGTTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCC CTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTG CAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAAC TATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCGATT CTAGGTGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCACATATGCC AGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTTACCAGAAATT TATCCTTAAGATCGTTTAAACTCGACTCTGGCTCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAA CAGCCGTGGCGCTCATTTGCTCGTCGGGCATCGAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGG CAGCGATCGCGGCTCCCGACATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCA TAACAGCAGCTTCAGCTACCTCTCAATTCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGG TTATGCTATCAATCGTTGCGTTACACACACAAAAAACCAACACACATCCATCTTCGATGGATAGCGA TTTTATTATCTAACTGCTGATCGAGTGTAGCCAGATCTAGTAATCAATTACGGGGTCATTAGTTCAT AGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACG ACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTG ACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCA AGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCT TATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATGCGGTT TTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCG CCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGT GAACCGTCAGATCAGATCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGCCGCTAATA CGACTCACTATAGGGAGAAGTACTGCCACCATGGGCAAGAAGCAAAATCGCAAGACGGGGAATTCCA AGACACAATCCGCTAGCCCACCACCTAAAGAGCGTTCTAGCTCCCCTGCTACTGAGCAGTCCTGGAT GGAAAACGACTTCGATGAACTCCGGGAAGAGGGATTTAGGCGATCCAACTATTCAGAACTCCGCGAA GATATCCAGACAAAGGGGAAGGAAGTCGAGAATTTCGAGAAGAACCTCGAGGAGTGCATCACCCGTA TCACAAACACTGAGAAATGTCTCAAAGAACTCATGGAACTTAAGACAAAAGCCAGGGAGCTTCGAGA GGAGTGTCGGAGTCTGAGATCCAGGTGTGACCAGCTCGAGGAGCGCGTGAGCGCGATGGAAGACGAG ATGAACGAGATGAAAAGAGAGGGCAAATTCAGGGAGAAGCGCATTAAGAGGAACGAACAGAGTCTGC AGGAGATTTGGGATTACGTCAAGAGGCCTAACCTGCGGTTGATCGGCGTCCCCGAGAGCGACGTAGA

[0820]

[0821]

	AAACGGGACTAAACTGGAGAATACACTTCAAGACATCATTCAAGAAAATTTTCCAAACCTGGCTCGG CAAGCTAATGTGCAAATCCAAGAGATCCAACGCACACCCCAGCGGTATAGCTCTCGGCGTGCCACCC CTAGGCATATTATCGTGCGCTTTACTAAGGTGGAGATGAAAGAGAAGATGCTGCGAGCCGCTCGGGA AAAGGGAAGGGTGACTTTGAAGGGCAAACCTATTCGGCTGACGGTTGACCTTAGCGCCGAGACACTC CAGGCACGCCGGGAATGGGGCCCCATCTTTAATATCCTGAAGGAGAAGAACTTCCAGCCACGAATCT CTTACCCTGCAAAGTTGAGTTTTATCTCCGAGGGTGAGATTAAGTATTTCATCGATAAACAGATGCT GCGAGACTTCGTGACAACTCGCCCAGCTCTCAAGGAACTGCTCAAAGAGGCTCTTAATATGGAGCGC AATAATAGATATCAACCCTTGCAGAACCACGCAAAGATGggctccggcgagggcaggggaagCcttc taacatgcggggacgtggaggaaaatcccggcccaGGTAGCGGCccaaagaagaagcggaaggtcgg cggcggcagcACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAACGGCCTCAACTCA GCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTTTGTTGCATCCAAG AGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGAAAGATTTATCAGGC GAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGACGGATTTCAAGCCCACC AAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATACAGCAGGAAGAACTTA CCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAACAGGTCCTGTCCGATCT GCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACACCATTGAGCACCCTGGAT CGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCACTGCATCAGGCAGATC TCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCTTCAGCGCCCCACATCA TACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAAGTGCAAGCGCACAGAG ATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGAATCAAGAACCTGACCC AGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATTGGGTCCACAATGAGAT GAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTACCTATCAAAACCTTTGG GATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTATAAAAGAAAACAAGAGA GATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGGAACAGACCCACTCCAA GGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGATCGAAACCCAAAAGACT CTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAGATAGACAGACCTCTGG CACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGATTGATACCATCAAGAACGACAAGGGCGACAT CACTACTGACCCGACCGAGATCCAGACCACTATTCGGGAGTATTATAAGCATTTGTATGCTAACAAG CTTGAGAACCTGGAAGAGATGGACACTTTTCTGGATACCTATACTCTGCCACGGCTTAATCAAGAGG AAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCGAGATTGTGGCCATAATTAACTCCCTGCCGAC AAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGAGTTTTATCAACGGTATATGGAAGAGCTTGTA CCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAGGAAGGCATCTTGCCCAATTCCTTCTACGAAG CTTCTATAATACTTATTCCCAAACCAGGACGCGATACCACAAAGAAGGAAAACTTCCGGCCCATTAG TCTCATGAATATCGACGCTAAAATATTGAACAAGATTCTCGCCAACAGAATCCAACAACATATTAAG AAATTGATACATCACGACCAGGTGGGGTTTATACCTGGCATGCAGGGCTGGTTTAACATCCGGAAGA GTATTAACGTCATTCAACACATTAATAGAGCTAAGGATAAGAATCATATGATCATCTCTATAGACGC GGAAAAGGCATTCGATAAGATTCAGCAGCCATTTATGCTCAAGACTCTGAACAAACTCGGCATCGAC GGAACATATTTTAAGATTATTCGCGCAATTTACGATAAGCCGACTGCTAACATTATCCTTAACGGCC AAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTTGTT TAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCGTCAAGAGAAAGAGATTAAAGGGATACAGCTC GGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGATGATATGATTGTGTACCTGGAGAATCCTATTG TGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACTTTAGCAAGGTCAGCGGCTATAAGATTAACGT CCAGAAATCTCAGGCCTTTCTGTACACAAATAATCGACAGACCGAATCCCAGATAATGGGTGAGCTT CCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCTCGGAATCCAGCTGACACGAGACGTTAAAGATT TGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGAGATTAAGGAAGATACTAATAAGTGGAAGAATAT CCCCTGTTCATGGGTTGGCAGAATCAACATAGTGAAGATGGCAATACTTCCTAAAGTGATATATCGC TTTAACGCCATCCCAATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAGAAAACAACCCTTAAAT TTATATGGAATCAAAAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGAAGAATAAGGCCGGTGG GATTACTTTGCCTGATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGACAGCCTGGTATTGGTAT CAGAATAGAGACATCGACCAGTGGAATCGGACCGAACCATCAGAGATAATGCCCCACATCTATAATT ACCTTATATTCGATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTGGTG TTGGGAGAATTGGCTGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTACACCCTACACTAAA ATCAACAGTAGGTGGATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGAATC TTGGGATCACAATACAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCATGGC CACTAAGGATAAGATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAGACC ACGATCAGAGTTAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATAAGG GGTTGATAAGCAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCATCAA GAAGTGGGCAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCATATG AAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACCTTA CCCCAGTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCGAGAT TGGCACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGGCGC TTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTATCCTA ACGAATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTTACGAT AGCTAAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTGGCATATT

	TATACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACCTGGATGA AGCTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCTTCTCTCT CATTGGTGGTAACGACTACAAAGACGATGACGACAAGTAAAGCGCTTCTAGAAGTTGTCTCCTCCTG CACTGACTGACTGATACAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTT GTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAAT GCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTG CTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTG ACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCC CCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTG TTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTG TTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCT TCCTTCCCGCtgagagacacaaaaaattccaacacactattgcaatgaaaataaatttcctttatta gccagaagtcagatgctcaaggggcttcatgatgtccccataatttttggcagagggaaaaagatct cagtggtatttgtgagccagggcattggccttctgataggcagcctgcacctgaggagtgcggccgc tttacttgtacagctcgtccatgccgagagtgatcccggcggcggtcacgaactccagcaggaccat gtgatcgcgcttctcgttggggtctttgctcagggcggactgggtgctcaggtagtggttgtcgggc agcagcacggggccgtcgccgatgggggtgttctgctggtagtggtcggcgagctgcacgctgccgt cctcgatgttgtggcggatcttgaagttcaccttgatgccgttcttctgcttgtcggccatgatata gacgttgtggctgttgtagttgtactccagcttgtgccccaggatgttgccgtcctccttgaagtcg atgcccttcagctcgatgcggttcaccagggtgtcgccctcgaacttcacctcggcgcgggtcttgt agttgccgtcgtccttgaagaagatggtgcgctcctggacgtagccttcgggcatggcggacttgaa gaagtcgtgctgcttcatgtggtcggggtagcggctgaagcactgcacgccgtaggtcagggtggtc acgagggtgggccagggcacgggcagcttgccggtggtgcagatgaacttcagggtcagcttgccgt aggtggcatcgccctcgccctcgccggacacgctgaacttgtggccgtttacgtcgccgtccagctc gaccaggatgggcaccaccccggtgaacagctcctcgcccttgctcaccatggtggcgggatctgac ggttcactaaaccagctctgcttatatagacctcccaccgtacacgcctaccgcccatttgcgtcaa tggggcggagttgttacgacattttggaaagtcccgttgattttggtgccaaaacaaactcccattg acgtcaatggggtggagacttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtact gccaaaaccgcatcaccatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtc ccataaggtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaataggggggc gtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgtaaatactccacccatt gacgtcaatggaaagtccctattggcgttactatgggaacatacgtcattattgacgtcaatgggcg ggggtcgttgggcggtcagccaggcgggccatttaccgtaagttatgtaacgGGCCTGCTGCCGGCT CTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCC CGCCTGTCTAGCTTGACTGACTGAGATACAGCGTACCTTCAGCTCACAGACATGATAAGATACATTG ATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGC TATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTT ATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTA TTGGCCCATCTCTATCGGTATCGTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTT TTGTGCCCCTCGGGCCGGATTGCTATCTACCGGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCT GCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTC CATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGGTCGTCAGCTATCCTGCAGGCGATCTCTCGA TTTCGATCAAGACATTCCTTTAATGGTCTTTTCTGGACACCACTAGGGGTCAGAAGTAGTTCATCAA ACTTTCTTCCCTCCCTAATCTCATTGGTTACCTTGGGCTATCGAAACTTAATTAAGCGATCTGCATC TCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTC CGCCCATTCTCCGCCCCATCGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCC TCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAGGAGGTAGCC AACATGATTGAACAAGATGGATTGCACGCAGGTTCTCCCGCCGCTTGGGTGGAGAGGCTATTCGGCT ATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCG CCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAGGACGAGGCAGCGCGG CTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAA GGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGA GAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTC GACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGG ATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGAT GCCCGACGGCGAGGATCTCGTCGTGACCCACGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAAT GGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGT TGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGG TATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTAGTATGTA AGCCCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGG AAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTG TCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGG CATGCTGGGGATGCGGTGGGCTCTATGGTTAATTAACCAGTCAAGTCAGCTACTTGGCGAGATCGAC TTGTCTGGGTTTCGACTACGCTCAGAATTGCGTCAGTCAAGTTCGATCTGGTCCTTGCTATTGCACC

[0822]

	CGTTCTCCGATTACGAGTTTCATTTAAATCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA AAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACG CTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCC CTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAA GCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA GGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGT ATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGC AAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAG GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTA AGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGT TTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCATTTAAATTTCCGAACTCTCCAAG GCCCTCGTCGGAAAATCTTCAAACCTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTA TCGCAAGTCTCT(서열번호 109)
ret-029-line1-orf1-t2a-orf2-c_뉴클레오플라스민_nls-gfp	ACCGAGAACGATCCTCTCAGTGCGAGTCTCGACGATCCATATCGTTGCTTGGCAGTCAGCCAGTCGG AATCCAGCTTGGGACCCAGGAAGTCCAATCGTCAGATATTGTACTCAAGCCTGGTCACGGCAGCGTA CCGATCTGTTTAAACCTAGATATTGATAGTCTGATCGGTCAACGTATAATCGAGTCCTAGCTTTTGC AAACATCTATCAAGAGACAGGATCAGCAGGAGGCTTTCGCATGAGTATTCAACATTTCCGTGTCGCC CTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAA AAGATGCTGAAGATCAGTTGGGTGCGCGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGAT CCTTGAGAGTTTTCGCCCCGAAGAACGCTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATG ACTTGGTTGAGTATTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATG CAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATTGGAGGACCG AAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGG AGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACCTT GCGTAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAGTTGATAGACTGGATGGAG GCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAAT CTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCG TATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAG ATAGGTGCCTCACTGATTAAGCATTGGTAACCGATTCTAGGTGCATTGGCGCAGAAAAAAATGCCTG ATGCGACGCTGCGCGTCTTATACTCCCACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACT TGCCCACTTCCATACGTGTCCTCCTTACCAGAAATTTATCCTTAAGATCGTTTAAACTCGACTCTGG CTCTATCGAATCTCCGTCGTTTCGAGCTTACGCGAACAGCCGTGGCGCTCATTTGCTCGTCGGGCAT CGAATCTCGTCAGCTATCGTCAGCTTACCTTTTTGGCAGCGATCGCGGCTCCCGACATCTTGGACCA TTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCATAACAGCAGCTTCAGCTACCTCTCAATTCAA AAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTATCAATCGTTGCGTTACACACACA AAAAACCAACACACATCCATCTTCGATGGATAGCGATTTTATTATCTAACTGCTGATCGAGTGTAGC CAGATCTAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAAC TTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTA TGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTA AATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA CGTATTAGTCATCGCTATTACCATGCTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGG TTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAA ATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCAGATCTTTGTCGATCCTA CCATCCACTCGACACACCCGCCAGCGGCCGCTAATACGACTCACTATAGGGAGAAGTACTGCCACCA TGGGCAAGAAGCAAAATCGCAAGACGGGGAATTCCAAGACACAATCCGCTAGCCCACCACCTAAAGA GCGTTCTAGCTCCCCTGCTACTGAGCAGTCCTGGATGGAAAACGACTTCGATGAACTCCGGGAAGAG GGATTTAGGCGATCCAACTATTCAGAACTCCGCGAAGATATCCAGACAAAGGGGAAGGAAGTCGAGA ATTTCGAGAAGAACCTCGAGGAGTGCATCACCCGTATCACAAACACTGAGAAATGTCTCAAAGAACT CATGGAACTTAAGACAAAAGCCAGGGAGCTTCGAGAGGAGTGTCGGAGTCTGAGATCCAGGTGTGAC CAGCTCGAGGAGCGCGTGAGCGCGATGGAAGACGAGATGAACGAGATGAAAAGAGAGGGCAAATTCA GGGAGAAGCGCATTAAGAGGAACGAACAGAGTCTGCAGGAGATTTGGGATTACGTCAAGAGGCCTAA CCTGCGGTTGATCGGCGTCCCCGAGAGCGACGTAGAAAACGGGACTAAACTGGAGAATACACTTCAA GACATCATTCAAGAAAATTTTCCAAACCTGGCTCGGCAAGCTAATGTGCAAATCCAAGAGATCCAAC GCACACCCCAGCGGTATAGCTCTCGGCGTGCCACCCCTAGGCATATTATCGTGCGCTTTACTAAGGT GGAGATGAAAGAGAAGATGCTGCGAGCCGCTCGGGAAAAGGGAAGGGTGACTTTGAAGGGCAAACCT ATTCGGCTGACGGTTGACCTTAGCGCCGAGACACTCCAGGCACGCCGGGAATGGGGCCCCATCTTTA ATATCCTGAAGGAGAAGAACTTCCAGCCACGAATCTCTTACCCTGCAAAGTTGAGTTTTATCTCCGA

[0823]

[0824]

	GGGTGAGATTAAGTATTTCATCGATAAACAGATGCTGCGAGACTTCGTGACAACTCGCCCAGCTCTC AAGGAACTGCTCAAAGAGGCTCTTAATATGGAGCGCAATAATAGATATCAACCCTTGCAGAACCACG CAAAGATGggctccggcgagggcaggggaagCcttctaacatgcggggacgtggaggaaaatcccgg cccaGGTAGCGGCACCGGCTCTAACTCACATATCACCATCCTTACACTTAACATTAACGGCCTCAAC TCAGCTATCAAGCGCCATCGGCTGGCCAGCTGGATCAAATCACAGGATCCAAGCGTTTGTTGCATCC AAGAGACCCACCTGACCTGTAGAGATACTCACCGCCTCAAGATCAAGGGATGGCGAAAGATTTATCA GGCGAACGGTAAGCAGAAGAAAGCCGGAGTCGCAATTCTGGTCTCAGACAAGACGGATTTCAAGCCC ACCAAAATTAAGCGTGATAAGGAAGGTCACTATATTATGGTGAAAGGCAGCATACAGCAGGAAGAAC TTACCATATTGAACATCTACGCGCCAAACACCGGCGCACCTCGCTTTATCAAACAGGTCCTGTCCGA TCTGCAGCGAGATCTGGATTCTCATACGTTGATTATGGGTGATTTCAATACACCATTGAGCACCCTG GATCGCAGCACCAGGCAAAAGGTAAATAAAGACACGCAAGAGCTCAATAGCGCACTGCATCAGGCAG ATCTCATTGATATTTATCGCACTCTTCATCCTAAGAGTACCGAGTACACATTCTTCAGCGCCCCACA TCATACATACTCAAAGATCGATCATATCGTCGGCTCAAAGGCTCTGCTGTCAAAGTGCAAGCGCACA GAGATAATTACAAATTACCTGTCAGATCATAGCGCGATCAAGCTCGAGCTGAGAATCAAGAACCTGA CCCAGAGCCGGAGTACCACTTGGAAGCTTAATAACCTGCTGCTCAACGATTATTGGGTCCACAATGA GATGAAGGCAGAGATTAAAATGTTCTTCGAAACAAATGAGAATAAGGATACTACCTATCAAAACCTT TGGGATGCCTTTAAGGCCGTCTGCAGAGGCAAGTTCATCGCCCTCAACGCCTATAAAAGAAAACAAG AGAGATCTAAGATCGATACTCTCACCTCTCAGCTGAAGGAGTTGGAGAAACAGGAACAGACCCACTC CAAGGCGTCAAGACGGCAGGAGATCACAAAGATTCGCGCCGAGTTGAAAGAGATCGAAACCCAAAAG ACTCTTCAGAAAATTAACGAGTCTCGTAGTTGGTTCTTCGAGCGGATTAATAAGATAGACAGACCTC TGGCACGACTGATTAAGAAGAAGCGCGAAAAGAACCAGATTGATACCATCAAGAACGACAAGGGCGA CATCACTACTGACCCGACCGAGATCCAGACCACTATTCGGGAGTATTATAAGCATTTGTATGCTAAC AAGCTTGAGAACCTGGAAGAGATGGACACTTTTCTGGATACCTATACTCTGCCACGGCTTAATCAAG AGGAAGTCGAGTCCCTCAACCGCCCAATTACAGGAAGCGAGATTGTGGCCATAATTAACTCCCTGCC GACAAAGAAATCTCCTGGTCCGGACGGGTTTACAGCTGAGTTTTATCAACGGTATATGGAAGAGCTT GTACCGTTTCTGCTCAAGCTCTTTCAGTCTATAGAAAAGGAAGGCATCTTGCCCAATTCCTTCTACG AAGCTTCTATAATACTTATTCCCAAACCAGGACGCGATACCACAAAGAAGGAAAACTTCCGGCCCAT TAGTCTCATGAATATCGACGCTAAAATATTGAACAAGATTCTCGCCAACAGAATCCAACAACATATT AAGAAATTGATACATCACGACCAGGTGGGGTTTATACCTGGCATGCAGGGCTGGTTTAACATCCGGA AGAGTATTAACGTCATTCAACACATTAATAGAGCTAAGGATAAGAATCATATGATCATCTCTATAGA CGCGGAAAAGGCATTCGATAAGATTCAGCAGCCATTTATGCTCAAGACTCTGAACAAACTCGGCATC GACGGAACATATTTTAAGATTATTCGCGCAATTTACGATAAGCCGACTGCTAACATTATCCTTAACG GCCAAAAGCTCGAGGCCTTTCCGCTCAAGACTGGAACCCGCCAAGGCTGTCCCCTCTCCCCGCTTTT GTTTAATATTGTACTCGAGGTGCTGGCTAGGGCTATTCGTCAAGAGAAAGAGATTAAAGGGATACAG CTCGGGAAGGAAGAGGTCAAGCTTTCCTTGTTCGCCGATGATATGATTGTGTACCTGGAGAATCCTA TTGTGTCTGCTCAGAACCTTCTTAAACTTATTTCTAACTTTAGCAAGGTCAGCGGCTATAAGATTAA CGTCCAGAAATCTCAGGCCTTTCTGTACACAAATAATCGACAGACCGAATCCCAGATAATGGGTGAG CTTCCGTTTGTCATAGCCAGCAAAAGGATAAAGTATCTCGGAATCCAGCTGACACGAGACGTTAAAG ATTTGTTTAAGGAAAATTACAAGCCTCTCCTGAAAGAGATTAAGGAAGATACTAATAAGTGGAAGAA TATCCCCTGTTCATGGGTTGGCAGAATCAACATAGTGAAGATGGCAATACTTCCTAAAGTGATATAT CGCTTTAACGCCATCCCAATTAAACTGCCTATGACCTTCTTTACGGAGCTCGAGAAAACAACCCTTA AATTTATATGGAATCAAAAGAGAGCAAGAATAGCGAAGTCCATCTTGAGCCAGAAGAATAAGGCCGG TGGGATTACTTTGCCTGATTTTAAGTTGTATTATAAAGCCACAGTAACTAAGACAGCCTGGTATTGG TATCAGAATAGAGACATCGACCAGTGGAATCGGACCGAACCATCAGAGATAATGCCCCACATCTATA ATTACCTTATATTCGATAAGCCAGAAAAGAATAAACAGTGGGGCAAAGACAGCCTCTTCAACAAGTG GTGTTGGGAGAATTGGCTGGCCATATGCCGGAAACTCAAGCTCGACCCCTTTCTTACACCCTACACT AAAATCAACAGTAGGTGGATCAAGGACTTGAATGTCAAGCCAAAGACTATAAAGACACTGGAAGAGA ATCTTGGGATCACAATACAAGATATAGGCGTCGGCAAAGATTTTATGTCAAAGACGCCCAAGGCCAT GGCCACTAAGGATAAGATTGATAAGTGGGACCTTATTAAGCTCAAAAGCTTCTGTACTGCCAAGGAG ACCACGATCAGAGTTAATAGGCAGCCCACTACATGGGAAAAGATTTTCGCCACTTATTCATCAGATA AGGGGTTGATAAGCAGAATATATAACGAGCTGAAGCAGATCTACAAGAAGAAAACGAATAATCCCAT CAAGAAGTGGGCAAAAGATATGAACAGGCATTTTAGCAAAGAGGATATCTACGCCGCGAAGAAGCAT ATGAAGAAGTGTAGTTCAAGCTTGGCCATTCGTGAGATGCAGATTAAGACGACCATGCGATACCACC TTACCCCAGTGAGGATGGCAATTATCAAGAAATCTGGCAATAATAGATGTTGGCGGGGCTGTGGCGA GATTGGCACCCTGCTCCATTGCTGGTGGGATTGCAAGCTGGTGCAGCCGCTTTGGAAATCAGTCTGG CGCTTTCTGAGGGACCTCGAGCTTGAGATTCCCTTCGATCCCGCAATTCCCTTGCTCGGAATCTATC CTAACGAATACAAGAGCTGTTGTTACAAGGATACGTGTACCCGGATGTTCATCGCGGCCTTGTTTAC GATAGCTAAGACGTGGAATCAGCCTAAGTGCCCCACAATGATCGATTGGATCAAGAAAATGTGGCAT ATTTATACCATGGAGTATTACGCAGCAATTAAGAATGACGAATTTATTTCCTTCGTTGGGACCTGGA TGAAGCTGGAGACTATTATTCTGAGCAAGCTGTCTCAGGAGCAAAAGACAAAGCATAGAATCTTCTC TCTCATTGGTGGTAACGACTACAAAGACGATGACGACAAGaaaaggccggcggccacgaaaaaggcc ggccaggcaaaaaagaaaaagTAAAGCGCTTCTAGAAGTTGTCTCCTCCTGCACTGACTGACTGATA CAATCGATTTCTGGATCCGCAGGCCTAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGG TATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCT

[0825]

ATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGG
AGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGG
TTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACG
GCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATT
CCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCT
GCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCtgagag
acacaaaaaattccaacacactattgcaatgaaaataaatttccttattagccagaagtcagatgc
tcaaggggcttcatgatgtccccataatttttggcagagggaaaaagatctcagtggtatttgtgag
ccagggcattggccttctgataggcagcctgcacctgaggagtgcggccgctttacttgtacagctc
gtccatgccgagagtgatcccggcggcggtcacgaactccagcaggaccatgtgatcgcgcttctcg
ttggggtctttgctcagggcggactgggtgctcaggtagtggttgtcgggcagcagcacggggccgt
cgccgatgggggtgttctgctggtagtggtcggcgagctgcacgctgccgtcctcgatgttgtggcg
gatcttgaagttcaccttgatgccgttcttctgcttgtcggccatgatatagacgttgtggctgttg
tagttgtactccagcttgtgccccaggatgttgccgtcctccttgaagtcgatgcccttcagctcga
tgcggttcaccagggtgtcgccctcgaacttcacctcggcgcgggtcttgtagttgccgtcgtcctt
gaagaagatggtgcgctcctggacgtagccttcgggcatggcggacttgaagaagtcgtgctgcttc
atgtggtcggggtagcggctgaagcactgcacgccgtaggtcagggtggtcacgagggtgggccagg
gcacgggcagcttgccggtggtgcagatgaacttcagggtcagcttgccgtaggtggcatcgccctc
gccctcgccggacacgctgaacttgtggccgtttacgtcgccgtccagctcgaccaggatgggcacc
accccggtgaacagctcctcgcccttgctcaccatggtggcgggatctgacggttcactaaaccagc
tctgcttatatagacctcccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgtta
cgacattttggaaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtgga
gacttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaaccgcatcac
catggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcccataaggtcatgtac
tgggcataatgccaggcgggccatttaccgtcattgacgtcaataggggggcgtacttggcatatgat
acacttgatgtactgccaagtgggcagtttaccgtaaatactccacccattgacgtcaatggaaagt
ccctattggcgttactatgggaacatacgtcattattgacgtcaatgggcgggggtcgttgggcggt
cagccaggcgggccatttaccgtaagttatgtaacgGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGC
GTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTCTAGCTTGA
CTGACTGAGATACAGCGTACCTTCAGCTCACAGACATGATAAGATACATTGATGAGTTTGGACAAAC
CACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTA
ACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGG
GGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATTGGCCCATCTCTATC
GGTATCGTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTGCCCCTCGGGCC
GGATTGCTATCTACCGGCATTGGCGCAGAAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCC
CACATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACGTGTCCTCCTT
ACCAGAAATTTATCCTTAAGGTCGTCAGCTATCCTGCAGGCGATCTCTCGATTTCGATCAAGACATT
CCTTTAATGGTCTTTTCTGGACACCACTAGGGGTCAGAAGTAGTTCATCAAACTTTCTTCCCTCCCT
AATCTCATTGGTTACCTTGGGCTATCGAAACTTAATTAAGCGATCTGCATCTCAATTAGTCAGCAAC
CATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCC
CATCGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGA
AGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAGGAGGTAGCCAACATGATTGAACAAG
ATGGATTGCACGCAGGTTCTCCCGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACA
GACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTC
AAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCA
CGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATT
GGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATG
GCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAAC
ATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGA
GCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGAT
CTCGTCGTGACCCACGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGAT
TCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATAT
TGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGAT
TCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTAGTATGTAAGCCCTGTGCCTTCTA
GTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCAC
TGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGG
GGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGG
TGGGCTCTATGGTTAATTAACCAGTCAAGTCAGCTACTTGGCGAGATCGACTTGTCTGGGTTTCGAC
TACGCTCAGAATTGCGTCAGTCAAGTTCGATCTGGTCCTTGCTATTGCACCCGTTCTCCGATTACGA
GTTTCATTTAAATCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCT
GGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGG
CGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTG
TTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCA
TAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAA

[0826]

CCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGAC
ACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGC
TACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCT
CTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG
GTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCC
TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG
AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAA
GTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGAT
CTGTCTATTTCGTTCATCCATAGTTGCATTTAAATTTCCGAACTCTCCAAGGCCCTCGTCGGAAAAT
CTTCAAACCTTTCGTCCGATCCATCTTGCAGGCTACCTCTCGAACGAACTATCGCAAGTCTCTTGGC
CGGCCTTGCGCCTTGGCTATTGCTTGGCAGCGCCTATCGCCAGGTATTACTCCAATCCCGAATATCC
GAGATCGGGATCACCCGAGAGAAGTTCAACCTACATCCTCAATCCCGATCTATCCGAGATCCGAGGA
ATATCGAAATCGGGGCGCGCCTGGTGT(서열번호 110)

[0827] **실시예 22. 인간 LINE1 구축물에 의한 역전위에 대한 핵 이출 신호 도입의 효과**

[0828] 이 실시예에서, 카고 서열의 번역을 개선하기 위해 핵 이출 서열의 추가를 시험한다. GFP의 C-말단 또는 N-말단에서의 NES의 추가를 발현의 증가에 대해 시험한다. GFP의 C-말단 또는 N-말단에서의 여러 NES의 추가를 발현의 증가에 대해 시험한다. NES가 특정 크기의 카고에 영향을 미치는지를 알아보기 위해 더 큰 카고(5 kB 코딩 서열보다 더 크거나 7.5 kB 코딩 서열보다 더 큼)을 시험할 것이다. 특히 트랜스-골지체 또는 ER 국소화를 요구하는

추가 카고, 예를 들어, 막횡단 도메인을 가진 카고를 시험할 수 있다.

[0829] **실시예 23. 다양한 인간 세포의 역전위를 위한 효율적인 LINE1 시스템에 대한 스크리닝**

[0830] 이 실시예에서, 인간 게놈 내에서 효율적으로 이동할 수 있는 하나 이상의 과활성 레트로트랜스포존 요소를 선택하기 위해 생물정보학을 사용하여 다수의 유기체들에 걸쳐 계통발생적 스크린을 수행한 후, 실험실 시험을 수행한다. 신규 도메인 변이를 가진 ORF1 및 ORF2 단백질을 비롯한 천만 개 이상의 LINE-1 서열들을 503개의 상이한 게놈으로부터 확인하였다(Ivancevic et al., 2016). 이들 중 '과활성' LINE-1 종은 인간, 래트 및 마우스의 역전위 활성보다 더 우수한 역전위 활성을 나타내는 것으로 밝혀졌다. 이 요소들을 LINE-1-GFPai 플라스미드에 클로닝하고 HEK293T 세포에서 스크리닝할 것이다. 본 명세서의 다른 곳에 기재된 바와 같이 레트로트랜스포존 복합체의 ORF1/2 유전자에 대해 역방향으로 삽입된 프로모터 및 폴리 A 서열과 함께 GFP를 코딩하는 서열을 포함하는 시험 mRNA 구축물을 세포의 전기천공에 사용한다. 통합의 효율을 유세포분석 및 PCR로 측정한다. 전술된 바와 같이, 유세포분석을 위한 게이트를 설정하기 위해 음성 대조군과 양성 대조군에서 GFP를 측정한다. 전술된 바와 같이, 사용된 예시적인 음성 대조군은 <0.1% GFP+로 설정될 것이다. 전술된 바와 같이, 사용된 예시적인 양성 대조군은 >90% GFP+로 설정될 것이다. PCR을 이용하여 통합을 시험할 것이다. 성공적인 통합을 표시하는 예시적인 측정은 유세포분석에 의해 측정될 때 역전위 후 10일째 날까지 세포의 >10%에서 GFP 발현을 측정하는 것이다.

[0831] **실시예 24. 카고 유전자 발현의 개선**

[0832] 카고 유전자 발현을 개선하기 위해, 다양한 카고 유전자 프로모터 및 폴리(A) 신호를 시험할 것이다. 카고 유전자 발현을 개선하기 위해, 짧고 항시적으로 스플라이싱되는 인트론의 다양한 삽입도 시험할 것이다. 카고 유전자 발현을 개선하기 위해, 코돈 최적화된 서열, 예컨대, 다양한 대안적인 코돈 알고리즘을 사용함으로써 생성된 서열도 시험할 것이다. 전술된 바와 같이, 유세포분석을 위한 게이트를 설정하기 위해 음성 대조군과 양성 대조군에서 GFP를 측정한다. 전술된 바와 같이, 사용된 예시적인 음성 대조군은 <0.1% GFP+로 설정될 것이다. 전술된 바와 같이, 사용된 예시적인 양성 대조군은 >90% GFP+로 설정될 것이다. 성공적인 카고 유전자 발현을 표시하는 예시적인 측정은 유세포분석에 의해 측정될 때 역전위 후 10일째 날까지 세포의 >10%에서 증가된 GFP 발현을 측정하는 것이다.

[0833] **실시예 25. 생체내 카고 유전자 전달을 위한 효율 최적화된 레트로트랜스포존 시스템의 적용**

[0834] 마우스에서 세포내 또는 분비된 단백질을 검출하기 위해 간 표적화 LNP를 사용한 생체내 전달을 위해 레트로트랜스포존 mRNA를 조작할 것이다. 카고 유전자 mRNA의 3' UTR에 있는 세포 유형 특이적 miRNA 표적은 원치 않는 세포 유형에서 분해되도록 이를 표적화하는 데 적용될 것이다. 생체내 효율적인 전달에 대해 LNP를 시험할 것이다. GFP 코딩 서열을 플랭킹하는 프로모터 및 폴리 A 서열과 함께 GFP를 코딩하는 서열을 포함하는 LNP 캡슐화된 mRNA 구축물을 마우스에게 투여할 것이다. GFP 유전자는 본 명세서의 다른 부분에 기재된 바와 같이 레트로트랜스포존 복합체의 ORF1/2 유전자에 대해 역방향으로 삽입된다. 단백질 발현을 영상화로 조사할 것이고, 시험 마우스를 희생시킬 것이고, 결정된 시점에서 조직학적 분석 및 PCR을 위해 조직을 채취할 것이다.

[0835] 세포 특이적 표적화는 LNP 캡슐화된 mRNA의 전신 투여 후에 달성되며, 이때 LNP는 표적 세포에서 동족 대응물을 인식하는 항체 또는 리간드를 표시하기 위한 표면 변형을 포함한다.

[0836] 세포 특이적 표적화는 특이적 구축물을 디자인한 후에 시험되고, 이때 UTR은 하나 이상의 세포 유형 특이적 miRNA를 코딩하도록 변형된다. 재조합 mRNA가 의도된 표적 세포가 아닌 세포에 의해 흡수될 때, 세포에 존재하는 miRNA는 mRNA에 결합하여 이를 파괴할 것이다. 따라서, mRNA는 의도된 세포 유형에서 발현될 수 있다.

[0837] 간으로의 성공적인 생체내 mRNA 전달을 표시하는 예시적인 측정은 간세포에서 유세포분석 또는 조직학에 의해 측정된 바와 같은 GFP 발현의 측정이다.

[0838] **실시예 26. 특정 게놈 위치로의 역전위의 표적화**

[0839] 이 실시예에서, 특정 게놈 위치에서 표적화의 특이성을 증가시키기 위해 역전위 디자인 변형을 수행한다. ORF2p의 표적화 특이성의 증가를 보장하기 위해 표적화 모이어티가 통합되어 있는 레트로트랜스포존 융합 구축물을 생성한다. 한 방법은 LINE1 요소를 카고로 재표적화하여, 높은 지속 가능한 전사 활성을 가진 게놈 안전 피난처(genomic safe harbor)(GHS) 부위들 중 한 부위 내로 특이적으로 통합시키는 것이다. AAVS1, CCR5 및 ROSA26은 GHS 부위들의 일부이다. 전술된 바와 같이, 특이적 LINE1 요소는 인간에서 쉽게 전위되는 효율적인 LINE1 요소를 확인하는 데 적합한 스크린 후에 선택된다. 선택된 LINE1은 GHS 유전자좌들 중 하나 내로 통합되는 LINE1일

수 있다.

[0840] 또 다른 방법은 ORF2와 Cas9 및 이의 돌연변이체의 융합체를 갖도록 레트로트랜스포존 ORF2를 조작하는 것이다. 엔도뉴클레아제 활성을 갖거나 갖지 않는 이종 DNA 결합 도메인을 가진 다른 후보 요소는 ORF2p의 엔도뉴클레아제 도메인을 대체하거나 보완할 Cpf1, 징크 핑거 요소, TAL 이펙터, Cas6-8 '캐스케이드', 제한 엔도뉴클레아제이다.

[0841] mRNA 카고 주위에 상이한 길이의 상동성 아암을 추가하는 것도 시험할 것이다.

[0842] ORF1; RT 및 dCas9; 안티센스 방향의 카고 유전자; 및 RT 활성을 프라이밍하기 위해 표적 부위로 안내하는 가이드 RNA를 코딩하는 서열을 포함하는 예시적인 구축물을 디자인한다. 한 예시적인 구축물에서, 표적화된 위치의 상류 및 하류에서 닉을 도입하는 것이 유전자 카고의 부위 특이적 통합 및 배향을 개선하는지를 확인하기 위해 dCas9 닉카제에 각각 융합된 2개의 가이드 RNA의 사용을 시험한다.

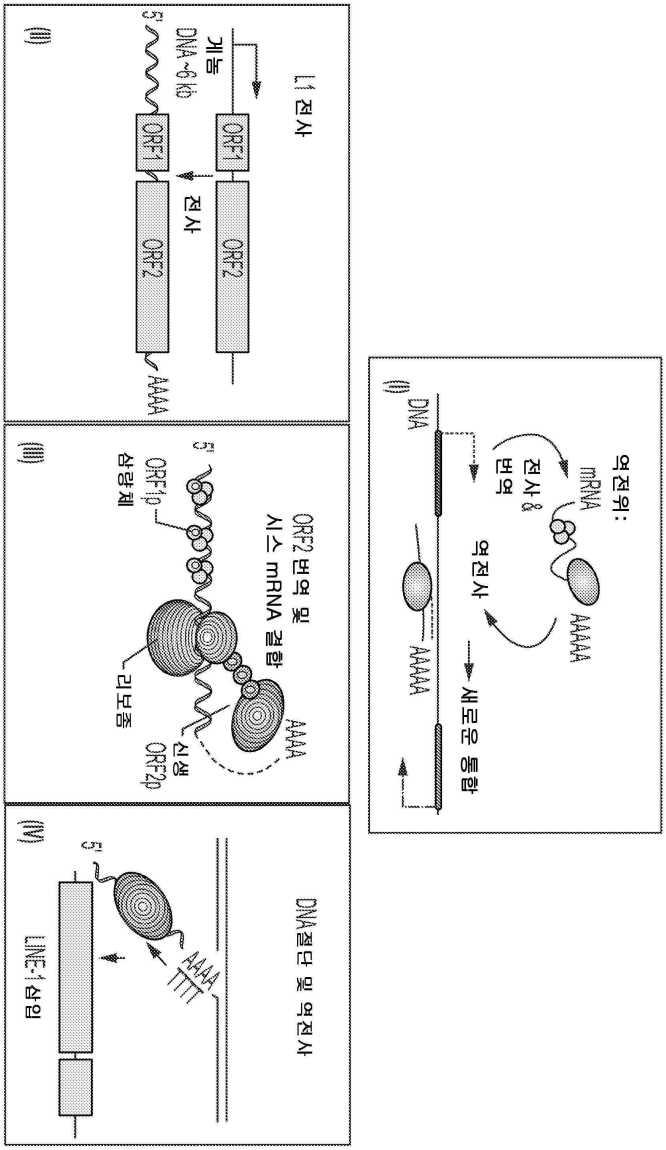
[0843] ORF2p가 EN을 결여하고(결실), 특이성을 유도하고 게놈 DNA에서 닉을 도입하는 Cas 9 또는 이의 돌연변이체, Cas 12a(Cpf1), Cas6-8 '캐스케이드', 메가뉴클레아제, 징크 핑거 뉴클레아제, TALEN 또는 제한 엔도뉴클레아제와 융합되어 있는 예시적인 구축물을 디자인하고 시험한다.

[0844] ORF2p가 감소되거나 변경된 활성을 가진 돌연변이체 EN을 포함하고, 특이성을 유도하고 게놈 DNA에서 닉을 도입하는 Cas 9 또는 이의 돌연변이체, Cas 12a(Cpf1), Cas6-8 '캐스케이드', 메가뉴클레아제, 징크 핑거 뉴클레아제, TALEN 또는 제한 엔도뉴클레아제와 융합되어 있는 예시적인 구축물을 디자인하고 시험한다. 예를 들어, ORF2 단백질과 Cas9 및 Cas9 돌연변이체의 융합체를 코딩하는 레트로트랜스포존 플라스미드가 조작될 구축물을 디자인하고 시험한다. ORF1을 함유하는 mRNA, 및 RT와 dCas9 및 안티센스 방향의 카고 유전자를 코딩하는 mRNA, 및 pegRNA를, RT 활성을 프라이밍하기 위해 표적 부위로 전달한다. 표적화된 위치의 상류 및 하류에서 닉을 도입하는 것이 유전자 카고의 부위 특이적 통합 및 배향을 개선하는지를 확인하기 위해 dCas9 닉카제에 각각 융합된 2개의 sgRNA의 사용을 평가할 수 있다. 변형된 LINE-1 RT 융합 단백질이 그 자신의 mRNA에 결합하여 핵으로 이동할 수 있는 경우, LINE-1 RT가 게놈 DNA 내로의 통합을 위해 mRNA 유전자 카고를 역전사할 수 있도록 pegRNA는 게놈 영역을 표적화할 수 있고 (dCas9 닉카제 활성을 통해) RT 활성을 프라이밍할 수 있다.

[0845] 통합의 효율을 유세포분석으로 측정할 것이다. 전술된 바와 같이, 유세포분석을 위한 게이트를 설정하기 위해 음성 대조군과 양성 대조군에서 GFP를 측정한다. 전술된 바와 같이, 사용된 예시적인 음성 대조군은 <0.1% GFP+로 설정될 것이다. 전술된 바와 같이, 사용된 예시적인 양성 대조군은 >90% GFP+로 설정될 것이다. 통합을 PCR로 확인할 것이다. NGS를 이용하여 게놈 통합 부위를 분석한다. 특정 게놈 위치에서 우선적으로 통합되는 임의의 구축물에 대해 여러 주기의 방향성 진화를 수행하여 그의 효율 및/또는 통합 정밀도를 개선한다. 성공적인 유전자 통합 및 발현을 표시하는 예시적인 측정은 특정 게놈 위치에서 검출된 90%의 카고 유전자 통합, 및 역전위 후 10일째 날까지 세포의 >2%에서 유세포분석에 의해 측정된 GFP 발현이다.

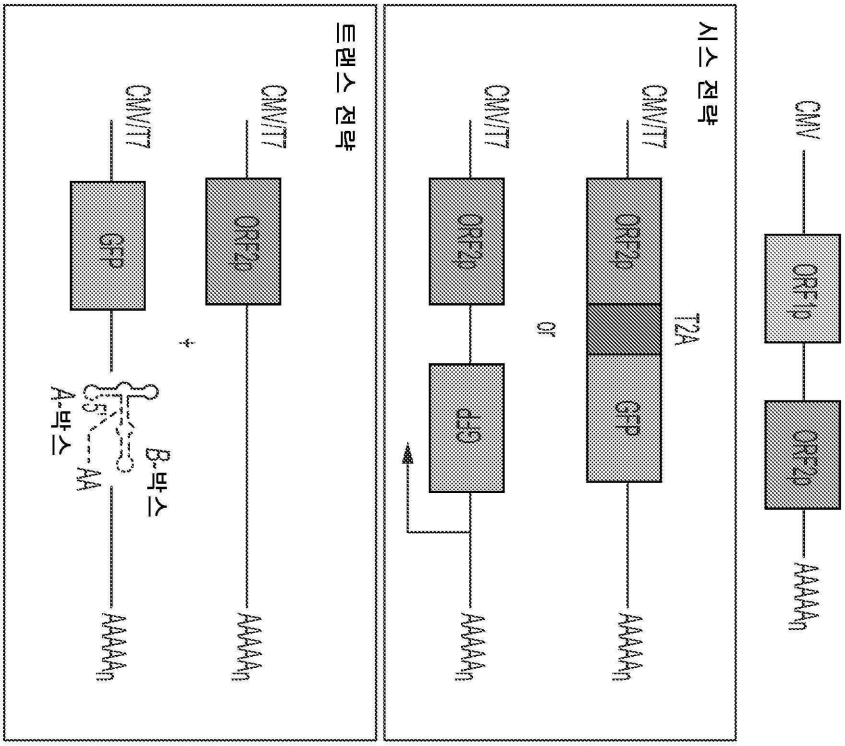
도면

도면1a

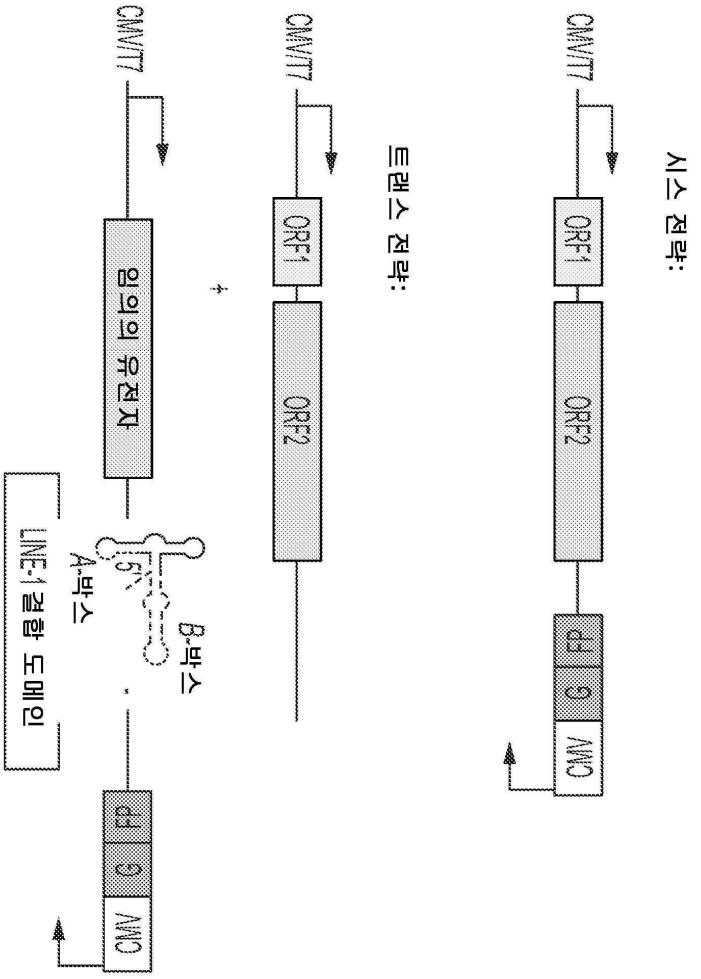

(I)
DNA
AAAAA
역전위:
mRNA
전사 &
번역
역전사
AAAAA
새로운 통합
(II)
게놈
DNA ~6 kb
5'
ORF1
ORF2
AAAA
전사
L1 전사
ORF1
ORF2
(III)
5'
ORF1p
삼량체
리보좀
AAAA
신생
ORF2p
ORF2 번역 및
시스 mRNA 결합
(IV)
5'
AAAA
TTTT
LINE-1 삽입
DNA절단 및 역전사

도면1b

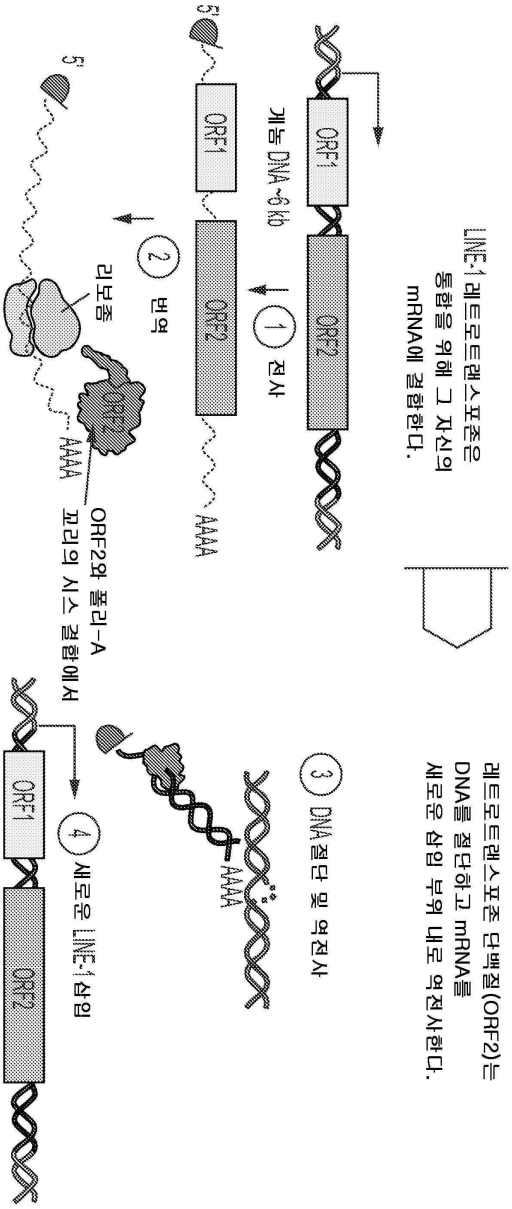
도면1c


CMV
ORF1p
ORF2p
AAAAAn
시스 전략
CMV/T7
ORF2p
T2A
GFP
AAAAAn
or
CMV/T7
ORF2p
dGP
AAAAAn
트랜스 전략
CMV/T7
ORF2p
AAAAAn
+
CMV/T7
GFP
B-박스
AA
A-박스
AAAAAn

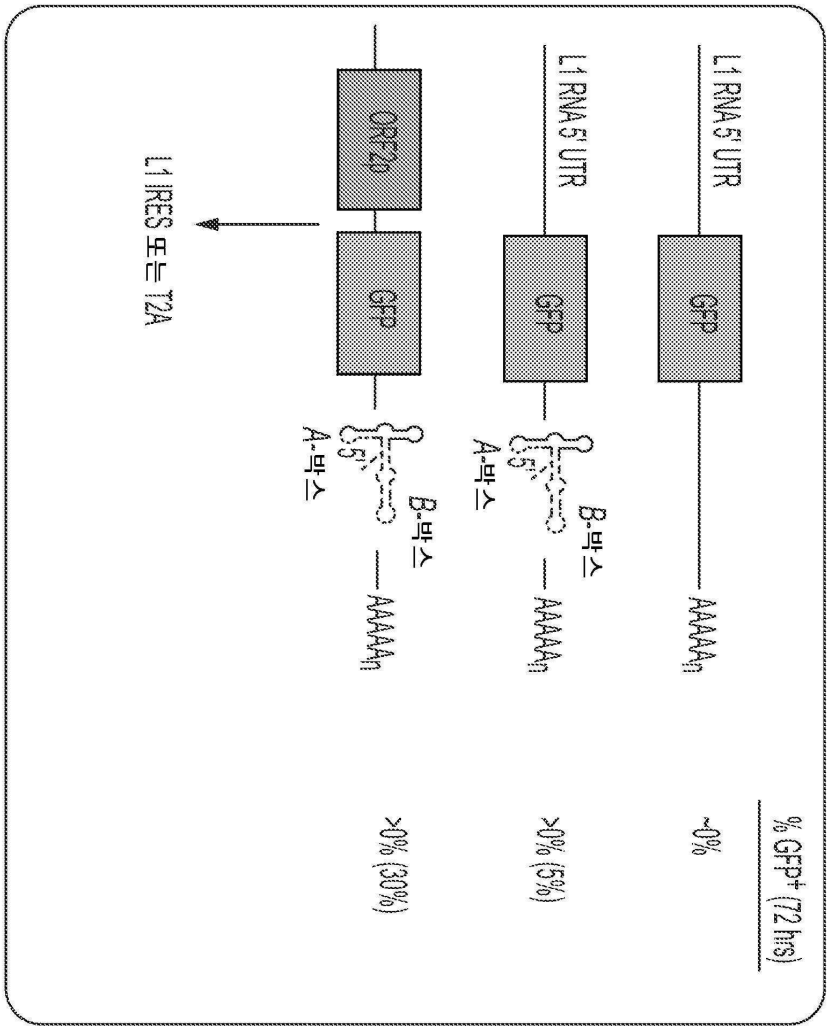
도면1d


시스 전략:
CMV/T7
ORF1
ORF2
FP
G
CMV
트랜스 전략:
CMV/T7
ORF1
ORF2
CMV/T7
임의의 유전자
B-박스
A-박스
LINE-1 결합 도메인
FP
G
CMV

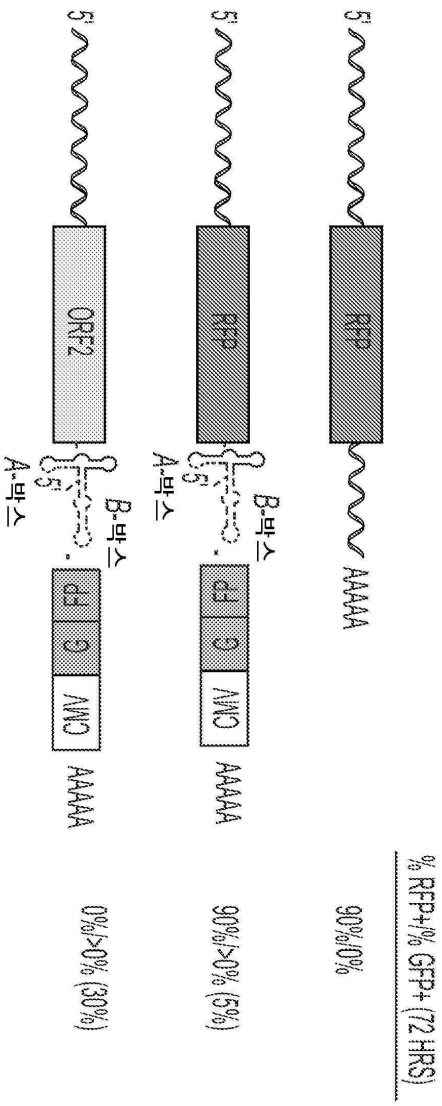
도면1e


LINE-1 레트로트랜스포존은
통합을 위해 그 자신의
mRNA에 결합한다.
ORF1
ORF2
게놈 DNA ~6 kb
① 전사
5'
AAAA
② 번역
리보좀
ORF2
AAAA
ORF2와 폴리-A
꼬리의 시스 결합에서
레트로트랜스포존 단백질(ORF2)는
DNA를 절단하고 mRNA를
새로운 삽입 부위 내로 역전사한다.
③ DNA절단 및 역전사
AAAA
④ 새로운 LINE-1삽입
ORF1
ORF2

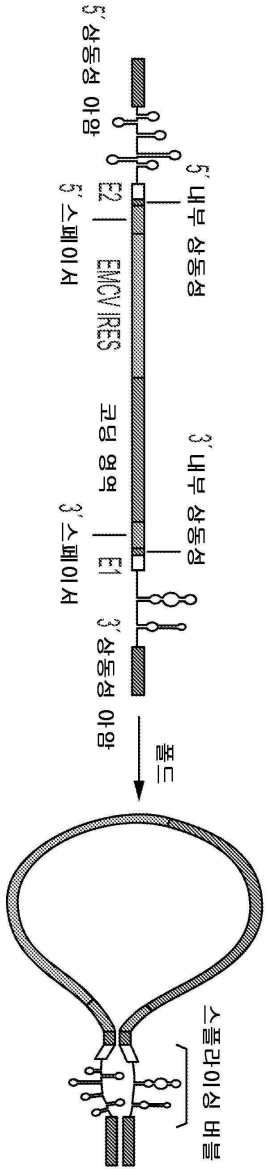
도면2a


L1 RNA 5'UTR
GFP
AAAAAn
% GFP+ (72 hrs)
~0%
L1 RNA 5'UTR
GFP
B-박스
A-박스
AAAAAn
>0% (5%)
ORF2p
GFP
L1 IRES 또는 T2A
B-박스
A-박스
AAAAAn
>0% (30%)

도면2b

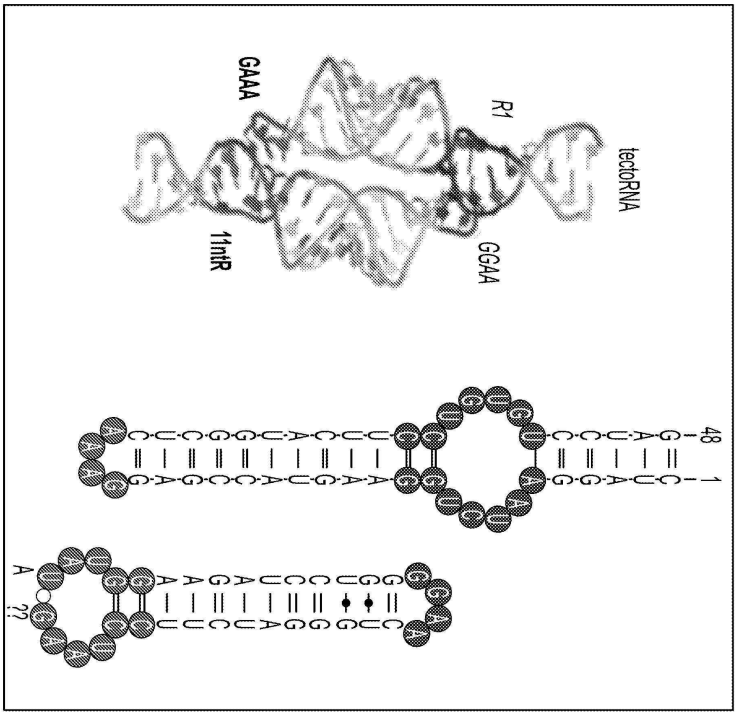

5'
RFP
AAAAA
% RFP+/% GFP+ (72 HRS)
90%/0%
5'
RFP
B-박스
A-박스
AAAAA
90%/>0% (5%)
5'
ORF2
B-박스
A-박스
AAAAA
0%/>0% (30%)

도면3a

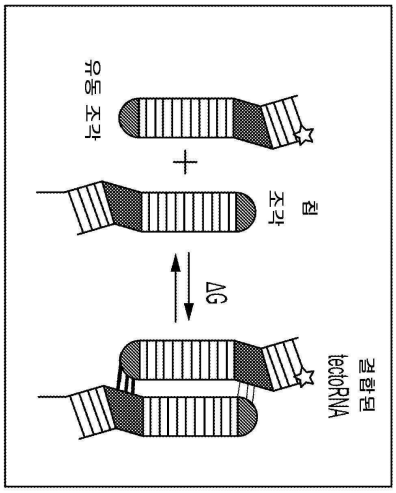

5' 상동성 아암
5' 내부 상동성
5' 스페이서
E2
EMCV IRES
코딩 영역
3' 내부 상동성
3' 스페이서
E1
3' 상동성 아암
폴드
스플라이싱 버블

도면3bc

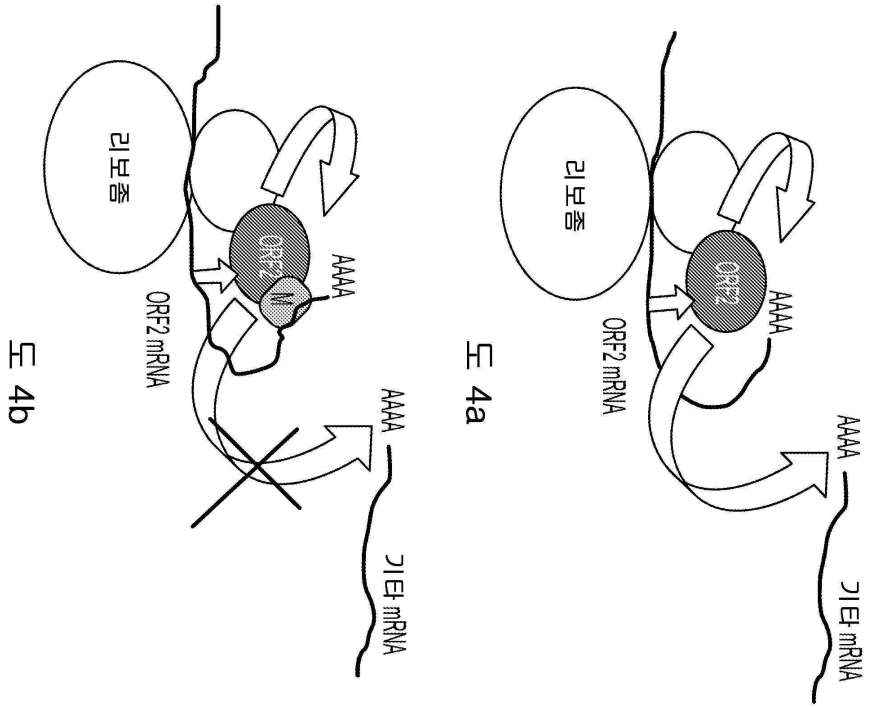
도 3b



도 3c



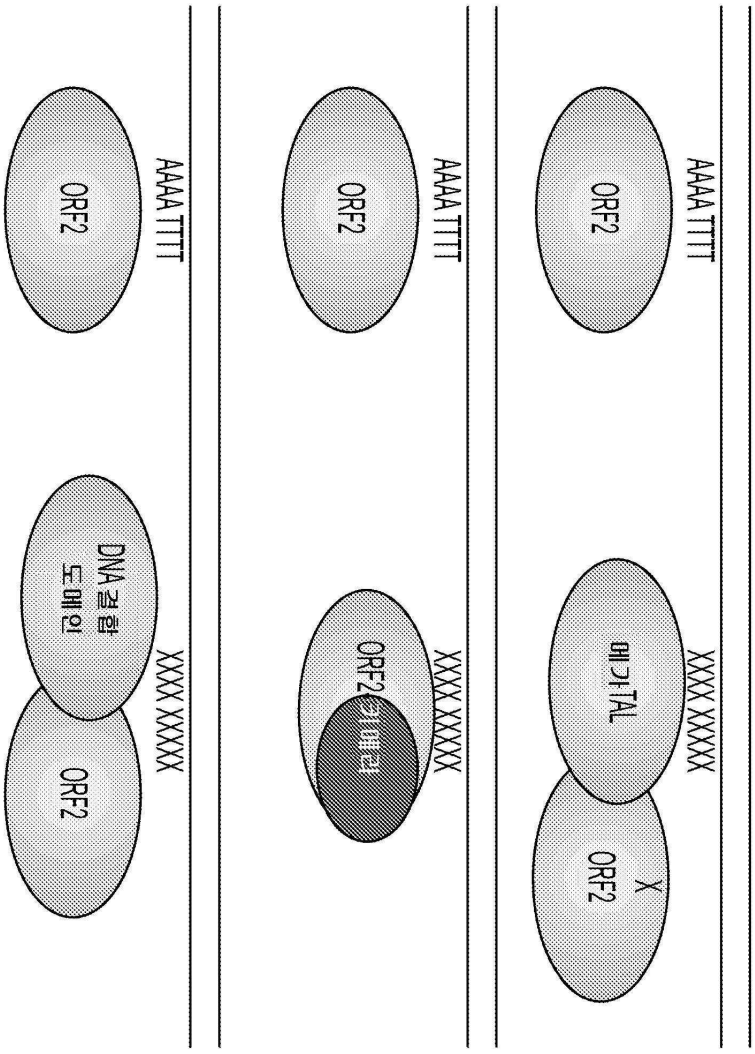
도면4ab



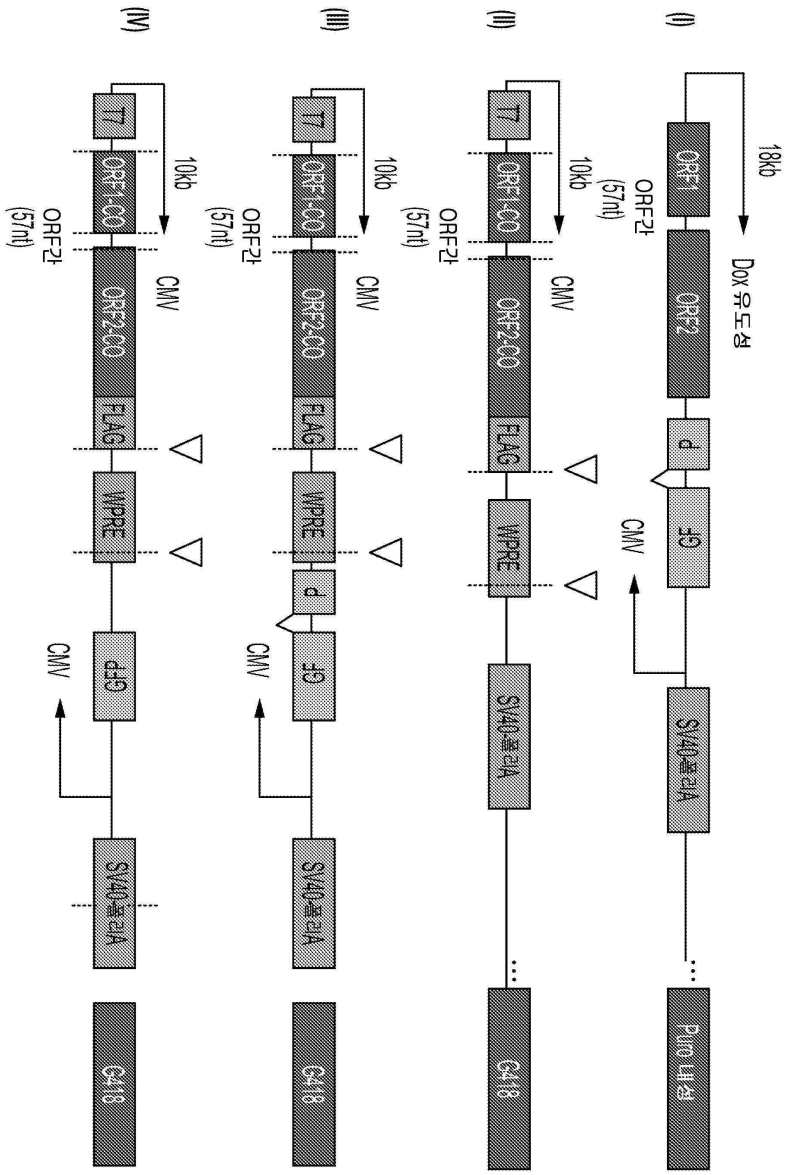
도 4a

도 4b

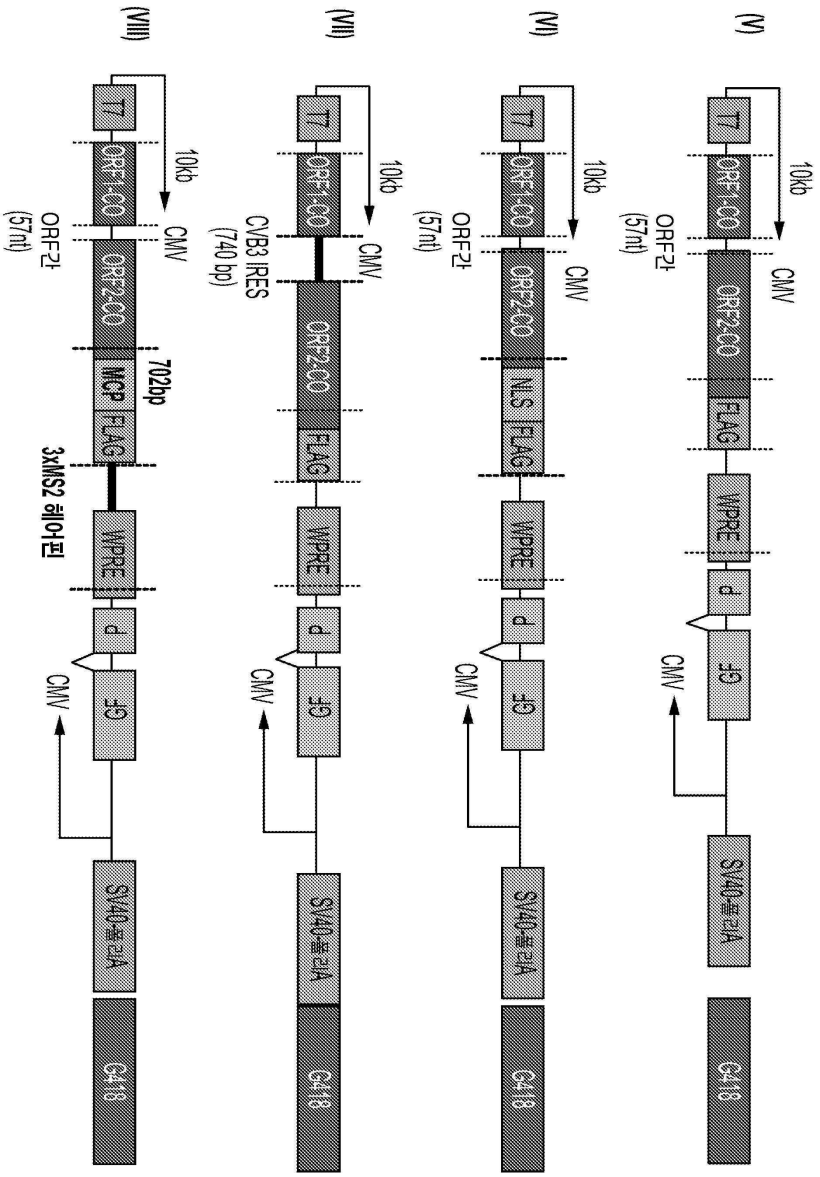
도면4c


AAAA TTTT
ORF2
XXXX XXXXX
메가TAL
X
ORF2
AAAA TTTT
ORF2
XXXX XXXXX
ORF2
AAAA TTTT
ORF2
XXXX XXXXX
DNA 결합
도메인
ORF2

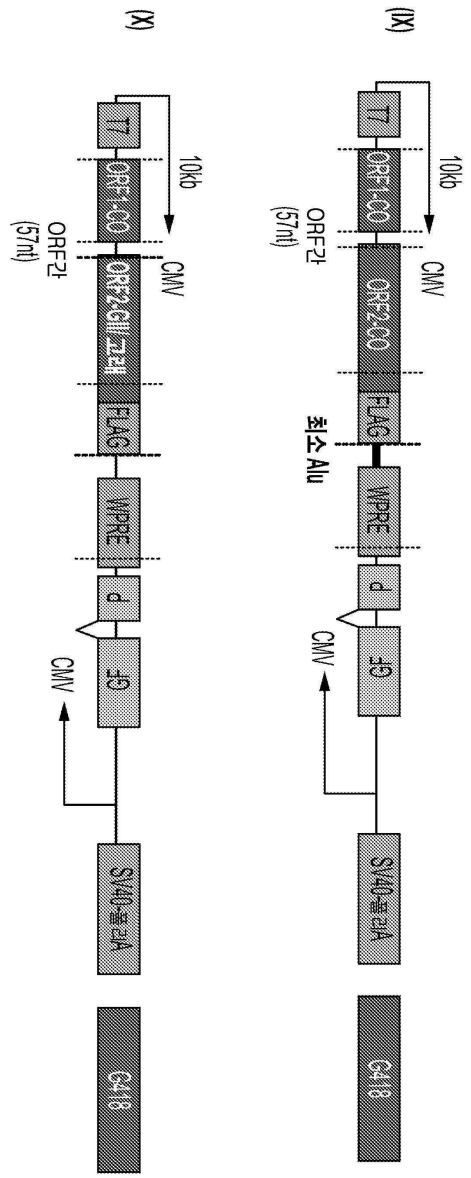
도면5a


(I)
18kb
ORF1
ORF간
(57nt)
ORF2
Dox 유도성
rtTA
CMV
SV40폴리A
Puro 내성
(II)
10kb
T7
ORF1-CO
ORF간
(57nt)
ORF2-CO
CMV
FLAG
WPRE
SV40폴리A
G418
(III)
10kb
T7
ORF1-CO
ORF간
(57nt)
ORF2-CO
CMV
FLAG
WPRE
rtTA
CMV
SV40폴리A
G418
(IV)
10kb
T7
ORF1-CO
ORF간
(57nt)
ORF2-CO
CMV
FLAG
WPRE
rtTA
CMV
SV40폴리A
G418

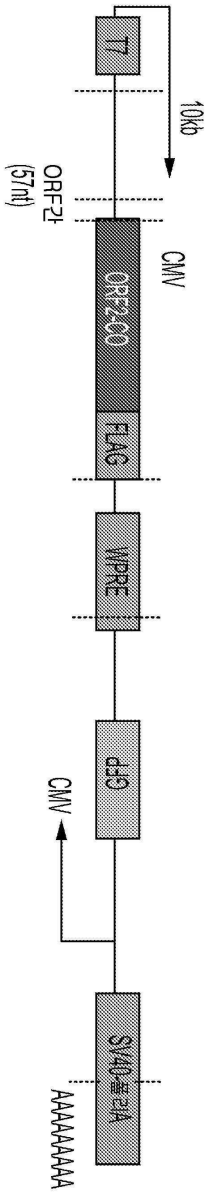
도면5b


(V)
(VI)
(VII)
(VIII)
T7
ORF1-CO
ORF2-CO
FLAG
WPRE
P
GF
SV40-폴리A
G418
10kb
CMV
ORF간
(57nt)
NLS
CVB3 IRES
(740 bp)
MCP
702bp
3xMS2 헤어핀

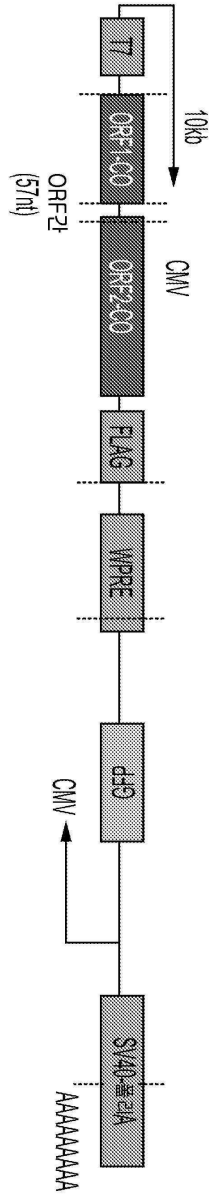
도면5c


(IX)
T7
ORF1-OO
10kb
ORF간
(57nt)
CMV
ORF2-OO
FLAG
최소 Alu
WPRE
p
GF
CMV
SV40폴리A
G418
(X)
T7
ORF1-OO
10kb
ORF간
(57nt)
CMV
ORF2-OO
FLAG
WPRE
p
GF
CMV
SV40폴리A
G418

도면6

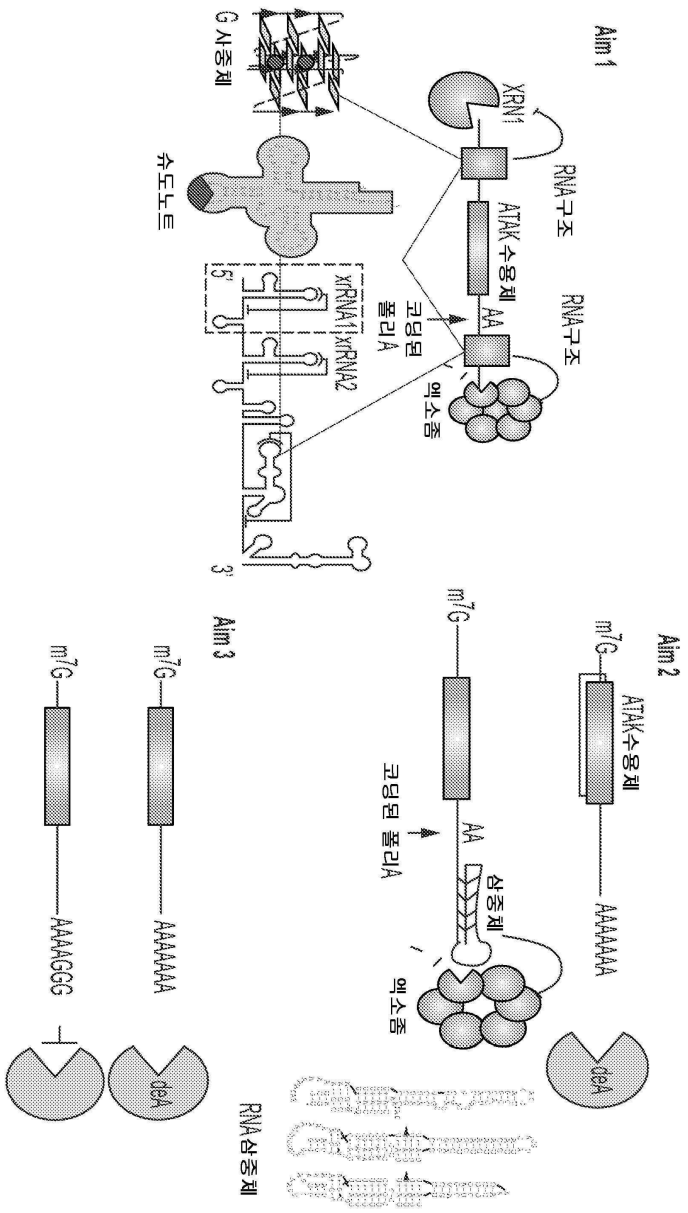

T7
10kb
ORF간
(57nt)
CMV
ORF2-CO
FLAG
WPRE
GFP
CMV
SV40-폴리A
AAAAAAAAA

도 6b

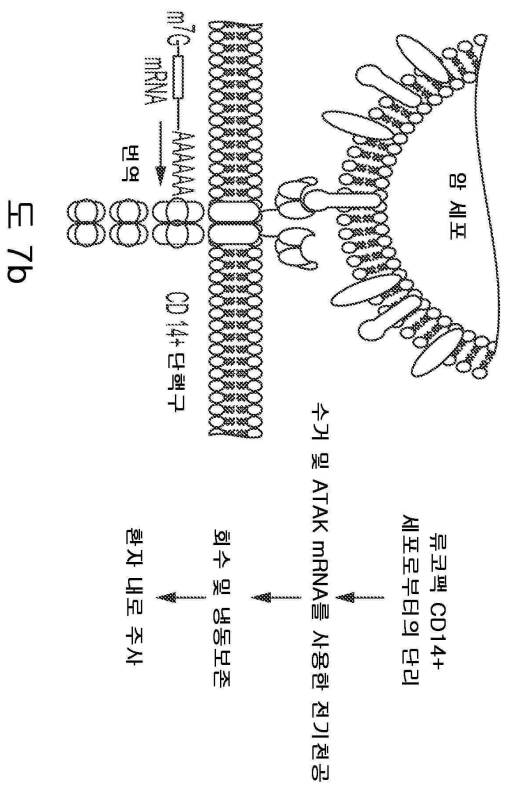

T7
10kb
ORF1-CO
ORF간
(57nt)
CMV
ORF2-CO
FLAG
WPRE
GFP
CMV
SV40-폴리A
AAAAAAAAA

도 6a

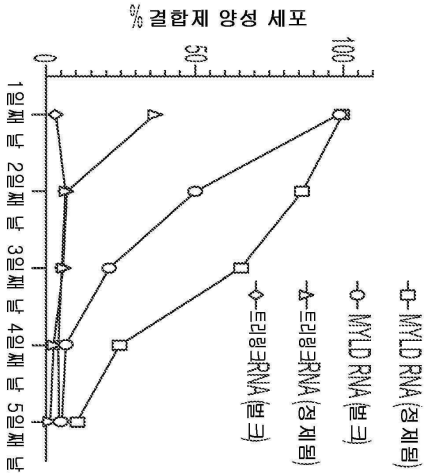
도면7a


Aim 1
XRN1
RNA구조
ATAK수용체
AA
RNA구조
코딩된
폴리A
엑소좀
G 시종제
슈도노트
xrRNA1
xrRNA2
5'
3'
Aim 2
m7G
ATAK수용체
AAAAAAA
코딩된
폴리A
AA
삼중제
엑소좀
dcA
Aim 3
m7G
AAAAGGG
AAAAAAA
dcA
RNA삼중제

도면7bc

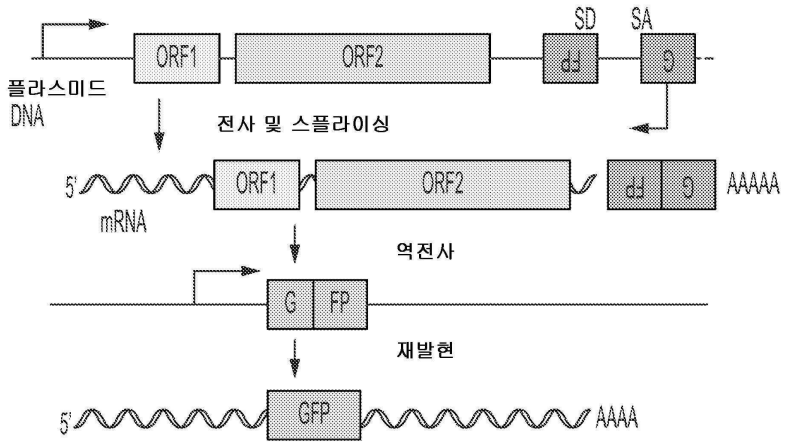


도 7b

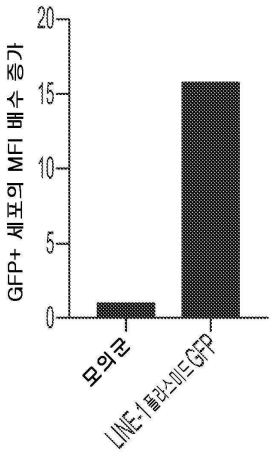


도 7c

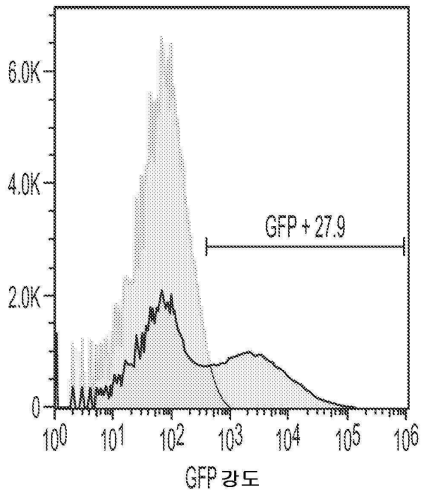
도면8



도 8a

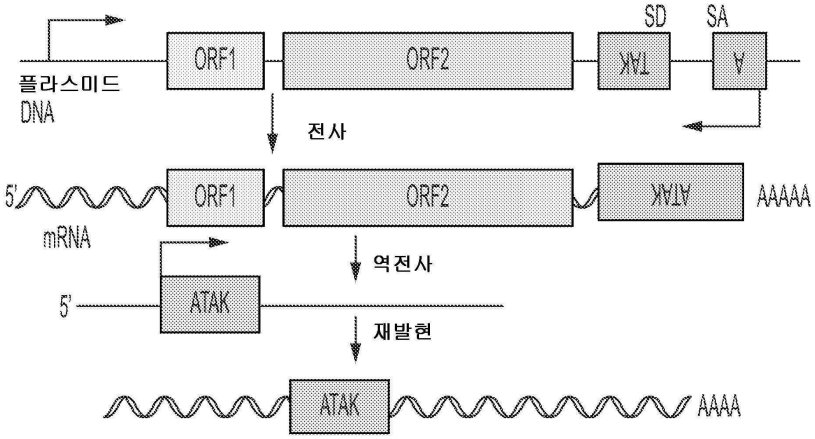


도 8b

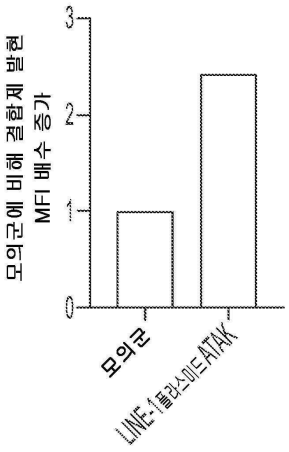


도 8c

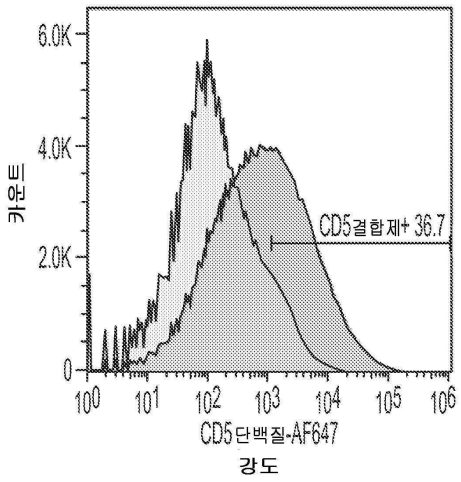
도면9



도 9a

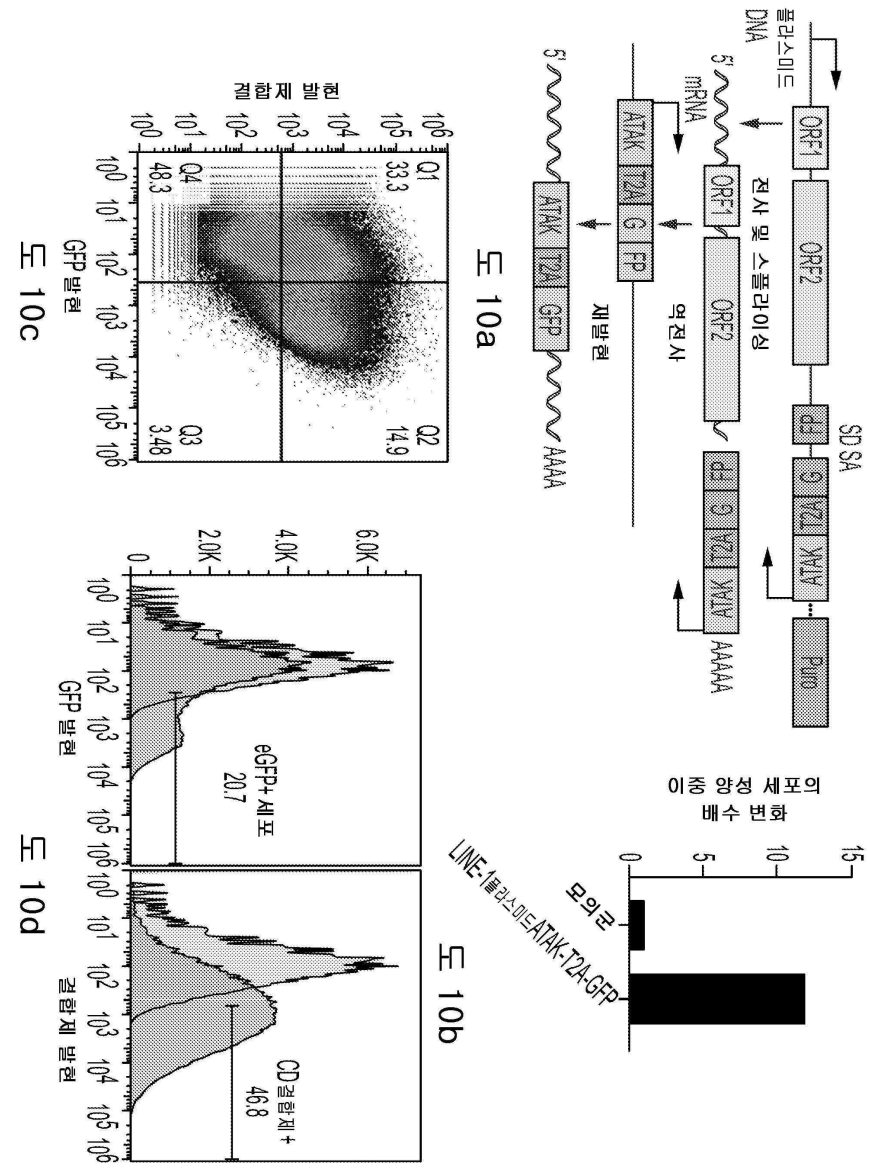


도 9b



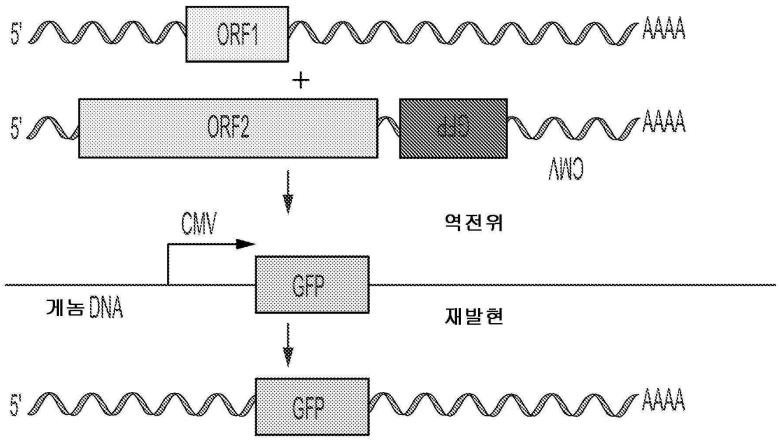
도 9c

도면10

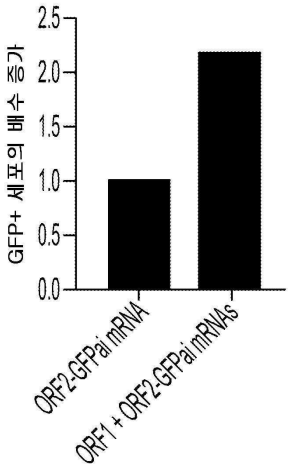


도 10a 도 10b 도 10c 도 10d

도면11

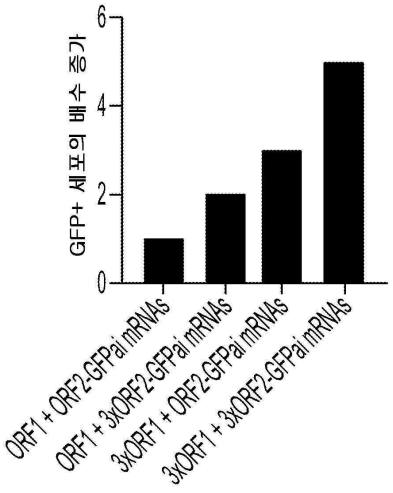


도 11a

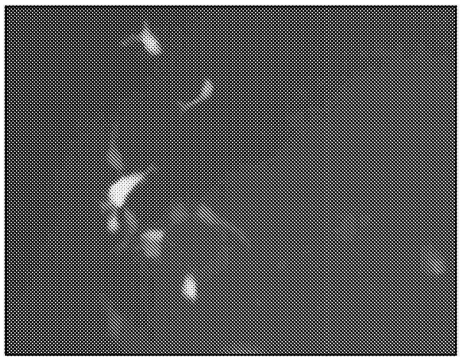


도 11b

도면12



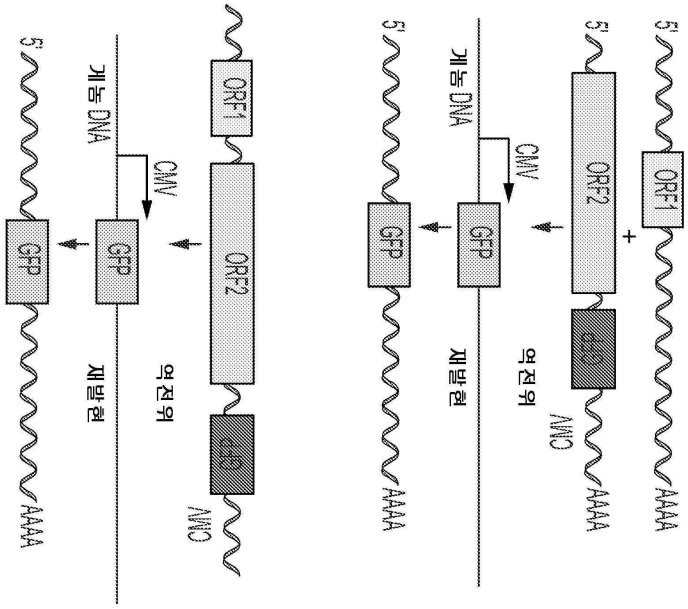
도 12a



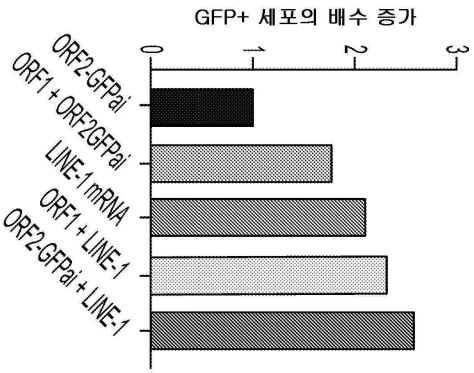
도 12b

도면13

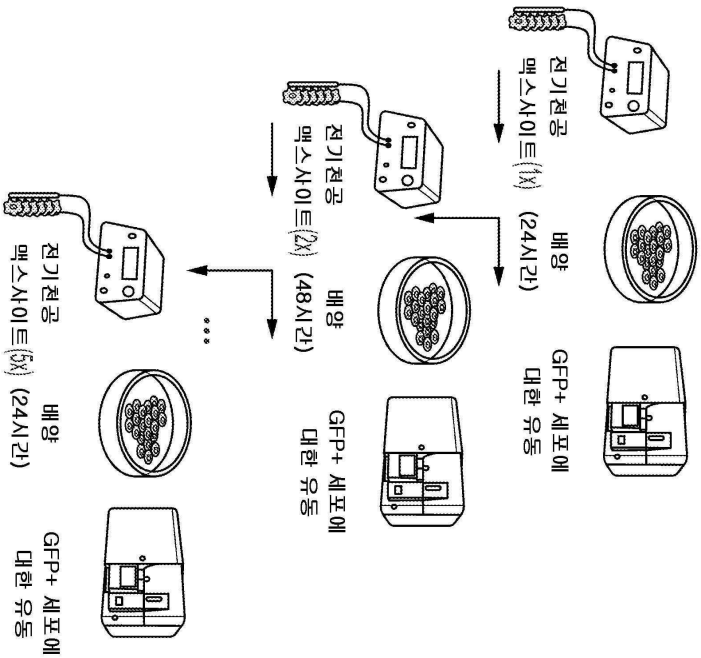
도 13a



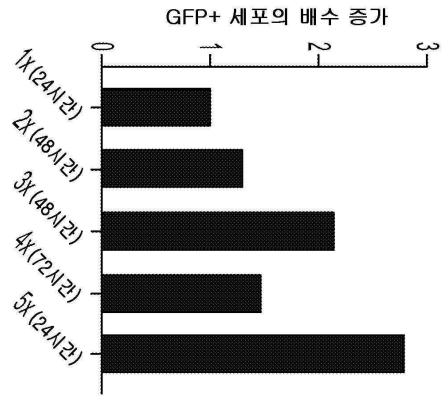
도 13b



도면14

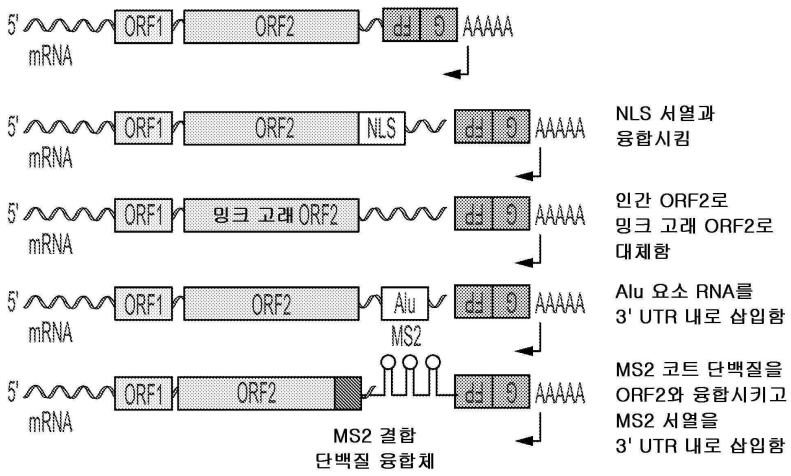


도 14a

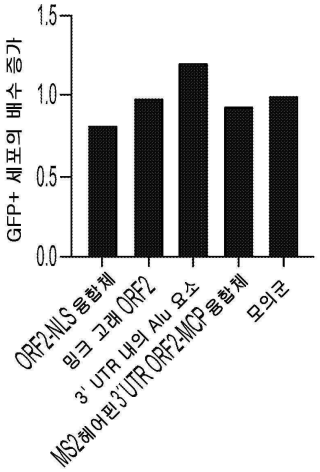


도 14b

도면15

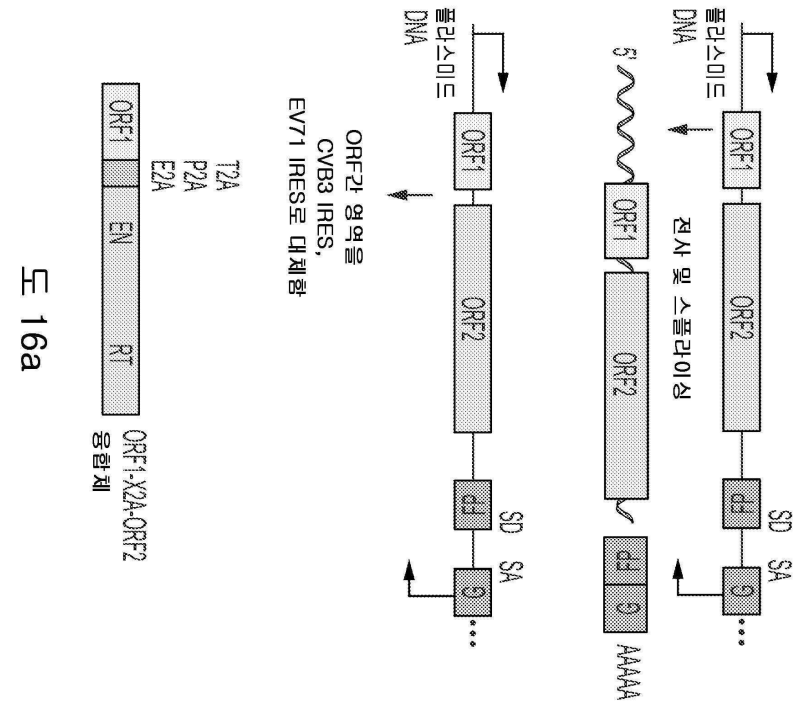


도 15a

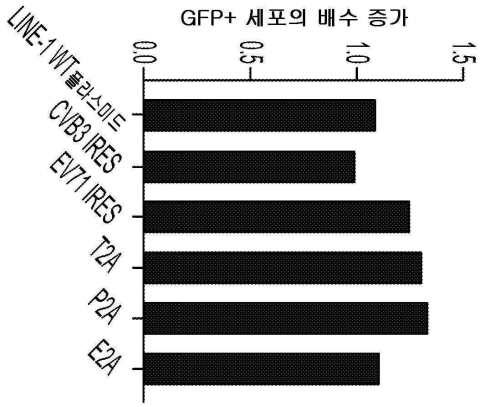


도 15b

도면16

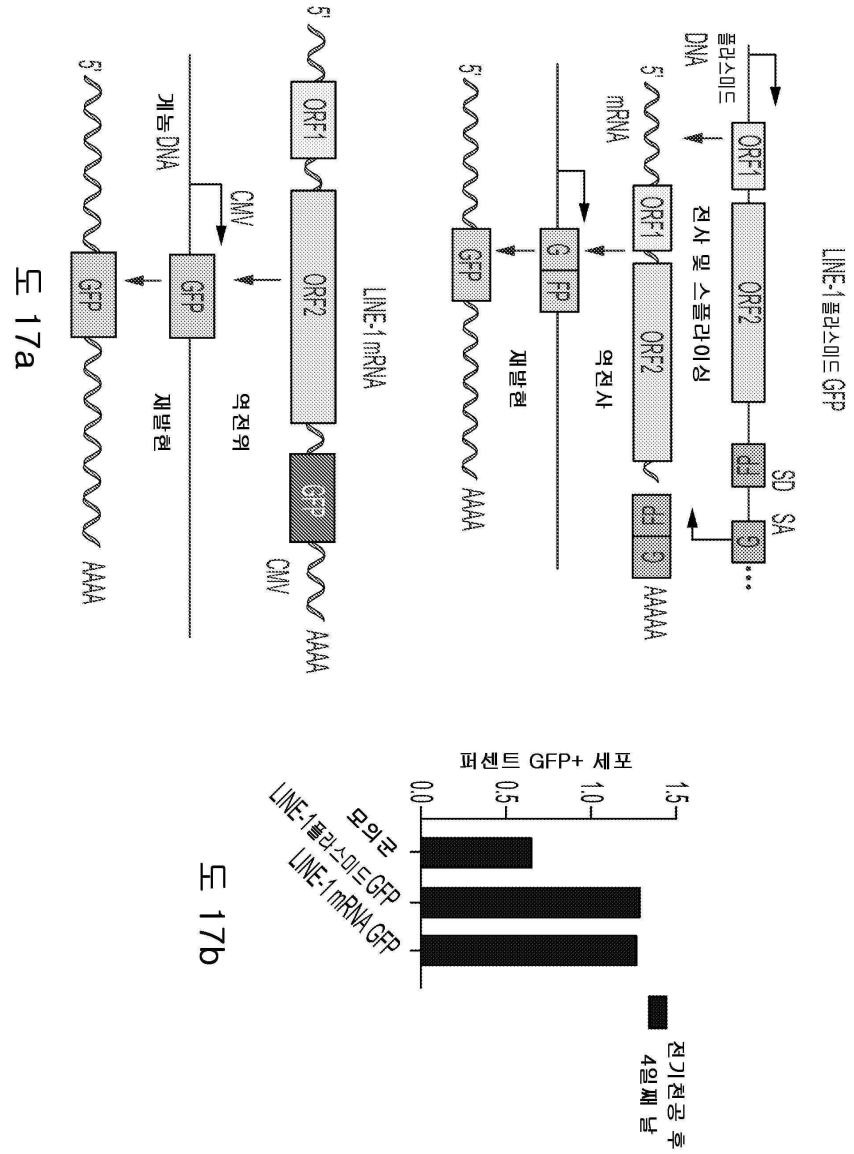


도 16a



도 16b

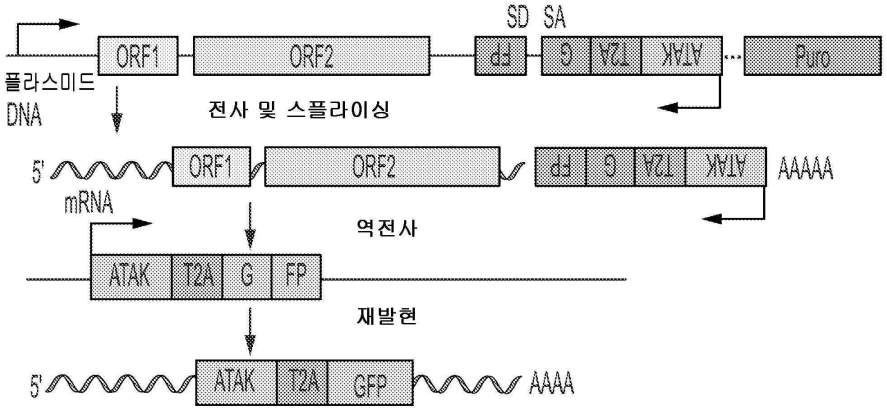
도면17



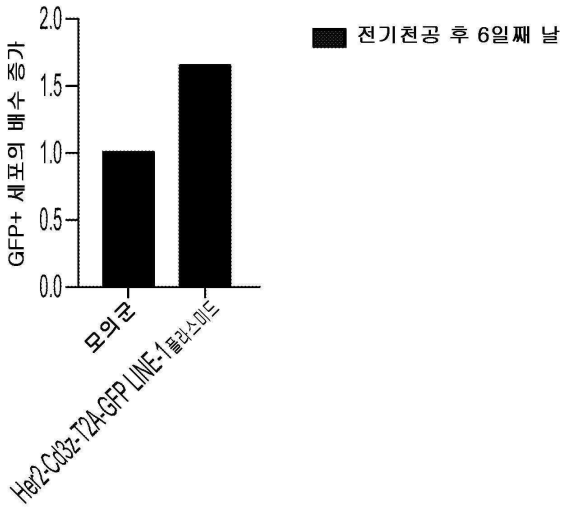
도 17a

도 17b

도면18

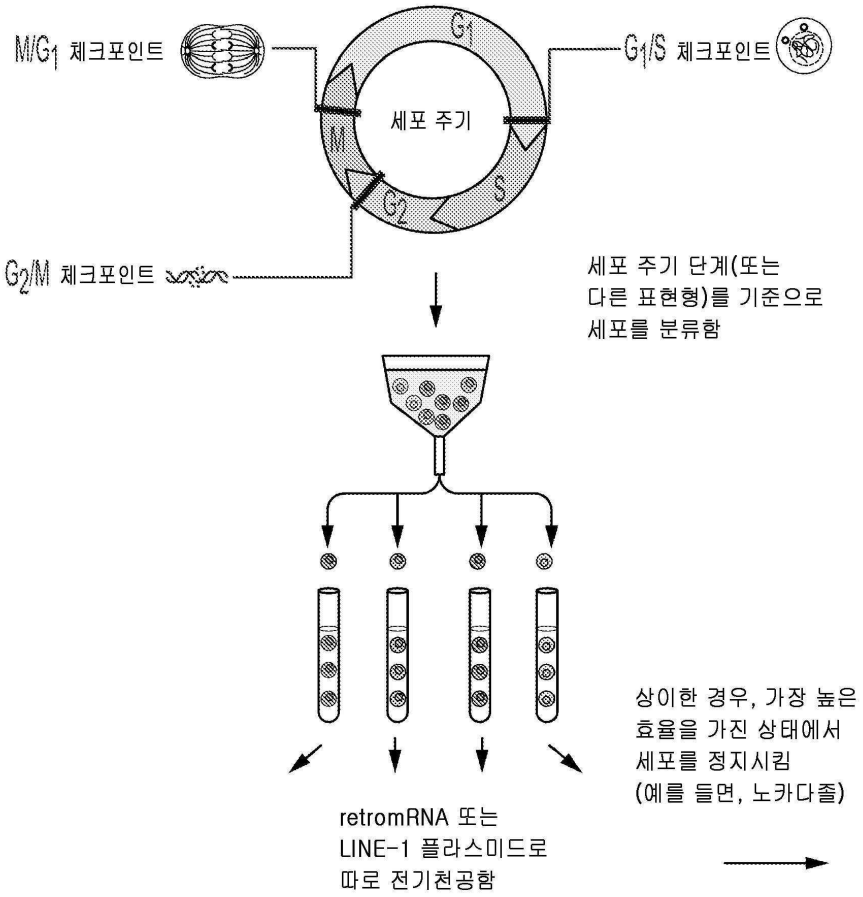


도 18a

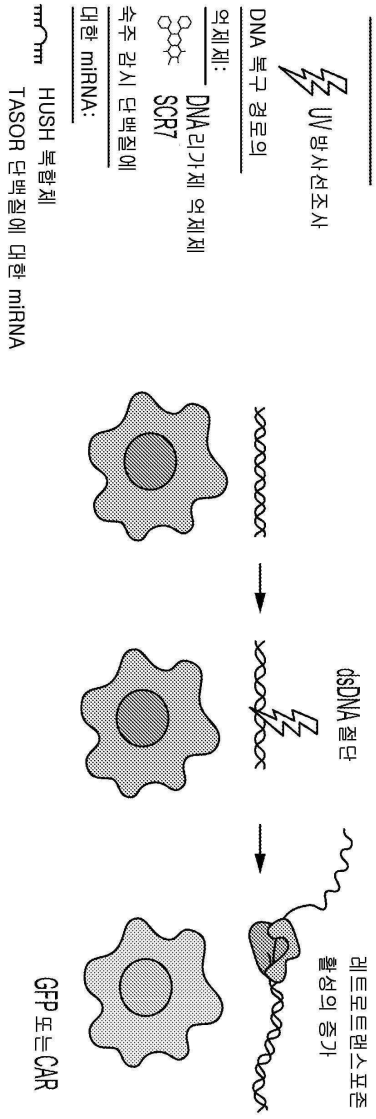


도 18b

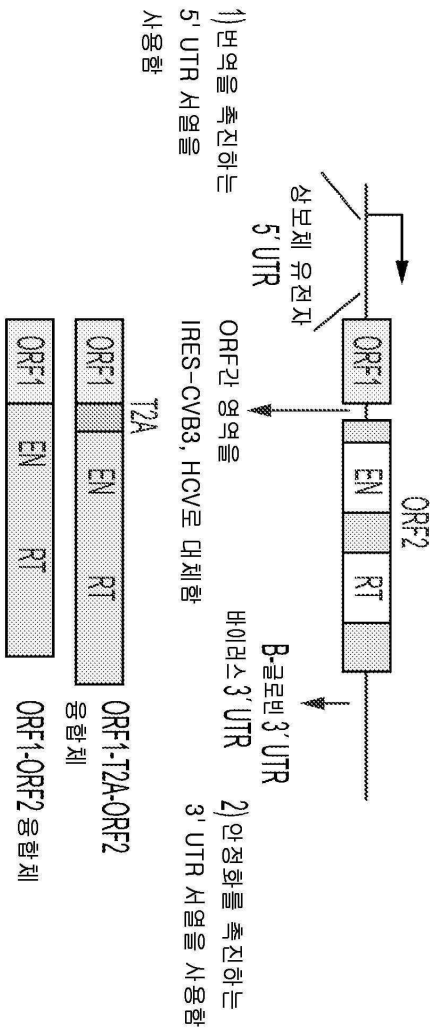
도면19


M/G1 체크포인트
G1/S 체크포인트
G2/M 체크포인트
G1
S
G2
M
세포 주기
세포 주기 단계(또는
다른 표현형)를 기준으로
세포를 분류함
상이한 경우, 가장 높은
효율을 가진 상태에서
세포를 정지시킴
(예를 들면, 노카다졸)
retromRNA 또는
LINE-1 플라스미드로
따로 전기천공함

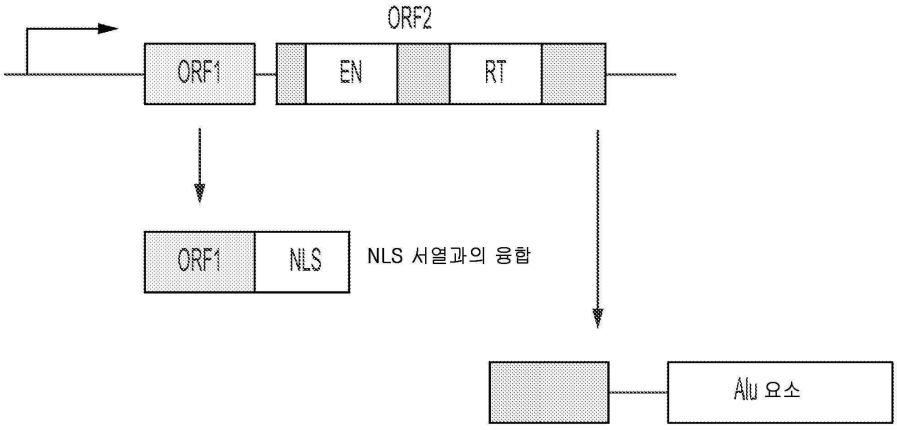
도면20


DNA손상 유도제:
UV 방사선조사
DNA 복구 경로의
억제제:
DNA리가제 억제제
SCR7
숙주 감시 단백질에
대한 miRNA:
HUSH 복합체
TASOR 단백질에 대한 miRNA
dsDNA 절단
레트로트랜스포존
활성의 증가
GFP 또는CAR

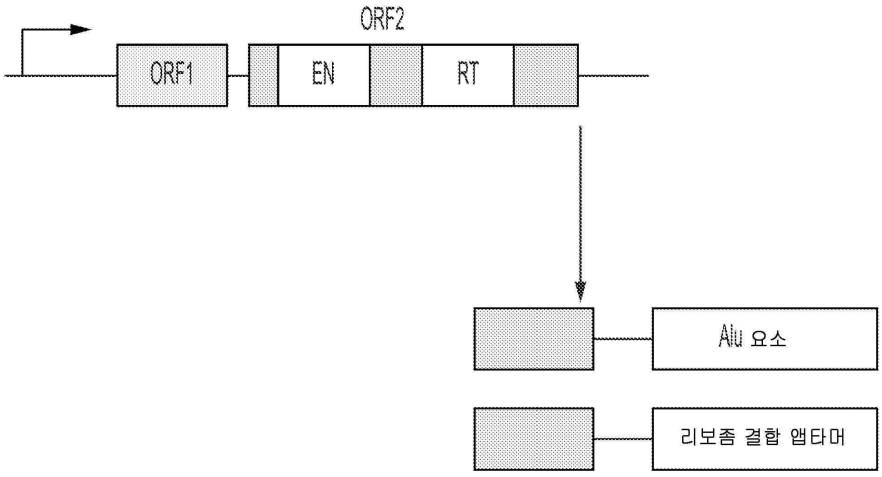
도면21


1) 번역을 촉진하는
5' UTR 서열을
사용함
상보체 유전자
5' UTR
ORF1
EN
RT
ORF2
ORF간 영역을
IRES-CVB3, HCV로 대체함
바이러스 3' UTR
B-글로빈 3' UTR
2) 안정화를 촉진하는
3' UTR 서열을 사용함
ORF1
EN
RT
T2A
ORF1-T2A-ORF2
융합체
ORF1-ORF2 융합체

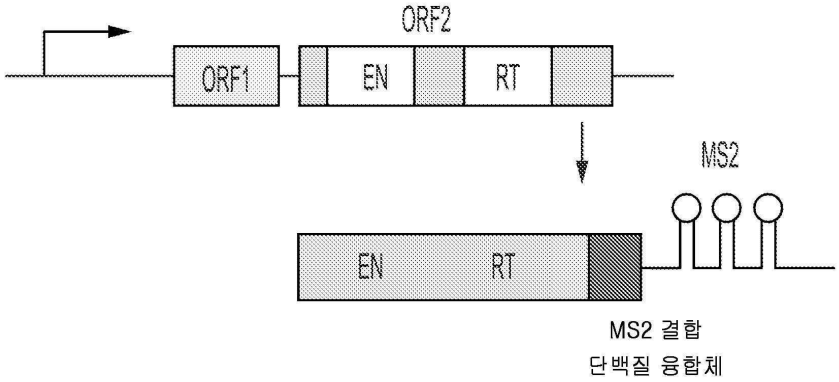
도면22


ORF2
ORF1
EN
RT
ORF1
NLS
NLS 서열과의 융합
Alu 요소

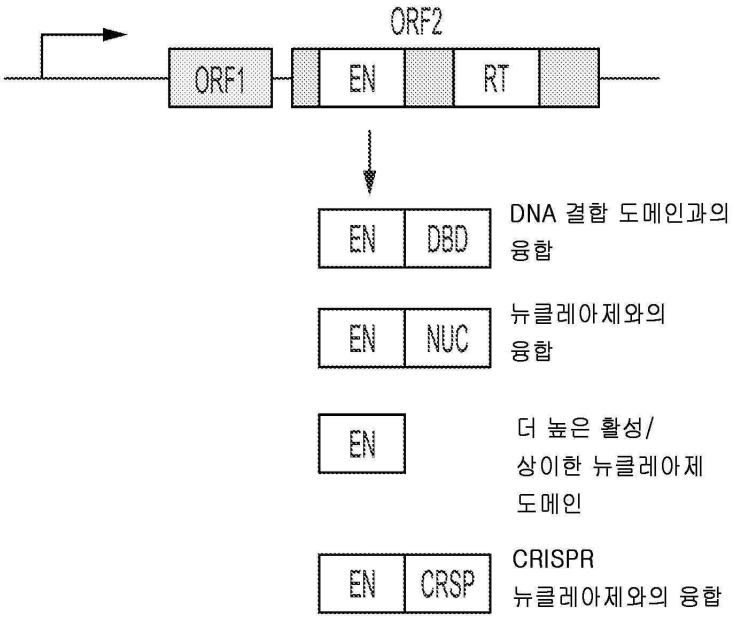
도면23



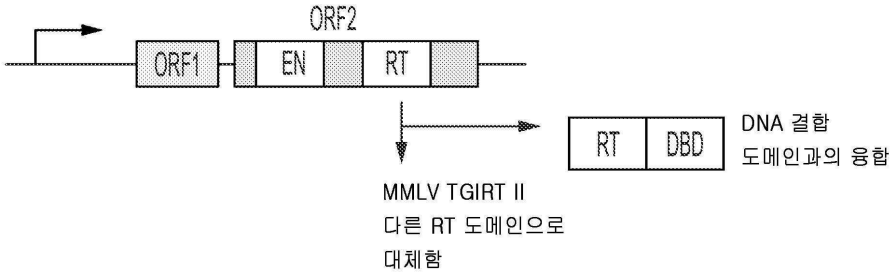
도면24



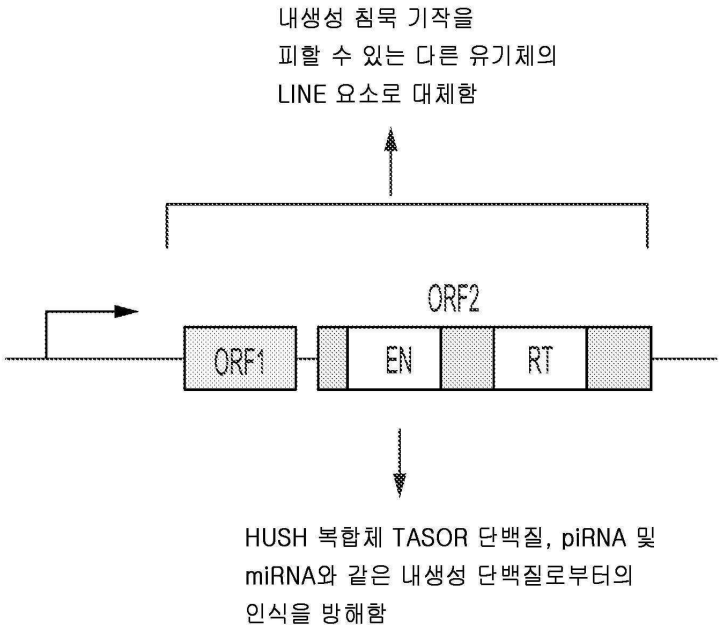
도면25



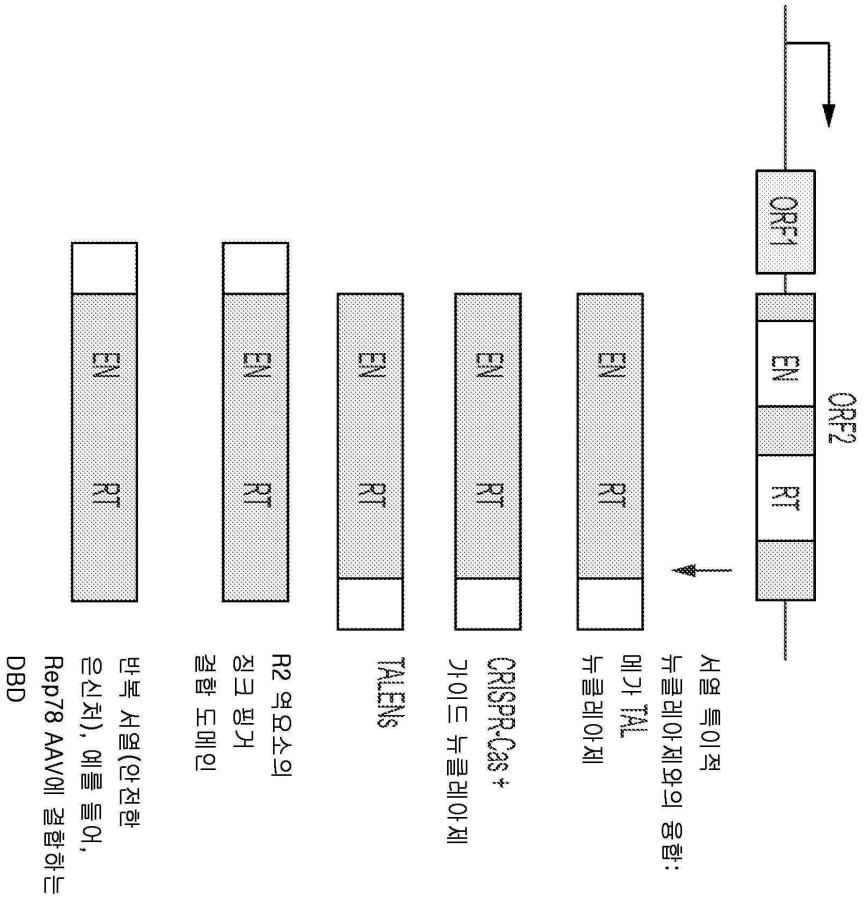
도면26


ORF2
ORF1
EN
RT
RT
DBD
DNA 결합
도메인과의 융합
MMLV TGIRT II
다른 RT 도메인으로
대체함

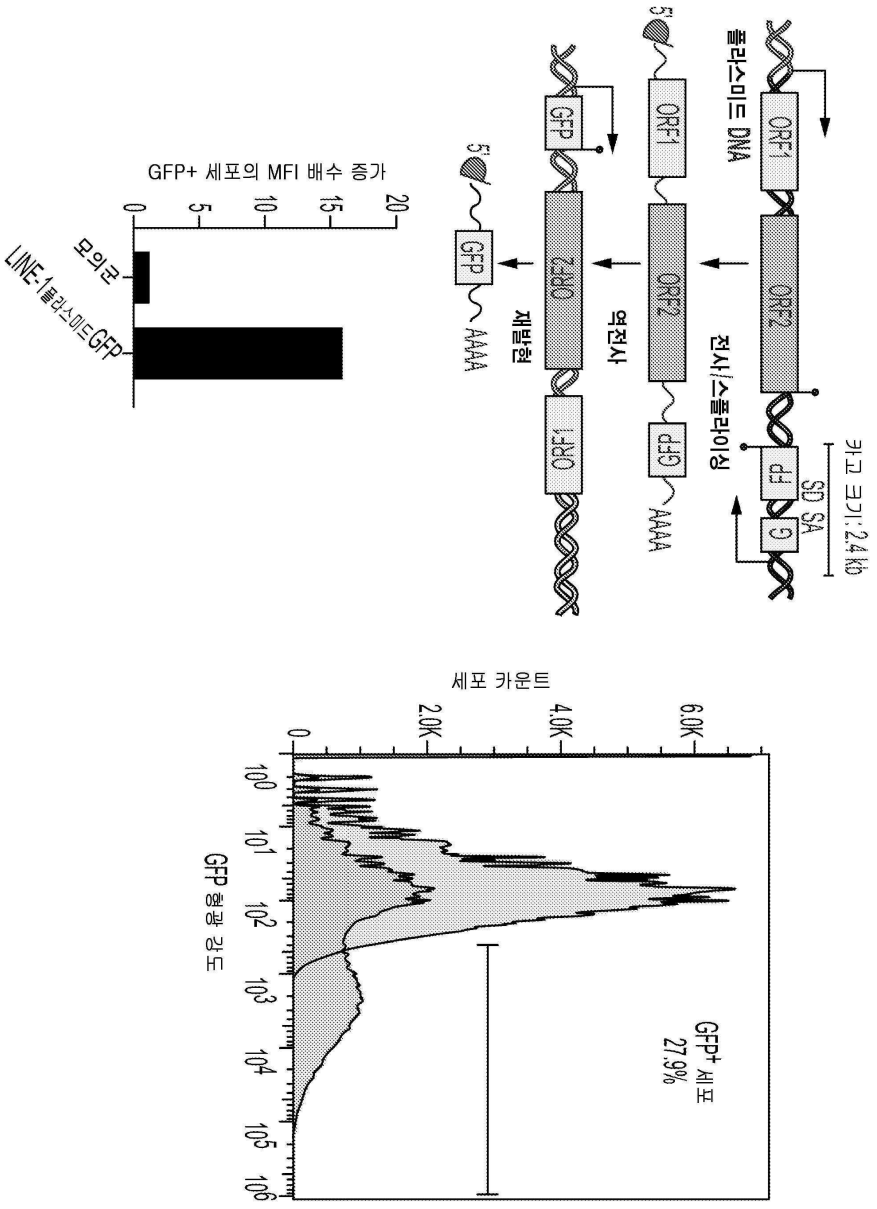
도면27


내생성 침묵 기작을
피할 수 있는 다른 유기체의
LINE 요소로 대체함
ORF2
ORF1
EN
RT
HUSH 복합체 TASOR 단백질, piRNA 및
miRNA와 같은 내생성 단백질로부터의
인식을 방해함

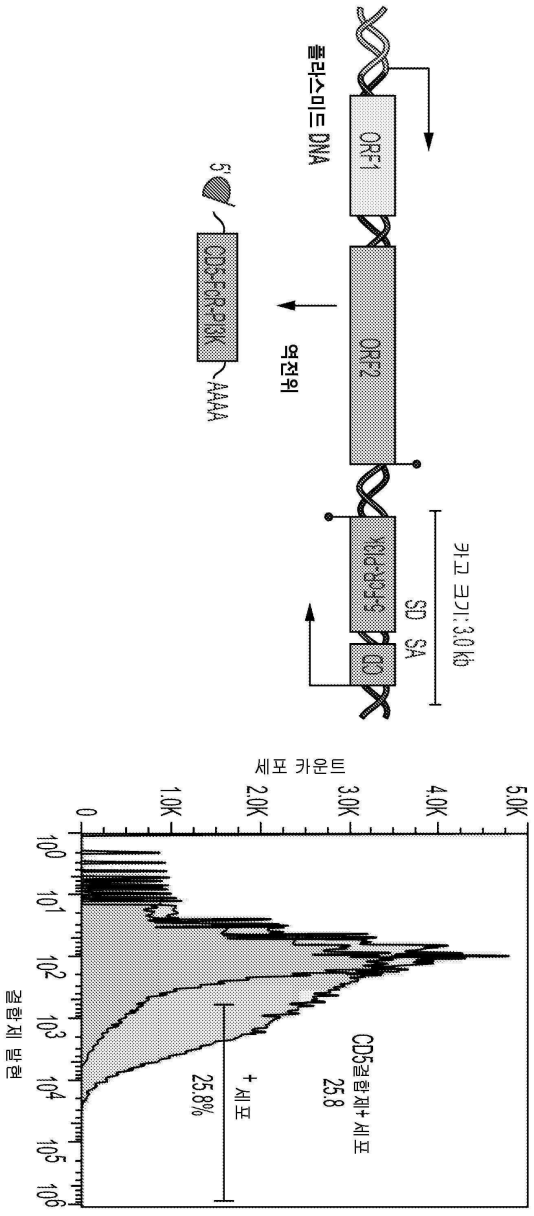
도면28


ORF1
EN
RT
ORF2
서열 특이적
뉴클레아제와의 융합:
메가 TAL
뉴클레아제
EN
RT
CRISPR-Cas +
가이드 뉴클레아제
EN
RT
TALENs
EN
RT
R2 역요소의
징크 핑거
결합 도메인
EN
RT
반복 서열(안전한
은신처), 예를 들어,
Rep78 AAV에 결합하는
DBD

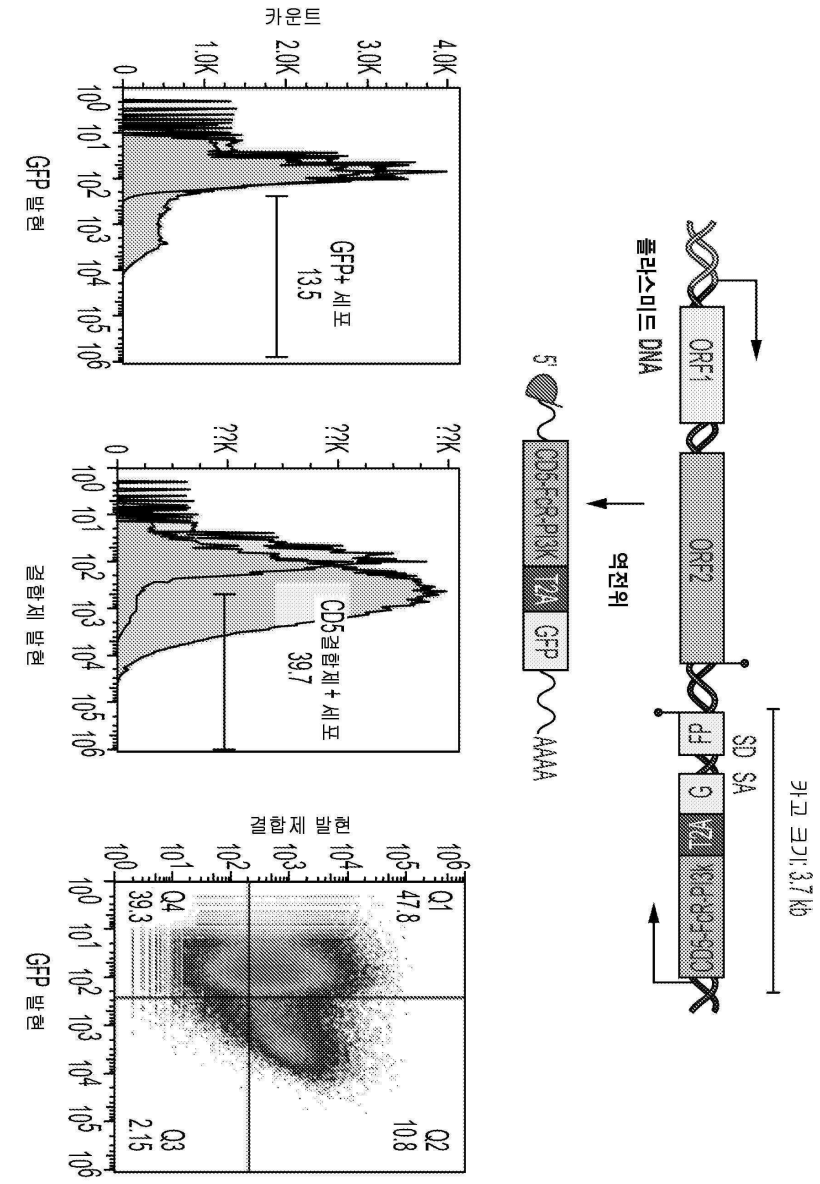
도면29


플라스미드 DNA
ORF1
ORF2
EP
SD SA
G
카고 크기: 2.4 kb
전사/스플라이싱
ORF1
ORF2
GFP
AAAA
역전사
ORF2
ORF1
GFP
5'
재통합
5'
GFP
AAAA
GFP+ 세포의 MFI 배수 증가
0
5
10
15
20
모의군
LINE-1플라스미드GFP
세포 카운트
0
2.0K
4.0K
6.0K
GFP 형광 강도
10^0
10^1
10^2
10^3
10^4
10^5
10^6
GFP+ 세포
27.9%

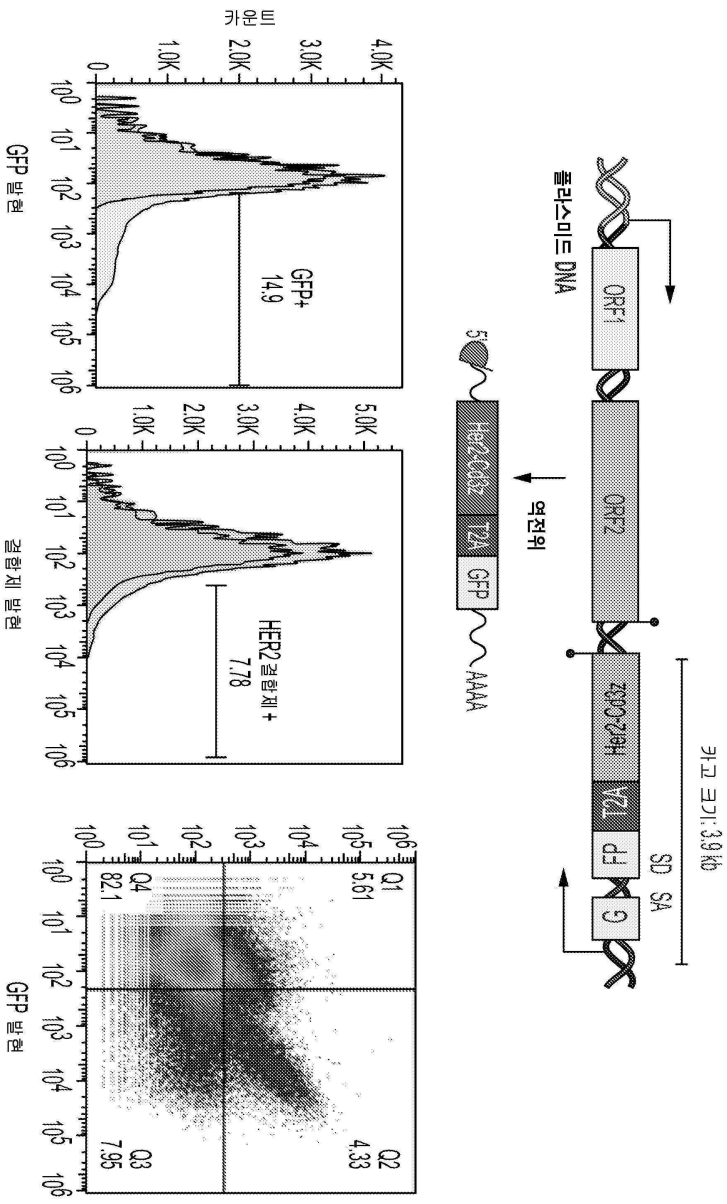
도면30


플라스미드 DNA
ORF1
ORF2
CD5-FGFR-PI3K
SD
SA
CD
카고 크기:3.0 kb
역전위
5'
AAAA
세포 카운트
0
1.0K
2.0K
3.0K
4.0K
5.0K
결합 제 발현
10^0
10^1
10^2
10^3
10^4
10^5
10^6
CD5결합제+세포
25.8
+ 세포
25.8%

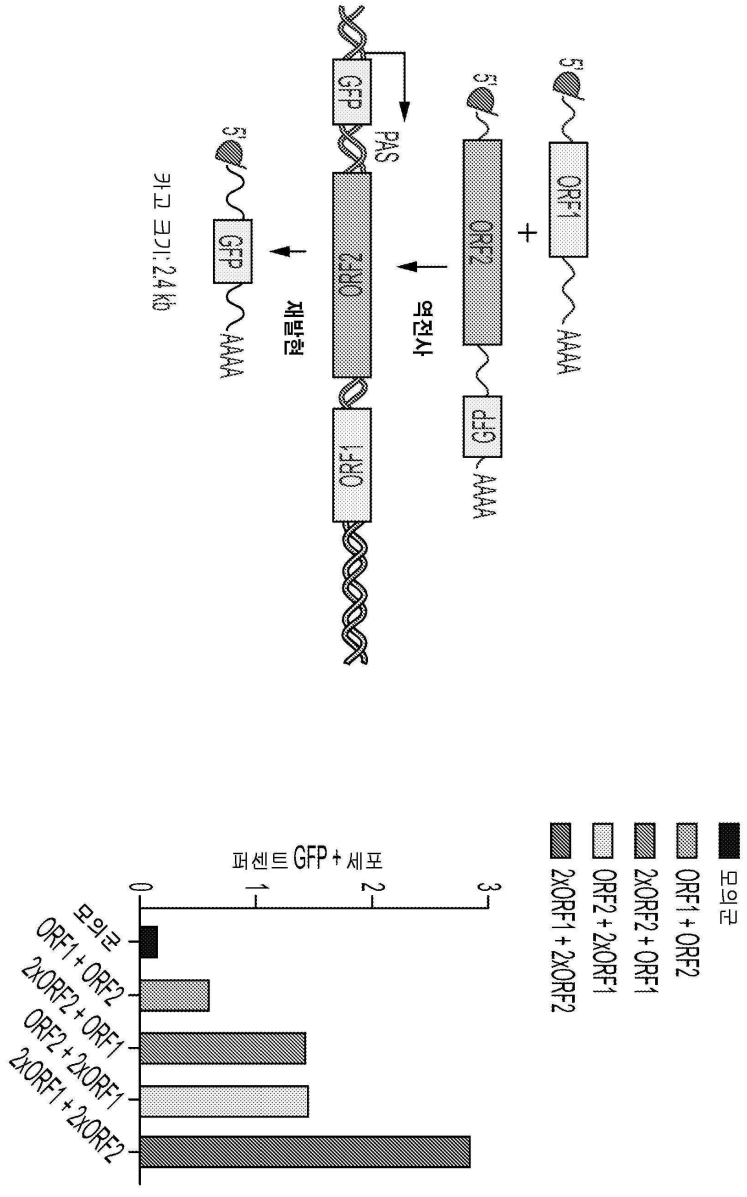
도면31


카운트
GFP 발현
GFP+ 세포
13.5
결합제 발현
CD5결합제 + 세포
39.7
결합제 발현
GFP 발현
Q1
47.8
Q2
10.8
Q3
2.15
Q4
39.3
플라스미드 DNA
ORF1
ORF2
역전위
5'
CD5-FcR-PI3K
T2A
GFP
AAAA
SD
SA
FP
G
T2A
CD5-FcR-PI3K
카고 크기: 3.7 kb

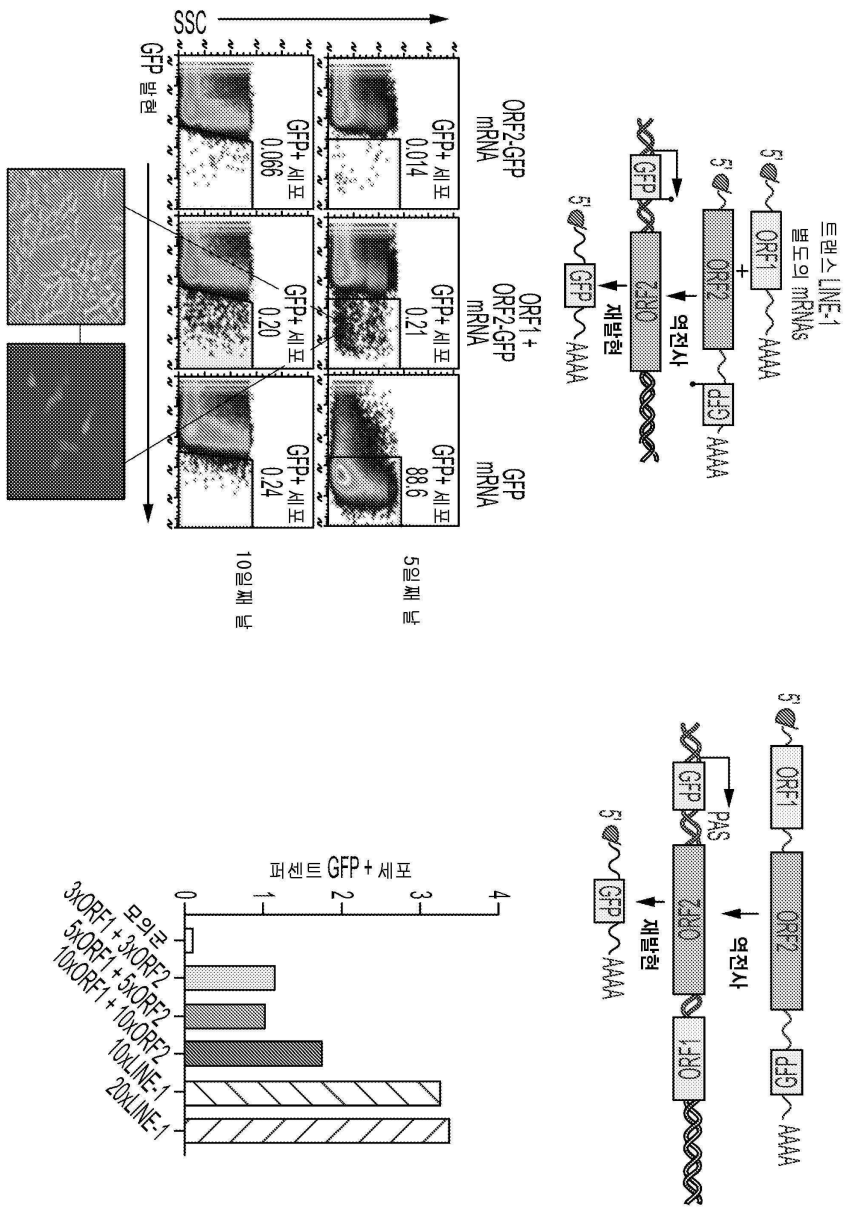
도면32


카운트
GFP 발현
GFP+
14.9
결합제 발현
HER2결합제+
7.78
GFP 발현
Q1
5.61
Q2
4.33
Q3
7.95
Q4
82.1
플라스미드 DNA
ORF1
ORF2
HER2결합제
T2A
FP
G
SD
SA
카고 크기:3.9 kb
역전위
5'
HER2결합제
T2A
GFP
AAAA

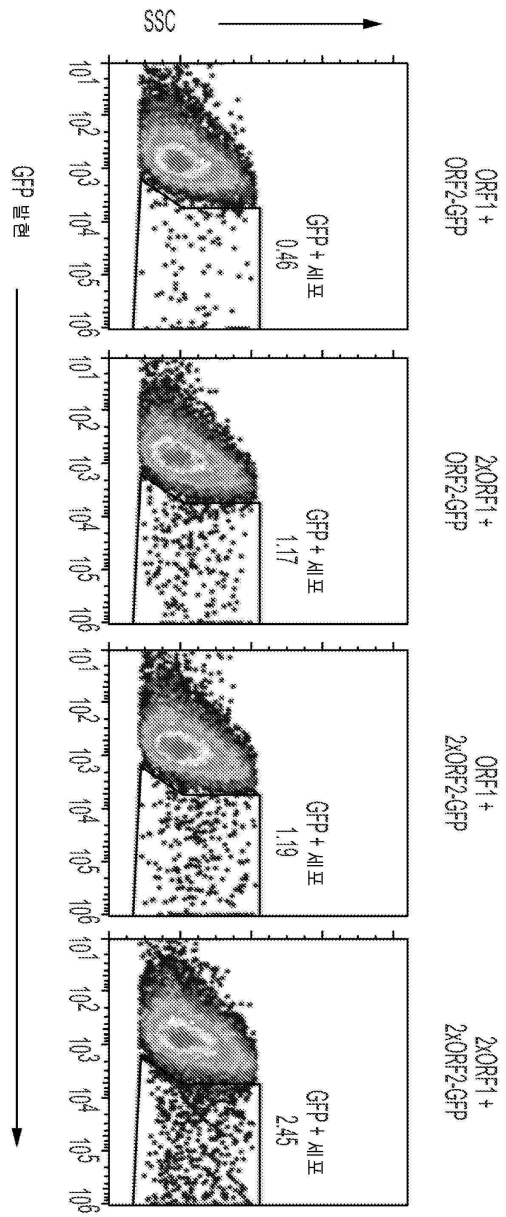
도면33a


5'
ORF1
AAAA
+
5'
ORF2
GFP
AAAA
전사
GFP
PAS
ORF2
ORF1
재결합
5'
GFP
AAAA
카고 크기: 2.4 kb
퍼센트 GFP + 세포
0
1
2
3
모의군
ORF1 + ORF2
2xORF2 + ORF1
ORF2 + 2xORF1
2xORF1 + 2xORF2
모의군
ORF1 + ORF2
2xORF2 + ORF1
ORF2 + 2xORF1
2xORF1 + 2xORF2

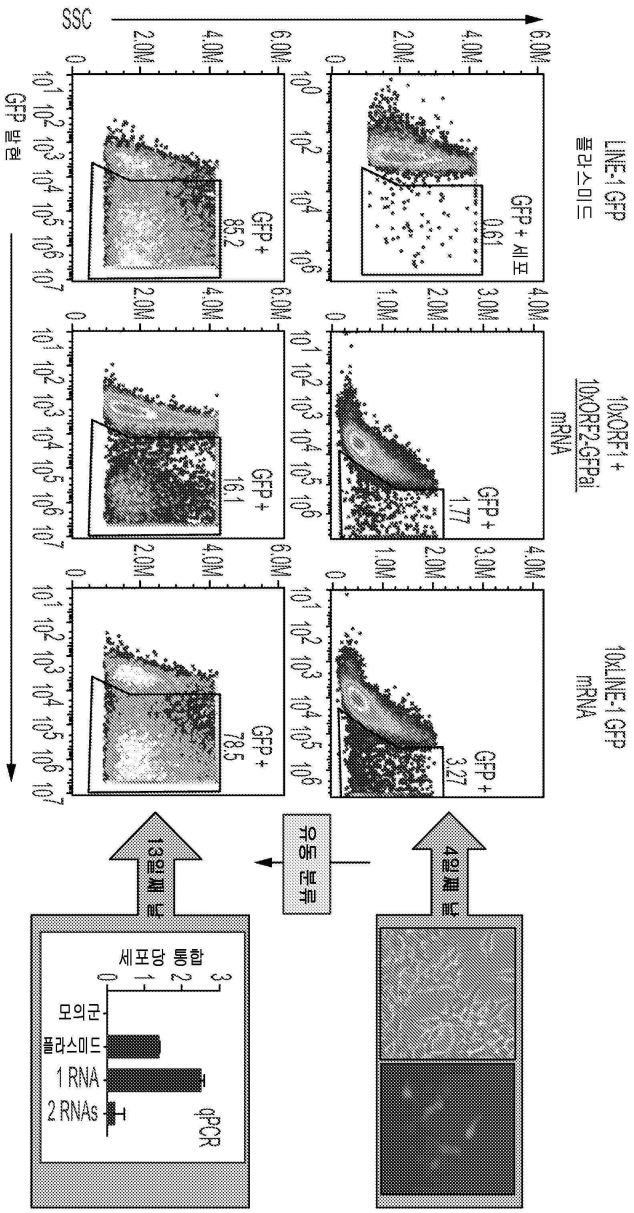
도면33b


SSC
GFP 발현
ORF2-GFP mRNA
GFP+ 세포 0.014
GFP+ 세포 0.066
ORF1+ ORF2-GFP mRNA
GFP+ 세포 0.21
GFP+ 세포 0.20
GFP mRNA
GFP+ 세포 88.6
GFP+ 세포 0.24
5일째 날
10일째 날
트랜스 LINE-1 별도의 mRNAs
ORF1
ORF2
GFP
AAAA
역전사
재발현
PAS
퍼센트 GFP + 세포
0
1
2
3
4
모의군
3xORF1 + 3xORF2
5xORF1 + 5xORF2
10xORF1 + 10xORF2
10xLINE-1
20xLINE-1

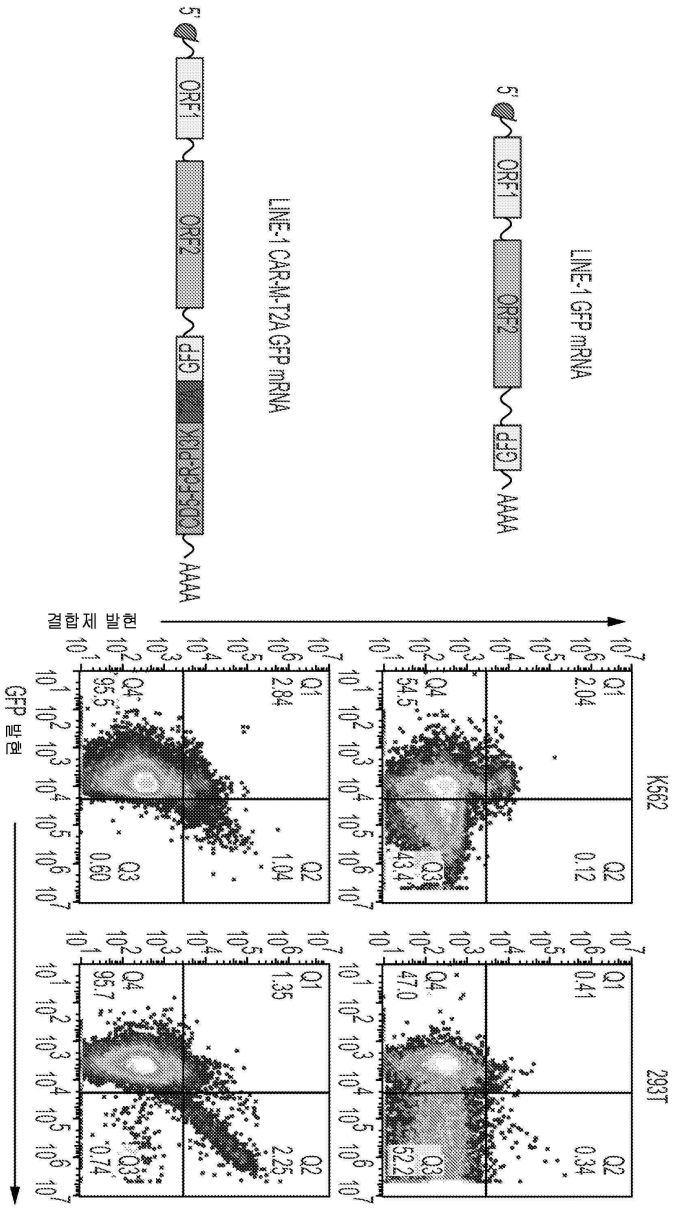
도면33c


SSC
GFP 발현
ORF1 +
ORF2-GFP
GFP + 세포
0.46
2xORF1 +
ORF2-GFP
GFP + 세포
1.17
ORF1 +
2xORF2-GFP
GFP + 세포
1.19
2xORF1 +
2xORF2-GFP
GFP + 세포
2.45

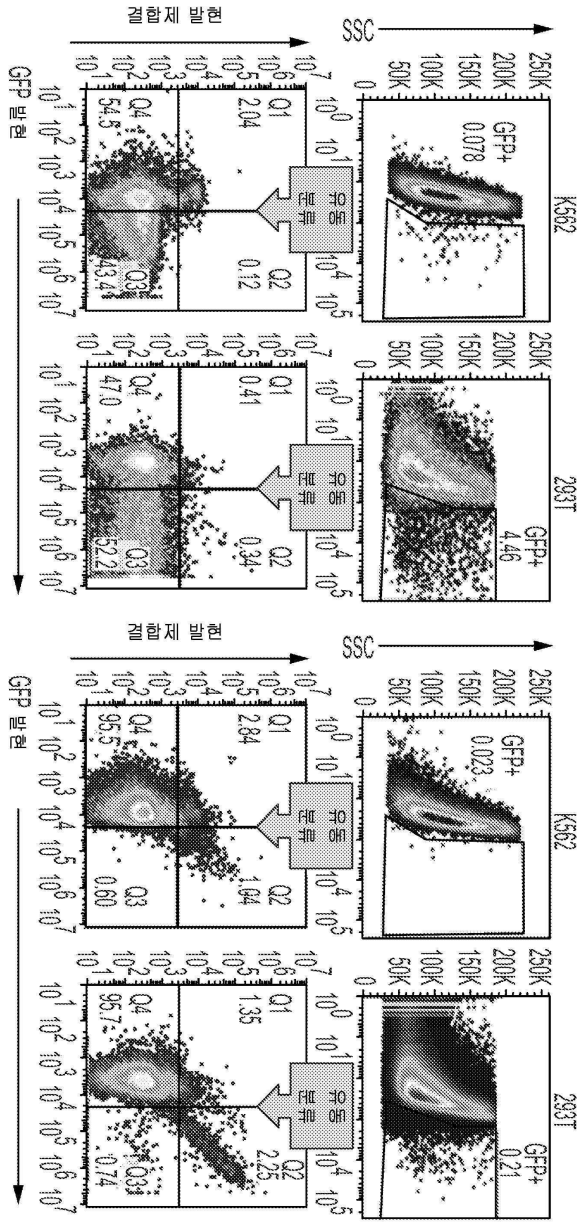
도면33d


SSC
GFP 발현
LINE-1 GFP
플라스미드
10xORF1 +
10xORF2-GFPai
mRNA
10xLINE-1 GFP
mRNA
GFP + 세포
0.61
GFP +
85.2
GFP +
1.77
GFP +
16.1
GFP +
3.27
GFP +
78.5
유동 분류
4일째 날
13일째 날
세포당 통합
모의군
플라스미드
1 RNA
2 RNAs
qPCR

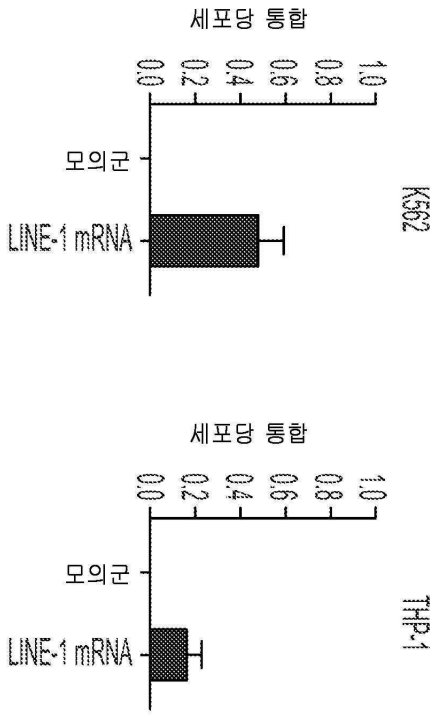
도면34


LINE-1 GFP mRNA
LINE-1 CAR-M-T2A GFP mRNA
ORF1
ORF2
GFP
AAAA
결합제 발현
GFP 발현
K562
293T
Q1 2.04
Q2 0.12
Q3 43.4
Q4 54.5
Q1 2.84
Q2 1.04
Q3 0.60
Q4 95.5
Q1 0.41
Q2 0.34
Q3 52.2
Q4 47.0
Q1 1.35
Q2 2.25
Q3 0.74
Q4 95.7

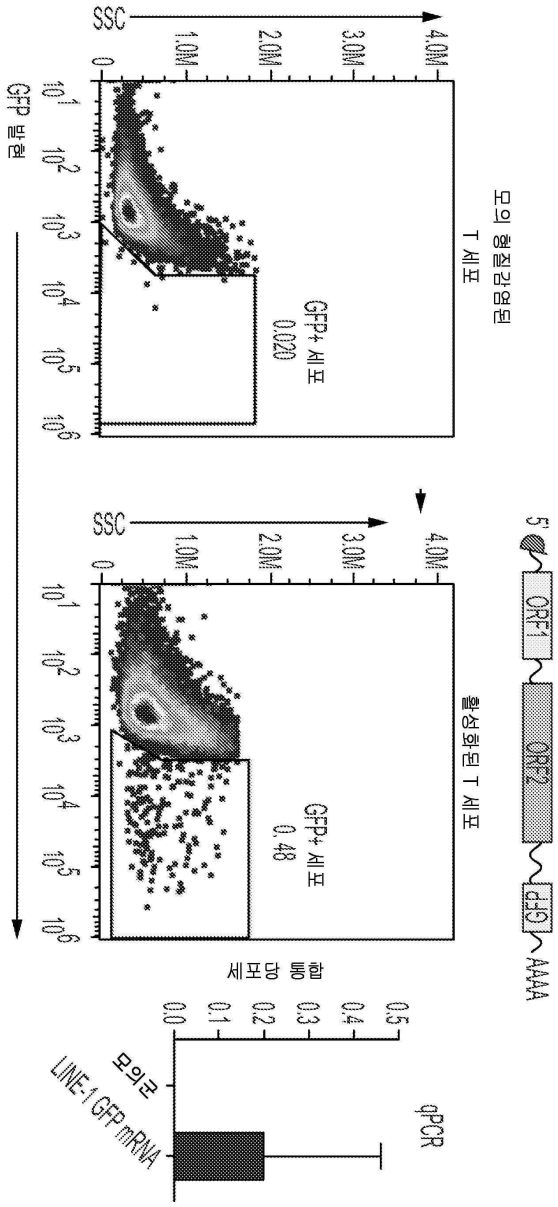
도면35a


결합제 발현
GFP 발현
SSC
K562
293T
GFP+
0.078
4.46
0.023
0.21
Q1
Q2
Q3
Q4
2.04
0.12
54.5
43.4
0.41
0.34
47.0
52.2
2.84
1.94
95.5
0.60
1.35
2.25
95.7
0.74
유동분류

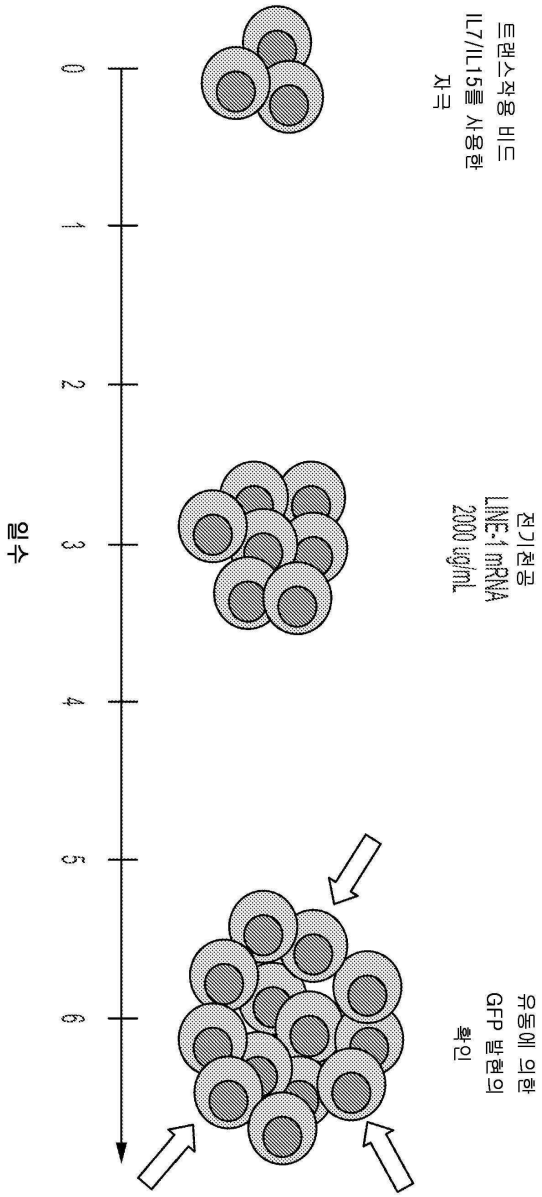
도면35b


세포당 통합
0.0
0.2
0.4
0.6
0.8
1.0
모의군
LINE-1 mRNA
K562
세포당 통합
0.0
0.2
0.4
0.6
0.8
1.0
모의군
LINE-1 mRNA
THP-1

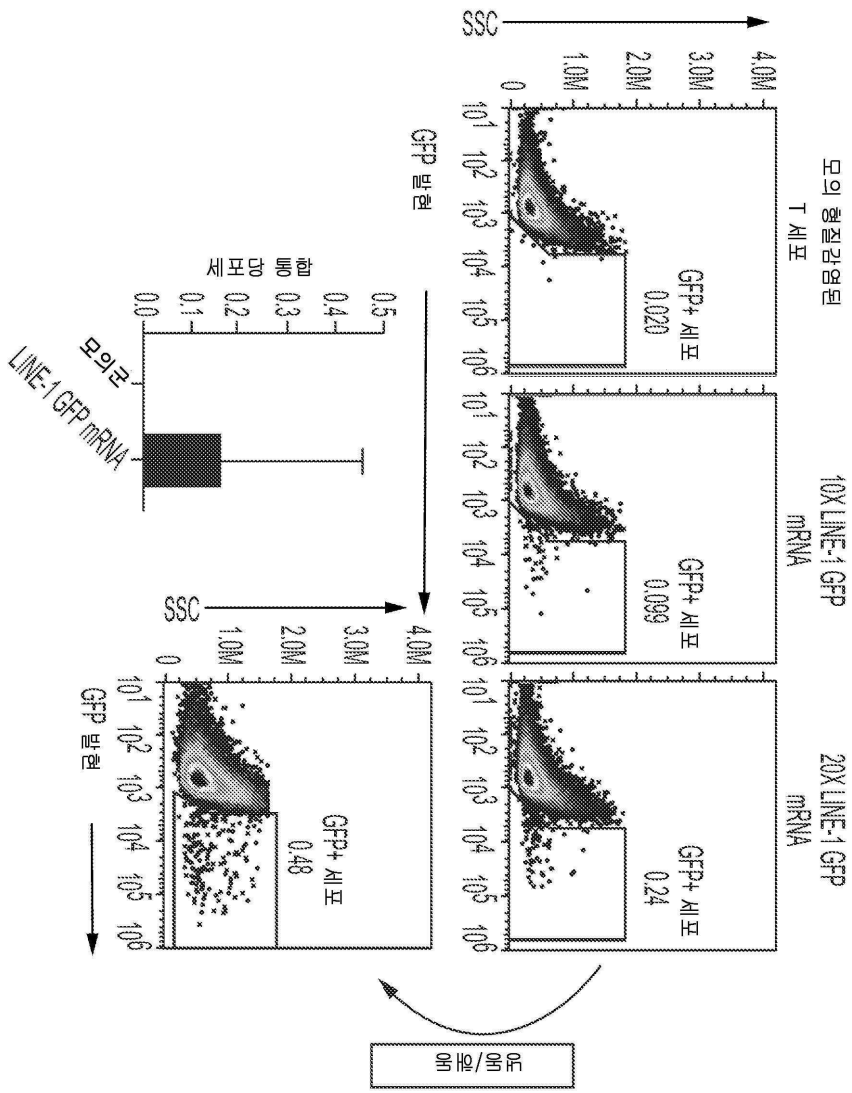
도면36


SSC
GFP 발현
모의 형질감염된 T 세포
GFP+ 세포
0.020
활성화된 T 세포
GFP+ 세포
0.48
세포당 통합
5'
ORF1
ORF2
EGFP
AAAA
모의로
LINE-1 GFP mRNA
qPCR

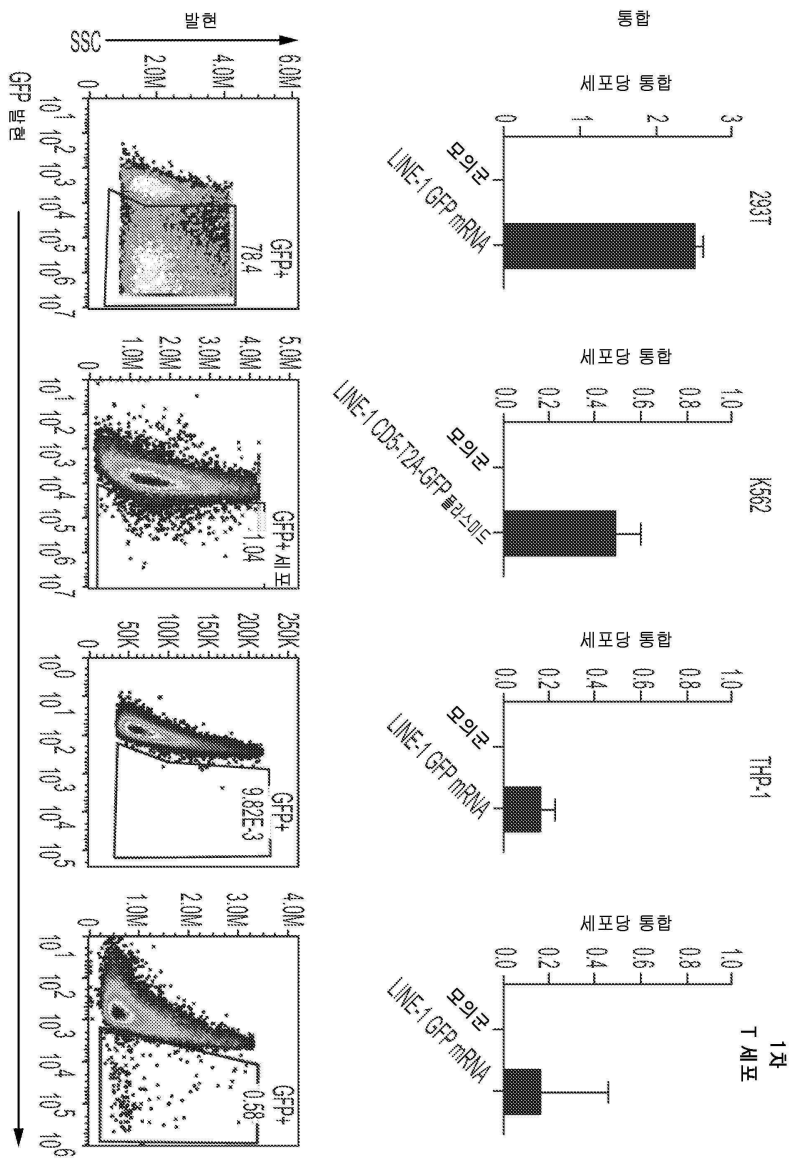
도면37a


일수
0
1
2
3
4
5
6
트랜스작용 비드
IL7/IL15를 사용한
자극
전기천공
LINE-1 mRNA
200 ug/mL
유동에 의한
GFP 발현의
확인

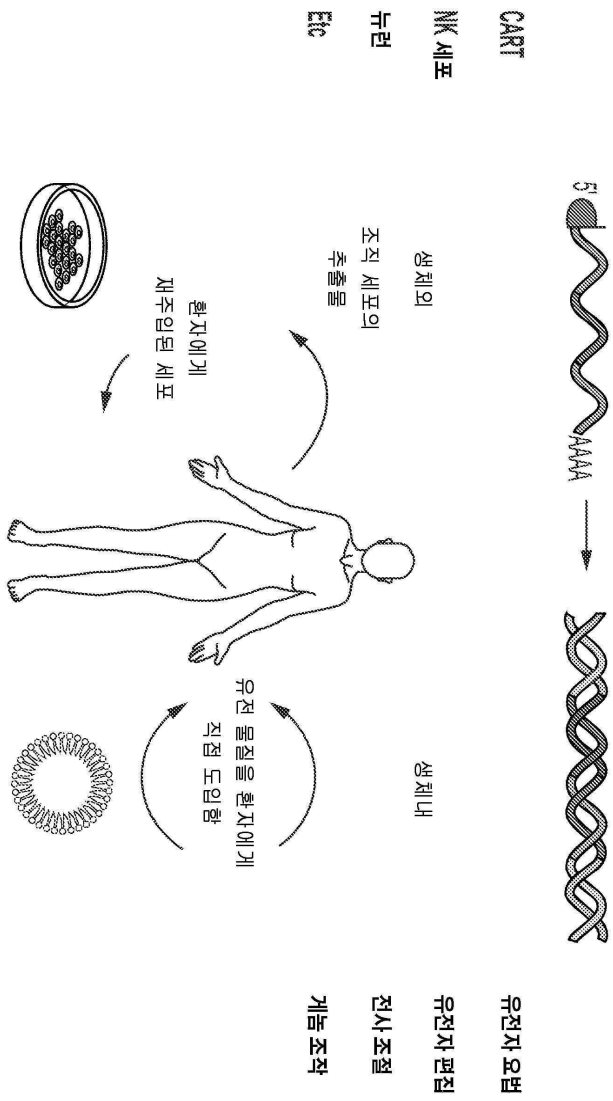
도면37b


SSC
GFP 발현
모의 형질감염된 T 세포
GFP+ 세포
0.020
10X LINE-1 GFP mRNA
GFP+ 세포
0.099
20X LINE-1 GFP mRNA
GFP+ 세포
0.24
세포당 통합
모의군
LINE-1 GFP mRNA
GFP+ 세포
0.48
낮음/하음

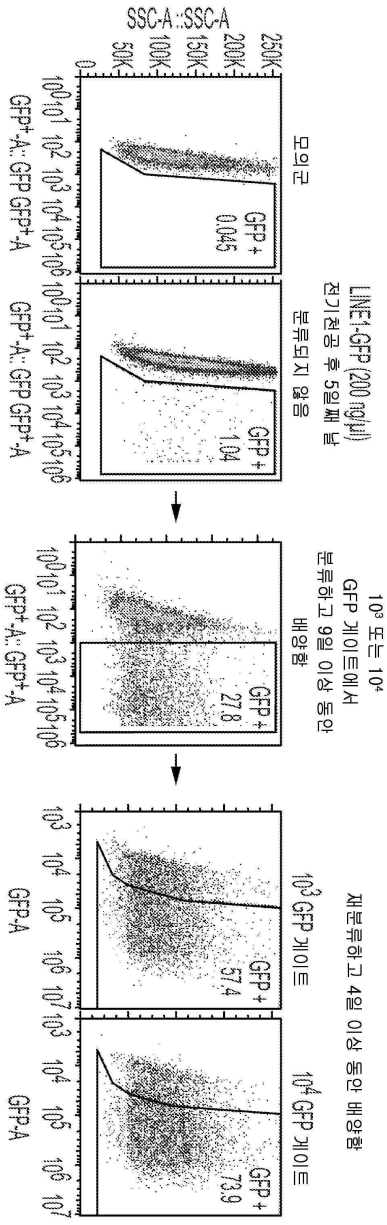
도면38


통합
세포당 통합
293T
K562
THP-1
1차 T 세포
모의군
LINE-1 GFP mRNA
LINE-1 CD5-T2A-GFP 플라스미드
발현
SSC
GFP 발현
GFP+
78.4
GFP+ 세포
1.04
9.82E-3
0.58

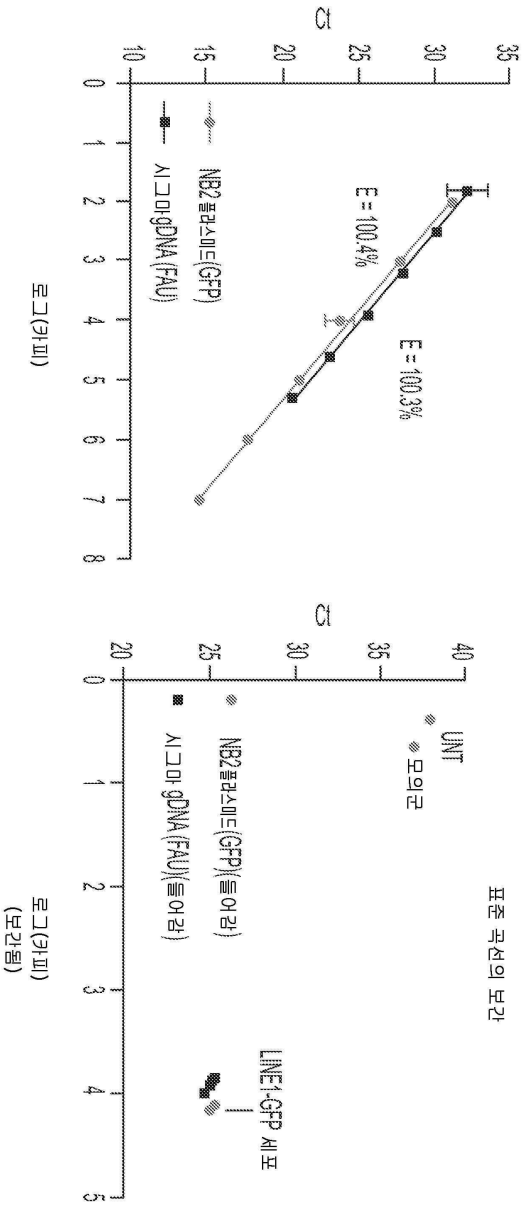
도면39


CART
NK 세포
뉴런
Etc
5'
AAAA
생체외
조직 세포의
추출물
환자에게
재주입된 세포
생체내
유전 물질을 환자에게
직접 도입함
유전자 요법
유전자 편집
전사 조절
게놈 조작

도면40


SSC-A::SSC-A
250K
200K
150K
100K
50K
0
GFP+
0.045
모의군
LINE1-GFP (200 ng/μl)
전기천공 후 5일째 날
분류되지 않음
GFP+
1.04
10³ 또는 10⁴
GFP 게이트에서
분류하고 9일 이상 동안
배양함
GFP+
27.8
재분류하고 4일 이상 동안 배양함
10³ GFP 게이트
GFP+
57.4
10⁴ GFP 게이트
GFP+
73.9
GFP+-A:: GFP
GFP-A

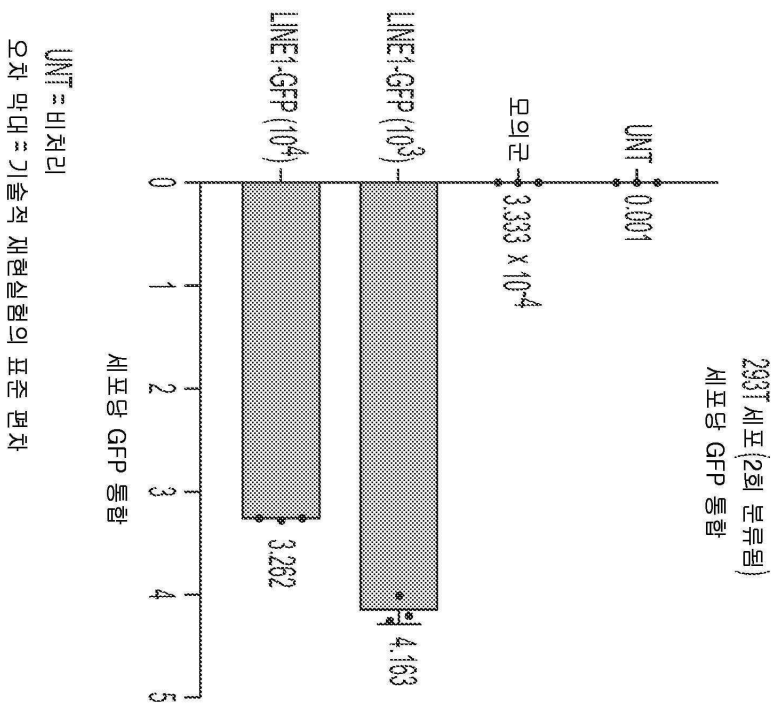
도면41ab



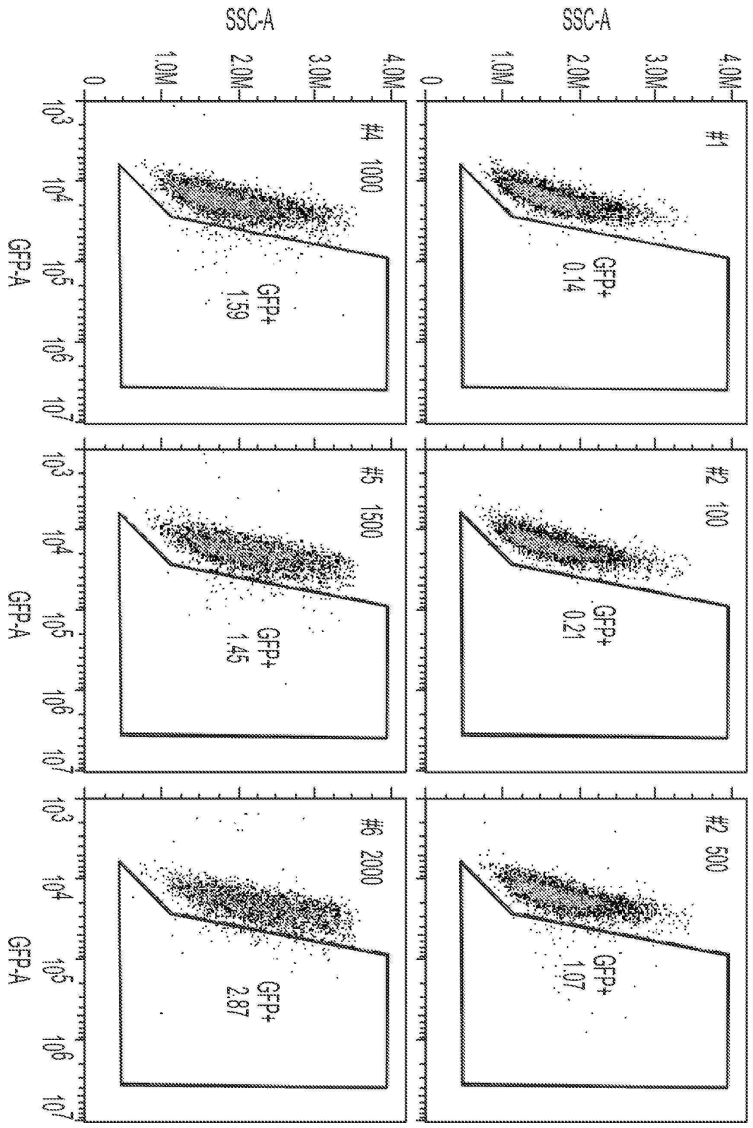
도 41a

도 41b

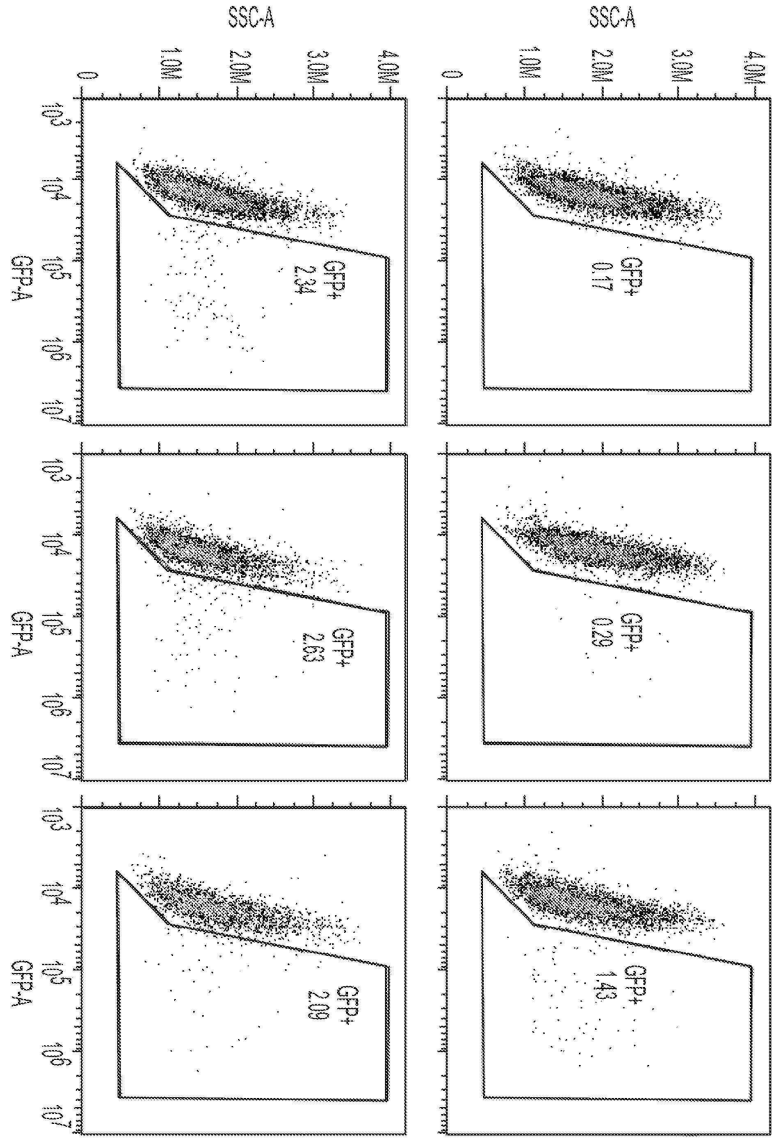
도면41c


293T 세포 (2회 분류됨)
세포당 GFP 통합
UNT
0.001
모의군
3.333 x 10-4
LINE1-GFP (103)
4.163
LINE1-GFP (104)
3.262
0
1
2
3
4
5
세포당 GFP 통합
UNT = 비처리
오차 막대 = 기술적 재현실험의 표준 편차

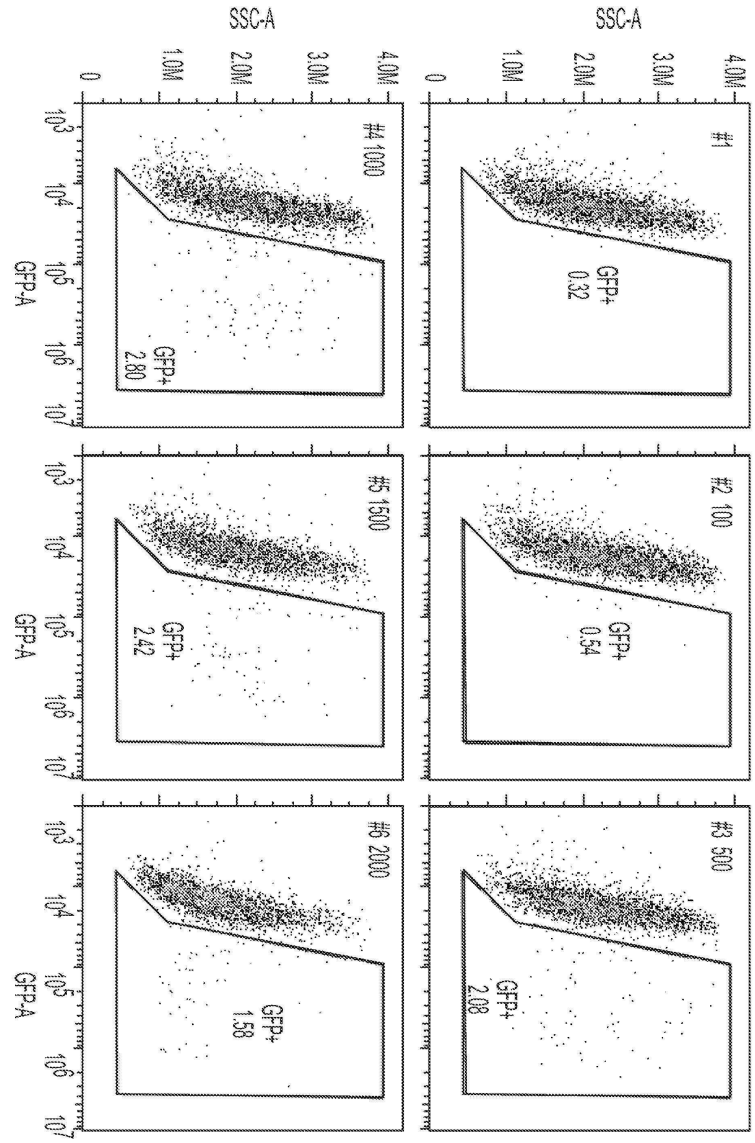
도면42


SSC-A
GFP-A
#1
GFP+
0.14
#2 100
GFP+
0.21
#2 500
GFP+
1.07
#4 1000
GFP+
1.59
#5 1500
GFP+
1.45
#6 2000
GFP+
2.87

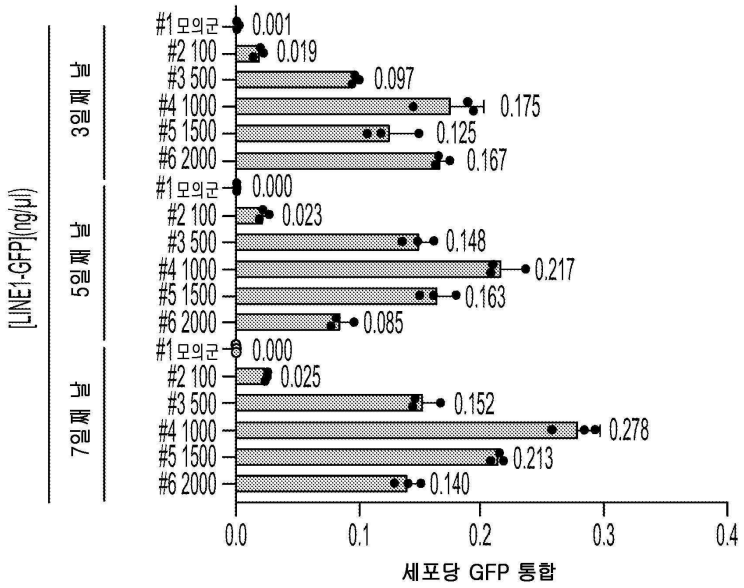
도면43

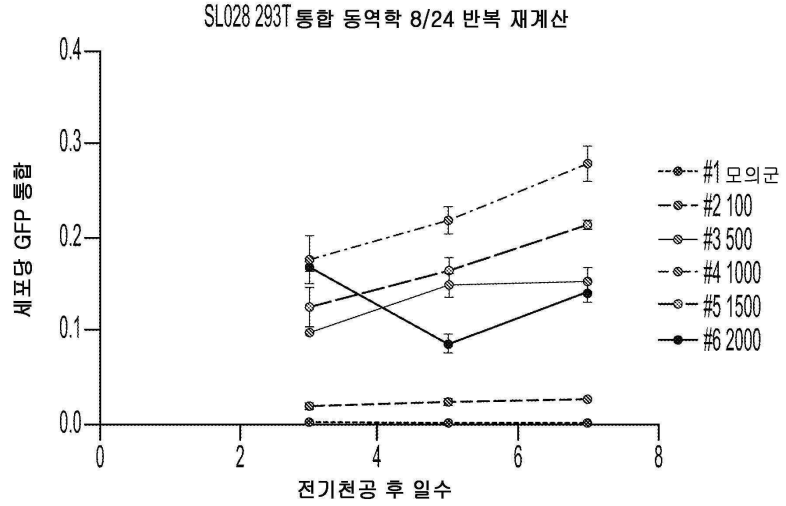

SSC-A
0
1.0M
2.0M
3.0M
4.0M
GFP-A
10^3
10^4
10^5
10^6
10^7
GFP+
0.17
GFP+
2.34
GFP+
0.29
GFP+
2.63
GFP+
1.43
GFP+
2.09

도면44

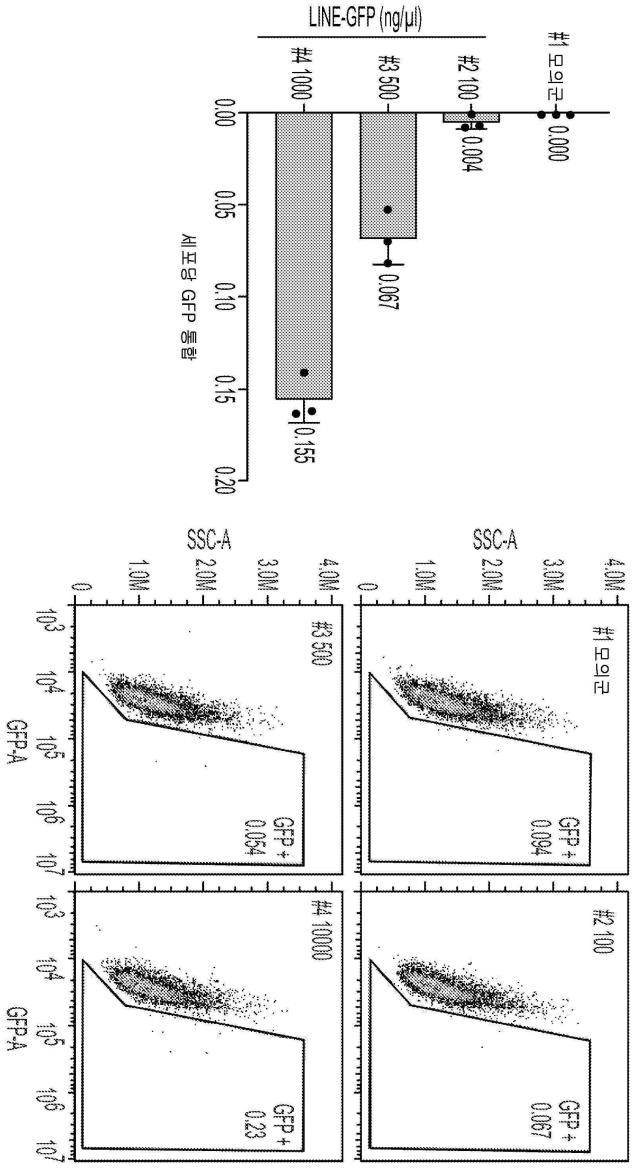

SSC-A
GFP-A
#1
GFP+
0.32
#2 100
GFP+
0.54
#3 500
GFP+
2.08
#4 1000
GFP+
2.80
#5 1500
GFP+
2.42
#6 2000
GFP+
1.58

도면45

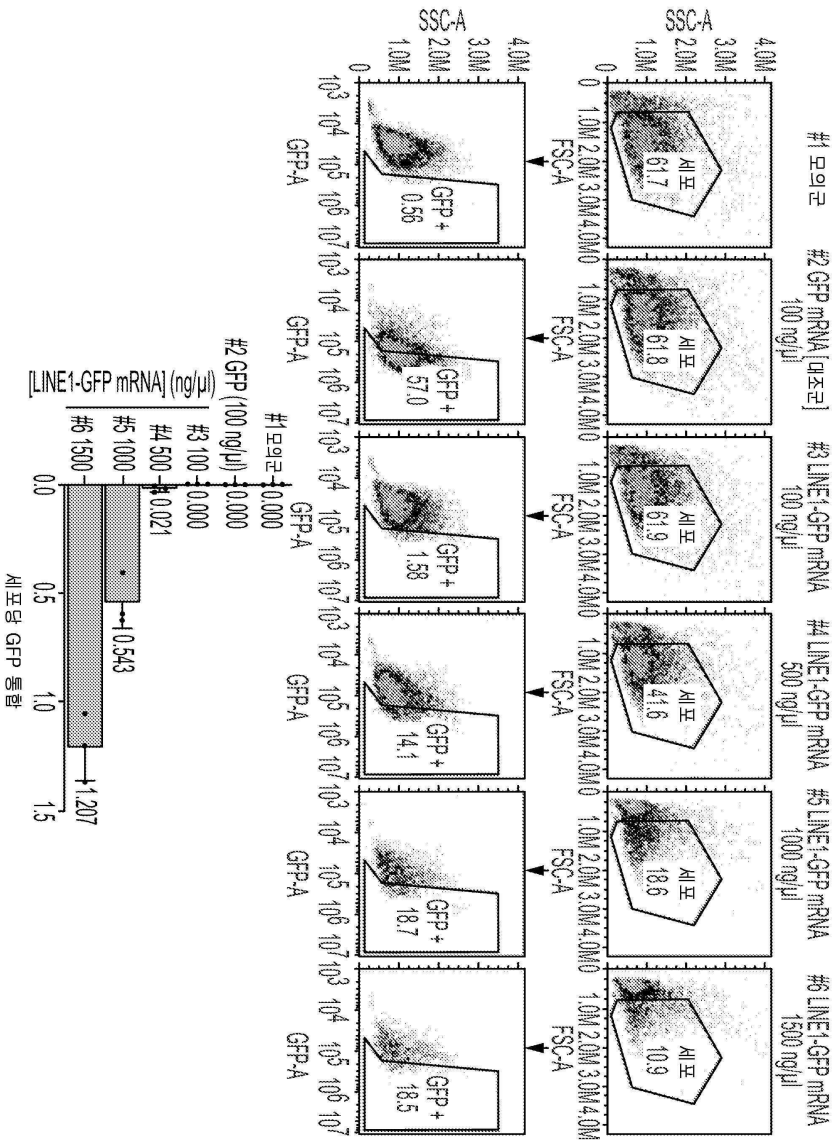

[LINE1-GFP](ng/μl)
3일째 날
5일째 날
7일째 날
#1 모의군 0.001
#2 100 0.019
#3 500 0.097
#4 1000 0.175
#5 1500 0.125
#6 2000 0.167
#1 모의군 0.000
#2 100 0.023
#3 500 0.148
#4 1000 0.217
#5 1500 0.163
#6 2000 0.085
#1 모의군 0.000
#2 100 0.025
#3 500 0.152
#4 1000 0.278
#5 1500 0.213
#6 2000 0.140
0.0
0.1
0.2
0.3
0.4
세포당 GFP 통합


SL028 293T 통합 동역학 8/24 반복 재계산
세포당 GFP 통합
0.0
0.1
0.2
0.3
0.4
0
2
4
6
8
전기천공 후 일수
#1 모의군
#2 100
#3 500
#4 1000
#5 1500
#6 2000

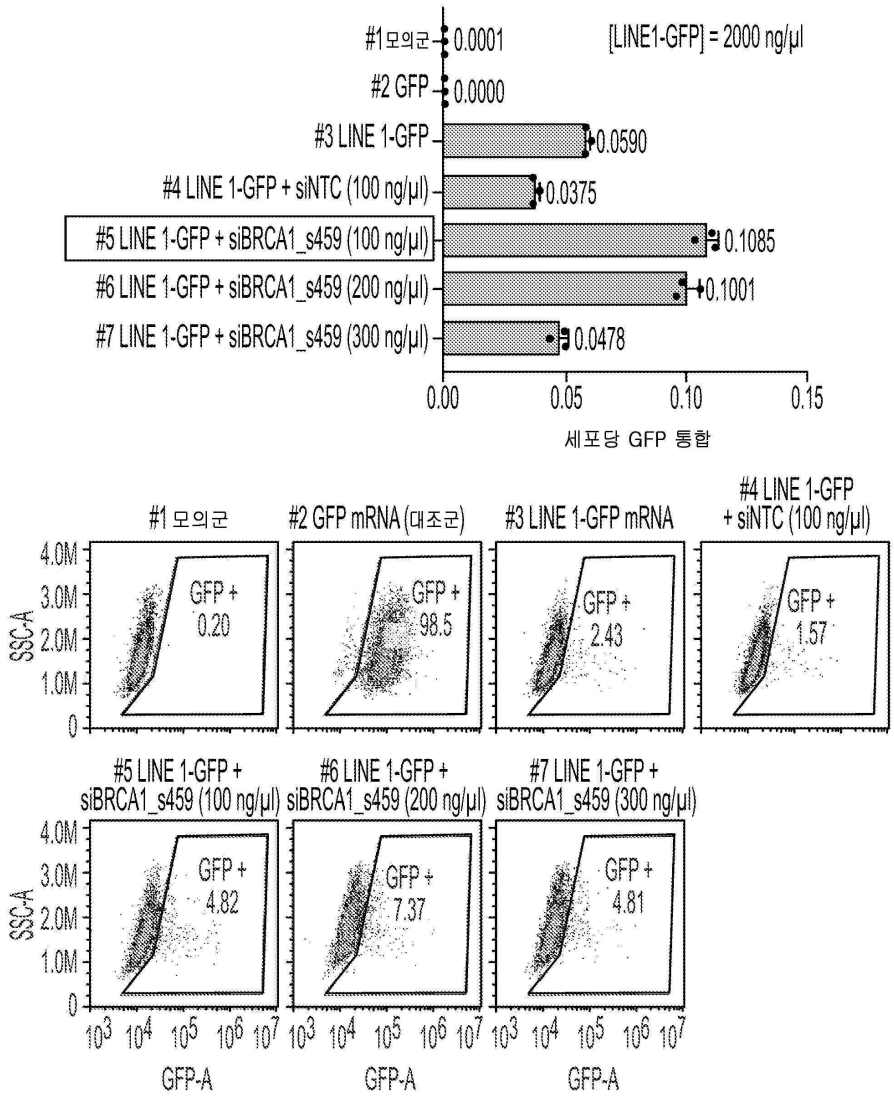
도면46


LINE-GFP (ng/μl)
#1 모의군
#2 100
#3 500
#4 1000
0.000
0.004
0.067
0.155
세포당 GFP 출현
0.00
0.05
0.10
0.15
0.20
SSC-A
GFP-A
#1 모의군
GFP+ 0.094
#2 100
GFP+ 0.067
#3 500
GFP+ 0.054
#4 10000
GFP+ 0.23

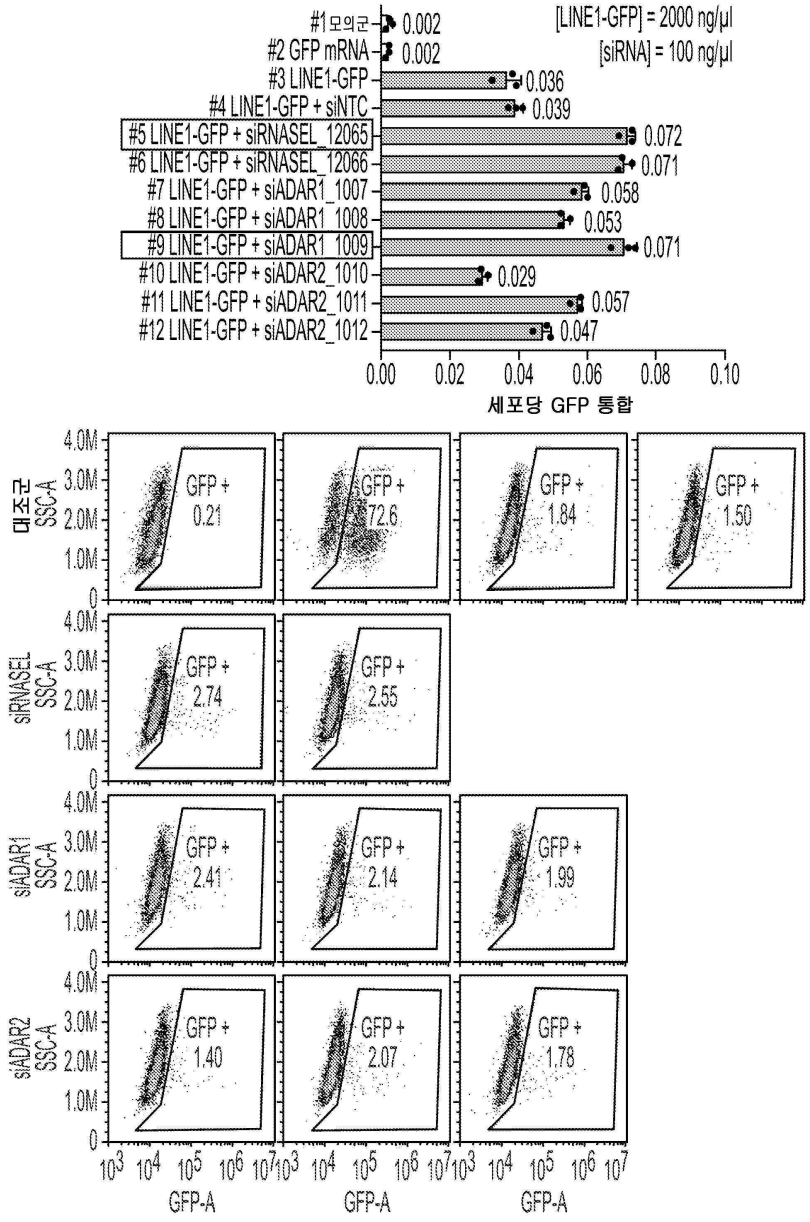
도면47


#1 모의군
#2 GFP mRNA [대조군] 100 ng/μl
#3 LINE1-GFP mRNA 100 ng/μl
#4 LINE1-GFP mRNA 500 ng/μl
#5 LINE1-GFP mRNA 1000 ng/μl
#6 LINE1-GFP mRNA 1500 ng/μl
SSC-A
FSC-A
GFP-A
세포
61.7
61.8
61.9
41.6
18.6
10.9
GFP+
0.56
57.0
1.58
14.1
18.7
18.5
[LINE1-GFP mRNA] (ng/μl)
#1 모의군 0.000
#2 GFP (100 ng/μl) 0.000
#3 100 0.000
#4 500 0.021
#5 1000 0.543
#6 1500 1.207
세포당 GFP 통합

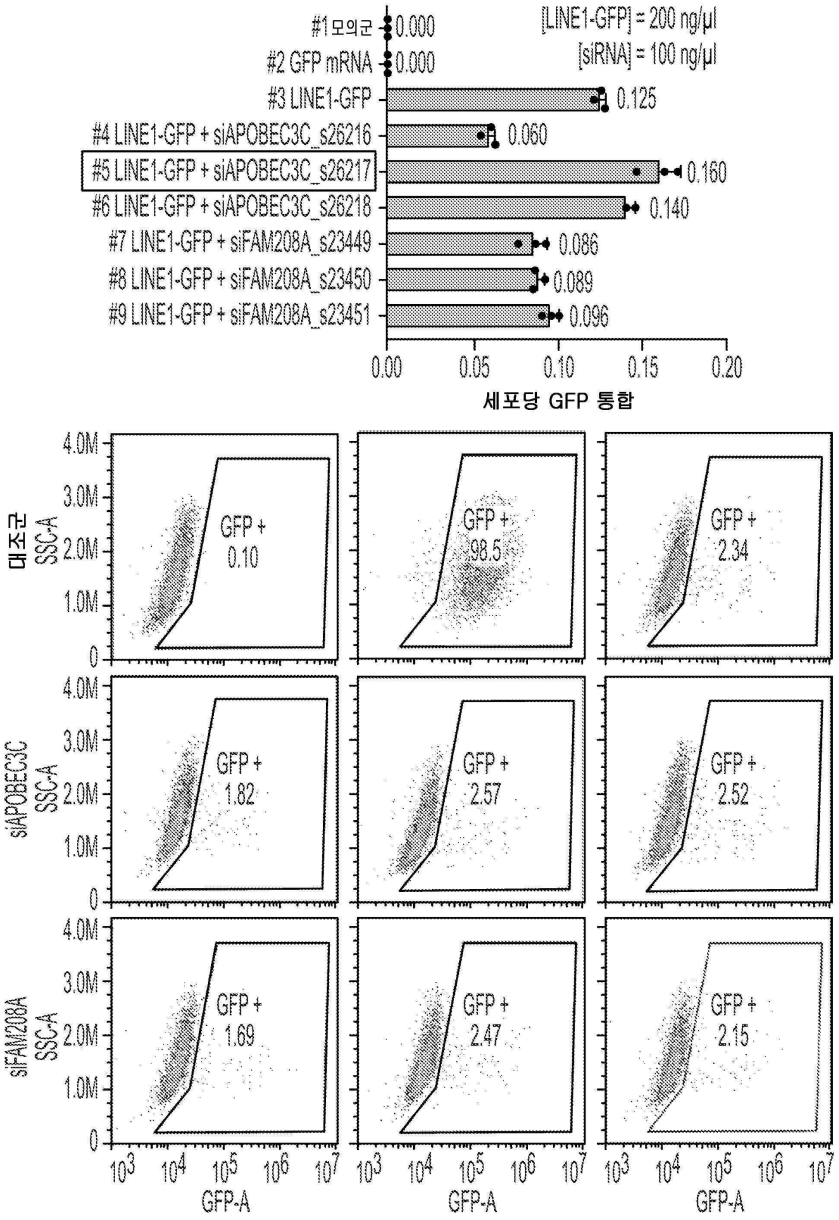
도면48


#1모의군
0.0001
[LINE1-GFP] = 2000 ng/μl
#2 GFP
0.0000
#3 LINE 1-GFP
0.0590
#4 LINE 1-GFP + siNTC (100 ng/μl)
0.0375
#5 LINE 1-GFP + siBRCA1_s459 (100 ng/μl)
0.1085
#6 LINE 1-GFP + siBRCA1_s459 (200 ng/μl)
0.1001
#7 LINE 1-GFP + siBRCA1_s459 (300 ng/μl)
0.0478
0.00
0.05
0.10
0.15
세포당 GFP 통합
#1 모의군
#2 GFP mRNA (대조군)
#3 LINE 1-GFP mRNA
#4 LINE 1-GFP + siNTC (100 ng/μl)
#5 LINE 1-GFP + siBRCA1_s459 (100 ng/μl)
#6 LINE 1-GFP + siBRCA1_s459 (200 ng/μl)
#7 LINE 1-GFP + siBRCA1_s459 (300 ng/μl)
SSC-A
4.0M
3.0M
2.0M
1.0M
0
GFP + 0.20
GFP + 98.5
GFP + 2.43
GFP + 1.57
GFP + 4.82
GFP + 7.37
GFP + 4.81
10^3 10^4 10^5 10^6 10^7
GFP-A

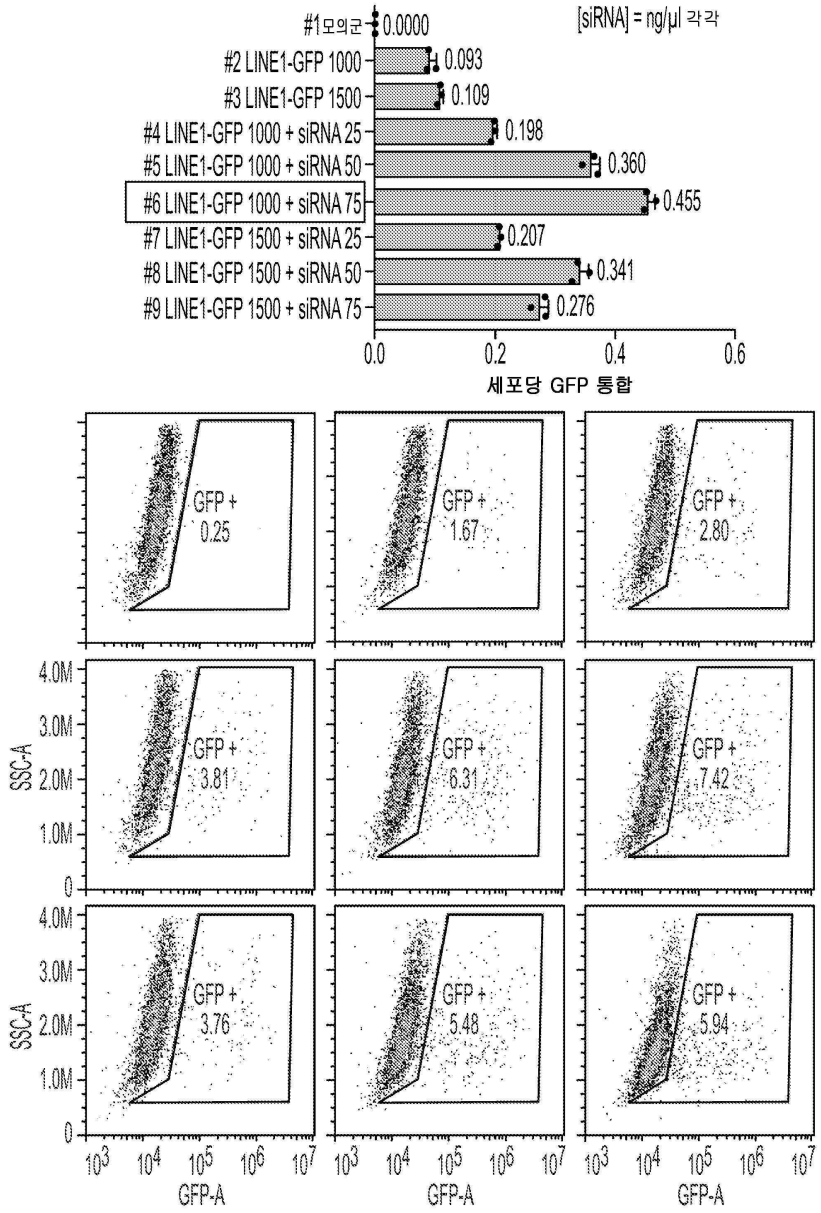
도면49


[LINE1-GFP] = 2000 ng/μl
[siRNA] = 100 ng/μl
#1 모의군 0.002
#2 GFP mRNA 0.002
#3 LINE1-GFP 0.036
#4 LINE1-GFP + siNTC 0.039
#5 LINE1-GFP + siRNASEL_12065 0.072
#6 LINE1-GFP + siRNASEL_12066 0.071
#7 LINE1-GFP + siADAR1_1007 0.058
#8 LINE1-GFP + siADAR1_1008 0.053
#9 LINE1-GFP + siADAR1_1009 0.071
#10 LINE1-GFP + siADAR2_1010 0.029
#11 LINE1-GFP + siADAR2_1011 0.057
#12 LINE1-GFP + siADAR2_1012 0.047
0.00 0.02 0.04 0.06 0.08 0.10
세포당 GFP 통합
대조군 SSC-A
siRNASEL SSC-A
siADAR1 SSC-A
siADAR2 SSC-A
GFP + 0.21
GFP + 72.6
GFP + 1.84
GFP + 1.50
GFP + 2.74
GFP + 2.55
GFP + 2.41
GFP + 2.14
GFP + 1.99
GFP + 1.40
GFP + 2.07
GFP + 1.78
GFP-A

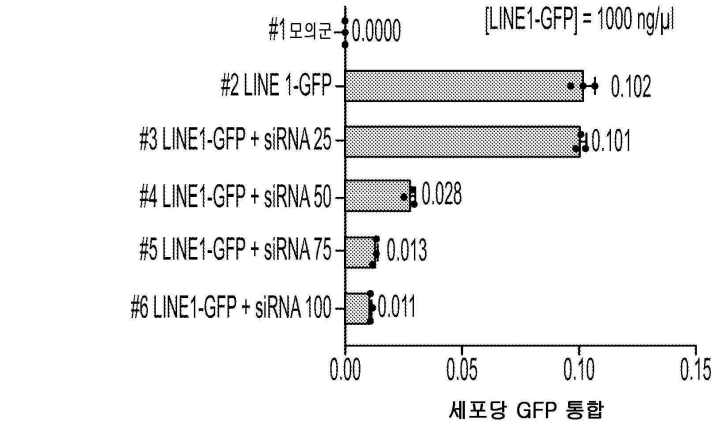
도면50

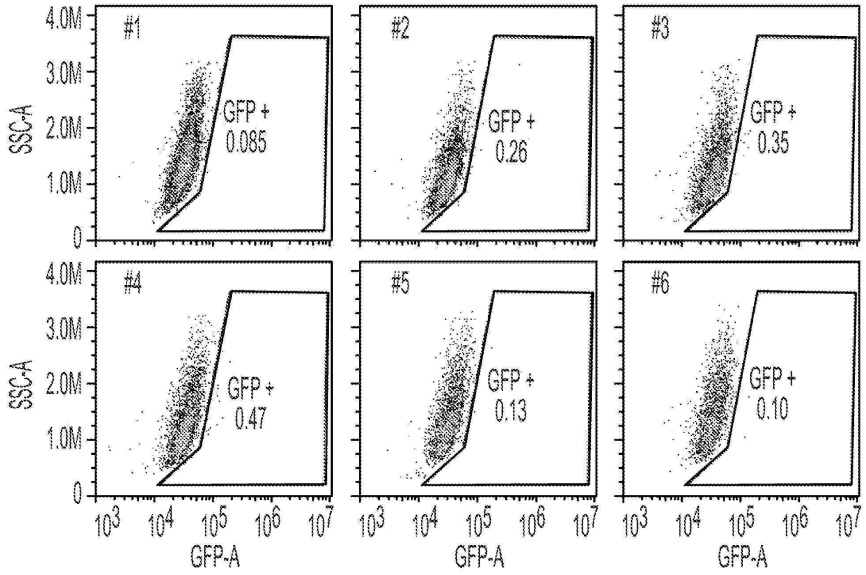

[LINE1-GFP] = 200 ng/μl
[siRNA] = 100 ng/μl
#1 모의군 0.000
#2 GFP mRNA 0.000
#3 LINE1-GFP 0.125
#4 LINE1-GFP + siAPOBEC3C_s26216 0.060
#5 LINE1-GFP + siAPOBEC3C_s26217 0.160
#6 LINE1-GFP + siAPOBEC3C_s26218 0.140
#7 LINE1-GFP + siFAM208A_s23449 0.086
#8 LINE1-GFP + siFAM208A_s23450 0.089
#9 LINE1-GFP + siFAM208A_s23451 0.096
0.00 0.05 0.10 0.15 0.20
세포당 GFP 통합
대조군
siAPOBEC3C
siFAM208A
SSC-A
GFP-A
GFP + 0.10
GFP + 98.5
GFP + 2.34
GFP + 1.82
GFP + 2.57
GFP + 2.52
GFP + 1.69
GFP + 2.47
GFP + 2.15

도면51

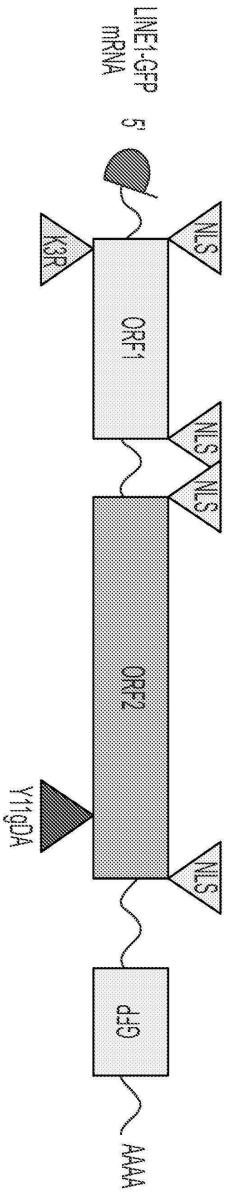

[siRNA] = ng/μl 각각
#1모의군 0.0000
#2 LINE1-GFP 1000 0.093
#3 LINE1-GFP 1500 0.109
#4 LINE1-GFP 1000 + siRNA 25 0.198
#5 LINE1-GFP 1000 + siRNA 50 0.360
#6 LINE1-GFP 1000 + siRNA 75 0.455
#7 LINE1-GFP 1500 + siRNA 25 0.207
#8 LINE1-GFP 1500 + siRNA 50 0.341
#9 LINE1-GFP 1500 + siRNA 75 0.276
0.0
0.2
0.4
0.6
세포당 GFP 통합
GFP + 0.25
GFP + 1.67
GFP + 2.80
GFP + 3.81
GFP + 6.31
GFP + 7.42
GFP + 3.76
GFP + 5.48
GFP + 5.94
SSC-A
4.0M
3.0M
2.0M
1.0M
0
10^3 10^4 10^5 10^6 10^7
GFP-A

도면52

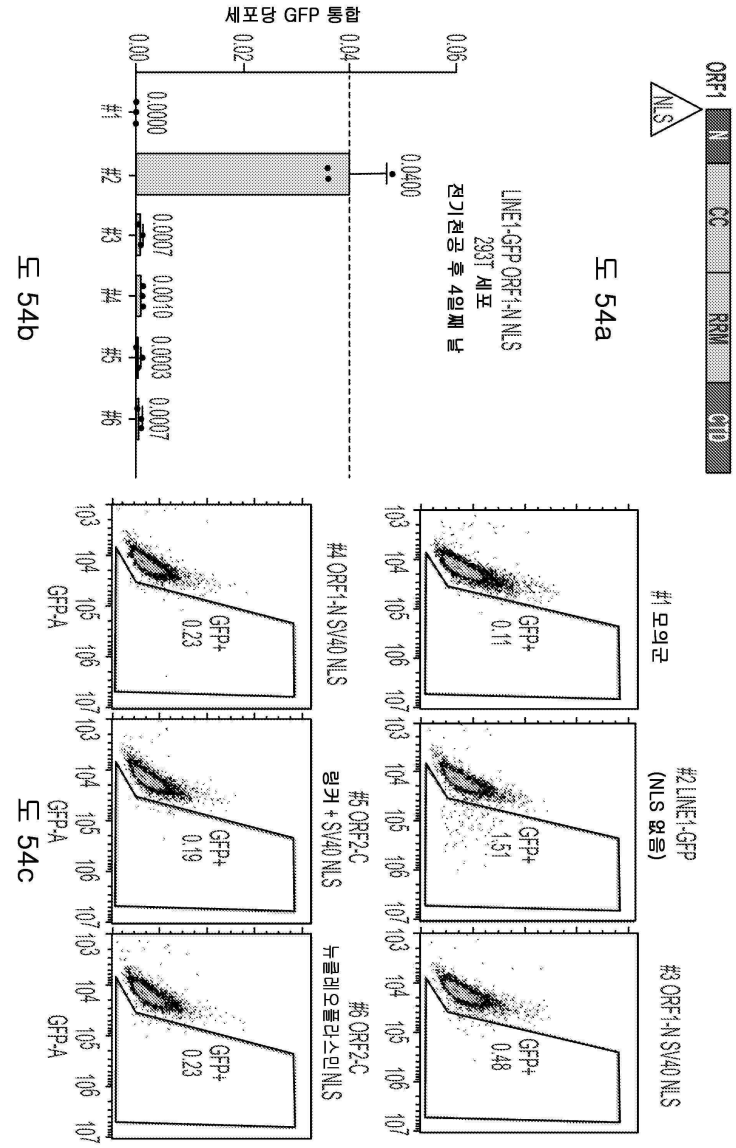

[LINE1-GFP] = 1000 ng/μl
#1 모의군
0.0000
#2 LINE 1-GFP
0.102
#3 LINE1-GFP + siRNA 25
0.101
#4 LINE1-GFP + siRNA 50
0.028
#5 LINE1-GFP + siRNA 75
0.013
#6 LINE1-GFP + siRNA 100
0.011
0.00
0.05
0.10
0.15
세포당 GFP 통합


#1
GFP +
0.085
#2
GFP +
0.26
#3
GFP +
0.35
#4
GFP +
0.47
#5
GFP +
0.13
#6
GFP +
0.10
SSC-A
4.0M
3.0M
2.0M
1.0M
0
GFP-A

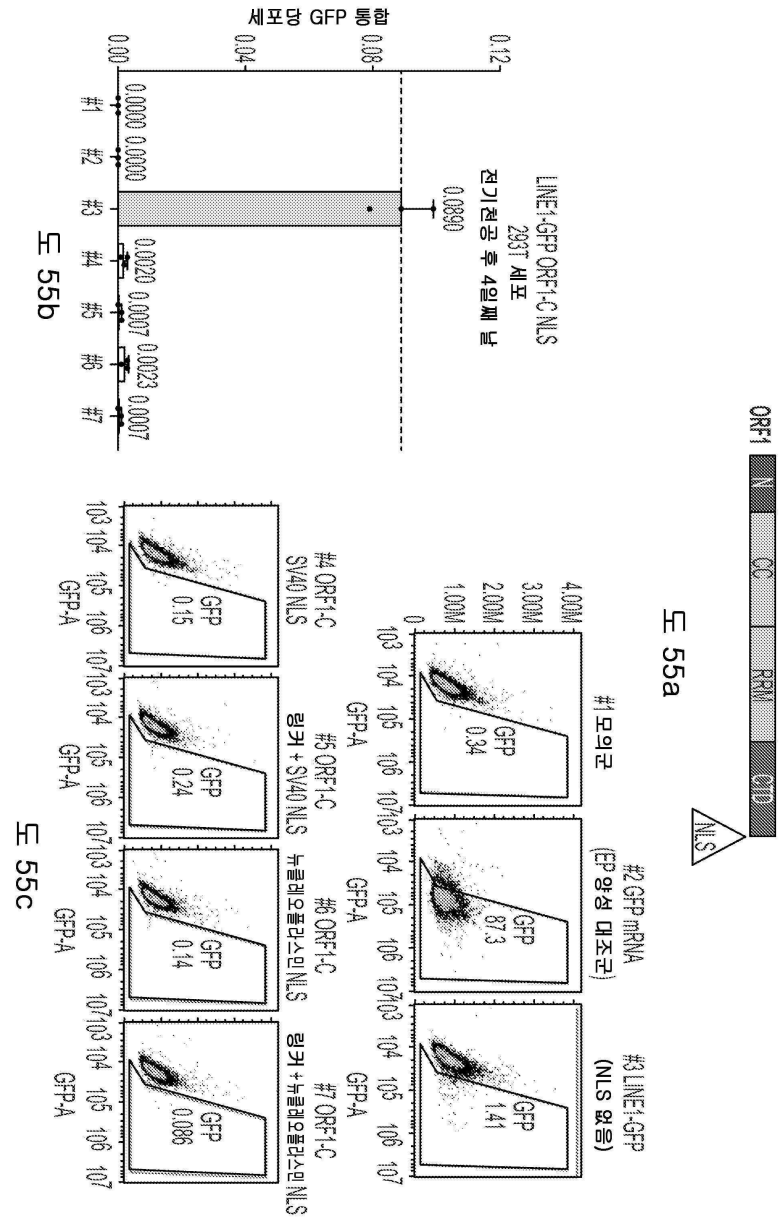
도면53



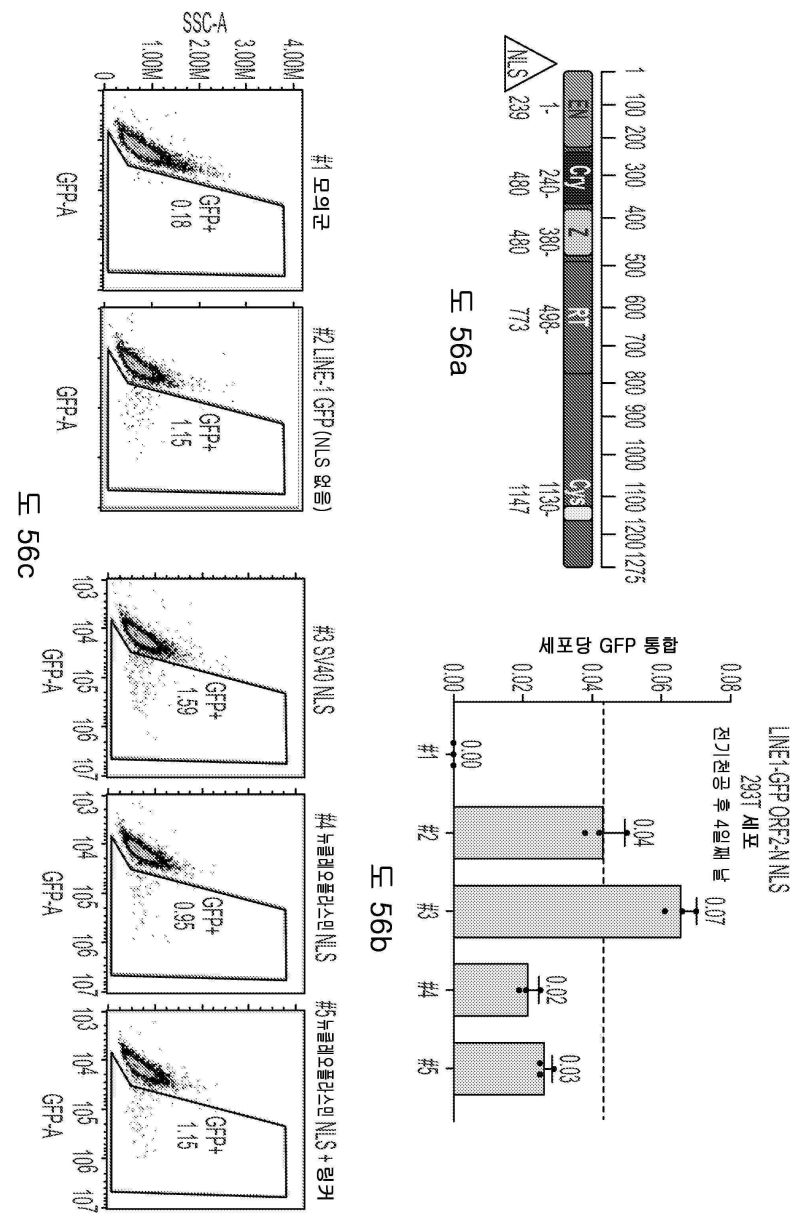
도면54


ORF1
NLS
N
CC
RRM
CTD
LINE1-GFP ORF1-N NLS
293T 세포
전기천공 후 4일째 날
세포당 GFP 통합
0.00
0.02
0.04
0.06
0.0000
0.0400
0.0007
0.0010
0.0003
0.0007
#1
#2
#3
#4
#5
#6
도 54a
도 54b
#1 모의군
#2 LINE1-GFP (NLS 없음)
#3 ORF1-N SV40 NLS
#4 ORF1-N SV40 NLS
#5 ORF2-C 링커 + SV40 NLS
#6 ORF2-C 뉴클레오플라스민 NLS
GFP+
0.11
1.51
0.48
0.23
0.19
0.23
GFP-A
도 54c

도면55


세포당 GFP 통합
0.12
0.08
0.04
0.00
#1 #2 #3 #4 #5 #6 #7
0.0000 0.0000 0.0890 0.0020 0.0007 0.0023 0.0007
LINE1-GFP ORF1-C-NLS
293T 세포
전기천공 후 4일째 날
도 55b
ORF1
N
CC
RRM
CTD
NLS
도 55a
#1 모의군
#2 GFP mRNA
(EP 양성 대조군)
#3 LINE1-GFP
(NLS 없음)
GFP 0.34
GFP 87.3
GFP 1.41
#4 ORF1-C
SV40 NLS
#5 ORF1-C
링커 + SV40 NLS
#6 ORF1-C
뉴클레오플라스민 NLS
#7 ORF1-C
링커 + 뉴클레오플라스민 NLS
GFP 0.15
GFP 0.24
GFP 0.14
GFP 0.086
GFP-A
도 55c

도면56

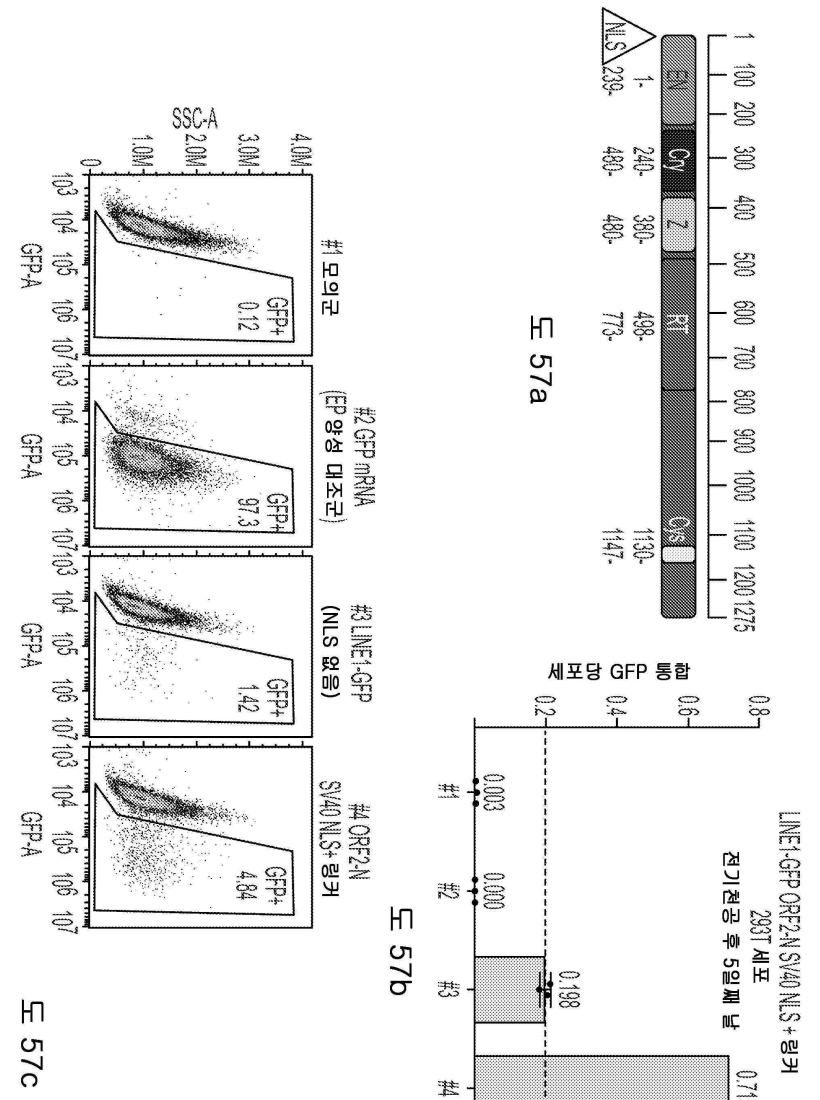


도 56c

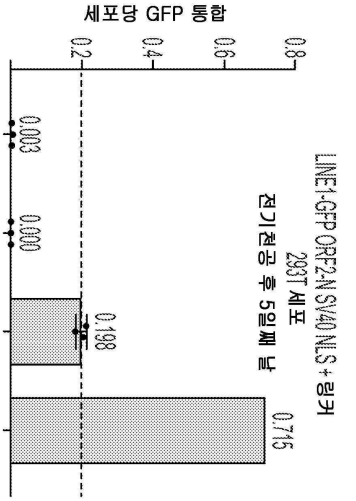
도 56b

도 56a

도면57



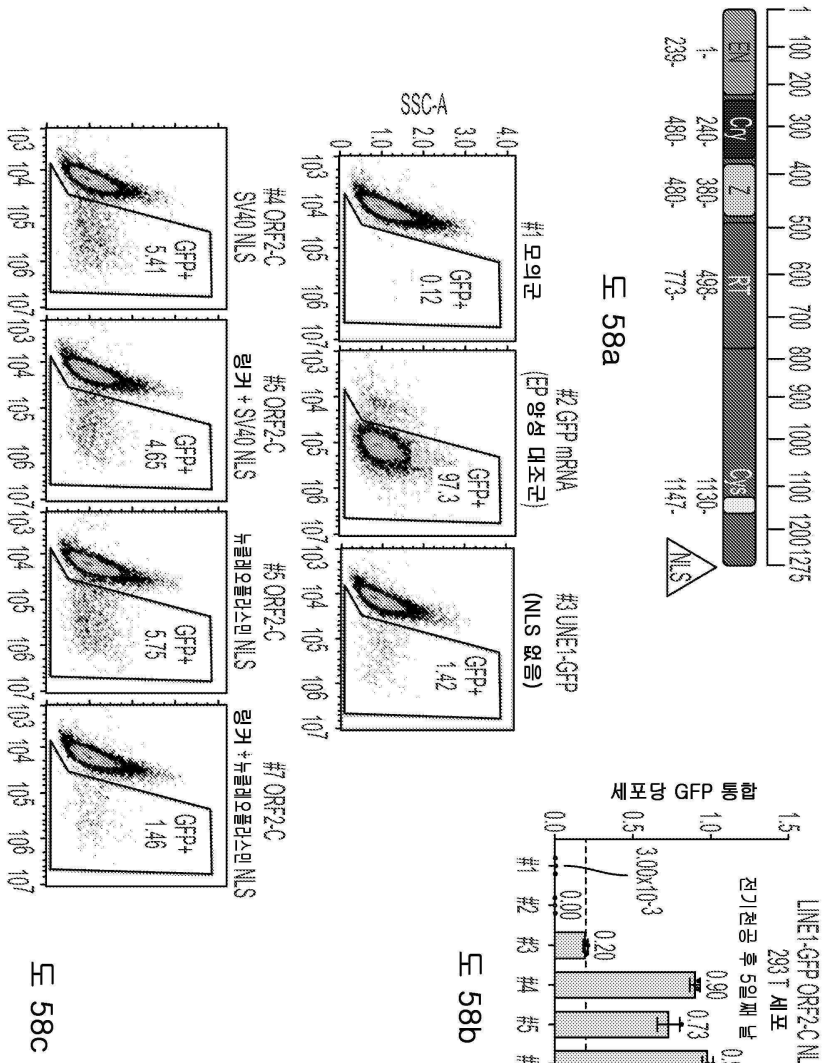
도 57a



도 57b

도 57c

도면58



서 열 목 록

SEQUENCE LISTING

<110> MYELOID THERAPEUTICS, INC.

<120> METHODS AND COMPOSITIONS FOR GENOMIC INTEGRATION

<130> 56371-723.601

<140> PCT/US2022/028831

<141> 2022-05-11

<150> 63/274,907

<151> 2021-11-02

<150> 63/254,791

<151> 2021-10-12

<150> 63/187,117

<151> 2021-05-11

<160> 113

<170> PatentIn version 3.5

<210> 1

<211> 116

<212> PRT

<213> Artificial Sequence

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

<400> 1

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

1 5 10 15

Ser Val Arg Ile Ser Cys Ala Ala Ser Gly Tyr Thr Phe Thr Asn Tyr

20 25 30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Met

35 40 45

Gly Trp Ile Asn Thr His Thr Gly Glu Pro Thr Tyr Ala Asp Ser Phe

50 55 60

Lys Gly Arg Phe Thr Phe Ser Leu Asp Asp Ser Lys Asn Thr Ala Tyr

65 70 75 80

Leu Gln Ile Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Phe Cys

85 90 95

Thr Arg Arg Gly Tyr Asp Trp Tyr Phe Asp Val Trp Gly Gln Gly Thr

100 105 110

Thr Val Thr Val

115

<210> 2

<211> 107

<212> PRT

<213> Artificial Sequence

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

<400

> 2

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

1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Ile Asn Ser Tyr

20 25 30

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

35 40 45

Tyr Arg Ala Asn Arg Leu Glu Ser Gly Val Pro Ser Arg Phe Ser Gly

50 55 60

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

65 70 75 80

Glu Asp Phe Gly Ile Tyr Tyr Cys Gln Gln Tyr Asp Glu Ser Pro Trp

85 90 95

Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys

100 105

<210> 3

<211> 46

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 3

Leu Tyr Cys Arg Arg Leu Lys Ile Gln Val Arg Lys Ala Ala Ile Thr

1 5 10 15

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

20 25 30

Gln Glu Thr Tyr Glu Thr Leu Lys His Glu Lys Pro Pro Gln

35 40 45

<210> 4

<211> 35

<212> PRT

<213> Artificial Sequence

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

<400> 4

Tyr Glu Asp Met Arg Gly Ile Leu Tyr Ala Ala Pro Gln Leu Arg Ser

1 5 10 15

Ile Arg Gly Gln Pro Gly Pro Asn His Glu Glu Asp Ala Asp Ser Tyr

20 25 30

Glu Asn Met

35

<210> 5

<211> 62

<212> PRT

<213> Artificial Sequence

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

<400> 5

Lys Lys Val Ala Lys Lys Pro Thr Asn Lys Ala Pro His Pro Lys Gln

1 5 10 15

Glu Pro Gln Glu Ile Asn Phe Pro Asp Asp Leu Pro Gly Ser Asn Thr

20 25 30

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

35 40 45

Glu Asp Gly Lys Glu Ser Arg Ile Ser Val Gln Glu Arg Gln

50 55 60

<210> 6

<211> 21

<212> PRT

<213> Artificial Sequence

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

<400> 6

Ile Tyr Ile Trp Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu

1 5 10 15

Ser Leu Val Ile Thr

20

<210> 7

<211> 62

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 7

Ala Leu Ser Asn Ser Ile Met Tyr Phe Ser His Phe Val Pro Val Phe

1 5 10 15

Leu Pro Ala Lys Pro Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro

20 25 30

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

35 40 45

Arg Pro Ala Ala Gly Gly Ala Val His Thr Arg Gly Leu Asp

50 55 60

<210> 8

<211> 130

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 8

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

1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Val Asn Thr Ala

20 25 30

Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile

35 40 45

Tyr Ser Ala Ser Phe Leu Tyr Ser Gly Val Pro Ser Arg Phe Ser Gly

50 55 60

Ser Arg Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro Pro
85 90 95
Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Gly Ser Thr
100 105 110
Ser Gly Ser Gly Lys Pro Gly Ser Gly Glu Gly Ser Glu Val Gln Leu
115 120 125
Val Glu
130

<210> 9
<211> 108
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 9
Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly
1 5 10 15
Phe Asn Ile Lys Asp Thr Tyr Ile His Trp Val Arg Gln Ala Pro Gly
20 25 30
Lys Gly Leu Glu Trp Val Ala Arg Ile Tyr Pro Thr Asn Gly Tyr Thr
35 40 45
Arg Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Ala Asp Thr
50 55 60

Ser Lys Asn Thr Ala Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp
65 70 75 80
Thr Ala Val Tyr Tyr Cys Ser Arg Trp Gly Gly Asp Gly Phe Tyr Ala
85 90 95
Met Asp Val Trp Gly Gln Gly Thr Leu Val Thr Val
100 105

<210> 10

<211> 17

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 10

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

1 5 10 15

Ser

<210> 11

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 11

Ser Gly Gly Gly Gly Ser Gly

1 5

<210> 12

<211> 4

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 12

Ser Gly Gly Gly

1

<210> 13

<211> 4

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 13

Gly Ser Gly Ser

1

<210> 14

<211> 432

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 14

Met Trp Leu Gln Ser Leu Leu Leu Leu Gly Thr Val Ala Cys Ser Ile

1 5 10 15

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

20 25 30

Gly Ser Val Arg Ile Ser Cys Ala Ala Ser Gly Tyr Thr Phe Thr Asn

35 40 45

Tyr Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp

50 55 60

Met Gly Trp Ile Asn Thr His Thr Gly Glu Pro Thr Tyr Ala Asp Ser

65 70 75 80

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

85 90 95

Tyr Leu Gln Ile Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Phe

100 105 110

Cys Thr Arg Arg Gly Tyr Asp Trp Tyr Phe Asp Val Trp Gly Gln Gly

115 120 125

Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly

130 135 140

Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser

145 150 155 160

Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser

165 170 175

Gln Asp Ile Asn Ser Tyr Leu Ser Trp Phe Gln Gln Lys Pro Gly Lys

180 185 190

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

195 200 205

Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Thr Leu Thr

210 215 220

Ile Ser Ser Leu Gln Tyr Glu Asp Phe Gly Ile Tyr Tyr Cys Gln Gln

225 230 235 240

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

245 250 255

Lys Ser Gly Gly Gly Gly Ser Gly Ala Leu Ser Asn Ser Ile Met Tyr

260 265 270

Phe Ser His Phe Val Pro Val Phe Leu Pro Ala Lys Pro Thr Thr Thr

275 280 285

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

290 295 300

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

305 310 315 320

His Thr Arg Gly Leu Asp Ile Tyr Ile Trp Ala Pro Leu Ala Gly Thr

325 330 335

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

340 345 350

Leu Lys Ile Gln Val Arg Lys Ala Ala Ile Thr Ser Tyr Glu Lys Ser

355 360 365

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

370 375 380

Thr Leu Lys His Glu Lys Pro Pro Gln Gly Ser Gly Ser Tyr Glu Asp

385 390 395 400

Met Arg Gly Ile Leu Tyr Ala Ala Pro Gln Leu Arg Ser Ile Arg Gly

405 410 415

Gln Pro Gly Pro Asn His Glu Glu Asp Ala Asp Ser Tyr Glu Asn Met
420 425 430

<210> 15

<211> 436

<212> PRT

<213> Artificial Sequence

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

<400> 15

Met Trp Leu Gln Ser Leu Leu Leu Leu Gly Thr Val Ala Cys Ser Ile
1 5 10 15
Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val
20 25 30
Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Val Asn Thr
35 40 45
Ala Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu
50 55 60
Ile Tyr Ser Ala Ser Phe Leu Tyr Ser Gly Val Pro Ser Arg Phe Ser
65 70 75 80
Gly Ser Arg Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln
85 90 95
Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro
100 105 110
Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Gly Ser
115 120 125
Thr Ser Gly Ser Gly Lys Pro Gly Ser Gly Glu Gly Ser Glu Val Gln
130 135 140
Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly Ser Leu Arg
145 150 155 160
Leu Ser Cys Ala Ala Ser Gly Phe Asn Ile Lys Asp Thr Tyr Ile His
165 170 175
Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ala Arg Ile

180 185 190

Tyr Pro Thr Asn Gly Tyr Thr Arg Tyr Ala Asp Ser Val Lys Gly Arg

195 200 205

Phe Thr Ile Ser Ala Asp Thr Ser Lys Asn Thr Ala Tyr Leu Gln Met

210 215 220

Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ser Arg Trp

225 230 235 240

Gly Gly Asp Gly Phe Tyr Ala Met Asp Val Trp Gly Gln Gly Thr Leu

245 250 255

Val Thr Val Ser Ser Ser Gly Gly Gly Gly Ser Gly Ala Leu Ser Asn

260 265 270

Ser Ile Met Tyr Phe Ser His Phe Val Pro Val Phe Leu Pro Ala Lys

275 280 285

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

290 295 300

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

305 310 315 320

Gly Gly Ala Val His Thr Arg Gly Leu Asp Ile Tyr Ile Trp Ala Pro

325 330 335

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

340 345 350

Tyr Cys Arg Arg Leu Lys Ile Gln Val Arg Lys Ala Ala Ile Thr Ser

355 360 365

Tyr Glu Lys Ser Asp Gly Val Tyr Thr Gly Leu Ser Thr Arg Asn Gln

370 375 380

Glu Thr Tyr Glu Thr Leu Lys His Glu Lys Pro Pro Gln Gly Ser Gly

385 390 395 400

Ser Tyr Glu Asp Met Arg Gly Ile Leu Tyr Ala Ala Pro Gln Leu Arg

405 410 415

Ser Ile Arg Gly Gln Pro Gly Pro Asn His Glu Glu Asp Ala Asp Ser

420 425 430

Tyr Glu Asn Met
435

<210> 16

<211> 454

<212> PRT

<213> Artificial Sequence

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

<400> 16

Met Trp Leu Gln Ser Leu Leu Leu Gly Thr Val Ala Cys Ser Ile
1 5 10 15

Ser Glu Ile Gln Leu Val Gln Ser Gly Gly Gly Leu Val Lys Pro Gly
20 25 30

Gly Ser Val Arg Ile Ser Cys Ala Ala Ser Gly Tyr Thr Phe Thr Asn
35 40 45

Tyr Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp
50 55 60

Met Gly Trp Ile Asn Thr His Thr Gly Glu Pro Thr Tyr Ala Asp Ser
65 70 75 80

Phe Lys Gly Arg Phe Thr Phe Ser Leu Asp Asp Ser Lys Asn Thr Ala
85 90 95

Tyr Leu Gln Ile Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Phe
100 105 110

Cys Thr Arg Arg Gly Tyr Asp Trp Tyr Phe Asp Val Trp Gly Gln Gly
115 120 125

Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
130 135 140

Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser
145 150 155 160

Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser
165 170 175

Gln Asp Ile Asn Ser Tyr Leu Ser Trp Phe Gln Gln Lys Pro Gly Lys

180 185 190

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

195 200 205

Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Thr Leu Thr

210 215 220

Ile Ser Ser Leu Gln Tyr Glu Asp Phe Gly Ile Tyr Tyr Cys Gln Gln

225 230 235 240

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

245 250 255

Lys Ser Gly Gly Gly Gly Ser Gly Ala Leu Ser Asn Ser Ile Met Tyr

260 265 270

Phe Ser His Phe Val Pro Val Phe Leu Pro Ala Lys Pro Thr Thr Thr

275 280 285

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

290 295 300

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

305 310 315 320

His Thr Arg Gly Leu Asp Ile Tyr Ile Trp Ala Pro Leu Ala Gly Thr

325 330 335

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

340 345 350

Lys Ile Gln Val Arg Lys Ala Ala Ile Thr Ser Tyr Glu Lys Ser Asp

355 360 365

Gly Val Tyr Thr Gly Leu Ser Thr Arg Asn Gln Glu Thr Tyr Glu Thr

370 375 380

Leu Lys His Glu Lys Pro Pro Gln Lys Lys Val Ala Lys Lys Pro Thr

385 390 395 400

Asn Lys Ala Pro His Pro Lys Gln Glu Pro Gln Glu Ile Asn Phe Pro

405 410 415

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

420 425 430

His Gly Cys Gln Pro Val Thr Gln Glu Asp Gly Lys Glu Ser Arg Ile

435 440 445

Ser Val Gln Glu Arg Gln

450

<210> 17

<211> 17

<212> PRT

<213> Artificial Sequence

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

<400> 17

Met Trp Leu Gln Ser Leu Leu Leu Leu Gly Thr Val Ala Cys Ser Ile

1 5 10 15

Ser

<210> 18

<211> 27

<212> PRT

<213> Artificial Sequence

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

<400> 18

Phe Trp Val Leu Val Val Val Gly Gly Val Leu Ala Cys Tyr Ser Leu

1 5 10 15

Leu Val Thr Val Ala Phe Ile Ile Phe Trp Val

20 25

<210> 19

<211> 25

<212> PRT

<213> Artificial Sequence

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

<400> 19

Ile Leu Leu Pro Leu Ile Ile Gly Leu Ile Leu Leu Gly Leu Leu Ala
1 5 10 15
Leu Val Leu Ile Ala Phe Cys Ile Ile
20 25

<210> 20
<211> 45
<212> PRT
<213> Artificial Sequence

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

<400> 20
Leu Tyr Cys Arg Leu Lys Ile Gln Val Arg Lys Ala Ala Ile Thr Ser
1 5 10 15
Tyr Glu Lys Ser Asp Gly Val Tyr Thr Gly Leu Ser Thr Arg Asn Gln
20 25 30
Glu Thr Tyr Glu Thr Leu Lys His Glu Lys Pro Pro Gln
35 40 45

<210> 21
<211> 219
<212> PRT
<213> Artificial Sequence

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

<400> 21
Gln Arg Trp Lys Ser Lys Leu Tyr Ser Ile Val Cys Gly Lys Ser Thr
1 5 10 15
Pro Glu Lys Glu Gly Glu Leu Glu Gly Thr Thr Thr Lys Pro Leu Ala
20 25 30
Pro Asn Pro Ser Phe Ser Pro Thr Pro Gly Phe Thr Pro Thr Leu Gly
35 40 45
Phe Ser Pro Val Pro Ser Ser Thr Phe Thr Ser Ser Ser Thr Tyr Thr

50 55 60

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

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

180 185 190
Asp Leu Leu Gly Cys Leu Glu Asp Ile Glu Glu Ala Leu Cys Gly Pro
195 200 205
Ala Ala Leu Pro Pro Ala Pro Ser Leu Leu Arg
210 215
<210> 22
<211> 171
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 22
Pro Leu Cys Leu Gln Arg Glu Ala Lys Val Pro His Leu Pro Ala Asp
1 5 10 15

Lys Ala Arg Gly Thr Gln Gly Pro Glu Gln Gln His Leu Leu Ile Thr
20 25 30
Ala Pro Ser Ser Ser Ser Ser Ser Leu Glu Ser Ser Ala Ser Ala Leu

35 40 45

Asp Arg Arg Ala Pro Thr Arg Asn Gln Pro Gln Ala Pro Gly Val Glu

50 55 60

Ala Ser Gly Ala Gly Glu Ala Arg Ala Ser Thr Gly Ser Ser Asp Ser

65 70 75 80

Ser Pro Gly Gly His Gly Thr Gln Val Asn Val Thr Cys Ile Val Asn

85 90 95

Val Cys Ser Ser Ser Asp His Ser Ser Gln Cys Ser Ser Gln Ala Ser

100 105 110

Ser Thr Met Gly Asp Thr Asp Ser Ser Pro Ser Glu Ser Pro Lys Asp

115 120 125

Glu Gln Val Pro Phe Ser Lys Glu Glu Cys Ala Phe Arg Ser Gln Leu

130 135 140

Glu Thr Pro Glu Thr Leu Leu Gly Ser Thr Glu Glu Lys Pro Leu Pro

145 150 155 160

Leu Gly Val Pro Asp Ala Gly Met Lys Pro Ser

165 170

<210> 23

<211> 211

<212> PRT

<213> Artificial Sequence

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

<400> 23

Met Ser Asn Gly Tyr Ser Thr Asp Glu Asn Phe Arg Tyr Leu Ile Ser

1 5 10 15

Cys Phe Arg Ala Arg Val Lys Met Tyr Ile Gln Val Glu Pro Val Leu

20 25 30

Asp Tyr Leu Thr Phe Leu Pro Ala Glu Val Lys Glu Gln Ile Gln Arg

35 40 45

Thr Val Ala Thr Ser Gly Asn Met Gln Ala Val Glu Leu Leu Leu Ser

50 55 60

Thr Leu Glu Lys Gly Val Trp His Leu Gly Trp Thr Arg Glu Phe Val
65 70 75 80
Glu Ala Leu Arg Arg Thr Gly Ser Pro Leu Ala Ala Arg Tyr Met Asn

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

145 150 155 160
Glu Ser Gly Val Arg Glu Leu Leu Lys Arg Ile Val Gln Lys Glu Asn
165 170 175
Trp Phe Ser Ala Phe Leu Asn Val Leu Arg Gln Thr Gly Asn Asn Glu
180 185 190
Leu Val Gln Glu Leu Thr Gly Ser Asp Cys Ser Glu Ser Asn Ala Glu
195 200 205
Ile Glu Asn
210
<210> 24
<211> 607

<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 24
Met Trp Leu Gln Ser Leu Leu Leu Leu Gly Thr Val Ala Cys Ser Ile
1 5 10 15
Ser Glu Ile Gln Leu Val Gln Ser Gly Gly Gly Leu Val Lys Pro Gly
20 25 30
Gly Ser Val Arg Ile Ser Cys Ala Ala Ser Gly Tyr Thr Phe Thr Asn

35 40 45

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

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

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

Tyr Asp Glu Ser Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile
245 250 255
Lys Ser Gly Gly Gly Gly Ser Gly Ala Leu Ser Asn Ser Ile Met Tyr
260 265 270
Phe Ser His Phe Val Pro Val Phe Leu Pro Ala Lys Pro Thr Thr Thr
275 280 285

Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala Ser Gln Pro
290 295 300

Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly Gly Ala Val
305 310 315 320
His Thr Arg Gly Leu Asp Ile Tyr Ile Trp Ala Pro Leu Ala Gly Thr
325 330 335
Cys Gly Val Leu Leu Leu Ser Leu Val Ile Thr Leu Tyr Cys Arg Leu
340 345 350
Lys Ile Gln Val Arg Lys Ala Ala Ile Thr Ser Tyr Glu Lys Ser Asp
355 360 365

Gly Val Tyr Thr Gly Leu Ser Thr Arg Asn Gln Glu Thr Tyr Glu Thr
370 375 380
Leu Lys His Glu Lys Pro Pro Gln Gly Ser Gly Ser Met Ser Asn Gly
385 390 395 400
Tyr Ser Thr Asp Glu Asn Phe Arg Tyr Leu Ile Ser Cys Phe Arg Ala
405 410 415
Arg Val Lys Met Tyr Ile Gln Val Glu Pro Val Leu Asp Tyr Leu Thr
420 425 430

Phe Leu Pro Ala Glu Val Lys Glu Gln Ile Gln Arg Thr Val Ala Thr
435 440 445
Ser Gly Asn Met Gln Ala Val Glu Leu Leu Leu Ser Thr Leu Glu Lys
450 455 460
Gly Val Trp His Leu Gly Trp Thr Arg Glu Phe Val Glu Ala Leu Arg
465 470 475 480
Arg Thr Gly Ser Pro Leu Ala Ala Arg Tyr Met Asn Pro Glu Leu Thr
485 490 495

Asp Leu Pro Ser Pro Ser Phe Glu Asn Ala His Asp Glu Tyr Leu Gln
500 505 510
Leu Leu Asn Leu Leu Gln Pro Thr Leu Val Asp Lys Leu Leu Val Arg
515 520 525
Asp Val Leu Asp Lys Cys Met Glu Glu Glu Leu Leu Thr Ile Glu Asp

530 535 540

Arg Asn Arg Ile Ala Ala Ala Glu Asn Asn Gly Asn Glu Ser Gly Val
545 550 555 560

Arg Glu Leu Leu Lys Arg Ile Val Gln Lys Glu Asn Trp Phe Ser Ala
565 570 575

Phe Leu Asn Val Leu Arg Gln Thr Gly Asn Asn Glu Leu Val Gln Glu
580 585 590

Leu Thr Gly Ser Asp Cys Ser Glu Ser Asn Ala Glu Ile Glu Asn
595 600 605

<210> 25

<211> 615

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 25

Met Trp Leu Gln Ser Leu Leu Leu Leu Gly Thr Val Ala Cys Ser Ile
1 5 10 15

Ser Glu Ile Gln Leu Val Gln Ser Gly Gly Gly Leu Val Lys Pro Gly
20 25 30

Gly Ser Val Arg Ile Ser Cys Ala Ala Ser Gly Tyr Thr Phe Thr Asn
35 40 45

Tyr Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp
50 55 60

Met Gly Trp Ile Asn Thr His Thr Gly Glu Pro Thr Tyr Ala Asp Ser
65 70 75 80

Phe Lys Gly Arg Phe Thr Phe Ser Leu Asp Asp Ser Lys Asn Thr Ala
85 90 95

Tyr Leu Gln Ile Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Phe
100 105 110

Cys Thr Arg Arg Gly Tyr Asp Trp Tyr Phe Asp Val Trp Gly Gln Gly

115 120 125

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

Ala Pro Lys Thr Leu Ile Tyr Arg Ala Asn Arg Leu Glu Ser Gly Val
195 200 205
Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Thr Leu Thr
210 215 220
Ile Ser Ser Leu Gln Tyr Glu Asp Phe Gly Ile Tyr Tyr Cys Gln Gln
225 230 235 240
Tyr Asp Glu Ser Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile
245 250 255

Lys Ser Gly Gly Gly Gly Ser Gly Ala Leu Ser Asn Ser Ile Met Tyr
260 265 270
Phe Ser His Phe Val Pro Val Phe Leu Pro Ala Lys Pro Thr Thr Thr
275 280 285
Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala Ser Gln Pro
290 295 300
Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly Gly Ala Val
305 310 315 320

His Thr Arg Gly Leu Asp Ile Tyr Ile Trp Ala Pro Leu Ala Gly Thr
325 330 335
Cys Gly Val Leu Leu Leu Ser Leu Val Ile Thr Leu Tyr Cys Arg Leu
340 345 350
Lys Ile Gln Val Arg Lys Ala Ala Ile Thr Ser Tyr Glu Lys Ser Asp
355 360 365

Gly Val Tyr Thr Gly Leu Ser Thr Arg Asn Gln Glu Thr Tyr Glu Thr
370 375 380

Leu Lys His Glu Lys Pro Pro Gln Gly Ser Gly Ser Gln Arg Trp Lys
385 390 395 400
Ser Lys Leu Tyr Ser Ile Val Cys Gly Lys Ser Thr Pro Glu Lys Glu
405 410 415
Gly Glu Leu Glu Gly Thr Thr Thr Lys Pro Leu Ala Pro Asn Pro Ser
420 425 430
Phe Ser Pro Thr Pro Gly Phe Thr Pro Thr Leu Gly Phe Ser Pro Val
435 440 445

Pro Ser Ser Thr Phe Thr Ser Ser Ser Thr Tyr Thr Pro Gly Asp Cys
450 455 460
Pro Asn Phe Ala Ala Pro Arg Arg Glu Val Ala Pro Pro Tyr Gln Gly
465 470 475 480
Ala Asp Pro Ile Leu Ala Thr Ala Leu Ala Ser Asp Pro Ile Pro Asn
485 490 495
Pro Leu Gln Lys Trp Glu Asp Ser Ala His Lys Pro Gln Ser Leu Asp
500 505 510

Thr Asp Asp Pro Ala Thr Leu Tyr Ala Val Val Glu Asn Val Pro Pro
515 520 525
Leu Arg Trp Lys Glu Phe Val Arg Arg Leu Gly Leu Ser Asp His Glu
530 535 540
Ile Asp Arg Leu Glu Leu Gln Asn Gly Arg Cys Leu Arg Glu Ala Gln
545 550 555 560
Tyr Ser Met Leu Ala Thr Trp Arg Arg Arg Thr Pro Arg Arg Glu Ala
565 570 575

Thr Leu Glu Leu Leu Gly Arg Val Leu Arg Asp Met Asp Leu Leu Gly
580 585 590
Cys Leu Glu Asp Ile Glu Glu Ala Leu Cys Gly Pro Ala Ala Leu Pro
595 600 605
Pro Ala Pro Ser Leu Leu Arg

610 615

<210> 26

<211> 567

<212> PRT

<213> Artificial Sequence

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

<400> 26

Met Trp Leu Gln Ser Leu Leu Leu Gly Thr Val Ala Cys Ser Ile

1 5 10 15

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

20 25 30

Gly Ser Val Arg Ile Ser Cys Ala Ala Ser Gly Tyr Thr Phe Thr Asn

35 40 45

Tyr Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp

50 55 60

Met Gly Trp Ile Asn Thr His Thr Gly Glu Pro Thr Tyr Ala Asp Ser

65 70 75 80

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

85 90 95

Tyr Leu Gln Ile Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Phe

100 105 110

Cys Thr Arg Arg Gly Tyr Asp Trp Tyr Phe Asp Val Trp Gly Gln Gly

115 120 125

Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly

130 135 140

Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser

145 150 155 160

Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser

165 170 175

Gln Asp Ile Asn Ser Tyr Leu Ser Trp Phe Gln Gln Lys Pro Gly Lys

180 185 190

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

195 200 205

Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Thr Leu Thr

210 215 220

Ile Ser Ser Leu Gln Tyr Glu Asp Phe Gly Ile Tyr Tyr Cys Gln Gln

225 230 235 240

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

245 250 255

Lys Ser Gly Gly Gly Gly Ser Gly Ala Leu Ser Asn Ser Ile Met Tyr

260 265 270

Phe Ser His Phe Val Pro Val Phe Leu Pro Ala Lys Pro Thr Thr Thr

275 280 285

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

290 295 300

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

305 310 315 320

His Thr Arg Gly Leu Asp Ile Tyr Ile Trp Ala Pro Leu Ala Gly Thr

325 330 335

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

340 345 350

Lys Ile Gln Val Arg Lys Ala Ala Ile Thr Ser Tyr Glu Lys Ser Asp

355 360 365

Gly Val Tyr Thr Gly Leu Ser Thr Arg Asn Gln Glu Thr Tyr Glu Thr

370 375 380

Leu Lys His Glu Lys Pro Pro Gln Gly Ser Gly Ser Pro Leu Cys Leu

385 390 395 400

Gln Arg Glu Ala Lys Val Pro His Leu Pro Ala Asp Lys Ala Arg Gly

405 410 415

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

420 425 430

Ser Ser Ser Ser Leu Glu Ser Ser Ala Ser Ala Leu Asp Arg Arg Ala

435 440 445

Pro Thr Arg Asn Gln Pro Gln Ala Pro Gly Val Glu Ala Ser Gly Ala

450 455 460

Gly Glu Ala Arg Ala Ser Thr Gly Ser Ser Asp Ser Ser Pro Gly Gly

465 470 475 480

His Gly Thr Gln Val Asn Val Thr Cys Ile Val Asn Val Cys Ser Ser

485 490 495

Ser Asp His Ser Ser Gln Cys Ser Ser Gln Ala Ser Ser Thr Met Gly

500 505 510

Asp Thr Asp Ser Ser Pro Ser Glu Ser Pro Lys Asp Glu Gln Val Pro

515 520 525

Phe Ser Lys Glu Glu Cys Ala Phe Arg Ser Gln Leu Glu Thr Pro Glu

530 535 540

Thr Leu Leu Gly Ser Thr Glu Glu Lys Pro Leu Pro Leu Gly Val Pro

545 550 555 560

Asp Ala Gly Met Lys Pro Ser

565

<210> 27

<211> 42

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 27

Arg Leu Lys Ile Gln Val Arg Lys Ala Ala Ile Thr Ser Tyr Glu Lys

1 5 10 15

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

20 25 30

Glu Thr Leu Lys His Glu Lys Pro Pro Gln

35 40

<210> 28

<211> 42

<212> PRT

<213> Artificial Sequence

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

<400> 28

Arg Leu Lys Ile Gln Val Arg Lys Ala Ala Ile Thr Ser Tyr Glu Lys
1 5 10 15

Ser Asp Gly Val Tyr Thr Gly Leu Ser Thr Arg Asn Gln Glu Thr Tyr
20 25 30

Glu Thr Leu Lys His Glu Lys Pro Pro Gln
35 40

<210> 29

<211> 24

<212> PRT

<213> Artificial Sequence

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

<400> 29

Ile Tyr Ile Trp Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu

1 5 10 15

Ser Leu Val Ile Thr Leu Tyr Cys
20

<210> 30

<211> 86

<212> PRT

<213> Artificial Sequence

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

<400> 30

Ala Leu Ser Asn Ser Ile Met Tyr Phe Ser His Phe Val Pro Val Phe
1 5 10 15

Leu Pro Ala Lys Pro Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro

20 25 30

Ala Pro Thr Ile Ala Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys
35 40 45
Arg Pro Ala Ala Gly Gly Ala Val His Thr Arg Gly Leu Asp Ile Tyr
50 55 60
Ile Trp Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu Ser Leu
65 70 75 80
Val Ile Thr Leu Tyr Cys
85

<210> 31

<211> 83

<212> PRT

<213> Artificial Sequence

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

<400> 31

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

50 55 60
Ile Trp Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu Ser Leu
65 70 75 80
Val Ile Thr

<210> 32

<211> 255

<212> PRT

<213> Artificial Sequence

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

<400> 32

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Val Asn Thr Ala
20 25 30

Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
35 40 45

Tyr Ser Ala Ser Phe Leu Tyr Ser Gly Val Pro Ser Arg Phe Ser Gly
50 55 60

Ser Arg Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro Pro
85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Gly Ser Thr
100 105 110

Ser Gly Ser Gly Lys Pro Gly Ser Gly Glu Gly Ser Glu Val Gln Leu
115 120 125

Val Glu Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
130 135 140

Gly Gly Ser Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala
145 150 155 160

Ala Ser Gly Phe Asn Ile Lys Asp Thr Tyr Ile His Trp Val Arg Gln
165 170 175

Ala Pro Gly Lys Gly Leu Glu Trp Val Ala Arg Ile Tyr Pro Thr Asn
180 185 190

Gly Tyr Thr Arg Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser
195 200 205

Ala Asp Thr Ser Lys Asn Thr Ala Tyr Leu Gln Met Asn Ser Leu Arg
210 215 220

Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ser Arg Trp Gly Gly Asp Gly
225 230 235 240
Phe Tyr Ala Met Asp Val Trp Gly Gln Gly Thr Leu Val Thr Val
245 250 255

<210> 33

<211> 240

<212> PRT

<213> Artificial Sequence

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

<400> 33

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

Lys Gly Arg Phe Thr Phe Ser Leu Asp Asp Ser Lys Asn Thr Ala Tyr
65 70 75 80
Leu Gln Ile Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Phe Cys
85 90 95
Thr Arg Arg Gly Tyr Asp Trp Tyr Phe Asp Val Trp Gly Gln Gly Thr
100 105 110
Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
115 120 125

Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu
130 135 140
Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln
145 150 155 160
Asp Ile Asn Ser Tyr Leu Ser Trp Phe Gln Gln Lys Pro Gly Lys Ala

165 170 175

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

180 185 190

Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Thr Leu Thr Ile

195 200 205

Ser Ser Leu Gln Tyr Glu Asp Phe Gly Ile Tyr Tyr Cys Gln Gln Tyr

210 215 220

Asp Glu Ser Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys

225 230 235 240

<210> 34

<211> 26

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 34

Ile Tyr Leu Ile Ile Gly Ile Cys Gly Gly Gly Ser Leu Leu Met Val

1 5 10 15

Phe Val Ala Leu Leu Val Phe Tyr Ile Thr

20 25

<210> 35

<211> 1075

<212> DNA

<213> Artificial Sequence

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

polynucleotide

<400> 35

taatacgact cactataggg agaaagacgc caccatgggc aagaagcaaa atcgcaagac 60

ggggaattcc aagacacaat ccgctagccc accacctaaa gagcgttcta gctcccctgc 120

tactgagcag tcctggatgg aaaacgactt cgatgaactc cgggaagagg gatttaggcg 180

atccaactat tcagaactcc gcgaagatat ccagacaaag gggaaggaag tcgagaattt 240

cgagaagaac ctcgaggagt gcatcacccg tatcacaaac actgagaaat gtctcaaaga 300

actcatggaa cttaagacaa aagccaggga gcttcgagag gagtgtcgga gtctgagatc 360

caggtgtgac cagctcgagg agcgcgtgag cgcgatggaa gacgagatga acgagatgaa 420

aagagagggc aaattcaggg agaagcgcat taagaggaac gaacagagtc tgcaggagat 480

ttgggattac gtcaagaggc ctaacctgcg gttgatcggc gtccccgaga gcgacgtaga 540

aaacgggact aaactggaga atacacttca agacatcatt caagaaaatt ttccaaacct 600

ggctcggcaa gctaatgtgc aaatccaaga gatccaacgc acaccccagc ggtatagctc 660

tcggcgtgcc acccctaggc atattatcgt gcgctttact aaggtggaga tgaaagagaa 720

gatgctgcga gccgctcggg aaaagggaag ggtgactttg aagggcaaac ctattcggct 780

gacggttgac cttagcgccg agacactcca ggcacgccgg gaatggggcc ccatctttaa 840

tatcctgaag gagaagaact tccagccacg aatctcttac cctgcaaagt tgagttttat 900

ctccgagggt gagattaagt atttcatcga taaacagatg ctgcgagact tcgtgacaac 960

tcgcccagct ctcaaggaac tgctcaaaga ggctcttaat atggagcgca ataatagata 1020

tcaacccttg cagaaccacg caaagatgga ttataaggat gacgatgata aatga 1075

<210> 36

<211> 5751

<212> DNA

<213> Artificial Sequence

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

<400> 36

taatacgact cactataggg agaaagacgc caccatgaca ggttcaaata gtcacattac 60

gattctcact ctgaatataa atgggctgaa ttctgcaatt aaacggcaca ggcttgcttc 120

ctggataaag tctcaagacc cctcagtgtg ctgtattcag gaaacgcatc tcacgtgcag 180

ggacacccat cggctgaaaa taaaaggctg gcggaagatc taccaagcca atggaaaaca 240

aaagaaggct ggggtggcga tacttgtaag cgataaaaca gactttaaac caactaagat 300

caaacgggac aaagagggcc attacatcat ggtaaagggt agtattcaac aagaggagct 360

gactatcctg aatatttatg cacctaatac tggagccccc agattcataa agcaagtgtt 420

gagtgacctt caacgcgacc tcgactccca cactctgatc atgggagact ttaacacccc 480

gctgtccact ctcgacagat ctactagaca gaaagtcaac aaggatacac aggaactgaa 540

cagtgctctc caccaagcgg accttatcga catctacaga acactccacc ccaaaagcac 600

agaatatacc ttcttttcag cccctcacca cacctattcc aaaattgacc acattgtggg 660

gagtaaagcc cttctctcca aatgtaaacg gaccgaaatt atcactaact atctctccga 720

ccacagtgca ataaaacttg aattgcgaat taagaatctc actcaaagta gatccacgac 780

atggaaactg aacaatctcc tcttgaatga ctactgggtg cataacgaaa tgaaggctga 840

aataaagatg ttctttgaga ccaacgaaaa caaagacacc acgtaccaga atctctggga 900

cgctttcaaa gcagtgtgtc gaggaaaatt tattgcactg aatgcttaca agcggaagca 960

ggaaagatcc aaaatagaca ccctgactag ccaacttaaa gaactggaaa agcaagagca 1020

aactcatagc aaagctagcc gtcgccaaga aattacgaaa atcagagctg aactgaagga 1080

aattgagaca cagaaaaccc tgcaaaagat aaatgaaagc cgcagctggt tctttgaacg 1140

catcaacaaa atcgataggc cacttgctcg ccttatcaag aagaaaaggg agaagaatca 1200

aatcgacact ataaagaatg ataaaggcga tataaccacc gatcccacag aaattcaaac 1260

aaccatacgc gaatactaca aacacctcta cgccaataaa ctcgaaaatc tcgaggaaat 1320

ggatacattc ctcgacacgt acacccttcc caggctgaac caggaagaag ttgaatcact 1380

gaatcggcct atcacgggga gtgaaatagt agctatcatc aattcactcc ctaccaagaa 1440

gtcacccgga cctgatggat tcaccgccga attctaccag agatacatgg aagaactggt 1500

gcccttcttg ctgaaacttt tccaaagtat tgagaaagag ggaatacttc caaactcatt 1560

ttatgaggca tccatcattc tgatcccgaa gcccggcagg gacacgacca agaaagagaa 1620

ttttcgacca atctcattga tgaacattga tgcaaagatc ctcaataaaa tactggcaaa 1680

tcggattcag cagcacataa agaagctgat ccaccatgat caagtaggct tcatccccgg 1740

tatgcaaggt tggttcaata tacgaaaatc aatcaatgtt atccagcata taaaccgggc 1800

caaagacaag aaccacatga ttattagtat cgatgctgag aaagcctttg acaaaataca 1860

acaacccttc atgctgaaaa cattgaataa gctgggaatt gatggcacct acttcaaaat 1920

catcagagcc atatatgaca aaccaacagc aaatatcatt ctgaatggtc agaaattgga 1980

agcattcccc ttgaaaaccg gcacacggca gggttgccct ctgtcaccac tcctcttcaa 2040

catcgtgttg gaagttcttg cccgcgcaat ccggcaggaa aaggaaatca agggcattca 2100

actgggcaaa gaggaagtta aattgagcct gtttgcagac gacatgatcg tctatttgga 2160

aaaccccata gttagtgcac aaaatctgct gaagttgatc agtaatttct ccaaagtgag 2220

tgggtacaaa atcaatgtgc aaaagagcca agctttcttg tacaccaaca acaggcaaac 2280

tgagtctcaa atcatgggcg aactcccctt cgtgattgca tccaagcgga tcaaatacct 2340

ggggattcaa ttgactcgtg atgtgaagga cctcttcaag gagaactaca aacccctgct 2400

caaggaaatc aaagaggaca caaacaaatg gaagaacatt ccatgctctt gggtgggaag 2460

gatcaatatc gtcaaaatgg ccatcctgcc caaggtaatt tacaggttca atgctatacc 2520

catcaagctc cccatgacat tcttcacaga acttgaaaag acgacgctga agttcatttg 2580

gaaccagaaa cgtgccagga ttgctaaatc tattctctcc caaaagaaca aagctggcgg 2640

aatcacactc ccagacttca aactttacta caaggcgacc gtgacgaaaa cggcttggta 2700

ctggtaccaa aacagggata tagatcaatg gaaccgaacg gagcccagcg aaattatgcc 2760

tcatatatac aactatctga tctttgacaa accggagaag aacaagcaat ggggaaagga 2820

tagtctgttt aataaatggt gctgggaaaa ctggctcgca atctgtagga agctgaaact 2880

ggatccattc ttgacgcctt atacaaagat aaattcccga tggattaaag atctcaacgt 2940

gaaacccaaa acaattaaaa ccctcgagga aaacctgggt attacgattc aggacattgg 3000

ggtgggaaag gacttcatgt ccaaaacccc aaaagcgatg gcaaccaaag acaaaatcga 3060

caaatgggat ctcataaaac ttaagtcatt ttgcacagct aaagaaacga caattagggt 3120

gaaccgacaa ccgaccactt gggagaaaat cttcgcaaca tacagttctg acaaaggcct 3180

gatttccagg atctacaatg aattgaaaca aatttacaag aagaagacga acaaccctat 3240

aaagaaatgg gccaaggaca tgaacagaca cttctctaag gaagacattt atgcagccaa 3300

gaaacacatg aagaaatgca gctcttcact ggcaatcagg gaaatgcaaa tcaaaacaac 3360

aatgagatat catctcacac ccgtcagaat ggccatcatt aagaagagcg gaaacaaccg 3420

gtgctggcgt ggttgcggag aaatcggtac tctccttcac tgttggtggg actgtaaact 3480

cgttcaacca ctgtggaagt ctgtgtggcg gttcctcaga gatctggaac tcgaaatccc 3540

atttgaccca gccatccctc tcctgggtat atacccgaat gagtataaat cctgctgcta 3600

taaagacacc tgcacaagga tgtttattgc agctctcttc acaatcgcga agacgtggaa 3660

ccaacccaaa tgtccgacta tgattgactg gattaagaag atgtggcaca tatacactat 3720

ggaatactat gctgcgatca agaacgatga gttcatatca tttgtgggca catggatgaa 3780

actcgaaacc atcatactct ctaaattgag tcaagaacag aaaactaaac accgtatatt 3840

ttccctgatc ggtgggaatt agctacaaag acgatgacga caaggaccat ggagacggtg 3900

agagacacaa aaaattccaa cacactattg caatgaaaat aaatttcctt tattagccag 3960

aagtcagatg ctcaaggggc ttcatgatgt ccccataatt tttggcagag ggaaaaagat 4020

ctcagtggta tttgtgagcc agggcattgg ccttctgata ggcagcctgc acctgaggag 4080

tgcggccgct ttacttgtac agctcgtcca tgccgagagt gatcccggcg gcggtcacga 4140

actccagcag gaccatgtga tcgcgcttct cgttggggtc tttgctcagg gcggactggg 4200

tgctcaggta gtggttgtcg ggcagcagca cggggccgtc gccgatgggg gtgttctgct 4260

ggtagtggtc ggcgagctgc acgctgccgt cctcgatgtt gtggcggatc ttgaagttca 4320

ccttgatgcc gttcttctgc ttgtcggcca tgatatagac gttgtggctg ttgtagttgt 4380

actccagctt gtgccccagg atgttgccgt cctccttgaa gtcgatgccc ttcagctcga 4440
tgcggttcac cagggtgtcg ccctcgaact tcacctcggc gcgggtcttg tagttgccgt 4500
cgtccttgaa gaagatggtg cgctcctgga cgtagccttc gggcatggcg gacttgaaga 4560
agtcgtgctg cttcatgtgg tcggggtagc ggctgaagca ctgcacgccg taggtcaggg 4620
tggtcacgag ggtgggccag ggcacgggca gcttgccggt ggtgcagatg aacttcaggg 4680
tcagcttgcc gtaggtggca tcgccctcgc cctcgccgga cacgctgaac ttgtggccgt 4740
ttacgtcgcc gtccagctcg accaggatgg gcaccacccc ggtgaacagc tcctcgccct 4800

tgctcaccat ggtggcggga tctgacggtt cactaaacca gctctgctta tatagacctc 4860
ccaccgtaca cgcctaccgc ccatttgcgt caatggggcg gagttgttac gacattttgg 4920
aaagtcccgt tgattttggt gccaaaacaa actcccattg acgtcaatgg ggtggagact 4980
tggaaatccc cgtgagtcaa accgctatcc acgcccattg atgtactgcc aaaaccgcat 5040
caccatggta atagcgatga ctaatacgta gatgtactgc caagtaggaa agtcccataa 5100
ggtcatgtac tgggcataat gccaggcggg ccatttaccg tcattgacgt caatggggg 5160
cgtacttggc atatgataca cttgatgtac tgccaagtgg gcagtttacc gtaaatactc 5220

cacccattga cgtcaatgga aagtccctat tggcgttact atgggaacat acgtcattat 5280
tgacgtcaat gggcgggggt cgttgggcgg tcagccaggc gggccattta ccgtaagtta 5340
tgtaacgacg tctcagctga caatgagatc acatggacac aggaagggga atatcacact 5400
ctggggactg tggtggggtc gggggagggg ggagggatag cattgggaga tatacctaat 5460
gctagatgac acattagtgg gtgcagcgca ccagcatggc acatgtatac atatgtaact 5520
aacctgcaca atgtgcacat gtaccctaaa acttagagta taatggatcc gcaggcctct 5580
gctagcttga ctgactgaga tacagcgtac cttcagctca cagacatgat aagatacatt 5640

gatgagtttg gacaaaccac aactagaatg cagtgaaaaa aatgctttat ttgtgaaatt 5700
tgtgatgcta ttgctttatt tgtaaccatt ataagctgca ataaacaagt t 5751

<210> 37

<211> 18285

<212> DNA

<213> Artificial Sequence

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

<400> 37

cggccgcggg gggaggagcc aagatggccg aataggaaca gctccggtct acagctccca 60

gcgtgagcga cgcagaagac ggtgatttct gcatttccat ctgaggtacc gggttcatct 120

cactagggag tgccagacag tgggcgcagg ccagtgtgtg tgcgcaccgt gcgcgagccg 180

aagcagggcg aggcattgcc tcacctggga agcgcaaggg gtcagggagt tccctttccg 240

agtcaaagaa aggggtgacg gacgcacctg gaaaatcggg tcactcccac ccgaatattg 300

cgcttttcag accggcttaa gaaacggcgc accacgagac tatatcccac acctggctcg 360

gagggtccta cgcccacgga atctcgctga ttgctagcac agcagtctga gatcaaactg 420

caaggcggca acgaggctgg gggaggggcg cccgccattg cccaggcttg cttaggtaaa 480

caaagcagca gggaagctcg aactgggtgg agcccaccac agctcaagga ggcctgcctg 540

cctctgtagg ctccacctct gggggcaggg cacagacaaa caaaaagaca gcagtaacct 600

ctgcagactt aagtgtccct gtctgacagc tttgaagaga gcagtggttc tcccagcacg 660

cagctggaga tctgagaacg ggcagactgc ctcctcaagt gggtccctga cccctgaccc 720

ccgagcagcc taactgggag gcacccccca gcaggggcac actgacacct cacacggcag 780

ggtattccaa cagacctgca gctgagggtc ctgtctgtta gaaggaaaac taacaaccag 840

aaaggacatc tacaccgaaa acccatctgt acatcaccat catcaaagac caaaagtaga 900

taaaaccaca aagatgggga aaaaacagaa cagaaaaact ggaaactcta aaacgcagag 960

cgcctctcct cctccaaagg aacgcagttc ctcaccagca acagaacaaa gctggatgga 1020

gaatgatttt gatgagctga gagaagaagg cttcagacga tcaaattact ctgagctacg 1080

ggaggacatt caaaccaaag gcaaagaagt tgaaaacttt gaaaaaaatt tagaagaatg 1140

tataactaga ataaccaata cagagaagtg cttaaaggag ctgatggagc tgaaaaccaa 1200

ggctcgagaa ctacgtgaag aatgcagaag cctcaggagc cgatgcgatc aactggaaga 1260

aagggtatca gcaatggaag atgaaatgaa tgaaatgaag cgagaaggga agtttagaga 1320

aaaaagaata aaaagaaatg agcaaagcct ccaagaaata tgggactatg tgaaaagacc 1380

aaatctacgt ctgattggtg tacctgaaag tgatgtggag aatggaacca agttggaaaa 1440

cactctgcag gatattatcc aggagaactt ccccaatcta gcaaggcagg ccaacgttca 1500

gattcaggaa atacagagaa cgccacaaag atactcctcg agaagagcaa ctccaagaca 1560

cataattgtc agattcacca aagttgaaat gaaggaaaaa atgttaaggg cagccagaga 1620

gaaaggtcgg gttaccctca aaggaaagcc catcagacta acagcggatc tctcggcaga 1680

aacccctacaa gccagaagag agtgggggcc aatattcaac attcttaaag aaaagaattt 1740

tcaacccaga atttcatatc cagccaaact aagcttcata agtgaaggag aaataaaata 1800

ctttatagac aagcaaatgt tgagagattt tgtcaccacc aggcctgccc taaaagagct 1860

cctgaaggaa gcgctaaaca tggaaaggaa caaccggtac cagccgctgc aaaatcatgc 1920
caaaatgtaa agaccatcaa gactaggaag aaactgcatc aactaatgag caaaatcacc 1980
agctaacatc ataatgacag gatcaacttc acacataaca atattaactt taaatataaa 2040
tggactaaat tctgcaatta aaagacacag actggcaagt tggataaaga gtcaagaccc 2100
atcagtgtgc tgtattcagg aaacccatct cacgtgcaga gacacacata ggctcaaaat 2160
aaaaggatgg aggaagatct accaagccaa tggaaaacaa aaaaaggcag gggttgcaat 2220
cctagtctct gataaaacag actttaaacc aacaaagatc aaaagagaca aagaaggcca 2280

ttacataatg gtaaagggat caattcaaca agaggagcta actatcctaa atatttatgc 2340
acccaataca ggagcaccca gattcataaa gcaagtcctc agtgacctac aaagagactt 2400
agactcccac acattaataa tgggagactt taacacccca ctgtcaacat tagacagatc 2460
aacgagacag aaagtcaaca aggataccca ggaattgaac tcagctctgc accaagcaga 2520
cctaatagac atctacagaa ctctccaccc caaatcaaca gaatatacat ttttttcagc 2580
accacaccac acctattcca aaattgacca catagttgga agtaaagctc tcctcagcaa 2640
atgtaaaaga acagaaatta taacaaacta tctctcagac cacagtgcaa tcaaactaga 2700

actcaggatt aagaatctca ctcaaagccg ctcaactaca tggaaactga acaacctgct 2760
cctgaatgac tactgggtac ataacgaaat gaaggcagaa ataaagatgt tctttgaaac 2820
caacgagaac aaagacacca cataccagaa tctctgggac gcattcaaag cagtgtgtag 2880
agggaaattt atagcactaa atgcctacaa gagaaagcag gaaagatcca aaattgacac 2940
cctaacatca caattaaaag aactagaaaa gcaagagcaa acacattcaa aagctagcag 3000
aaggcaagaa ataactaaaa tcagagcaga actgaaggaa atagagacac aaaaaaccct 3060
tcaaaaaatc aatgaatcca ggagctggtt ttttgaaagg atcaacaaaa ttgatagacc 3120

gctagcaaga ctaataaaga aaaaaagaga gaagaatcaa atagacacaa taaaaaatga 3180
taaaggggat atcaccaccg atcccacaga aatacaaact accatcagag aatactacaa 3240
acacctctac gcaaataaac tagaaaatct agaagaaatg gatacattcc tcgacacata 3300
cactctccca agactaaacc aggaagaagt tgaatctctg aatcgaccaa taacaggctc 3360
tgaaattgtg gcaataatca atagtttacc aaccaaaaag agtccaggac cagatggatt 3420
cacagccgaa ttctaccaga ggtacaagga ggaactggta ccattccttc tgaaactatt 3480
ccaatcaata gaaaaagagg gaatcctccc taactcattt tatgaggcca gcatcattct 3540

gataccaaag ccgggcagag acacaaccaa aaaagagaat tttagaccaa tatccttgat 3600

gaacattgat gcaaaaatcc tcaataaaat actggcaaac cgaatccagc agcacatcaa 3660

aaagcttatc caccatgatc aagtgggctt catccctggg atgcaaggct ggttcaatat 3720

acgcaaatca ataaatgtaa tccagcatat aaacagagcc aaagacaaaa accacatgat 3780

tatctcaata gatgcagaaa aagcctttga caaaattcaa caacccttca tgctaaaaac 3840

tctcaataaa ttaggtattg atgggacgta tttcaaaata ataagagcta tctatgacaa 3900

acccacagcc aatatcatac tgaatgggca aaaactggaa gcattccctt tgaaaaccgg 3960

cacaagacag ggatgccctc tctcaccgct cctattcaac atagtgttgg aagttctggc 4020

cagggcaatc aggcaggaga aggaaataaa gggtattcaa ttaggaaaag aggaagtcaa 4080

attgtccctg tttgcagacg acatgattgt ttatctagaa aaccccatcg tctcagccca 4140

aaatctcctt aagctgataa gcaacttcag caaagtctca ggatacaaaa tcaatgtaca 4200

aaaatcacaa gcattcttat acaccaacaa cagacaaaca gagagccaaa tcatgggtga 4260

actcccattc acaattgctt caaagagaat aaaataccta ggaatccaac ttacaaggga 4320

tgtgaaggac ctcttcaagg agaactacaa accactgctc aaggaaataa aagaggagac 4380

aaacaaatgg aagaacattc catgctcatg ggtaggaaga atcaatatcg tgaaaatggc 4440

catactgccc aaggtaattt acagattcaa tgccatcccc atcaagctac caatgacttt 4500

cttcacagaa ttggaaaaaa ctactttaaa gttcatatgg aaccaaaaaa gagcccgcat 4560

tgccaagtca atcctaagcc aaaagaacaa agctggaggc atcacactac ctgacttcaa 4620

actatactac aaggctacag taaccaaaac agcatggtac tggtaccaaa acagagatat 4680

agatcaatgg aacagaacag agccctcaga aataatgccg catatctaca actatctgat 4740

ctttgacaaa cctgagaaaa acaagcaatg gggaaaggat tccctattta ataaatggtg 4800

ctgggaaaac tggctagcca tatgtagaaa gctgaaactg gatcccttcc ttacacctta 4860

tacaaaaatc aattcaagat ggattaaaga tttaaacgtt aaacctaaaa ccataaaaac 4920

cctagaagaa aacctaggca ttaccattca ggacataggc gtgggcaagg acttcatgtc 4980

caaaacacca aaagcaatgg caacaaaaga caaaattgac aaatgggatc taattaaact 5040

aaagagcttc tgcacagcaa aagaaactac catcagagtg aacaggcaac ctacaacatg 5100

ggagaaaatt tttgcaacct actcatctga caaagggcta atatccagaa tctacaatga 5160

actcaaacaa atttacaaga aaaaaacaaa caacccccatc aaaaagtggg cgaaggacat 5220

gaacagacac ttctcaaaag aagacattta tgcagccaaa aaacacatga agaaatgctc 5280

atcatcactg gccatcagag aaatgcaaat caaaaccact atgagatatc atctcacacc 5340

agttagaatg gcaatcatta aaaagtcagg aaacaacagg tgctggagag gatgcggaga 5400

aataggaaca cttttacact gttggtggga ctgtaaacta gttcaaccat tgtggaagtc 5460

agtgtggcga ttcctcaggg atctagaact agaaatacca tttgacccag ccatcccatt 5520

actgggtata tacccaaatg agtataaatc atgctgctat aaagacacat gcacacgtat 5580

gtttattgcg gcactattca caatagcaaa gacttggaac caacccaaat gtccaacaat 5640

gatagactgg attaagaaaa tgtggcacat atacaccatg gaatactatg cagccataaa 5700

aaatgatgag ttcatatcct ttgtagggac atggatgaaa ttggaaacca tcattctcag 5760

taaactatcg caagaacaaa aaaccaaaca ccgcatattc tcactcatag gtgggaattg 5820

aacaatgaga tcacatggac acaggaaggg gaatatcaca ctctggggac tgtggtgggg 5880

tcgggggagg ggggagggat agcattggga gatataccta atgctagatg acacattagt 5940

gggtgcagcg caccagcatg gcacatgtat acggatccga attctcgacg gatcgatccg 6000

aacaaacgac ccaacacccg tgcgttttat tctgtctttt tattgccgat cccctcagaa 6060

gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta 6120

aagcacgagg aagcggtcag cccattcgcc gccaagctct tcagcaatat cacgggtagc 6180

caacgctatg tcctgatagc ggtcggccgc tttacttgta cagctcgtcc atgccgagag 6240

tgatcccggc ggcggtcacg aactccagca ggaccatgtg atcgcgcttc tcgttggggt 6300

ctttgctcag ggcggactgg gtgctcaggt agtggttgtc gggcagcagc acggggccgt 6360

cgccgatggg ggtgttctgc tggtagtggt cggccaggtg agtccaggag atgtttcagc 6420

actgttgcct ttagtctcga ggcaacttag acaactgagt attgatctga gcacagcagg 6480

gtgtgagctg tttgaagata ctggggttgg gggtgaagaa actgcagagg actaactggg 6540

ctgagaccca gtggcaatgt tttagggcct aaggaatgcc tctgaaaatc tagatggaca 6600

actttgactt tgagaaaaga gaggtggaaa tgaggaaaat gacttttctt tattagattt 6660

cggtagaaag aactttcatc tttcccctat ttttgttatt cgttttaaaa catctatctg 6720

gaggcaggac aagtatggtc attaaaaaga tgcaggcaga aggcatatat tggctcagtc 6780

aaagtgggga actttggtgg ccaaacatac attgctaagg ctattcctat atcagctgga 6840

cacatataaa atgctgctaa tgcttcatta caaacttata tcctttaatt ccagatgggg 6900

gcaaagtatg tccaggggtg aggaacaatt gaaacatttg ggctggagta gattttgaaa 6960

gtcagctctg tgtgtgtgtg tgtgtgtgtg tgtgtgagag cgtgtgtttc ttttaacgtt 7020

ttcagcctac agcatacagg gttcatggtg gcaagaagat aacaagattt aaattatggc 7080

cagtgactag tgctgcaaga agaacaacta cctgcattta atgggaaagc aaaatctcag 7140

gctttgaggg aagttaacat aggcttgatt ctgggtggaa gctgggtgtg tagttatctg 7200

gaggccaggc tggagctctc agctcactat gggttcatct ttattgtctc ctttcatctc 7260

aacagctgca cgctgccgtc ctcgatgttg tggcggatct tgaagttcac cttgatgccg 7320

ttcttctgct tgtcggccat gatatagacg ttgtggctgt tgtagttgta ctccagcttg 7380
tgccccagga tgttgccgtc ctccttgaag tcgatgccct tcagctcgat gcggttcacc 7440
agggtgtcgc cctcgaactt cacctcggcg cgggtcttgt agttgccgtc gtccttgaag 7500
aagatggtgc gctcctggac gtagccttcg ggcatggcgg acttgaagaa gtcgtgctgc 7560
ttcatgtggt cggggtagcg gctgaagcac tgcacgccgt aggtcagggt ggtcacgagg 7620
gtgggccagg gcacgggcag cttgccggtg gtgcagatga acttcagggt cagcttgccg 7680
taggtggcat cgccctcgcc ctcgccggac acgctgaact tgtggccgtt tacgtcgccg 7740

tccagctcga ccaggatggg caccaccccg gtgaacagct cctcgccctt gctcaccatg 7800
gtggcgaatt cgaagcttga gctcgagatc tgagtccggt agcgctagcg gatctgacgg 7860
ttcactaaac cagctctgct tatatagacc tcccaccgta cacgcctacc gcccatttgc 7920
gtcaatgggg cggagttgtt acgacatttt ggaaagtccc gttgattttg gtgccaaaac 7980
aaactcccat tgacgtcaat ggggtggaga cttggaaatc cccgtgagtc aaaccgctat 8040
ccacgcccat tgatgtactg ccaaaaccgc atcaccatgg taatagcgat gactaatacg 8100
tagatgtact gccaagtagg aaagtcccat aaggtcatgt actgggcata atgccaggcg 8160

ggccatttac cgtcattgac gtcaataggg ggcgtacttg gcatatgata cacttgatgt 8220
actgccaagt gggcagttta ccgtaaatac tccacccatt gacgtcaatg gaaagtccct 8280
attggcgtta ctatgggaac atacgtcatt attgacgtca atgggcgggg gtcgttgggc 8340
ggtcagccag gcgggccatt taccgtaagt tatgtaacgc ggaactccat atatgggcta 8400
tgaactaatg accccgtaat tgattactat tagcccgggg gatccagaca tgataagata 8460
cattgatgag tttggacaaa ccacaactag aatgcagtga aaaaaatgct ttatttgtga 8520
aatttgtgat gctattgctt tatttgtaac cattataagc tgcaataaac aagttaacaa 8580

caacaattgc attcatttta tgtttcaggt tcaggggggag gtgtgggagg ttttttaaag 8640
caagtaaaac ctctacaaat gtggtatggc tgattatgat ccggctgcct cgcgcgtttc 8700
ggtgatgacg gtgaaaacct ctgacacatg cagctcccgg agacggtcac agcttgtctg 8760
taagcggatg ccgggagcag acaagcccgt cagggcgcgt cagcgggtgt tggcgggtgt 8820
cggggcgcag ccatgaggtc gatcgactct agaggatcga tccccgcccc ggacgaacta 8880
aacctgacta cgacatctct gccccttctt cgcggggcag tgcatgtaat cccttcagtt 8940
ggttggtaca acttgccaac tgggccctgt tccacatgtg acacgggggg ggaccaaaca 9000

caaaggggtt ctctgactgt agttgacatc cttataaatg gatgtgcaca tttgccaaca 9060

ctgagtggct ttcatcctgg agcagacttt gcagtctgtg gactgcaaca caacattgcc 9120

tttatgtgta actcttggct gaagctctta caccaatgct gggggacatg tacctcccag 9180

gggcccagga agactacggg aggctacacc aacgtcaatc agaggggcct gtgtagctac 9240

cgataagcgg accctcaaga gggcattagc aatagtgttt ataaggcccc cttgttaacc 9300

ctaaacgggt agcatatgct tcccgggtag tagtatatac tatccagact aaccctaatt 9360

caatagcata tgttacccaa cgggaagcat atgctatcga attagggtta gtaaaagggt 9420

cctaaggaac agcgatatct cccaccccat gagctgtcac ggttttattt acatggggtc 9480

aggattccac gagggtagtg aaccatttta gtcacaaggg cagtggctga agatcaagga 9540

gcgggcagtg aactctcctg aatcttcgcc tgcttcttca ttctccttcg tttagctaat 9600

agaataactg ctgagttgtg aacagtaagg tgtatgtgag gtgctcgaaa acaaggtttc 9660

aggtgacgcc cccagaataa aatttggacg gggggttcag tggtggcatt gtgctatgac 9720

accaatataa ccctcacaaa ccccttgggc aataaatact agtgtaggaa tgaaacattc 9780

tgaatatctt taacaataga aatccatggg gtggggacaa gccgtaaaga ctggatgtcc 9840

atctcacacg aatttatggc tatgggcaac acataatcct agtgcaatat gatactgggg 9900

ttattaagat gtgtcccagg cagggaccaa gacaggtgaa ccatgttgtt acactctatt 9960

tgtaacaagg ggaaagagag tggacgccga cagcagcgga ctccactggt tgtctctaac 10020

acccccgaaa attaaacggg gctccacgcc aatggggccc ataaacaaag acaagtggcc 10080

actcttttt ttgaaattgt ggagtggggg cacgcgtcag cccccacacg ccgccctgcg 10140

gttttggact gtaaaataag ggtgtaataa cttggctgat tgtaaccccg ctaaccactg 10200

cggtcaaacc acttgcccac aaaaccacta atggcacccc ggggaatacc tgcataagta 10260

ggtgggcggg ccaagatagg ggcgcgattg ctgcgatctg gaggacaaat tacacacact 10320

tgcgcctgag cgccaagcac agggttgttg gtcctcatat tcacgaggtc gctgagagca 10380

cggtgggcta atgttgccat gggtagcata tactacccaa atatctggat agcatatgct 10440

atcctaatct atatctgggt agcataggct atcctaatct atatctgggt agcatatgct 10500

atcctaatct atatctgggt agtatatgct atcctaattt atatctgggt agcataggct 10560

atcctaatct atatctgggt agcatatgct atcctaatct atatctgggt agtatatgct 10620

atcctaatct gtatccgggt agcatatgct atcctaatag agattagggt agtatatgct 10680

atcctaattt atatctgggt agcatatact acccaaatat ctggatagca tatgctatcc 10740

taatctatat ctgggtagca tatgctatcc taatctatat ctgggtagca taggctatcc 10800

taatctatat ctgggtagca tatgctatcc taatctatat ctgggtagta tatgctatcc 10860

taatttatat ctgggtagca taggctatcc taatctatat ctgggtagca tatgctatcc 10920

taatctatat ctgggtagta tatgctatcc taatctgtat ccgggtagca tatgctatcc 10980

tcatgcatat acagtcagca tatgataccc agtagtagag tgggagtgct atcctttgca 11040

tatgccgcca cctcccaagg gggcgtgaat tttcgctgct tgtccttttc ctgcatgctg 11100

gttgctccca ttcttaggtg aatttaagga ggccaggcta aagccgtcgc atgtctgatt 11160

gctcaccagg taaatgtcgc taatgttttc caacgcgaga aggtgttgag cgcggagctg 11220

agtgacgtga caacatgggt atgcccaatt gccccatgtt gggaggacga aaatggtgac 11280

aagacagatg gccagaaata caccaacagc acgcatgatg tctactgggg atttattctt 11340

tagtgcgggg gaatacacgg cttttaatac gattgagggc gtctcctaac aagttacatc 11400

actcctgccc ttcctcaccc tcatctccat cacctccttc atctccgtca tctccgtcat 11460

caccctccgc ggcagcccct tccaccatag gtggaaacca gggaggcaaa tctactccat 11520

cgtcaaagct gcacacagtc accctgatat tgcaggtagg agcgggcttt gtcataacaa 11580

ggtccttaat cgcatccttc aaaacctcag caaatatatg agtttgtaaa aagaccatga 11640

aataacagac aatggactcc cttagcgggc caggttgtgg gccgggtcca ggggccattc 11700

caaaggggag acgactcaat ggtgtaagac gacattgtgg aatagcaagg gcagttcctc 11760

gccttaggtt gtaaagggag gtcttactac ctccatatac gaacacaccg gcgacccaag 11820

ttccttcgtc ggtagtcctt tctacgtgac tcctagccag gagagctctt aaaccttctg 11880

caatgttctc aaatttcggg ttggaacctc cttgaccacg atgctttcca aaccaccctc 11940

ctttttgcg cctgcctcca tcaccctgac cccggggtcc agtgcttggg ccttctcctg 12000

ggtcatctgc ggggccctgc tctatcgctc ccgggggcac gtcaggctca ccatctgggc 12060

caccttcttg gtggtattca aaataatcgg cttcccctac agggtggaaa aatggccttc 12120

tacctggagg gggcctgcgc ggtggagacc cggatgatga tgactgacta ctgggactcc 12180

tgggcctctt ttctccacgt ccacgacctc tccccctggc tctttcacga cttccccccc 12240

tggctctttc acgtcctcta ccccggcggc ctccactacc tcctcgaccc cggcctccac 12300

tacctcctcg accccggcct ccactgcctc ctcgaccccg gcctccacct cctgctcctg 12360

cccctcctgc tcctgcccct cctcctgctc ctgcccctcc tgcccctcct gctcctgccc 12420

ctcctgcccc tcctgctcct gcccctcctg cccctcctgc tcctgcccct cctgcccctc 12480

ctcctgctcc tgcccctcct gcccctcctc ctgctcctgc ccctcctgcc cctcctgctc 12540

ctgcccctcc tgcccctcct gctcctgccc ctcctgcccc tcctgctcct gcccctcctg 12600

ctcctgcccc tcctgctcct gcccctcctg ctcctgcccc tcctgcccct cctgcccctc 12660

ctcctgctcc tgcccctcct gctcctgccc ctcctgcccc tcctgcccct cctgctcctg 12720

ccctcctcc tgctcctgcc cctcctgccc ctcctgcccc tcctcctgct cctgcccctc 12780

ctgcccctcc tcctgctcct gcccctcctc ctgctcctgc ccctcctgcc cctcctgccc 12840
ctcctcctgc tcctgcccct cctgcccctc ctcctgctcc tgcccctcct cctgctcctg 12900
ccctcctgc ccctcctgcc cctcctcctg ctcctgcccc tcctcctgct cctgcccctc 12960
ctgcccctcc tgcccctcct gcccctcctc ctgctcctgc ccctcctcct gctcctgccc 13020
ctcctgctcc tgcccctccc gctcctgctc ctgctcctgt tccaccgtgg gtccctttgc 13080
agccaatgca acttggacgt ttttggggtc tccggacacc atctctatgt cttggccctg 13140
atcctgagcc gcccggggct cctggtcttc cgcctcctcg tcctcgtcct cttccccgtc 13200

ctcgtccatg gttatcaccc cctcttcttt gaggtccact gccgccggag ccttctggtc 13260
cagatgtgtc tcccttctct cctaggccat ttccaggtcc tgtacctggc ccctcgtcag 13320
acatgattca cactaaaaga gatcaataga catctttatt agacgacgct cagtgaatac 13380
agggagtgca gactcctgcc ccctccaaca gccccccccac cctcatcccc ttcatggtcg 13440
ctgtcagaca gatccaggtc tgaaaattcc ccatcctccg aaccatcctc gtcctcatca 13500
ccaattactc gcagcccgga aaactcccgc tgaacatcct caagatttgc gtcctgagcc 13560
tcaagccagg cctcaaattc ctcgtccccc ttttgctgg acggtaggga tggggattct 13620

cgggacccct cctcttcctc ttcaaggtca ccagacagag atgctactgg ggcaacggaa 13680
gaaaagctgg gtgcggcctg tgaggatcag cttatcgatg ataagctgtc aaacatgaga 13740
attcttgaag acgaaagggc ctcgtgatac gcctattttt ataggttaat gtcatgataa 13800
taatggtttc ttagacgtca ggtggcactt ttcggggaaa tgtgcgcgga acccctattt 13860
gtttatttttt ctaaatacat tcaaatatgt atccgctcat gagacaataa ccctgataaa 13920
tgcttcaata atattgaaaa aggaagagta tgagtattca acatttccgt gtcgcccttа 13980
ttccctttttt tgcggcattt tgccttcctg tttttgctca cccagaaacg ctggtgaaag 14040

taaaagatgc tgaagatcag ttgggtgcac gagtgggtta catcgaactg gatctcaaca 14100
gcggtaagat ccttgagagt tttcgccccg aagaacgttt tccaatgatg agcacttttа 14160
aagttctgct atgtggcgcg gtattatccc gtgttgacgc cgggcaagag caactcggtc 14220
gccgcataca ctattctcag aatgacttgg ttgagtactc accagtcaca gaaaagcatc 14280
ttacggatgg catgacagta agagaattat gcagtgctgc cataaccatg agtgataaca 14340
ctgcggccaa cttacttctg acaacgatcg gaggaccgaa ggagctaacc gctttttttgc 14400
acaacatggg ggatcatgta actcgccttg atcgttggga accggagctg aatgaagcca 14460

taccaaacga cgagcgtgac accacgatgc ctgcagcaat ggcaacaacg ttgcgcaaac 14520

tattaactgg cgaactactt actctagctt cccggcaaca attaatagac tggatggagg 14580

cggataaagt tgcaggacca cttctgcgct cggcccttcc ggctggctgg tttattgctg 14640

ataaatctgg agccggtgag cgtgggtctc gcggtatcat tgcagcactg gggccagatg 14700

gtaagccctc ccgtatcgta gttatctaca cgacggggag tcaggcaact atggatgaac 14760

gaaatagaca gatcgctgag ataggtgcct cactgattaa gcattggtaa ctgtcagacc 14820

aagtttactc atatatactt tagattgatt taaaacttca tttttaattt aaaaggatct 14880

aggtgaagat cctttttgat aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc 14940

actgagcgtc agaccccgta gaaaagatca aaggatcttc ttgagatcct ttttttctgc 15000

gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc agcggtggtt tgtttgccgg 15060

atcaagagct accaactctt tttccgaagg taactggctt cagcagagcg cagataccaa 15120

atactgtcct tctagtgtag ccgtagttag gccaccactt caagaactct gtagcaccgc 15180

ctacatacct cgctctgcta atcctgttac cagtggctgc tgccagtggc gataagtcgt 15240

gtcttaccgg gttggactca agacgatagt taccggataa ggcgcagcgg tcgggctgaa 15300

cgggggggttc gtgcacacag cccagcttgg agcgaacgac ctacaccgaa ctgagatacc 15360

tacagcgtga gctatgagaa agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc 15420

cggtaagcgg cagggtcgga acaggagagc gcacgaggga gcttccaggg ggaaacgcct 15480

ggtatcttta tagtcctgtc gggtttcgcc acctctgact tgagcgtcga tttttgtgat 15540

gctcgtcagg ggggcggagc ctatggaaaa acgccagcaa cgcggccttt ttacggttcc 15600

tggccttttg ctggcctttt gctcacatgt tctttcctgc gttatcccct gattctgtgg 15660

ataaccgtat taccgccttt gagtgagctg ataccgctcg ccgcagccga acgaccgagc 15720

gcagcgagtc agtgagcgag gaagcggaag agcgcctgat gcggtatttt ctccttacgc 15780

atctgtgcgg tatttcacac cgcatatggt gcactctcag tacaatctgc tctgatgccg 15840

catagttaag ccagctgtgg aatgtgtgtc agttagggtg tggaaagtcc ccaggctccc 15900

cagcaggcag aagtatgcaa agcatgcatc tcaattagtc agcaaccagg tgtggaaagt 15960

ccccaggctc cccagcaggc agaagtatgc aaagcatgca tctcaattag tcagcaacca 16020

tagtcccgcc cctaactccg cccatcccgc ccctaactcc gcccagttcc gcccattctc 16080

cgccccatgg ctgactaatt ttttttattt atgcagaggc cgaggccgcc tcggcctctg 16140

agctattcca gaagtagtga ggaggctttt ttggaggcct aggcttttgc aaaaagcttg 16200

catgcctgca ggtcggccgc cacgaccggt gccgccacca tcccctgacc cacgcccctg 16260

acccctcaca aggagacgac cttccatgac cgagtacaag cccacggtgc gcctcgccac 16320

ccgcgacgac gtcccccggg ccgtacgcac cctcgccgcc gcgttcgccg actaccccgc 16380

cacgcgccac accgtcgacc cggaccgcca catcgagcgg gtcaccgagc tgcaagaact 16440

cttcctcacg cgcgtcgggc tcgacatcgg caaggtgtgg gtcgcggacg acggcgccgc 16500

ggtggcggtc tggaccacgc cggagagcgt cgaagcgggg gcggtgttcg ccgagatcgg 16560

cccgcgcatg gccgagttga gcggttcccg gctggccgcg cagcaacaga tggaaggcct 16620

cctggcgccg caccggccca aggagcccgc gtggttcctg gccaccgtcg gcgtctcgcc 16680

cgaccaccag ggcaagggtc tgggcagcgc cgtcgtgctc cccggagtgg aggcggccga 16740

gcgcgccggg gtgcccgcct tcctggagac ctccgcgccc cgcaacctcc ccttctacga 16800

gcggctcggc ttcaccgtca ccgccgacgt cgaggtgccc gaaggaccgc gcacctggtg 16860

catgacccgc aagcccggtg cctgacgccc gccccacgac ccgcagcgcc cgaccgaaag 16920

gagcgcacga ccccatggct ccgaccgaag ccgacccggg cggccccgcc gaccccgcac 16980

ccgccccccga ggcccaccga ctctagagga tcataatcag ccataccaca tttgtagagg 17040

ttttacttgc tttaaaaaac ctcccacacc tccccctgaa cctgaaacat aaaatgaatg 17100

caattgttgt tgttaacttg tttattgcag cttataatgg ttacaaataa agcaatagca 17160

tcacaaattt cacaaataaa gcatttttt cactgcattc tagttgtggt ttgtccaaac 17220

tcatcaatgt atcttatcat gtctggatca ctcgccgata gtggaaaccg acgccccagc 17280

actcgtccga gggcaaagga ataggggaga tgggggaggc taactgaaac acggaaggag 17340

acaataccgg aaggaacccg cgctatgacg gcaataaaaa gacagaataa aacgcacggg 17400

tgttgggtcg tttgttcata aacgcggggt tcggtcccag ggctggcact ctgtcgatac 17460

cccaccgaga ccccattggg gccaatacgc ccgcgtttct tccttttccc caccccaccc 17520

cccaagttcg ggtgaaggcc cagggctcgc agccaacgtc ggggcggcag gccctgccat 17580

agccactggc cccgtgggtt agggacgggg tcccccatgg ggaatggttt atggttcgtg 17640

ggggttatta ttttgggcgt tgcgtggggt ctggtccacg actggactga gcagacagac 17700

ccatggtttt tggatggcct gggcatggac cgcatgtact ggcgcgacac gaacaccggg 17760

cgtctgtggc tgccaaacac ccccgacccc caaaaaccac cgcgcggatt tctggcgtgc 17820

caagctagtc gaccaattct catgtttgac agcttatcat cgcagatccg ggcaacgttg 17880

ttgcattgct gcaggcgcag aactggtagg tatggaagat ctctagaagc tgggtaccag 17940

ctgctagcaa gcttgctagc ggccggctcg agtttactcc ctatcagtga tagagaacgt 18000

atgtcgagtt tactccctat cagtgataga gaacgatgtc gagtttactc cctatcagtg 18060

atagagaacg tatgtcgagt ttactcccta tcagtgatag agaacgtatg tcgagtttac 18120

tccctatcag tgatagagaa cgtatgtcga gtttatccct atcagtgata gagaacgtat 18180

gtcgagttta ctccctatca gtgatagaga acgtatgtcg aggtaggcgt gtacggtggg 18240

aggcctatat aagcagagct cgtttagtga accgtcagat cgccg 18285

<210> 38

<211> 7264

<212> DNA

<213> Artificial Sequence

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

<400> 38

taatacgact cactataggg agaagtactg ccaccatggg caagaagcaa aatcgcaaga 60

cggggaattc caagacacaa tccgctagcc caccacctaa agagcgttct agctcccctg 120

ctactgagca gtcctggatg gaaaacgact tcgatgaact ccgggaagag ggatttaggc 180

gatccaacta ttcagaactc cgcgaagata tccagacaaa ggggaaggaa gtcgagaatt 240

tcgagaagaa cctcgaggag tgcatcaccc gtatcacaaa cactgagaaa tgtctcaaag 300

aactcatgga acttaagaca aaagccaggg agcttcgaga ggagtgtcgg agtctgagat 360

ccaggtgtga ccagctcgag gagcgcgtga gcgcgatgga agacgagatg aacgagatga 420

aaagagaggg caaattcagg gagaagcgca ttaagaggaa cgaacagagt ctgcaggaga 480

tttgggatta cgtcaagagg cctaacctgc ggttgatcgg cgtccccgag agcgacgtag 540

aaaacgggac taaactggag aatacacttc aagacatcat tcaagaaaat tttccaaacc 600

tggctcggca agctaatgtg caaatccaag agatccaacg cacaccccag cggtatagct 660

ctcggcgtgc cacccctagg catattatcg tgcgctttac taaggtggag atgaaagaga 720

agatgctgcg agccgctcgg gaaaagggaa gggtgacttt gaagggcaaa cctattcggc 780

tgacggttga ccttagcgcc gagacactcc aggcacgccg ggaatggggc cccatcttta 840

atatcctgaa ggagaagaac ttccagccac gaatctctta ccctgcaaag ttgagtttta 900

tctccgaggg tgagattaag tatttcatcg ataaacagat gctgcgagac ttcgtgacaa 960

ctcgcccagc tctcaaggaa ctgctcaaag aggctcttaa tatggagcgc aataatagat 1020

atcaaccctt gcagaaccac gcaaagatgt gagacagccg tcagaccatc aagactagga 1080

agaaactgca tcaactaatg agcaaaatca ccagctaaca tcatagtata catgaccggc 1140

tctaactcac atatcaccat ccttacactt aacattaacg gcctcaactc agctatcaag 1200

cgccatcggc tggccagctg gatcaaatca caggatccaa gcgtttgttg catccaagag 1260

acccacctga cctgtagaga tactcaccgc ctcaagatca agggatggcg aaagatttat 1320

caggcgaacg gtaagcagaa gaaagccgga gtcgcaattc tggtctcaga caagacggat 1380

ttcaagccca ccaaaattaa gcgtgataag gaaggtcact atattatggt gaaaggcagc 1440

atacagcagg aagaacttac catattgaac atctacgcgc caaacaccgg cgcacctcgc 1500

tttatcaaac aggtcctgtc cgatctgcag cgagatctgg attctcatac gttgattatg 1560

ggtgatttca atacaccatt gagcaccctg gatcgcagca ccaggcaaaa ggtaaataaa 1620

gacacgcaag agctcaatag cgcactgcat caggcagatc tcattgatat ttatcgcact 1680

cttcatccta agagtaccga gtacacattc ttcagcgccc cacatcatac atactcaaag 1740

atcgatcata tcgtcggctc aaaggctctg ctgtcaaagt gcaagcgcac agagataatt 1800

acaaattacc tgtcagatca tagcgcgatc aagctcgagc tgagaatcaa gaacctgacc 1860

cagagccgga gtaccacttg gaagcttaat aacctgctgc tcaacgatta ttgggtccac 1920

aatgagatga aggcagagat taaaatgttc ttcgaaacaa atgagaataa ggatactacc 1980

tatcaaaacc tttgggatgc ctttaaggcc gtctgcagag gcaagttcat cgccctcaac 2040

gcctataaaa gaaaacaaga gagatctaag atcgatactc tcacctctca gctgaaggag 2100

ttggagaaac aggaacagac ccactccaag gcgtcaagac ggcaggagat cacaaagatt 2160

cgcgccgagt tgaaagagat cgaaacccaa aagactcttc agaaaattaa cgagtctcgt 2220

agttggttct tcgagcggat taataagata gacagacctc tggcacgact gattaagaag 2280

aagcgcgaaa agaaccagat tgataccatc aagaacgaca agggcgacat cactactgac 2340

ccgaccgaga tccagaccac tattcgggag tattataagc atttgtatgc taacaagctt 2400

gagaacctgg aagagatgga cacttttctg gatacctata ctctgccacg gcttaatcaa 2460

gaggaagtcg agtccctcaa ccgcccaatt acaggaagcg agattgtggc cataattaac 2520

tccctgccga caaagaaatc tcctggtccg gacgggttta cagctgagtt ttatcaacgg 2580

tatatggaag agcttgtacc gtttctgctc aagctctttc agtctataga aaaggaaggc 2640

atcttgccca attccttcta cgaagcttct ataatactta ttcccaaacc aggacgcgat 2700

accacaaaga aggaaaactt ccggcccatt agtctcatga atatcgacgc taaaatattg 2760

aacaagattc tcgccaacag aatccaacaa catattaaga aattgataca tcacgaccag 2820

gtggggttta tacctggcat gcagggctgg tttaacatcc ggaagagtat taacgtcatt 2880

caacacatta atagagctaa ggataagaat catatgatca tctctataga cgcggaaaag 2940

gcattcgata agattcagca gccatttatg ctcaagactc tgaacaaact cggcatcgac 3000

ggaacatatt ttaagattat tcgcgcaatt tacgataagc cgactgctaa cattatcctt 3060

aacggccaaa agctcgaggc ctttccgctc aagactggaa cccgccaagg ctgtcccctc 3120

tccccgcttt tgtttaatat tgtactcgag gtgctggcta gggctattcg tcaagagaaa 3180

gagattaaag ggatacagct cgggaaggaa gaggtcaagc tttccttgtt cgccgatgat 3240
atgattgtgt acctggagaa tcctattgtg tctgctcaga accttcttaa acttatttct 3300
aactttagca aggtcagcgg ctataagatt aacgtccaga aatctcaggc ctttctgtac 3360
acaaataatc gacagaccga atcccagata atgggtgagc ttccgtttgt catagccagc 3420
aaaaggataa agtatctcgg aatccagctg acacgagacg ttaaagattt gtttaaggaa 3480
aattacaagc ctctcctgaa agagattaag gaagatacta ataagtggaa gaatatcccc 3540
tgttcatggg ttggcagaat caacatagtg aagatggcaa tacttcctaa agtgatatat 3600

cgctttaacg ccatcccaat taaactgcct atgaccttct ttacggagct cgagaaaaca 3660
acccttaaat ttatatggaa tcaaaagaga gcaagaatag cgaagtccat cttgagccag 3720
aagaataagg ccggtgggat tactttgcct gattttaagt tgtattataa agccacagta 3780
actaagacag cctggtattg gtatcagaat agagacatcg accagtggaa tcggaccgaa 3840
ccatcagaga taatgcccca catctataat taccttatat tcgataagcc agaaaagaat 3900
aaacagtggg gcaaagacag cctcttcaac aagtggtgtt gggagaattg gctggccata 3960
tgccggaaac tcaagctcga ccccttttctt acaccctaca ctaaaatcaa cagtaggtgg 4020

atcaaggact tgaatgtcaa gccaaagact ataaagacac tggaagagaa tcttgggatc 4080
acaatacaag atataggcgt cggcaaagat tttatgtcaa agacgcccaa ggccatggcc 4140
actaaggata agattgataa gtgggacctt attaagctca aaagcttctg tactgccaag 4200
gagaccacga tcagagttaa taggcagccc actacatggg aaaagatttt cgccacttat 4260
tcatcagata aggggttgat aagcagaata tataacgagc tgaagcagat ctacaagaag 4320
aaaacgaata atcccatcaa gaagtgggca aaagatatga acaggcattt tagcaaagag 4380
gatatctacg ccgcgaagaa gcatatgaag aagtgtagtt caagcttggc cattcgtgag 4440

atgcagatta agacgaccat gcgataccac cttaccccag tgaggatggc aattatcaag 4500
aaatctggca ataatagatg ttggcggggc tgtggcgaga ttggcaccct gctccattgc 4560
tggtgggatt gcaagctggt gcagccgctt tggaaatcag tctggcgctt tctgagggac 4620
ctcgagcttg agattccctt cgatcccgca attcccttgc tcggaatcta tcctaacgaa 4680
tacaagagct gttgttacaa ggatacgtgt acccggatgt tcatcgcggc cttgtttacg 4740
atagctaaga cgtggaatca gcctaagtgc cccacaatga tcgattggat caagaaaatg 4800
tggcatattt ataccatgga gtattacgca gcaattaaga atgacgaatt tatttccttc 4860

gttgggacct ggatgaagct ggagactatt attctgagca agctgtctca ggagcaaaag 4920

acaaagcata gaatcttctc tctcattggt ggtaacgact acaaagacga tgacgacaag 4980

taaagcgctt ctagaagttg tctcctcctg cactgactga ctgatacaat cgatttctgg 5040

atccgcaggc ctaatcaacc tctggattac aaaatttgtg aaagattgac tggtattctt 5100

aactatgttg ctccttttac gctatgtgga tacgctgctt taatgccttt gtatcatgct 5160

attgcttccc gtatggcttt cattttctcc tccttgtata aatcctggtt gctgtctctt 5220

tatgaggagt tgtggcccgt tgtcaggcaa cgtggcgtgg tgtgcactgt gtttgctgac 5280

gcaacccccа ctggttgggg cattgccacc acctgtcagc tcctttccgg gactttcgct 5340

ttcccctcc ctattgccac ggcggaactc atcgccgcct gccttgcccg ctgctggaca 5400

ggggctcggc tgttgggcac tgacaattcc gtggtgttgt cggggaagct gacgtccttt 5460

ccatggctgc tcgcctgtgt tgccacctgg attctgcgcg ggacgtcctt ctgctacgtc 5520

ccttcggccc tcaatccagc ggaccttcct tcccgctgag agacacaaaa aattccaaca 5580

cactattgca atgaaaataa atttccttta ttagccagaa gtcagatgct caaggggctt 5640

catgatgtcc ccataatttt tggcagaggg aaaaagatct cagtggtatt tgtgagccag 5700

ggcattggcc ttctgatagg cagcctgcac ctgaggagtg cggccgcttt acttgtacag 5760

ctcgtccatg ccgagagtga tcccggcggc ggtcacgaac tccagcagga ccatgtgatc 5820

gcgcttctcg ttggggtctt tgctcagggc ggactgggtg ctcaggtagt ggttgtcggg 5880

cagcagcacg gggccgtcgc cgatgggggt gttctgctgg tagtggtcgg cgagctgcac 5940

gctgccgtcc tcgatgttgt ggcggatctt gaagttcacc ttgatgccgt tcttctgctt 6000

gtcggccatg atatagacgt tgtggctgtt gtagttgtac tccagcttgt gccccaggat 6060

gttgccgtcc tccttgaagt cgatgcccct cagctcgatg cggttcacca gggtgtcgcc 6120

ctcgaacttc acctcggcgc gggtcttgta gttgccgtcg tccttgaaga agatggtgcg 6180

ctcctggacg tagccttcgg gcatggcgga cttgaagaag tcgtgctgct tcatgtggtc 6240

ggggtagcgg ctgaagcact gcacgccgta ggtcagggtg gtcacgaggg tgggccaggg 6300

cacgggcagc ttgccggtgg tgcagatgaa cttcagggtc agcttgccgt aggtggcatc 6360

gccctcgccc tcgccggaca cgctgaactt gtggccgttt acgtcgccgt ccagctcgac 6420

caggatgggc accaccccgg tgaacagctc ctcgcccttg ctcaccatgg tggcgggatc 6480

tgacggttca ctaaaccagc tctgcttata tagacctccc accgtacacg cctaccgccc 6540

atttgcgtca atggggcgga gttgttacga cattttggaa agtcccgttg attttggtgc 6600

caaaacaaac tcccattgac gtcaatgggg tggagacttg gaaatccccg tgagtcaaac 6660

cgctatccac gcccattgat gtactgccaa aaccgcatca ccatggtaat agcgatgact 6720

aatacgtaga tgtactgcca agtaggaaag tcccataagg tcatgtactg ggcataatgc 6780

caggcgggcc atttaccgtc attgacgtca ataggggggcg tacttggcat atgatacact 6840

tgatgtactg ccaagtgggc agtttaccgt aaatactcca cccattgacg tcaatggaaa 6900

gtccctattg gcgttactat gggaacatac gtcattattg acgtcaatgg gcgggggtcg 6960

ttgggcggtc agccaggcgg gccatttacc gtaagttatg taacgggcct gctgccggct 7020

ctgcggcctc ttccgcgtct tcgccttcgc cctcagacga gtcggatctc cctttgggcc 7080

gcctccccgc ctgtctagct tgactgactg agatacagcg taccttcagc tcacagacat 7140

gataagatac attgatgagt ttggacaaac cacaactaga atgcagtgaa aaaaatgctt 7200

tatttgtgaa atttgtgatg ctattgcttt atttgtaacc attataagct gcaataaaca 7260

agtt 7264

<210> 39

<211> 18852

<212> DNA

<213> Artificial Sequence

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

<400> 39

cggccgcggg gggaggagcc aagatggccg aataggaaca gctccggtct acagctccca 60

gcgtgagcga cgcagaagac ggtgatttct gcatttccat ctgaggtacc gggttcatct 120

cactagggag tgccagacag tgggcgcagg ccagtgtgtg tgcgcaccgt gcgcgagccg 180

aagcagggcg aggcattgcc tcacctggga agcgcaaggg gtcagggagt tccctttccg 240

agtcaaagaa aggggtgacg gacgcacctg gaaaatcggg tcactcccac ccgaatattg 300

cgcttttcag accggcttaa gaaacggcgc accacgagac tatatcccac acctggctcg 360

gagggtccta cgcccacgga atctcgctga ttgctagcac agcagtctga gatcaaactg 420

caaggcggca acgaggctgg gggaggggcg cccgccattg cccaggcttg cttaggtaaa 480

caaagcagca gggaagctcg aactgggtgg agcccaccac agctcaagga ggcctgcctg 540

cctctgtagg ctccacctct gggggcaggg cacagacaaa caaaaagaca gcagtaacct 600

ctgcagactt aagtgtccct gtctgacagc tttgaagaga gcagtggttc tcccagcacg 660

cagctggaga tctgagaacg ggcagactgc ctcctcaagt gggtccctga cccctgaccc 720

ccgagcagcc taactgggag gcacccccca gcaggggcac actgacacct cacacggcag 780

ggtattccaa cagacctgca gctgagggtc ctgtctgtta gaaggaaaac taacaaccag 840

aaaggacatc tacaccgaaa acccatctgt acatcaccat catcaaagac caaaagtaga 900

taaaaccaca aagatgggga aaaaacagaa cagaaaaact ggaaactcta aaacgcagag 960

cgcctctcct cctccaaagg aacgcagttc ctcaccagca acagaacaaa gctggatgga 1020

gaatgatttt gatgagctga gagaagaagg cttcagacga tcaaattact ctgagctacg 1080

ggaggacatt caaaccaaag gcaaagaagt tgaaaacttt gaaaaaaatt tagaagaatg 1140

tataactaga ataaccaata cagagaagtg cttaaaggag ctgatggagc tgaaaaccaa 1200

ggctcgagaa ctacgtgaag aatgcagaag cctcaggagc cgatgcgatc aactggaaga 1260

aagggtatca gcaatggaag atgaaatgaa tgaaatgaag cgagaaggga agtttagaga 1320

aaaaagaata aaaagaaatg agcaaagcct ccaagaaata tgggactatg tgaaaagacc 1380

aaatctacgt ctgattggtg tacctgaaag tgatgtggag aatggaacca agttggaaaa 1440

cactctgcag gatattatcc aggagaactt ccccaatcta gcaaggcagg ccaacgttca 1500

gattcaggaa atacagagaa cgccacaaag atactcctcg agaagagcaa ctccaagaca 1560

cataattgtc agattcacca aagttgaaat gaaggaaaaa atgttaaggg cagccagaga 1620

gaaaggtcgg gttaccctca aaggaaagcc catcagacta acagcggatc tctcggcaga 1680

aaccctacaa gccagaagag agtgggggcc aatattcaac attcttaaag aaaagaattt 1740

tcaacccaga atttcatatc cagccaaact aagcttcata agtgaaggag aaataaaata 1800

ctttatagac aagcaaatgt tgagagattt tgtcaccacc aggcctgccc taaaagagct 1860

cctgaaggaa gcgctaaaca tggaaaggaa caaccggtac cagccgctgc aaaatcatgc 1920

caaaatgtaa agaccatcaa gactaggaag aaactgcatc aactaatgag caaaatcacc 1980

agctaacatc ataatgacag gatcaacttc acacataaca atattaactt taaatataaa 2040

tggactaaat tctgcaatta aaagacacag actggcaagt tggataaaga gtcaagaccc 2100

atcagtgtgc tgtattcagg aaacccatct cacgtgcaga gacacacata ggctcaaaat 2160

aaaaggatgg aggaagatct accaagccaa tggaaaacaa aaaaaggcag gggttgcaat 2220

cctagtctct gataaaacag actttaaacc aacaaagatc aaaagagaca aagaaggcca 2280

ttacataatg gtaaagggat caattcaaca agaggagcta actatcctaa atatttatgc 2340

acccaataca ggagcaccca gattcataaa gcaagtcctc agtgacctac aaagagactt 2400

agactcccac acattaataa tgggagactt taacacccca ctgtcaacat tagacagatc 2460

aacgagacag aaagtcaaca aggataccca ggaattgaac tcagctctgc accaagcaga 2520

cctaatagac atctacagaa ctctccaccc caaatcaaca gaatatacat ttttttcagc 2580

accacaccac acctattcca aaattgacca catagttgga agtaaagctc tcctcagcaa 2640

atgtaaaaga acagaaatta taacaaacta tctctcagac cacagtgcaa tcaaactaga 2700

actcaggatt aagaatctca ctcaaagccg ctcaactaca tggaaactga acaacctgct 2760

cctgaatgac tactgggtac ataacgaaat gaaggcagaa ataaagatgt tctttgaaac 2820

caacgagaac aaagacacca cataccagaa tctctgggac gcattcaaag cagtgtgtag 2880

agggaaattt atagcactaa atgcctacaa gagaaagcag gaaagatcca aaattgacac 2940

cctaacatca caattaaaag aactagaaaa gcaagagcaa acacattcaa aagctagcag 3000

aaggcaagaa ataactaaaa tcagagcaga actgaaggaa atagagacac aaaaaaccct 3060

tcaaaaaatc aatgaatcca ggagctggtt ttttgaaagg atcaacaaaa ttgatagacc 3120

gctagcaaga ctaataaaga aaaaaagaga gaagaatcaa atagacacaa taaaaaatga 3180

taaaggggat atcaccaccg atcccacaga aatacaaact accatcagag aatactacaa 3240

acacctctac gcaaataaac tagaaaatct agaagaaatg gatacattcc tcgacacata 3300

cactctccca agactaaacc aggaagaagt tgaatctctg aatcgaccaa taacaggctc 3360

tgaaattgtg gcaataatca atagtttacc aaccaaaaag agtccaggac cagatggatt 3420

cacagccgaa ttctaccaga ggtacaagga ggaactggta ccattccttc tgaaactatt 3480

ccaatcaata gaaaaagagg gaatcctccc taactcattt tatgaggcca gcatcattct 3540

gataccaaag ccgggcagag acacaaccaa aaaagagaat tttagaccaa tatccttgat 3600

gaacattgat gcaaaaatcc tcaataaaat actggcaaac cgaatccagc agcacatcaa 3660

aaagcttatc caccatgatc aagtgggctt catccctggg atgcaaggct ggttcaatat 3720

acgcaaatca ataaatgtaa tccagcatat aaacagagcc aaagacaaaa accacatgat 3780

tatctcaata gatgcagaaa aagcctttga caaaattcaa caacccttca tgctaaaaac 3840

tctcaataaa ttaggtattg atgggacgta tttcaaaata ataagagcta tctatgacaa 3900

acccacagcc aatatcatac tgaatgggca aaaactggaa gcattccctt tgaaaaccgg 3960

cacaagacag ggatgccctc tctcaccgct cctattcaac atagtgttgg aagttctggc 4020

cagggcaatc aggcaggaga aggaaataaa gggtattcaa ttaggaaaag aggaagtcaa 4080

attgtccctg tttgcagacg acatgattgt ttatctagaa aaccccatcg tctcagccca 4140

aaatctcctt aagctgataa gcaacttcag caaagtctca ggatacaaaa tcaatgtaca 4200

aaaatcacaa gcattcttat acaccaacaa cagacaaaca gagagccaaa tcatgggtga 4260

actcccattc acaattgctt caaagagaat aaaataccta ggaatccaac ttacaaggga 4320

tgtgaaggac ctcttcaagg agaactacaa accactgctc aaggaaataa aagaggagac 4380

aaacaaatgg aagaacattc catgctcatg ggtaggaaga atcaatatcg tgaaaatggc 4440

catactgccc aaggtaattt acagattcaa tgccatcccc atcaagctac caatgacttt 4500

cttcacagaa ttggaaaaaa ctactttaaa gttcatatgg aaccaaaaaa gagcccgcat 4560

tgccaagtca atcctaagcc aaaagaacaa agctggaggc atcacactac ctgacttcaa 4620
actatactac aaggctacag taaccaaaac agcatggtac tggtaccaaa acagagatat 4680
agatcaatgg aacagaacag agccctcaga aataatgccg catatctaca actatctgat 4740
ctttgacaaa cctgagaaaa acaagcaatg gggaaaggat tccctattta ataaatggtg 4800
ctgggaaaac tggctagcca tatgtagaaa gctgaaactg gatcccttcc ttacacctta 4860
tacaaaaatc aattcaagat ggattaaaga tttaaacgtt aaacctaaaa ccataaaaac 4920
cctagaagaa aacctaggca ttaccattca ggacataggc gtgggcaagg acttcatgtc 4980

caaaacacca aaagcaatgg caacaaaaga caaaattgac aaatgggatc taattaaact 5040
aaagagcttc tgcacagcaa aagaaactac catcagagtg aacaggcaac ctacaacatg 5100
ggagaaaatt tttgcaacct actcatctga caaagggcta atatccagaa tctacaatga 5160
actcaaacaa atttacaaga aaaaaacaaa caaccccatc aaaaagtggg cgaaggacat 5220
gaacagacac ttctcaaaag aagacattta tgcagccaaa aaacacatga agaaatgctc 5280
atcatcactg gccatcagag aaatgcaaat caaaaccact atgagatatc atctcacacc 5340
agttagaatg gcaatcatta aaaagtcagg aaacaacagg tgctggagag gatgcggaga 5400

aataggaaca cttttacact gttggtggga ctgtaaacta gttcaaccat tgtggaagtc 5460
agtgtggcga ttcctcaggg atctagaact agaaatacca tttgacccag ccatcccatt 5520
actgggtata tacccaaatg agtataaatc atgctgctat aaagacacat gcacacgtat 5580
gtttattgcg gcactattca caatagcaaa gacttggaac caacccaaat gtccaacaat 5640
gatagactgg attaagaaaa tgtggcacat atacaccatg gaatactatg cagccataaa 5700
aaatgatgag ttcatatcct ttgtagggac atggatgaaa ttggaaacca tcattctcag 5760
taaactatcg caagaacaaa aaaccaaaca ccgcatattc tcactcatag gtgggaattg 5820

aacaatgaga tcacatggac acaggaaggg gaatatcaca ctctggggac tgtggtgggg 5880
tcgggggagg ggggagggat agcattggga gatataccta atgctagatg acacattagt 5940
gggtgcagcg caccagcatg gcacatgtat acggatccga attctcgacg gatcgatccg 6000
aacaaacgac ccaacacccg tgcgttttat tctgtctttt tattgccgat cccctcagaa 6060
gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta 6120
aagcacgagg aagcggtcag cccattcgcc gccaagctct tcagcaatat cacgggtagc 6180
caacgctatg tcctgatagc ggtccggccgc tcatgttctc gtaggagtcg gcgtcctctt 6240

cgtggttagg tccaggttgg cctctgatag accgcagctg aggagcggcg tacagaatgc 6300

ctctcatgtc ctcatagctg ccgctgcctt gtggaggctt ctcgtgcttc agtgtctcgt 6360

atgtctcttg attccgggtg ctcaggccgg tgtacacgcc atcagatttc tcgtagctgg 6420

tgatggcggc cttccgcact tggatcttca gccgtctgca gtacagggtg atgaccagag 6480

acagcagcag gacaccacat gtgccagcca gaggggccca aatgtagata tccaggcctc 6540

tggtatgcac agctccgcct gcagcaggtc tacaggcttc aggtctgaga gacagaggct 6600

ggctggcgat tgtaggagct ggtgtaggtg gtctaggagc gggtgttgtt gtaggcttgg 6660

cgggcagaaa cacgggcacg aagtggctga agtacatgat gctattgctc agggctccgc 6720

ttcctccgcc gcctgatttg atttccagct tggtgcctcc gccaaatgtc caagggctct 6780

cgtcgtactg ctggcagtag tagatgccga agtcctcgta ctgcaggctg ctgattgtca 6840

gggtgtagtc ggtgccagag ccgctgccag aaaatctgct tggcacgccg ctttccagtc 6900

tgttggcccg gtagatcagt gtcttagggg ccttgccagg cttctgctgg aaccagctca 6960

ggtagctgtt gatgtcctgg ctggctctac aggtgatggt cactctatcg cccacagagg 7020

cagacaggct gctagggctc tgtgtcatct ggatatcaga gccaccaccg ccagatccac 7080

cgccacctga tcctccgcct ccgctagaaa ctgtcactgt ggtgccctgg ccccacacat 7140

cgaagtacca gtcgtagcct cttctggtgc agaagtacac ggcggtatcc tcggctctca 7200

ggctgttgat ctgcaggtag gcggtgttct tgctgtcgtc caggctgaag gtgaatctgc 7260

ccttaaagct atcggcgtag gttggctcgc cggtgtgggt attgatccag cccatccact 7320

caaggccagg tgagtccagg agatgtttca gcactgttgc ctttagtctc gaggcaactt 7380

agacaactga gtattgatct gagcacagca gggtgtgagc tgtttgaaga tactggggtt 7440

gggggtgaag aaactgcaga ggactaactg ggctgagacc cagtggcaat gttttagggc 7500

ctaaggaatg cctctgaaaa tctagatgga caactttgac tttgagaaaa gagaggtgga 7560

aatgaggaaa atgacttttc tttattagat ttcggtagaa agaactttca tctttcccct 7620

atttttgtta ttcgttttaa aacatctatc tggaggcagg acaagtatgg tcattaaaaa 7680

gatgcaggca gaaggcatat attggctcag tcaaagtggg gaactttggt ggccaaacat 7740

acattgctaa ggctattcct atatcagctg gacacatata aaatgctgct aatgcttcat 7800

tacaaactta tatcctttaa ttccagatgg gggcaaagta tgtccagggg tgaggaacaa 7860

ttgaaacatt tgggctggag tagattttga aagtcagctc tgtgtgtgtg tgtgtgtgtg 7920

tgtgtgtgag agcgtgtgtt tcttttaacg ttttcagcct acagcataca gggttcatgg 7980

tggcaagaag ataacaagat ttaaattatg gccagtgact agtgctgcaa gaagaacaac 8040

tacctgcatt taatgggaaa gcaaaatctc aggctttgag ggaagttaac ataggcttga 8100

ttctgggtgg aagctgggtg tgtagttatc tggaggccag gctggagctc tcagctcact 8160

atgggttcat ctttattgtc tccttttcc aggggcctgt cggacccagt tcatgccgta 8220

gttggtgaag gtgtagccgc tggcggcaca gctgattctg acagatccgc caggtttcac 8280

aagtccgccg ccagactgaa ccagctggat ctcagagatg ctacaggcca ctgttcccag 8340

cagcagcaga gactgcagcc acatctggtg gcgaattcga agcttgagct cgagatctga 8400

gtccggtagc gctagcggat ctgacggttc actaaaccag ctctgcttat atagacctcc 8460

caccgtacac gcctaccgcc catttgcgtc aatggggcgg agttgttacg acattttgga 8520

aagtcccgtt gattttggtg ccaaaacaaa ctcccattga cgtcaatggg gtggagactt 8580

ggaaatcccc gtgagtcaaa ccgctatcca cgcccattga tgtactgcca aaaccgcatc 8640

accatggtaa tagcgatgac taatacgtag atgtactgcc aagtaggaaa gtcccataag 8700

gtcatgtact gggcataatg ccaggcgggc catttaccgt cattgacgtc aataggggggc 8760

gtacttggca tatgatacac ttgatgtact gccaagtggg cagtttaccg taaatactcc 8820

acccattgac gtcaatggaa agtccctatt ggcgttacta tgggaacata cgtcattatt 8880

gacgtcaatg ggcgggggtc gttgggcggt cagccaggcg ggccatttac cgtaagttat 8940

gtaacgcgga actccatata tgggctatga actaatgacc ccgtaattga ttactattag 9000

cccgggggat ccagacatga taagatacat tgatgagttt ggacaaacca caactagaat 9060

gcagtgaaaa aaatgcttta tttgtgaaat ttgtgatgct attgctttat ttgtaaccat 9120

tataagctgc aataaacaag ttaacaacaa caattgcatt cattttatgt ttcaggttca 9180

gggggaggtg tgggaggttt tttaaagcaa gtaaaacctc tacaaatgtg gtatggctga 9240

ttatgatccg gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg acacatgcag 9300

ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca agcccgtcag 9360

ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgaggtcgat cgactctaga 9420

ggatcgatcc ccgccccgga cgaactaaac ctgactacga catctctgcc ccttcttcgc 9480

ggggcagtgc atgtaatccc ttcagttggt tggtacaact tgccaactgg gccctgttcc 9540

acatgtgaca cgggggggga ccaaacacaa aggggttctc tgactgtagt tgacatcctt 9600

ataaatggat gtgcacattt gccaacactg agtggctttc atcctggagc agactttgca 9660

gtctgtggac tgcaacacaa cattgccttt atgtgtaact cttggctgaa gctcttacac 9720

caatgctggg ggacatgtac ctcccagggg cccaggaaga ctacgggagg ctacaccaac 9780

gtcaatcaga ggggcctgtg tagctaccga taagcggacc ctcaagaggg cattagcaat 9840

agtgtttata aggccccctt gttaacccta aacgggtagc atatgcttcc cgggtagtag 9900

tatatactat ccagactaac cctaattcaa tagcatatgt tacccaacgg gaagcatatg 9960

ctatcgaatt agggttagta aaagggtcct aaggaacagc gatatctccc accccatgag 10020

ctgtcacggt tttatttaca tggggtcagg attccacgag ggtagtgaac cattttagtc 10080
acaagggcag tggctgaaga tcaaggagcg ggcagtgaac tctcctgaat cttcgcctgc 10140
ttcttcattc tccttcgttt agctaataga ataactgctg agttgtgaac agtaaggtgt 10200
atgtgaggtg ctcgaaaaca aggtttcagg tgacgccccc agaataaaat ttggacgggg 10260
ggttcagtgg tggcattgtg ctatgacacc aatataaccc tcacaaaccc cttgggcaat 10320
aaatactagt gtaggaatga aacattctga atatctttaa caatagaaat ccatggggtg 10380
gggacaagcc gtaaagactg gatgtccatc tcacacgaat ttatggctat gggcaacaca 10440

taatcctagt gcaatatgat actggggtta ttaagatgtg tcccaggcag ggaccaagac 10500
aggtgaacca tgttgttaca ctctatttgt aacaagggga aagagagtgg acgccgacag 10560
cagcggactc cactggttgt ctctaacacc cccgaaaatt aaacggggct ccacgccaat 10620
ggggcccata aacaaagaca agtggccact ctttttttg aaattgtgga gtgggggcac 10680
gcgtcagccc ccacacgccg ccctgcggtt ttggactgta aaataagggt gtaataactt 10740
ggctgattgt aaccccgcta accactgcgg tcaaaccact tgcccacaaa accactaatg 10800
gcaccccggg gaatacctgc ataagtaggt gggcgggcca agataggggc gcgattgctg 10860

cgatctggag gacaaattac acacacttgc gcctgagcgc caagcacagg gttgttggtc 10920
ctcatattca cgaggtcgct gagagcacgg tgggctaatg ttgccatggg tagcatatac 10980
tacccaaata tctggatagc atatgctatc ctaatctata tctgggtagc ataggctatc 11040
ctaatctata tctgggtagc atatgctatc ctaatctata tctgggtagt atatgctatc 11100
ctaatttata tctgggtagc ataggctatc ctaatctata tctgggtagc atatgctatc 11160
ctaatctata tctgggtagt atatgctatc ctaatctgta tccgggtagc atatgctatc 11220
ctaatagaga ttagggtagt atatgctatc ctaatttata tctgggtagc atatactacc 11280

caaatatctg gatagcatat gctatcctaa tctatatctg ggtagcatat gctatcctaa 11340
tctatatctg ggtagcatag gctatcctaa tctatatctg ggtagcatat gctatcctaa 11400
tctatatctg ggtagtatat gctatcctaa tttatatctg ggtagcatag gctatcctaa 11460
tctatatctg ggtagcatat gctatcctaa tctatatctg ggtagtatat gctatcctaa 11520
tctgtatccg ggtagcatat gctatcctca tgcatataca gtcagcatat gatacccagt 11580
agtagagtgg gagtgctatc ctttgcatat gccgccacct cccaaggggg cgtgaatttt 11640
cgctgcttgt ccttttcctg catgctggtt gctcccattc ttaggtgaat ttaaggaggc 11700

caggctaaag ccgtcgcatg tctgattgct caccaggtaa atgtcgctaa tgttttccaa 11760

cgcgagaagg tgttgagcgc ggagctgagt gacgtgacaa catgggtatg cccaattgcc 11820

ccatgttggg aggacgaaaa tggtgacaag acagatggcc agaaatacac caacagcacg 11880

catgatgtct actggggatt tattctttag tgcgggggaa tacacggctt ttaatacgat 11940

tgagggcgtc tcctaacaag ttacatcact cctgcccttc ctcaccctca tctccatcac 12000

ctccttcatc tccgtcatct ccgtcatcac cctccgcggc agccccttcc accataggtg 12060

gaaaccaggg aggcaaatct actccatcgt caaagctgca cacagtcacc ctgatattgc 12120

aggtaggagc gggctttgtc ataacaaggt ccttaatcgc atccttcaaa acctcagcaa 12180

atatatgagt ttgtaaaaag accatgaaat aacagacaat ggactccctt agcgggccag 12240

gttgtgggcc gggtccaggg gccattccaa aggggagacg actcaatggt gtaagacgac 12300

attgtggaat agcaagggca gttcctcgcc ttaggttgta aagggaggtc ttactacctc 12360

catatacgaa cacaccggcg acccaagttc cttcgtcggt agtcctttct acgtgactcc 12420

tagccaggag agctcttaaa ccttctgcaa tgttctcaaa tttcgggttg gaacctcctt 12480

gaccacgatg ctttccaaac caccctcctt ttttgcgcct gcctccatca ccctgacccc 12540

ggggtccagt gcttgggcct tctcctgggt catctgcggg gccctgctct atcgctcccg 12600

ggggcacgtc aggctcacca tctgggccac cttcttggtg gtattcaaaa taatcggctt 12660

cccctacagg gtggaaaaat ggccttctac ctggaggggg cctgcgcggt ggagacccgg 12720

atgatgatga ctgactactg ggactcctgg gcctcttttc tccacgtcca cgacctctcc 12780

ccctggctct ttcacgactt ccccccctgg ctctttcacg tcctctaccc cggcggcctc 12840

cactacctcc tcgaccccgg cctccactac ctcctcgacc ccggcctcca ctgcctcctc 12900

gaccccggcc tccacctcct gctcctgccc ctcctgctcc tgcccctcct cctgctcctg 12960

cccctcctgc ccctcctgct cctgcccctc ctgcccctcc tgctcctgcc cctcctgccc 13020

ctcctgctcc tgcccctcct gcccctcctc ctgctcctgc ccctcctgcc cctcctcctg 13080

ctcctgcccc tcctgcccct cctgctcctg cccctcctgc ccctcctgct cctgcccctc 13140

ctgcccctcc tgctcctgcc cctcctgctc ctgcccctcc tgctcctgcc cctcctgctc 13200

ctgcccctcc tgcccctcct gcccctcctc ctgctcctgc ccctcctgct cctgcccctc 13260

ctgcccctcc tgcccctcct gctcctgccc ctcctcctgc tcctgcccct cctgcccctc 13320

ctgcccctcc tcctgctcct gcccctcctg cccctcctcc tgctcctgcc cctcctcctg 13380

ctcctgcccc tcctgcccct cctgcccctc ctcctgctcc tgcccctcct gcccctcctc 13440

ctgctcctgc ccctcctcct gctcctgccc ctcctgcccc tcctgcccct cctcctgctc 13500

ctgcccctcc tcctgctcct gcccctcctg cccctcctgc ccctcctgcc cctcctcctg 13560

ctcctgcccc tcctcctgct cctgcccctc ctgctcctgc ccctcccgct cctgctcctg 13620

ctcctgttcc accgtgggtc cctttgcagc caatgcaact tggacgtttt tggggtctcc 13680

ggacaccatc tctatgtctt ggccctgatc ctgagccgcc cggggctcct ggtcttccgc 13740

ctcctcgtcc tcgtcctctt ccccgtcctc gtccatggtt atcacccccct cttctttgag 13800

gtccactgcc gccggagcct tctggtccag atgtgtctcc cttctctcct aggccatttc 13860

caggtcctgt acctggcccc tcgtcagaca tgattcacac taaaagagat caatagacat 13920

ctttattaga cgacgctcag tgaatacagg gagtgcagac tcctgccccc tccaacagcc 13980

cccccaccct catccccttc atggtcgctg tcagacagat ccaggtctga aaattcccca 14040

tcctccgaac catcctcgtc ctcatcacca attactcgca gcccggaaaa ctcccgctga 14100

acatcctcaa gatttgcgtc ctgagcctca agccaggcct caaattcctc gtcccccttt 14160

ttgctggacg gtagggatgg ggattctcgg gacccctcct cttcctcttc aaggtcacca 14220

gacagagatg ctactggggc aacggaagaa aagctgggtg cggcctgtga ggatcagctt 14280

atcgatgata agctgtcaaa catgagaatt cttgaagacg aaagggcctc gtgatacgcc 14340

tatttttata ggttaatgtc atgataataa tggtttctta gacgtcaggt ggcacttttc 14400

ggggaaatgt gcgcggaacc cctatttgtt tatttttcta aatacattca aatatgtatc 14460

cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga 14520

gtattcaaca tttccgtgtc gcccttattc cctttttgc ggcattttgc cttcctgttt 14580

ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag 14640

tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag 14700

aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgtg 14760

ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg 14820

agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca 14880

gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca acgatcggag 14940

gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc 15000

gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg 15060

cagcaatggc aacaacgttg cgcaaactat taactggcga actacttact ctagcttccc 15120

ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg 15180

cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt gggtctcgcg 15240

gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga 15300

cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 15360

tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag attgatttaa 15420

aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat ctcatgacca 15480

aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag 15540
gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac 15600
cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa 15660
ctggcttcag cagagcgcag ataccaaata ctgtccttct agtgtagccg tagttaggcc 15720
accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatc ctgttaccag 15780
tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac 15840
cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttggagc 15900

gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc 15960
ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca 16020
cgagggagct tccagggggа aacgcctggt atctttatag tcctgtcggg tttcgccacc 16080
tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg 16140
ccagcaacgc ggccttttta cggttcctgg ccttttgctg gccttttgct cacatgttct 16200
ttcctgcgtt atcccctgat tctgtggata accgtattac cgcctttgag tgagctgata 16260
ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa gcggaagagc 16320

gcctgatgcg gtattttctc cttacgcatc tgtgcggtat ttcacaccgc atatggtgca 16380
ctctcagtac aatctgctct gatgccgcat agttaagcca gctgtggaat gtgtgtcagt 16440
tagggtgtgg aaagtcccca ggctccccag caggcagaag tatgcaaagc atgcatctca 16500
attagtcagc aaccaggtgt ggaaagtccc caggctcccc agcaggcaga agtatgcaaa 16560
gcatgcatct caattagtca gcaaccatag tcccgcccct aactccgccc atcccgcccc 16620
taactccgcc cagttccgcc cattctccgc cccatggctg actaattttt tttatttatg 16680
cagaggccga ggccgcctcg gcctctgagc tattccagaa gtagtgagga ggcttttttg 16740

gaggcctagg cttttgcaaa aagcttgcat gcctgcaggt cggccgccac gaccggtgcc 16800
gccaccatcc cctgacccac gcccctgacc cctcacaagg agacgacctt ccatgaccga 16860
gtacaagccc acggtgcgcc tcgccacccg cgacgacgtc ccccgggccg tacgcaccct 16920
cgccgccgcg ttcgccgact accccgccac gcgccacacc gtcgacccgg accgccacat 16980
cgagcgggtc accgagctgc aagaactctt cctcacgcgc gtcgggctcg acatcggcaa 17040
ggtgtgggtc gcggacgacg gcgccgcggt ggcggtctgg accacgccgg agagcgtcga 17100
agcgggggcg gtgttcgccg agatcggccc gcgcatggcc gagttgagcg gttcccggct 17160

ggccgcgcag caacagatgg aaggcctcct ggcgccgcac cggcccaagg agcccgcgtg 17220

gttcctggcc accgtcggcg tctcgcccga ccaccagggc aagggtctgg gcagcgccgt 17280

cgtgctcccc ggagtggagg cggccgagcg cgccggggtg cccgccttcc tggagacctc 17340

cgcgccccgc aacctcccct tctacgagcg gctcggcttc accgtcaccg ccgacgtcga 17400

ggtgcccgaa ggaccgcgca cctggtgcat gacccgcaag cccggtgcct gacgcccgcc 17460

ccacgacccg cagcgcccga ccgaaaggag cgcacgaccc catggctccg accgaagccg 17520

acccgggcgg ccccgccgac cccgcacccg cccccgaggc ccaccgactc tagaggatca 17580

taatcagcca taccacattt gtagaggttt tacttgcttt aaaaaacctc ccacacctcc 17640

ccctgaacct gaaacataaa atgaatgcaa ttgttgttgt taacttgttt attgcagctt 17700

ataatggtta caaataaagc aatagcatca caaatttcac aaataaagca ttttttttcac 17760

tgcattctag ttgtggtttg tccaaactca tcaatgtatc ttatcatgtc tggatcactc 17820

gccgatagtg gaaaccgacg ccccagcact cgtccgaggg caaaggaata ggggagatgg 17880

gggaggctaa ctgaaacacg gaaggagaca ataccggaag gaacccgcgc tatgacggca 17940

ataaaaagac agaataaaac gcacgggtgt tgggtcgttt gttcataaac gcggggttcg 18000

gtcccagggc tggcactctg tcgatacccc accgagaccc cattggggcc aatacgcccg 18060

cgtttcttcc ttttccccac cccacccccc aagttcgggt gaaggcccag ggctcgcagc 18120

caacgtcggg gcggcaggcc ctgccatagc cactggcccc gtgggttagg gacggggtcc 18180

cccatgggga atggtttatg gttcgtgggg gttattattt tgggcgttgc gtggggtctg 18240

gtccacgact ggactgagca gacagaccca tggtttttgg atggcctggg catggaccgc 18300

atgtactggc gcgacacgaa caccgggcgt ctgtggctgc caaacacccc cgacccccaa 18360

aaaccaccgc gcggatttct ggcgtgccaa gctagtcgac caattctcat gtttgacagc 18420

ttatcatcgc agatccgggc aacgttgttg cattgctgca ggcgcagaac tggtaggtat 18480

ggaagatctc tagaagctgg gtaccagctg ctagcaagct tgctagcggc cggctcgagt 18540

ttactcccta tcagtgatag agaacgtatg tcgagtttac tccctatcag tgatagagaa 18600

cgatgtcgag tttactccct atcagtgata gagaacgtat gtcgagttta ctccctatca 18660

gtgatagaga acgtatgtcg agtttactcc ctatcagtga tagagaacgt atgtcgagtt 18720

tatccctatc agtgatagag aacgtatgtc gagtttactc cctatcagtg atagagaacg 18780

tatgtcgagg taggcgtgta cggtgggagg cctatataag cagagctcgt ttagtgaacc 18840

gtcagatcgc cg 18852

<210> 40

<211> 19625

<212> DNA

<213> Artificial Sequence

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

<400> 40

cggccgcggg gggaggagcc aagatggccg aataggaaca gctccggtct acagctccca 60

gcgtgagcga cgcagaagac ggtgatttct gcatttccat ctgaggtacc gggttcatct 120

cactagggag tgccagacag tgggcgcagg ccagtgtgtg tgcgcaccgt gcgcgagccg 180

aagcagggcg aggcattgcc tcacctggga agcgcaaggg gtcagggagt tccctttccg 240

agtcaaagaa aggggtgacg gacgcacctg gaaaatcggg tcactcccac ccgaatattg 300

cgcttttcag accggcttaa gaaacggcgc accacgagac tatatcccac acctggctcg 360

gagggtccta cgcccacgga atctcgctga ttgctagcac agcagtctga gatcaaactg 420

caaggcggca acgaggctgg gggaggggcg cccgccattg cccaggcttg cttaggtaaa 480

caaagcagca gggaagctcg aactgggtgg agcccaccac agctcaagga ggcctgcctg 540

cctctgtagg ctccacctct gggggcaggg cacagacaaa caaaaagaca gcagtaacct 600

ctgcagactt aagtgtccct gtctgacagc tttgaagaga gcagtggttc tcccagcacg 660

cagctggaga tctgagaacg ggcagactgc ctcctcaagt gggtccctga cccctgaccc 720

ccgagcagcc taactgggag gcacccccca gcaggggcac actgacacct cacacggcag 780

ggtattccaa cagacctgca gctgagggtc ctgtctgtta gaaggaaaac taacaaccag 840

aaaggacatc tacaccgaaa acccatctgt acatcaccat catcaaagac caaaagtaga 900

taaaaccaca aagatgggga aaaaacagaa cagaaaaact ggaaactcta aaacgcagag 960

cgcctctcct cctccaaagg aacgcagttc ctcaccagca acagaacaaa gctggatgga 1020

gaatgatttt gatgagctga gagaagaagg cttcagacga tcaaattact ctgagctacg 1080

ggaggacatt caaaccaaag gcaaagaagt tgaaaacttt gaaaaaaatt tagaagaatg 1140

tataactaga ataaccaata cagagaagtg cttaaaggag ctgatggagc tgaaaaccaa 1200

ggctcgagaa ctacgtgaag aatgcagaag cctcaggagc cgatgcgatc aactggaaga 1260

aagggtatca gcaatggaag atgaaatgaa tgaaatgaag cgagaaggga agtttagaga 1320

aaaaagaata aaaagaaatg agcaaagcct ccaagaaata tgggactatg tgaaaagacc 1380

aaatctacgt ctgattggtg tacctgaaag tgatgtggag aatggaacca agttggaaaa 1440

cactctgcag gatattatcc aggagaactt ccccaatcta gcaaggcagg ccaacgttca 1500

gattcaggaa atacagagaa cgccacaaag atactcctcg agaagagcaa ctccaagaca 1560

cataattgtc agattcacca aagttgaaat gaaggaaaaa atgttaaggg cagccagaga 1620

gaaaggtcgg gttaccctca aaggaaagcc catcagacta acagcggatc tctcggcaga 1680

aaccctacaa gccagaagag agtgggggcc aatattcaac attcttaaag aaaagaattt 1740

tcaacccaga atttcatatc cagccaaact aagcttcata agtgaaggag aaataaaata 1800

ctttatagac aagcaaatgt tgagagattt tgtcaccacc aggcctgccc taaaagagct 1860

cctgaaggaa gcgctaaaca tggaaaggaa caaccggtac cagccgctgc aaaatcatgc 1920

caaaatgtaa agaccatcaa gactaggaag aaactgcatc aactaatgag caaaatcacc 1980

agctaacatc ataatgacag gatcaacttc acacataaca atattaactt taaatataaa 2040

tggactaaat tctgcaatta aaagacacag actggcaagt tggataaaga gtcaagaccc 2100

atcagtgtgc tgtattcagg aaacccatct cacgtgcaga gacacacata ggctcaaaat 2160

aaaaggatgg aggaagatct accaagccaa tggaaaacaa aaaaaggcag gggttgcaat 2220

cctagtctct gataaaacag actttaaacc aacaaagatc aaaagagaca aagaaggcca 2280

ttacataatg gtaaagggat caattcaaca agaggagcta actatcctaa atatttatgc 2340

acccaataca ggagcaccca gattcataaa gcaagtcctc agtgacctac aaagagactt 2400

agactcccac acattaataa tgggagactt taacacccca ctgtcaacat tagacagatc 2460

aacgagacag aaagtcaaca aggataccca ggaattgaac tcagctctgc accaagcaga 2520

cctaatagac atctacagaa ctctccaccc caaatcaaca gaatatacat ttttttcagc 2580

accacaccac acctattcca aaattgacca catagttgga agtaaagctc tcctcagcaa 2640

atgtaaaaga acagaaatta taacaaacta tctctcagac cacagtgcaa tcaaactaga 2700

actcaggatt aagaatctca ctcaaagccg ctcaactaca tggaaactga acaacctgct 2760

cctgaatgac tactgggtac ataacgaaat gaaggcagaa ataaagatgt tctttgaaac 2820

caacgagaac aaagacacca cataccagaa tctctgggac gcattcaaag cagtgtgtag 2880

agggaaattt atagcactaa atgcctacaa gagaaagcag gaaagatcca aaattgacac 2940

cctaacatca caattaaaag aactagaaaa gcaagagcaa acacattcaa aagctagcag 3000

aaggcaagaa ataactaaaa tcagagcaga actgaaggaa atagagacac aaaaaaccct 3060

tcaaaaaatc aatgaatcca ggagctggtt ttttgaaagg atcaacaaaa ttgatagacc 3120

gctagcaaga ctaataaaga aaaaaagaga gaagaatcaa atagacacaa taaaaaatga 3180

taaaggggat atcaccaccg atcccacaga aatacaaact accatcagag aatactacaa 3240

acacctctac gcaaataaac tagaaaatct agaagaaatg gatacattcc tcgacacata 3300

cactctccca agactaaacc aggaagaagt tgaatctctg aatcgaccaa taacaggctc 3360

tgaaattgtg gcaataatca atagtttacc aaccaaaaag agtccaggac cagatggatt 3420

cacagccgaa ttctaccaga ggtacaagga ggaactggta ccattccttc tgaaactatt 3480

ccaatcaata gaaaaagagg gaatcctccc taactcattt tatgaggcca gcatcattct 3540

gataccaaag ccgggcagag acacaaccaa aaaagagaat tttagaccaa tatccttgat 3600

gaacattgat gcaaaaatcc tcaataaaat actggcaaac cgaatccagc agcacatcaa 3660

aaagcttatc caccatgatc aagtgggctt catccctggg atgcaaggct ggttcaatat 3720

acgcaaatca ataaatgtaa tccagcatat aaacagagcc aaagacaaaa accacatgat 3780

tatctcaata gatgcagaaa aagcctttga caaaattcaa caacccttca tgctaaaaac 3840

tctcaataaa ttaggtattg atgggacgta tttcaaaata ataagagcta tctatgacaa 3900

acccacagcc aatatcatac tgaatgggca aaaactggaa gcattccctt tgaaaaccgg 3960

cacaagacag ggatgccctc tctcaccgct cctattcaac atagtgttgg aagttctggc 4020

cagggcaatc aggcaggaga aggaaataaa gggtattcaa ttaggaaaag aggaagtcaa 4080

attgtccctg tttgcagacg acatgattgt ttatctagaa aaccccatcg tctcagccca 4140

aaatctcctt aagctgataa gcaacttcag caaagtctca ggatacaaaa tcaatgtaca 4200

aaaatcacaa gcattcttat acaccaacaa cagacaaaca gagagccaaa tcatgggtga 4260

actcccattc acaattgctt caaagagaat aaaataccta ggaatccaac ttacaaggga 4320

tgtgaaggac ctcttcaagg agaactacaa accactgctc aaggaaataa aagaggagac 4380

aaacaaatgg aagaacattc catgctcatg ggtaggaaga atcaatatcg tgaaaatggc 4440

catactgccc aaggtaattt acagattcaa tgccatcccc atcaagctac caatgacttt 4500

cttcacagaa ttggaaaaaa ctactttaaa gttcatatgg aaccaaaaaa gagcccgcat 4560

tgccaagtca atcctaagcc aaaagaacaa agctggaggc atcacactac ctgacttcaa 4620

actatactac aaggctacag taaccaaaac agcatggtac tggtaccaaa acagagatat 4680

agatcaatgg aacagaacag agccctcaga aataatgccg catatctaca actatctgat 4740

ctttgacaaa cctgagaaaa acaagcaatg gggaaaggat tccctattta ataaatggtg 4800

ctgggaaaac tggctagcca tatgtagaaa gctgaaactg gatcccttcc ttacacctta 4860

tacaaaaatc aattcaagat ggattaaaga tttaaacgtt aaacctaaaa ccataaaaac 4920

cctagaagaa aacctaggca ttaccattca ggacataggc gtgggcaagg acttcatgtc 4980

caaaacacca aaagcaatgg caacaaaaga caaaattgac aaatgggatc taattaaact 5040

aaagagcttc tgcacagcaa aagaaactac catcagagtg aacaggcaac ctacaacatg 5100

ggagaaaatt tttgcaacct actcatctga caaagggcta atatccagaa tctacaatga 5160

actcaaacaa atttacaaga aaaaaacaaa caacccatc aaaaagtggg cgaaggacat 5220

gaacagacac ttctcaaaag aagacattta tgcagccaaa aaacacatga agaaatgctc 5280

atcatcactg gccatcagag aaatgcaaat caaaaccact atgagatatc atctcacacc 5340
agttagaatg gcaatcatta aaaagtcagg aaacaacagg tgctggagag gatgcggaga 5400
aataggaaca cttttacact gttggtggga ctgtaaacta gttcaaccat tgtggaagtc 5460
agtgtggcga ttcctcaggg atctagaact agaaatacca tttgacccag ccatcccatt 5520
actgggtata tacccaaatg agtataaatc atgctgctat aaagacacat gcacacgtat 5580
gtttattgcg gcactattca caatagcaaa gacttggaac caacccaaat gtccaacaat 5640
gatagactgg attaagaaaa tgtggcacat atacaccatg gaatactatg cagccataaa 5700

aaatgatgag ttcatatcct ttgtagggac atggatgaaa ttggaaacca tcattctcag 5760
taaactatcg caagaacaaa aaaccaaaca ccgcatattc tcactcatag gtgggaattg 5820
aacaatgaga tcacatggac acaggaaggg gaatatcaca ctctggggac tgtggtgggg 5880
tcgggggagg ggggagggat agcattggga gatataccta atgctagatg acacattagt 5940
gggtgcagcg caccagcatg gcacatgtat acggatccga attctcgacg gatcgatccg 6000
aacaaacgac ccaacacccg tgcgttttat tctgtctttt tattgccgat cccctcagaa 6060
gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta 6120

aagcacgagg aagcggtcag cccattcgcc gccaagctct tcagcaatat cacgggtagc 6180
caacgctatg tcctgatagc ggtcggccgc tttacttgta cagctcgtcc atgccgagag 6240
tgatcccggc ggcggtcacg aactccagca ggaccatgtg atcgcgcttc tcgttggggt 6300
ctttgctcag ggcggactgg gtgctcaggt agtggttgtc gggcagcagc acggggccgt 6360
cgccgatggg ggtgttctgc tggtagtggt cggccaggtg agtccaggag atgtttcagc 6420
actgttgcct ttagtctcga ggcaacttag acaactgagt attgatctga gcacagcagg 6480
gtgtgagctg tttgaagata ctggggttgg gggtgaagaa actgcagagg actaactggg 6540

ctgagaccca gtggcaatgt tttagggcct aaggaatgcc tctgaaaatc tagatggaca 6600
actttgactt tgagaaaaga gaggtggaaa tgaggaaaat gacttttctt tattagattt 6660
cggtagaaag aactttcatc tttcccctat ttttgttatt cgttttaaaa catctatctg 6720
gaggcaggac aagtatggtc attaaaaaga tgcaggcaga aggcatatat tggctcagtc 6780
aaagtgggga actttggtgg ccaaacatac attgctaagg ctattcctat atcagctgga 6840
cacatataaa atgctgctaa tgcttcatta caaacttata tcctttaatt ccagatgggg 6900
gcaaagtatg tccaggggtg aggaacaatt gaaacatttg ggctggagta gattttgaaa 6960

gtcagctctg tgtgtgtgtg tgtgtgtgtg tgtgtgagag cgtgtgtttc ttttaacgtt 7020

ttcagcctac agcatacagg gttcatggtg gcaagaagat aacaagattt aaattatggc 7080

cagtgactag tgctgcaaga agaacaacta cctgcattta atgggaaagc aaaatctcag 7140

gctttgaggg aagttaacat aggcttgatt ctgggtggaa gctgggtgtg tagttatctg 7200

gaggccaggc tggagctctc agctcactat gggttcatct ttattgtctc ctttcatctc 7260

aacagctgca cgctgccgtc ctcgatgttg tggcggatct tgaagttcac cttgatgccg 7320

ttcttctgct tgtcggccat gatatagacg ttgtggctgt tgtagttgta ctccagcttg 7380

tgccccagga tgttgccgtc ctccttgaag tcgatgccct tcagctcgat gcggttcacc 7440

agggtgtcgc cctcgaactt cacctcggcg cgggtcttgt agttgccgtc gtccttgaag 7500

aagatggtgc gctcctggac gtagccttcg ggcatggcgg acttgaagaa gtcgtgctgc 7560

ttcatgtggt cggggtagcg gctgaagcac tgcacgccgt aggtcagggt ggtcacgagg 7620

gtgggccagg gcacgggcag cttgccggtg gtgcagatga acttcagggt cagcttgccg 7680

taggtggcat cgccctcgcc ctcgccggac acgctgaact tgtggccgtt tacgtcgccg 7740

tccagctcga ccaggatggg caccaccccg gtgaacagct cctcgccctt gctcaccata 7800

gggccgggat tctcctccac gtcaccgcat gttagaagac ttcctctgcc ctccatgttc 7860

tcgtaggagt cggcgtcctc ttcgtggtta ggtccaggtt ggcctctgat agaccgcagc 7920

tgaggagcgg cgtacagaat gcctctcatg tcctcatagc tgccgctgcc ttgtggaggc 7980

ttctcgtgct tcagtgtctc gtatgtctct tgattccggg tgctcaggcc ggtgtacacg 8040

ccatcagatt tctcgtagct ggtgatggcg gccttccgca cttggatctt cagccgtctg 8100

cagtacaggg tgatgaccag agacagcagc aggacaccac atgtgccagc cagaggggcc 8160

caaatgtaga tatccaggcc tctggtatgc acagctccgc ctgcagcagg tctacaggct 8220

tcaggtctga gagacagagg ctggctggcg attgtaggag ctggtgtagg tggtctagga 8280

gcgggtgttg ttgtaggctt ggcgggcaga aacacgggca cgaagtggct gaagtacatg 8340

atgctattgc tcagggctcc gcttcctccg ccgcctgatt tgatttccag cttggtgcct 8400

ccgccaaatg tccaagggct ctcgtcgtac tgctggcagt agtagatgcc gaagtcctcg 8460

tactgcaggc tgctgattgt cagggtgtag tcggtgccag agccgctgcc agaaaatctg 8520

cttggcacgc cgctttccag tctgttggcc cggtagatca gtgtcttagg ggccttgcca 8580

ggcttctgct ggaaccagct caggtagctg ttgatgtcct ggctggctct acaggtgatg 8640

gtcactctat cgcccacaga ggcagacagg ctgctagggc tctgtgtcat ctggatatca 8700

gagccaccac cgccagatcc accgccacct gatcctccgc ctccgctaga aactgtcact 8760

gtggtgccct ggccccacac atcgaagtac cagtcgtagc ctcttctggt gcagaagtac 8820

acggcggtat cctcggctct caggctgttg atctgcaggt aggcggtgtt cttgctgtcg 8880

tccaggctga aggtgaatct gcccttaaag ctatcggcgt aggttggctc gccggtgtgg 8940

gtattgatcc agcccatcca ctcaaggcct tttccagggg cctgtcggac ccagttcatg 9000

ccgtagttgg tgaaggtgta gccgctggcg gcacagctga ttctgacaga tccgccaggt 9060

ttcacaagtc cgccgccaga ctgaaccagc tggatctcag agatgctaca ggccactgtt 9120

cccagcagca gcagagactg cagccacatt cgaagcttga gctcgagatc tgagtccggt 9180

agcgctagcg gatctgacgg ttcactaaac cagctctgct tatatagacc tcccaccgta 9240

cacgcctacc gcccatttgc gtcaatgggg cggagttgtt acgacatttt ggaaagtccc 9300

gttgattttg gtgccaaaac aaactcccat tgacgtcaat ggggtggaga cttggaaatc 9360

cccgtgagtc aaaccgctat ccacgcccat tgatgtactg ccaaaaccgc atcaccatgg 9420

taatagcgat gactaatacg tagatgtact gccaagtagg aaagtcccat aaggtcatgt 9480

actgggcata atgccaggcg ggccatttac cgtcattgac gtcaataggg ggcgtacttg 9540

gcatatgata cacttgatgt actgccaagt gggcagttta ccgtaaatac tccacccatt 9600

gacgtcaatg gaaagtccct attggcgtta ctatgggaac atacgtcatt attgacgtca 9660

atgggcgggg gtcgttgggc ggtcagccag gcgggccatt taccgtaagt tatgtaacgc 9720

ggaactccat atatgggcta tgaactaatg accccgtaat tgattactat tagccccggg 9780

gatccagaca tgataagata cattgatgag tttggacaaa ccacaactag aatgcagtga 9840

aaaaaatgct ttatttgtga aatttgtgat gctattgctt tatttgtaac cattataagc 9900

tgcaataaac aagttaacaa caacaattgc attcatttta tgtttcaggt tcaggggggag 9960

gtgtgggagg ttttttaaag caagtaaaac ctctacaaat gtggtatggc tgattatgat 10020

ccggctgcct cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg cagctcccgg 10080

agacggtcac agcttgtctg taagcggatg ccgggagcag acaagcccgt cagggcgcgt 10140

cagcgggtgt tggcgggtgt cggggcgcag ccatgaggtc gatcgactct agaggatcga 10200

tccccgcccc ggacgaacta aacctgacta cgacatctct gcccccttctt cgcggggcag 10260

tgcatgtaat cccttcagtt ggttggtaca acttgccaac tgggccctgt tccacatgtg 10320

acacgggggg ggaccaaaca caaaggggtt ctctgactgt agttgacatc cttataaatg 10380

gatgtgcaca tttgccaaca ctgagtggct ttcatcctgg agcagacttt gcagtctgtg 10440

gactgcaaca caacattgcc tttatgtgta actcttggct gaagctctta caccaatgct 10500

gggggacatg tacctcccag gggcccagga agactacggg aggctacacc aacgtcaatc 10560

agaggggcct gtgtagctac cgataagcgg accctcaaga gggcattagc aatagtgttt 10620

ataaggcccc cttgttaacc ctaaacgggt agcatatgct tcccgggtag tagtatatac 10680

tatccagact aaccctaatt caatagcata tgttacccaa cgggaagcat atgctatcga 10740

attagggtta gtaaaagggt cctaaggaac agcgatatct cccaccccat gagctgtcac 10800
ggttttattt acatggggtc aggattccac gagggtagtg aaccatttta gtcacaaggg 10860
cagtggctga agatcaagga gcgggcagtg aactctcctg aatcttcgcc tgcttcttca 10920
ttctccttcg tttagctaat agaataactg ctgagttgtg aacagtaagg tgtatgtgag 10980
gtgctcgaaa acaaggtttc aggtgacgcc cccagaataa aatttggacg gggggttcag 11040
tggtggcatt gtgctatgac accaatataa ccctcacaaa ccccttgggc aataaatact 11100
agtgtaggaa tgaaacattc tgaatatctt taacaataga aatccatggg gtggggacaa 11160

gccgtaaaga ctggatgtcc atctcacacg aatttatggc tatgggcaac acataatcct 11220
agtgcaatat gatactgggg ttattaagat gtgtcccagg cagggaccaa gacaggtgaa 11280
ccatgttgtt acactctatt tgtaacaagg ggaaagagag tggacgccga cagcagcgga 11340
ctccactggt tgtctctaac acccccgaaa attaaacggg gctccacgcc aatggggccc 11400
ataaacaaag acaagtggcc actctttttt ttgaaattgt ggagtggggg cacgcgtcag 11460
cccccacacg ccgccctgcg gttttggact gtaaaataag ggtgtaataa cttggctgat 11520
tgtaaccccg ctaaccactg cggtcaaacc acttgcccac aaaaccacta atggcacccc 11580

ggggaatacc tgcataagta ggtgggcggg ccaagatagg ggcgcgattg ctgcgatctg 11640
gaggacaaat tacacacact tgcgcctgag cgccaagcac agggttgttg gtcctcatat 11700
tcacgaggtc gctgagagca cggtgggcta atgttgccat gggtagcata tactacccaa 11760
atatctggat agcatatgct atcctaatct atatctgggt agcataggct atcctaatct 11820
atatctgggt agcatatgct atcctaatct atatctgggt agtatatgct atcctaattt 11880
atatctgggt agcataggct atcctaatct atatctgggt agcatatgct atcctaatct 11940
atatctgggt agtatatgct atcctaatct gtatccgggt agcatatgct atcctaatag 12000

agattagggt agtatatgct atcctaattt atatctgggt agcatatact acccaaatat 12060
ctggatagca tatgctatcc taatctatat ctgggtagca tatgctatcc taatctatat 12120
ctgggtagca taggctatcc taatctatat ctgggtagca tatgctatcc taatctatat 12180
ctgggtagta tatgctatcc taatttatat ctgggtagca taggctatcc taatctatat 12240
ctgggtagca tatgctatcc taatctatat ctgggtagta tatgctatcc taatctgtat 12300
ccgggtagca tatgctatcc tcatgcatat acagtcagca tatgataccc agtagtagag 12360
tgggagtgct atcctttgca tatgccgcca cctcccaagg gggcgtgaat tttcgctgct 12420

tgtccttttc ctgcatgctg gttgctccca ttcttaggtg aatttaagga ggccaggcta 12480

aagccgtcgc atgtctgatt gctcaccagg taaatgtcgc taatgttttc caacgcgaga 12540

aggtgttgag cgcggagctg agtgacgtga caacatgggt atgcccaatt gccccatgtt 12600

gggaggacga aaatggtgac aagacagatg gccagaaata caccaacagc acgcatgatg 12660

tctactgggg atttattctt tagtgcgggg gaatacacgg cttttaatac gattgagggc 12720

gtctcctaac aagttacatc actcctgccc ttcctcaccc tcatctccat cacctccttc 12780

atctccgtca tctccgtcat caccctccgc ggcagcccct tccaccatag gtggaaacca 12840

gggaggcaaa tctactccat cgtcaaagct gcacacagtc accctgatat tgcaggtagg 12900

agcgggcttt gtcataacaa ggtccttaat cgcatccttc aaaacctcag caaatatatg 12960

agtttgtaaa aagaccatga aataacagac aatggactcc cttagcgggc caggttgtgg 13020

gccgggtcca ggggccattc caaaggggag acgactcaat ggtgtaagac gacattgtgg 13080

aatagcaagg gcagttcctc gccttaggtt gtaaagggag gtcttactac ctccatatac 13140

gaacacaccg gcgacccaag ttccttcgtc ggtagtcctt tctacgtgac tcctagccag 13200

gagagctctt aaaccttctg caatgttctc aaatttcggg ttggaacctc cttgaccacg 13260

atgctttcca aaccaccctc ctttttgcg cctgcctcca tcaccctgac cccggggtcc 13320

agtgcttggg ccttctcctg ggtcatctgc ggggccctgc tctatcgctc ccgggggcac 13380

gtcaggctca ccatctgggc caccttcttg gtggtattca aaataatcgg cttcccctac 13440

agggtggaaa aatggccttc tacctggagg gggcctgcgc ggtggagacc cggatgatga 13500

tgactgacta ctgggactcc tgggcctctt ttctccacgt ccacgacctc tccccctggc 13560

tctttcacga cttccccccc tggctctttc acgtcctcta ccccggcggc ctccactacc 13620

tcctcgaccc cggcctccac tacctcctcg accccggcct ccactgcctc ctcgaccccg 13680

gcctccacct cctgctcctg cccctcctgc tcctgcccct cctcctgctc ctgcccctcc 13740

tgcccctcct gctcctgccc ctcctgcccc tcctgctcct gcccctcctg cccctcctgc 13800

tcctgcccct cctgcccctc ctcctgctcc tgcccctcct gcccctcctc ctgctcctgc 13860

ccctcctgcc cctcctgctc ctgcccctcc tgcccctcct gctcctgccc ctcctgcccc 13920

tcctgctcct gcccctcctg ctcctgcccc tcctgctcct gcccctcctg ctcctgcccc 13980

tcctgcccct cctgcccctc ctcctgctcc tgcccctcct gctcctgccc ctcctgcccc 14040

tcctgcccct cctgctcctg cccctcctcc tgctcctgcc cctcctgccc ctcctgcccc 14100

tcctcctgct cctgcccctc ctgcccctcc tcctgctcct gcccctcctc ctgctcctgc 14160

ccctcctgcc cctcctgccc ctcctcctgc tcctgcccct cctgcccctc ctcctgctcc 14220

tgcccctcct cctgctcctg cccctcctgc ccctcctgcc cctcctcctg ctcctgcccc 14280

tcctcctgct cctgcccctc ctgcccctcc tgcccctcct gcccctcctc ctgctcctgc 14340

ccctcctcct gctcctgccc ctcctgctcc tgcccctccc gctcctgctc ctgctcctgt 14400

tccaccgtgg gtccctttgc agccaatgca acttggacgt ttttggggtc tccggacacc 14460

atctctatgt cttggccctg atcctgagcc gcccggggct cctggtcttc cgcctcctcg 14520

tcctcgtcct cttccccgtc ctcgtccatg gttatcaccc cctcttcttt gaggtccact 14580

gccgccggag ccttctggtc cagatgtgtc tcccttctct cctaggccat ttccaggtcc 14640

tgtacctggc ccctcgtcag acatgattca cactaaaaga gatcaataga catctttatt 14700

agacgacgct cagtgaatac agggagtgca gactcctgcc ccctccaaca gcccccccac 14760

cctcatcccc ttcatggtcg ctgtcagaca gatccaggtc tgaaaattcc ccatcctccg 14820

aaccatcctc gtcctcatca ccaattactc gcagcccgga aaactcccgc tgaacatcct 14880

caagatttgc gtcctgagcc tcaagccagg cctcaaattc ctcgtccccc ttttgctgg 14940

acggtaggga tggggattct cgggacccct cctcttcctc ttcaaggtca ccagacagag 15000

atgctactgg ggcaacggaa gaaaagctgg gtgcggcctg tgaggatcag cttatcgatg 15060

ataagctgtc aaacatgaga attcttgaag acgaaagggc ctcgtgatac gcctatttt 15120

ataggttaat gtcatgataa taatggtttc ttagacgtca ggtggcactt ttcggggaaa 15180

tgtgcgcgga acccctattt gtttatttt ctaaatacat tcaaatatgt atccgctcat 15240

gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta tgagtattca 15300

acatttccgt gtcgccctta ttccctttt tgcggcattt tgccttcctg tttttgctca 15360

cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac gagtgggtta 15420

catcgaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg aagaacgttt 15480

tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc gtgttgacgc 15540

cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg ttgagtactc 15600

accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat gcagtgctgc 15660

cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg gaggaccgaa 15720

ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg atcgttggga 15780

accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc ctgcagcaat 15840

ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt cccggcaaca 15900

attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct cggcccttcc 15960

ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc gcggtatcat 16020

tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca cgacggggag 16080

tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct cactgattaa 16140

gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt taaaacttca 16200

ttttaattt aaaaggatct aggtgaagat ccttttttgat aatctcatga ccaaaatccc 16260
ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca aaggatcttc 16320
ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc 16380
agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg taactggctt 16440
cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag gccaccactt 16500
caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac cagtggctgc 16560
tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt taccggataa 16620

ggcgcagcgg tcgggctgaa cgggggttc gtgcacacag cccagcttgg agcgaacgac 16680
ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc ttcccgaagg 16740
gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc gcacgaggga 16800
gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact 16860
tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa acgccagcaa 16920
cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt tctttcctgc 16980
gttatcccct gattctgtgg ataaccgtat taccgccttt gagtgagctg ataccgctcg 17040

ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag agcgcctgat 17100
gcggtatttt ctccttacgc atctgtgcgg tatttcacac cgcatatggt gcactctcag 17160
tacaatctgc tctgatgccg catagttaag ccagctgtgg aatgtgtgtc agttagggtg 17220
tggaaagtcc ccaggctccc cagcaggcag aagtatgcaa agcatgcatc tcaattagtc 17280
agcaaccagg tgtggaaagt ccccaggctc cccagcaggc agaagtatgc aaagcatgca 17340
tctcaattag tcagcaacca tagtcccgcc cctaactccg cccatcccgc ccctaactcc 17400
gcccagttcc gcccattctc cgccccatgg ctgactaatt ttttttattt atgcagaggc 17460

cgaggccgcc tcggcctctg agctattcca gaagtagtga ggaggctttt ttggaggcct 17520
aggcttttgc aaaaagcttg catgcctgca ggtcggccgc cacgaccggt gccgccacca 17580
tcccctgacc cacgcccctg acccctcaca aggagacgac cttccatgac cgagtacaag 17640
cccacggtgc gcctcgccac ccgcgacgac gtcccccggg ccgtacgcac cctcgccgcc 17700
gcgttcgccg actaccccgc cacgcgccac accgtcgacc cggaccgcca catcgagcgg 17760
gtcaccgagc tgcaagaact cttcctcacg cgcgtcgggc tcgacatcgg caaggtgtgg 17820
gtcgcggacg acggcgccgc ggtggcggtc tggaccacgc cggagagcgt cgaagcgggg 17880

gcggtgttcg ccgagatcgg cccgcgcatg gccgagttga gcggttcccg gctggccgcg 17940

cagcaacaga tggaaggcct cctggcgccg caccggccca aggagcccgc gtggttcctg 18000

gccaccgtcg gcgtctcgcc cgaccaccag ggcaagggtc tgggcagcgc cgtcgtgctc 18060

cccggagtgg aggcggccga gcgcgccggg gtgcccgcct tcctggagac ctccgcgccc 18120

cgcaacctcc ccttctacga gcggctcggc ttcaccgtca ccgccgacgt cgaggtgccc 18180

gaaggaccgc gcacctggtg catgacccgc aagcccggtg cctgacgccc gccccacgac 18240

ccgcagcgcc cgaccgaaag gagcgcacga ccccatggct ccgaccgaag ccgacccggg 18300

cggccccgcc gaccccgcac ccgcccccga ggcccaccga ctctagagga tcataatcag 18360

ccataccaca tttgtagagg ttttacttgc tttaaaaaac ctcccacacc tccccctgaa 18420

cctgaaacat aaaatgaatg caattgttgt tgttaacttg tttattgcag cttataatgg 18480

ttacaaataa agcaatagca tcacaaattt cacaaataaa gcatttttt cactgcattc 18540

tagttgtggt ttgtccaaac tcatcaatgt atcttatcat gtctggatca ctcgccgata 18600

gtggaaaccg acgccccagc actcgtccga gggcaaagga tagggggaga tgggggaggc 18660

taactgaaac acggaaggag acaataccgg aaggaacccg cgctatgacg gcaataaaaa 18720

gacagaataa aacgcacggg tgttgggtcg tttgttcata aacgcggggt tcggtcccag 18780

ggctggcact ctgtcgatac cccaccgaga ccccattggg gccaatacgc ccgcgtttct 18840

tccttttccc caccccaccc cccaagttcg ggtgaaggcc cagggctcgc agccaacgtc 18900

ggggcggcag gccctgccat agccactggc cccgtgggtt agggacgggg tcccccatgg 18960

ggaatggttt atggttcgtg ggggttatta ttttgggcgt tgcgtggggt ctggtccacg 19020

actggactga gcagacagac ccatggtttt tggatggcct gggcatggac cgcatgtact 19080

ggcgcgacac gaacaccggg cgtctgtggc tgccaaacac ccccgacccc caaaaaccac 19140

cgcgcggatt tctggcgtgc caagctagtc gaccaattct catgtttgac agcttatcat 19200

cgcagatccg ggcaacgttg ttgcattgct gcaggcgcag aactggtagg tatggaagat 19260

ctctagaagc tgggtaccag ctgctagcaa gcttgctagc ggccggctcg agtttactcc 19320

ctatcagtga tagagaacgt atgtcgagtt tactccctat cagtgataga gaacgatgtc 19380

gagtttactc cctatcagtg atagagaacg tatgtcgagt ttactcccta tcagtgatag 19440

agaacgtatg tcgagtttac tccctatcag tgatagagaa cgtatgtcga gtttatccct 19500

atcagtgata gagaacgtat gtcgagttta ctccctatca gtgatagaga acgtatgtcg 19560

aggtaggcgt gtacggtggg aggcctatat aagcagagct cgtttagtga accgtcagat 19620

cgccg 19625

<210> 41

<211> 19730

<212> DNA

<213> Artificial Sequence

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

<400> 41

cggccgcggg gggaggagcc aagatggccg aataggaaca gctccggtct acagctccca 60

gcgtgagcga cgcagaagac ggtgatttct gcatttccat ctgaggtacc gggttcatct 120

cactagggag tgccagacag tgggcgcagg ccagtgtgtg tgcgcaccgt gcgcgagccg 180

aagcagggcg aggcattgcc tcacctggga agcgcaaggg gtcagggagt tccctttccg 240

agtcaaagaa aggggtgacg gacgcacctg gaaaatcggg tcactcccac ccgaatattg 300

cgcttttcag accggcttaa gaaacggcgc accacgagac tatatcccac acctggctcg 360

gagggtccta cgcccacgga atctcgctga ttgctagcac agcagtctga gatcaaactg 420

caaggcggca acgaggctgg gggaggggcg cccgccattg cccaggcttg cttaggtaaa 480

caaagcagca gggaagctcg aactgggtgg agcccaccac agctcaagga ggcctgcctg 540

cctctgtagg ctccacctct gggggcaggg cacagacaaa caaaaagaca gcagtaacct 600

ctgcagactt aagtgtccct gtctgacagc tttgaagaga gcagtggttc tcccagcacg 660

cagctggaga tctgagaacg ggcagactgc ctcctcaagt gggtccctga cccctgaccc 720

ccgagcagcc taactgggag gcacccccca gcaggggcac actgacacct cacacggcag 780

ggtattccaa cagacctgca gctgagggtc ctgtctgtta gaaggaaaac taacaaccag 840

aaaggacatc tacaccgaaa acccatctgt acatcaccat catcaaagac caaaagtaga 900

taaaaccaca aagatgggga aaaaacagaa cagaaaaact ggaaactcta aaacgcagag 960

cgcctctcct cctccaaagg aacgcagttc ctcaccagca acagaacaaa gctggatgga 1020

gaatgatttt gatgagctga gagaagaagg cttcagacga tcaaattact ctgagctacg 1080

ggaggacatt caaaccaaag gcaaagaagt tgaaaacttt gaaaaaaatt tagaagaatg 1140

tataactaga ataaccaata cagagaagtg cttaaaggag ctgatggagc tgaaaaccaa 1200

ggctcgagaa ctacgtgaag aatgcagaag cctcaggagc cgatgcgatc aactggaaga 1260

aagggtatca gcaatggaag atgaaatgaa tgaaatgaag cgagaaggga agtttagaga 1320

aaaaagaata aaaagaaatg agcaaagcct ccaagaaata tgggactatg tgaaaagacc 1380

aaatctacgt ctgattggtg tacctgaaag tgatgtggag aatggaacca agttggaaaa 1440

cactctgcag gatattatcc aggagaactt ccccaatcta gcaaggcagg ccaacgttca 1500

gattcaggaa atacagagaa cgccacaaag atactcctcg agaagagcaa ctccaagaca 1560

cataattgtc agattcacca aagttgaaat gaaggaaaaa atgttaaggg cagccagaga 1620

gaaaggtcgg gttaccctca aaggaaagcc catcagacta acagcggatc tctcggcaga 1680

aaccctacaa gccagaagag agtgggggcc aatattcaac attcttaaag aaaagaattt 1740

tcaacccaga atttcatatc cagccaaact aagcttcata agtgaaggag aaataaaata 1800

ctttatagac aagcaaatgt tgagagattt tgtcaccacc aggcctgccc taaaagagct 1860

cctgaaggaa gcgctaaaca tggaaaggaa caaccggtac cagccgctgc aaaatcatgc 1920

caaaatgtaa agaccatcaa gactaggaag aaactgcatc aactaatgag caaaatcacc 1980

agctaacatc ataatgacag gatcaacttc acacataaca atattaactt taaatataaa 2040

tggactaaat tctgcaatta aaagacacag actggcaagt tggataaaga gtcaagaccc 2100

atcagtgtgc tgtattcagg aaacccatct cacgtgcaga gacacacata ggctcaaaat 2160

aaaaggatgg aggaagatct accaagccaa tggaaaacaa aaaaaggcag gggttgcaat 2220

cctagtctct gataaaacag actttaaacc aacaaagatc aaaagagaca aagaaggcca 2280

ttacataatg gtaaagggat caattcaaca agaggagcta actatcctaa atatttatgc 2340

acccaataca ggagcaccca gattcataaa gcaagtcctc agtgacctac aaagagactt 2400

agactcccac acattaataa tgggagactt taacacccca ctgtcaacat tagacagatc 2460

aacgagacag aaagtcaaca aggataccca ggaattgaac tcagctctgc accaagcaga 2520

cctaatagac atctacagaa ctctccaccc caaatcaaca gaatatacat ttttttcagc 2580

accacaccac acctattcca aaattgacca catagttgga agtaaagctc tcctcagcaa 2640

atgtaaaaga acagaaatta taacaaacta tctctcagac cacagtgcaa tcaaactaga 2700

actcaggatt aagaatctca ctcaaagccg ctcaactaca tggaaactga acaacctgct 2760

cctgaatgac tactgggtac ataacgaaat gaaggcagaa ataaagatgt tctttgaaac 2820

caacgagaac aaagacacca cataccagaa tctctgggac gcattcaaag cagtgtgtag 2880

agggaaattt atagcactaa atgcctacaa gagaaagcag gaaagatcca aaattgacac 2940

cctaacatca caattaaaag aactagaaaa gcaagagcaa acacattcaa aagctagcag 3000

aaggcaagaa ataactaaaa tcagagcaga actgaaggaa atagagacac aaaaaaccct 3060

tcaaaaaatc aatgaatcca ggagctggtt ttttgaaagg atcaacaaaa ttgatagacc 3120

gctagcaaga ctaataaaga aaaaaagaga gaagaatcaa atagacacaa taaaaaatga 3180

taaaggggat atcaccaccg atcccacaga aatacaaact accatcagag aatactacaa 3240

acacctctac gcaaataaac tagaaaatct agaagaaatg gatacattcc tcgacacata 3300

cactctccca agactaaacc aggaagaagt tgaatctctg aatcgaccaa taacaggctc 3360

tgaaattgtg gcaataatca atagtttacc aaccaaaaag agtccaggac cagatggatt 3420

cacagccgaa ttctaccaga ggtacaagga ggaactggta ccattccttc tgaaactatt 3480

ccaatcaata gaaaaagagg gaatcctccc taactcattt tatgaggcca gcatcattct 3540

gataccaaag ccgggcagag acacaaccaa aaaagagaat tttagaccaa tatccttgat 3600

gaacattgat gcaaaaatcc tcaataaaat actggcaaac cgaatccagc agcacatcaa 3660

aaagcttatc caccatgatc aagtgggctt catccctggg atgcaaggct ggttcaatat 3720

acgcaaatca ataaatgtaa tccagcatat aaacagagcc aaagacaaaa accacatgat 3780

tatctcaata gatgcagaaa aagcctttga caaaattcaa caacccttca tgctaaaaac 3840

tctcaataaa ttaggtattg atgggacgta tttcaaaata ataagagcta tctatgacaa 3900

acccacagcc aatatcatac tgaatgggca aaaactggaa gcattccctt tgaaaaccgg 3960

cacaagacag ggatgccctc tctcaccgct cctattcaac atagtgttgg aagttctggc 4020

cagggcaatc aggcaggaga aggaaataaa gggtattcaa ttaggaaaag aggaagtcaa 4080

attgtccctg tttgcagacg acatgattgt ttatctagaa aaccccatcg tctcagccca 4140

aaatctcctt aagctgataa gcaacttcag caaagtctca ggatacaaaa tcaatgtaca 4200

aaaatcacaa gcattcttat acaccaacaa cagacaaaca gagagccaaa tcatgggtga 4260

actcccattc acaattgctt caaagagaat aaaataccta ggaatccaac ttacaaggga 4320

tgtgaaggac ctcttcaagg agaactacaa accactgctc aaggaaataa aagaggagac 4380

aaacaaatgg aagaacattc catgctcatg ggtaggaaga atcaatatcg tgaaaatggc 4440

catactgccc aaggtaattt acagattcaa tgccatcccc atcaagctac caatgacttt 4500

cttcacagaa ttggaaaaaa ctactttaaa gttcatatgg aaccaaaaaa gagcccgcat 4560

tgccaagtca atcctaagcc aaaagaacaa agctggaggc atcacactac ctgacttcaa 4620

actatactac aaggctacag taaccaaaac agcatggtac tggtaccaaa acagagatat 4680

agatcaatgg aacagaacag agccctcaga aataatgccg catatctaca actatctgat 4740

ctttgacaaa cctgagaaaa acaagcaatg gggaaaggat tccctattta ataaatggtg 4800

ctgggaaaac tggctagcca tatgtagaaa gctgaaactg gatcccttcc ttacacctta 4860

tacaaaaatc aattcaagat ggattaaaga tttaaacgtt aaacctaaaa ccataaaaac 4920

cctagaagaa aacctaggca ttaccattca ggacataggc gtgggcaagg acttcatgtc 4980

caaaacacca aaagcaatgg caacaaaaga caaaattgac aaatgggatc taattaaact 5040

aaagagcttc tgcacagcaa aagaaactac catcagagtg aacaggcaac ctacaacatg 5100

ggagaaaatt tttgcaacct actcatctga caaagggcta atatccagaa tctacaatga 5160

actcaaacaa atttacaaga aaaaaacaaa caaccccatc aaaaagtggg cgaaggacat 5220

gaacagacac ttctcaaaag aagacattta tgcagccaaa aaacacatga agaaatgctc 5280
atcatcactg gccatcagag aaatgcaaat caaaaccact atgagatatc atctcacacc 5340
agttagaatg gcaatcatta aaaagtcagg aaacaacagg tgctggagag gatgcggaga 5400
aataggaaca cttttacact gttggtggga ctgtaaacta gttcaaccat tgtggaagtc 5460
agtgtggcga ttcctcaggg atctagaact agaaatacca tttgacccag ccatcccatt 5520
actgggtata tacccaaatg agtataaatc atgctgctat aaagacacat gcacacgtat 5580
gtttattgcg gcactattca caatagcaaa gacttggaac caacccaaat gtccaacaat 5640

gatagactgg attaagaaaa tgtggcacat atacaccatg gaatactatg cagccataaa 5700
aaatgatgag ttcatatcct ttgtagggac atggatgaaa ttggaaacca tcattctcag 5760
taaactatcg caagaacaaa aaaccaaaca ccgcatattc tcactcatag gtgggaattg 5820
aacaatgaga tcacatggac acaggaaggg gaatatcaca ctctggggac tgtggtgggg 5880
tcgggggagg ggggagggat agcattggga gatataccta atgctagatg acacattagt 5940
gggtgcagcg caccagcatg gcacatgtat acggatccga attctcgacg gatcgatccg 6000
aacaaacgac ccaacacccg tgcgttttat tctgtctttt tattgccgat cccctcagaa 6060

gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta 6120
aagcacgagg aagcggtcag cccattcgcc gccaagctct tcagcaatat cacgggtagc 6180
caacgctatg tcctgatagc ggtcggccgc tttacttgta cagctcgtcc atgccgagag 6240
tgatcccggc ggcggtcacg aactccagca ggaccatgtg atcgcgcttc tcgttggggt 6300
ctttgctcag ggcggactgg gtgctcaggt agtggttgtc gggcagcagc acggggccgt 6360
cgccgatggg ggtgttctgc tggtagtggt cggccaggtg agtccaggag atgtttcagc 6420
actgttgcct ttagtctcga ggcaacttag acaactgagt attgatctga gcacagcagg 6480

gtgtgagctg tttgaagata ctggggttgg gggtgaagaa actgcagagg actaactggg 6540
ctgagaccca gtggcaatgt tttagggcct aaggaatgcc tctgaaaatc tagatggaca 6600
actttgactt tgagaaaaga gaggtggaaa tgaggaaaat gacttttctt tattagattt 6660
cggtagaaag aactttcatc tttcccctat ttttgttatt cgttttaaaa catctatctg 6720
gaggcaggac aagtatggtc attaaaaaga tgcaggcaga aggcatatat tggctcagtc 6780
aaagtgggga actttggtgg ccaaacatac attgctaagg ctattcctat atcagctgga 6840
cacatataaa atgctgctaa tgcttcatta caaacttata tcctttaatt ccagatgggg 6900

gcaaagtatg tccaggggtg aggaacaatt gaaacatttg ggctggagta gattttgaaa 6960

gtcagctctg tgtgtgtgtg tgtgtgtgtg tgtgtgagag cgtgtgtttc ttttaacgtt 7020

ttcagcctac agcatacagg gttcatggtg gcaagaagat aacaagattt aaattatggc 7080

cagtgactag tgctgcaaga agaacaacta cctgcattta atgggaaagc aaaatctcag 7140

gctttgaggg aagttaacat aggcttgatt ctgggtggaa gctgggtgtg tagttatctg 7200

gaggccaggc tggagctctc agctcactat gggttcatct ttattgtctc ctttcatctc 7260

aacagctgca cgctgccgtc ctcgatgttg tggcggatct tgaagttcac cttgatgccg 7320

ttcttctgct tgtcggccat gatatagacg ttgtggctgt tgtagttgta ctccagcttg 7380

tgccccagga tgttgccgtc ctccttgaag tcgatgccct tcagctcgat gcggttcacc 7440

agggtgtcgc cctcgaactt cacctcggcg cgggtcttgt agttgccgtc gtccttgaag 7500

aagatggtgc gctcctggac gtagccttcg ggcatggcgg acttgaagaa gtcgtgctgc 7560

ttcatgtggt cggggtagcg gctgaagcac tgcacgccgt aggtcagggt ggtcacgagg 7620

gtgggccagg gcacgggcag cttgccggtg gtgcagatga acttcagggt cagcttgccg 7680

taggtggcat cgccctcgcc ctcgccggac acgctgaact tgtggccgtt tacgtcgccg 7740

tccagctcga ccaggatggg caccaccccg gtgaacagct cctcgccctt gctcaccata 7800

gggccgggat tctcctccac gtcaccgcat gttagaagac ttcctctgcc ctctcttgga 7860

ggcagggcct gcatgtgcag ggcatcgtag gtatccttgg tggctgtgct cagtccctgg 7920

tacagtccat cgtggccctt gcctcttctt ctctcgccct tcatgccgat ctcgctgtag 7980

gcctcggcca tcttgtcttt ctgcagctca ttatacaggc cctcttgagg attctttctc 8040

cgctggggct tgccgcccat ctcaggatct ctgcctctcc gcttatccag cacgtcgtac 8100

tcttctcttc tccccaggtt cagctcgttg tacagctgat tctggccctg ctggtaagca 8160

ggagcgtcgg cggatctgct gaacttcact ctgcagtaca gggtgatgac cagagagagc 8220

agcagaacgc cacatgtgcc agccagaggg gcccaaatgt agatatccag gcctctggta 8280

tgcacagctc cgccagctgc aggtctacag gcttcaggtc tgagagacag aggctggctg 8340

gcgattgtag gagctggtgt aggtggtcta ggagcgggtg ttgttgtagg cttggcgggc 8400

agaaacacgg gcacgaagtg gctgaagtac atgatgctat tgctcagggc tccgcttcct 8460

ccgcctccgc tagaagaaac tgtgaccagg gtgccctgtc cccaaacatc catggcgtag 8520

aagccgtcgc ctccccatct agaacagtag tacacggcgg tgtcctcggc tctcaggctg 8580

ttcatctgca ggtaggcggt gttcttgctg gtgtcggcgc tgatggtgaa tctgcccttc 8640

acgctatcgg cgtatctggt gtagccgttg gtggggtaga ttctggcgac ccattcaagt 8700

ccctttccag gggcctgtcg gacccagtgg atgtaggtgt ccttgatgtt gaagccgctg 8760

gcggcacaag acagtctcag agagccgcca ggctgaacaa gtcctccgcc agattcaacc 8820

agctgcacct cagatccttc gccagatcca ggctttccag agccgctggt gctgcctgtt 8880

ctcttgattt ccaccttggt gccctggcca aaggttggag gtgtggtgta gtgctgctgg 8940

cagtagtagg tggcgaagtc ctcaggctgc aggctagaga tggtcagggt gaagtcggtg 9000

ccagatctgc tgccgctgaa tctgcttggc acgccgctgt acagaaagct ggcgctgtag 9060

atcagcagct taggggcttt tccaggcttc tgctgatacc aggccacggc ggtattcaca 9120

tcctggctgg ctctacaggt gatggtcact ctatcgccca cagaggcaga caggctgcta 9180

gggctctgtg tcatctggat gtcgctgatg ctgcaggcca ctgttcccag cagcagcaga 9240

gactgcagcc acattcgaag cttgagctcg agatctgagt ccggtagcgc tagcggatct 9300

gacggttcac taaaccagct ctgcttatat agacctccca ccgtacacgc ctaccgccca 9360

tttgcgtcaa tggggcggag ttgttacgac attttggaaa gtcccgttga ttttggtgcc 9420

aaaacaaact cccattgacg tcaatggggt ggagacttgg aaatccccgt gagtcaaacc 9480

gctatccacg cccattgatg tactgccaaa accgcatcac catggtaata gcgatgacta 9540

atacgtagat gtactgccaa gtaggaaagt cccataaggt catgtactgg gcataatgcc 9600

aggcgggcca tttaccgtca ttgacgtcaa taggggggcgt acttggcata tgatacactt 9660

gatgtactgc caagtgggca gtttaccgta aatactccac ccattgacgt caatggaaag 9720

tccctattgg cgttactatg ggaacatacg tcattattga cgtcaatggg cgggggtcgt 9780

tgggcggtca gccaggcggg ccatttaccg taagttatgt aacgcggaac tccatatatg 9840

ggctatgaac taatgacccc gtaattgatt actattagcc cgggggatcc agacatgata 9900

agatacattg atgagtttgg acaaaccaca actagaatgc agtgaaaaaa atgctttatt 9960

tgtgaaattt gtgatgctat tgctttattt gtaaccatta taagctgcaa taaacaagtt 10020

aacaacaaca attgcattca ttttatgttt caggttcagg gggaggtgtg ggaggttttt 10080

taaagcaagt aaaacctcta caaatgtggt atggctgatt atgatccggc tgcctcgcgc 10140

gtttcggtga tgacggtgaa aacctctgac acatgcagct cccggagacg gtcacagctt 10200

gtctgtaagc ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg ggtgttggcg 10260

ggtgtcgggg cgcagccatg aggtcgatcg actctagagg atcgatcccc gccccggacg 10320

aactaaacct gactacgaca tctctgcccc ttcttcgcgg ggcagtgcat gtaatccctt 10380

cagttggttg gtacaacttg ccaactgggc cctgttccac atgtgacacg gggggggacc 10440

aaacacaaag gggttctctg actgtagttg acatccttat aaatggatgt gcacatttgc 10500

caacactgag tggctttcat cctggagcag actttgcagt ctgtggactg caacacaaca 10560

ttgcctttat gtgtaactct tggctgaagc tcttacacca atgctggggg acatgtacct 10620

cccaggggcc caggaagact acgggaggct acaccaacgt caatcagagg ggcctgtgta 10680

gctaccgata agcggaccct caagagggca ttagcaatag tgtttataag gccccttgt 10740

taaccctaaa cgggtagcat atgcttcccg ggtagtagta tatactatcc agactaaccc 10800

taattcaata gcatatgtta cccaacggga agcatatgct atcgaattag ggttagtaaa 10860

agggtcctaa ggaacagcga tatctcccac cccatgagct gtcacggttt tatttacatg 10920

gggtcaggat tccacgaggg tagtgaacca ttttagtcac aagggcagtg gctgaagatc 10980

aaggagcggg cagtgaactc tcctgaatct tcgcctgctt cttcattctc cttcgtttag 11040

ctaatagaat aactgctgag ttgtgaacag taaggtgtat gtgaggtgct cgaaaacaag 11100

gtttcaggtg acgcccccag aataaaattt ggacgggggg ttcagtggtg gcattgtgct 11160

atgacaccaa tataaccctc acaaacccct tgggcaataa atactagtgt aggaatgaaa 11220

cattctgaat atctttaaca atagaaatcc atggggtggg gacaagccgt aaagactgga 11280

tgtccatctc acacgaattt atggctatgg gcaacacata atcctagtgc aatatgatac 11340

tggggttatt aagatgtgtc ccaggcaggg accaagacag gtgaaccatg ttgttacact 11400

ctatttgtaa caaggggaaa gagagtggac gccgacagca gcggactcca ctggttgtct 11460

ctaacacccc cgaaaattaa acggggctcc acgccaatgg ggcccataaa caaagacaag 11520

tggccactct ttttttttgaa attgtggagt gggggcacgc gtcagccccc acacgccgcc 11580

ctgcggtttt ggactgtaaa ataagggtgt aataacttgg ctgattgtaa ccccgctaac 11640

cactgcggtc aaaccacttg cccacaaaac cactaatggc accccgggga atacctgcat 11700

aagtaggtgg gcgggccaag ataggggcgc gattgctgcg atctggagga caaattacac 11760

acacttgcgc ctgagcgcca agcacagggt tgttggtcct catattcacg aggtcgctga 11820

gagcacggtg ggctaatgtt gccatgggta gcatatacta cccaaatatc tggatagcat 11880

atgctatcct aatctatatc tgggtagcat aggctatcct aatctatatc tgggtagcat 11940

atgctatcct aatctatatc tgggtagtat atgctatcct aatttatatc tgggtagcat 12000

aggctatcct aatctatatc tgggtagcat atgctatcct aatctatatc tgggtagtat 12060

atgctatcct aatctgtatc cgggtagcat atgctatcct aatagagatt agggtagtat 12120

atgctatcct aatttatatc tgggtagcat atactaccca aatatctgga tagcatatgc 12180

tatcctaatc tatatctggg tagcatatgc tatcctaatc tatatctggg tagcataggc 12240

tatcctaatc tatatctggg tagcatatgc tatcctaatc tatatctggg tagtatatgc 12300

tatcctaatt tatatctggg tagcataggc tatcctaatc tatatctggg tagcatatgc 12360

tatcctaatc tatatctggg tagtatatgc tatcctaatc tgtatccggg tagcatatgc 12420

tatcctcatg catatacagt cagcatatga tacccagtag tagagtggga gtgctatcct 12480

ttgcatatgc cgccacctcc caaggggggcg tgaattttcg ctgcttgtcc ttttcctgca 12540

tgctggttgc tcccattctt aggtgaattt aaggaggcca ggctaaagcc gtcgcatgtc 12600

tgattgctca ccaggtaaat gtcgctaatg ttttccaacg cgagaaggtg ttgagcgcgg 12660

agctgagtga cgtgacaaca tgggtatgcc caattgcccc atgttgggag gacgaaaatg 12720

gtgacaagac agatggccag aaatacacca acagcacgca tgatgtctac tggggattta 12780

ttctttagtg cgggggaata cacggctttt aatacgattg agggcgtctc ctaacaagtt 12840

acatcactcc tgcccttcct caccctcatc tccatcacct ccttcatctc cgtcatctcc 12900

gtcatcaccc tccgcggcag ccccttccac cataggtgga aaccagggag gcaaatctac 12960

tccatcgtca aagctgcaca cagtcaccct gatattgcag gtaggagcgg gctttgtcat 13020

aacaaggtcc ttaatcgcat ccttcaaaac ctcagcaaat atatgagttt gtaaaaagac 13080

catgaaataa cagacaatgg actcccttag cgggccaggt tgtgggccgg gtccaggggc 13140

cattccaaag gggagacgac tcaatggtgt aagacgacat tgtggaatag caagggcagt 13200

tcctcgcctt aggttgtaaa gggaggtctt actacctcca tatacgaaca caccggcgac 13260

ccaagttcct tcgtcggtag tcctttctac gtgactccta gccaggagag ctcttaaacc 13320

ttctgcaatg ttctcaaatt tcgggttgga acctccttga ccacgatgct ttccaaacca 13380

ccctcctttt ttgcgcctgc ctccatcacc ctgaccccgg ggtccagtgc ttgggccttc 13440

tcctgggtca tctgcggggc cctgctctat cgctcccggg ggcacgtcag gctcaccatc 13500

tgggccacct tcttggtggt attcaaaata atcggcttcc cctacagggt ggaaaaatgg 13560

ccttctacct ggaggggggcc tgcgcggtgg agacccggat gatgatgact gactactggg 13620

actcctgggc ctcttttctc cacgtccacg acctctcccc ctggctcttt cacgacttcc 13680

ccccctggct ctttcacgtc ctctaccccg gcggcctcca ctacctcctc gaccccggcc 13740

tccactacct cctcgacccc ggcctccact gcctcctcga ccccggcctc cacctcctgc 13800

tcctgcccct cctgctcctg cccctcctcc tgctcctgcc cctcctgccc ctcctgctcc 13860

tgcccctcct gcccctcctg ctcctgcccc tcctgcccct cctgctcctg cccctcctgc 13920

ccctcctcct gctcctgccc ctcctgcccc tcctcctgct cctgccccctc ctgcccctcc 13980

tgctcctgcc cctcctgccc ctcctgctcc tgcccctcct gcccctcctg ctcctgcccc 14040

tcctgctcct gcccctcctg ctcctgcccc tcctgctcct gcccctcctg cccctcctgc 14100

ccctcctcct gctcctgccc ctcctgctcc tgcccctcct gcccctcctg cccctcctgc 14160

tcctgcccct cctcctgctc ctgcccctcc tgcccctcct gcccctcctc ctgctcctgc 14220

ccctcctgcc cctcctcctg ctcctgcccc tcctcctgct cctgcccctc ctgcccctcc 14280

tgcccctcct cctgctcctg cccctcctgc ccctcctcct gctcctgccc ctcctcctgc 14340

tcctgcccct cctgcccctc ctgcccctcc tcctgctcct gcccctcctc ctgctcctgc 14400

ccctcctgcc cctcctgccc ctcctgcccc tcctcctgct cctgcccctc ctcctgctcc 14460

tgcccctcct gctcctgccc ctcccgctcc tgctcctgct cctgttccac cgtgggtccc 14520

tttgcagcca atgcaacttg gacgtttttg gggtctccgg acaccatctc tatgtcttgg 14580

ccctgatcct gagccgcccg gggctcctgg tcttccgcct cctcgtcctc gtcctcttcc 14640

ccgtcctcgt ccatggttat caccccctct tctttgaggt ccactgccgc cggagccttc 14700

tggtccagat gtgtctccct tctctcctag gccatttcca ggtcctgtac ctggcccctc 14760

gtcagacatg attcacacta aaagagatca atagacatct ttattagacg acgctcagtg 14820

aatacaggga gtgcagactc ctgccccctc caacagcccc cccaccctca tccccttcat 14880

ggtcgctgtc agacagatcc aggtctgaaa attccccatc ctccgaacca tcctcgtcct 14940

catcaccaat tactcgcagc ccggaaaact cccgctgaac atcctcaaga tttgcgtcct 15000

gagcctcaag ccaggcctca aattcctcgt ccccctttt gctggacggt agggatgggg 15060

attctcggga cccctcctct tcctcttcaa ggtcaccaga cagagatgct actggggcaa 15120

cggaagaaaa gctgggtgcg gcctgtgagg atcagcttat cgatgataag ctgtcaaaca 15180

tgagaattct tgaagacgaa agggcctcgt gatacgccta tttttatagg ttaatgtcat 15240

gataataatg gtttcttaga cgtcaggtgg cacttttcgg ggaaatgtgc gcggaacccc 15300

tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcatgagac aataaccctg 15360

ataaatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc 15420

ccttattccc ttttttgcgg cattttgcct tcctgttttt gctcacccag aaacgctggt 15480

gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg ggttacatcg aactggatct 15540

caacagcggt aagatccttg agagttttcg ccccgaagaa cgttttccaa tgatgagcac 15600

ttttaaagtt ctgctatgtg gcgcggtatt atcccgtgtt gacgccgggc aagagcaact 15660

cggtcgccgc atacactatt ctcagaatga cttggttgag tactcaccag tcacagaaaa 15720

gcatcttacg gatggcatga cagtaagaga attatgcagt gctgccataa ccatgagtga 15780

taacactgcg gccaacttac ttctgacaac gatcggagga ccgaaggagc taaccgcttt 15840

tttgcacaac atgggggatc atgtaactcg ccttgatcgt tgggaaccgg agctgaatga 15900

agccatacca aacgacgagc gtgacaccac gatgcctgca gcaatggcaa caacgttgcg 15960

caaactatta actggcgaac tacttactct agcttcccgg caacaattaa tagactggat 16020

ggaggcggat aaagttgcag gaccacttct gcgctcggcc cttccggctg gctggtttat 16080

tgctgataaa tctggagccg gtgagcgtgg gtctcgcggt atcattgcag cactggggcc 16140

agatggtaag ccctcccgta tcgtagttat ctacacgacg gggagtcagg caactatgga 16200
tgaacgaaat agacagatcg ctgagatagg tgcctcactg attaagcatt ggtaactgtc 16260
agaccaagtt tactcatata tactttagat tgatttaaaa cttcattttt aatttaaaag 16320
gatctaggtg aagatccttt ttgataatct catgaccaaa atcccttaac gtgagttttc 16380
gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atccttttt 16440
tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt 16500
gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca gagcgcagat 16560

accaaatact gtccttctag tgtagccgta gttaggccac cacttcaaga actctgtagc 16620
accgcctaca tacctcgctc tgctaatcct gttaccagtg gctgctgcca gtggcgataa 16680
gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg 16740
ctgaacgggg ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag 16800
atacctacag cgtgagctat gagaaagcgc cacgcttccc gaagggagaa aggcggacag 16860
gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa 16920
cgcctggtat ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc gtcgatttt 16980

gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg 17040
gttcctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc 17100
tgtggataac cgtattaccg cctttgagtg agctgatacc gctcgccgca gccgaacgac 17160
cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc ctgatgcggt attttctcct 17220
tacgcatctg tgcggtattt cacaccgcat atggtgcact ctcagtacaa tctgctctga 17280
tgccgcatag ttaagccagc tgtggaatgt gtgtcagtta gggtgtggaa agtccccagg 17340
ctccccagca ggcagaagta tgcaaagcat gcatctcaat tagtcagcaa ccaggtgtgg 17400

aaagtcccca ggctccccag caggcagaag tatgcaaagc atgcatctca attagtcagc 17460
aaccatagtc ccgcccctaa ctccgcccat cccgccccta actccgccca gttccgccca 17520
ttctccgccc catggctgac taatttttt tatttatgca gaggccgagg ccgcctcggc 17580
ctctgagcta ttccagaagt agtgaggagg ctttttgga ggcctaggct tttgcaaaaa 17640
gcttgcatgc ctgcaggtcg gccgccacga ccggtgccgc caccatcccc tgacccacgc 17700
ccctgacccc tcacaaggag acgaccttcc atgaccgagt acaagcccac ggtgcgcctc 17760
gccacccgcg acgacgtccc ccgggccgta cgcaccctcg ccgccgcgtt cgccgactac 17820

cccgccacgc gccacaccgt cgacccggac cgccacatcg agcgggtcac cgagctgcaa 17880

gaactcttcc tcacgcgcgt cgggctcgac atcggcaagg tgtgggtcgc ggacgacggc 17940

gccgcggtgg cggtctggac cacgccggag agcgtcgaag cgggggcggt gttcgccgag 18000

atcggcccgc gcatggccga gttgagcggt tcccggctgg ccgcgcagca acagatggaa 18060

ggcctcctgg cgccgcaccg gcccaaggag cccgcgtggt tcctggccac cgtcggcgtc 18120

tcgcccgacc accagggcaa gggtctgggc agcgccgtcg tgctccccgg agtggaggcg 18180

gccgagcgcg ccggggtgcc cgccttcctg gagacctccg cgccccgcaa cctccccttc 18240

tacgagcggc tcggcttcac cgtcaccgcc gacgtcgagg tgcccgaagg accgcgcacc 18300

tggtgcatga cccgcaagcc cggtgcctga cgcccgcccc acgacccgca gcgcccgacc 18360

gaaaggagcg cacgacccca tggctccgac cgaagccgac ccgggcggcc ccgccgaccc 18420

cgcacccgcc cccgaggccc accgactcta gaggatcata atcagccata ccacatttgt 18480

agaggtttta cttgctttaa aaaacctccc acacctcccc ctgaacctga aacataaaat 18540

gaatgcaatt gttgttgtta acttgtttat tgcagcttat aatggttaca aataaagcaa 18600

tagcatcaca aatttcacaa ataaagcatt tttttcactg cattctagtt gtggtttgtc 18660

caaactcatc aatgtatctt atcatgtctg gatcactcgc cgatagtgga aaccgacgcc 18720

ccagcactcg tccgagggca aaggaatagg ggagatgggg gaggctaact gaaacacgga 18780

aggagacaat accggaagga acccgcgcta tgacggcaat aaaaagacag aataaaacgc 18840

acgggtgttg ggtcgtttgt tcataaacgc ggggttcggt cccagggctg gcactctgtc 18900

gataccccac cgagacccca ttggggccaa tacgcccgcg tttcttcctt ttccccaccc 18960

cacccccaa gttcgggtga aggcccaggg ctcgcagcca acgtcggggc ggcaggccct 19020

gccatagcca ctggccccgt gggttaggga cggggtcccc catggggaat ggtttatggt 19080

tcgtgggggt tattattttg ggcgttgcgt ggggtctggt ccacgactgg actgagcaga 19140

cagacccatg gtttttggat ggcctgggca tggaccgcat gtactggcgc gacacgaaca 19200

ccgggcgtct gtggctgcca aacacccccg acccccaaaa accaccgcgc ggatttctgg 19260

cgtgccaagc tagtcgacca attctcatgt ttgacagctt atcatcgcag atccgggcaa 19320

cgttgttgca ttgctgcagg cgcagaactg gtaggtatgg aagatctcta gaagctgggt 19380

accagctgct agcaagcttg ctagcggccg gctcgagttt actccctatc agtgatagag 19440

aacgtatgtc gagtttactc cctatcagtg atagagaacg atgtcgagtt tactccctat 19500

cagtgataga gaacgtatgt cgagtttact ccctatcagt gatagagaac gtatgtcgag 19560

tttactccct atcagtgata gagaacgtat gtcgagttta tccctatcag tgatagagaa 19620

cgtatgtcga gtttactccc tatcagtgat agagaacgta tgtcgaggta ggcgtgtacg 19680

gtgggaggcc tatataagca gagctcgttt agtgaaccgt cagatcgccg 19730

<210> 42

<211> 7291

<212> DNA

<213> Artificial Sequence

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

<400> 42

taatacgact cactataggg agaagtactg ccaccatggg caagaagcaa aatcgcaaga 60

cggggaattc caagacacaa tccgctagcc caccacctaa agagcgttct agctcccctg 120

ctactgagca gtcctggatg gaaaacgact tcgatgaact ccgggaagag ggatttaggc 180

gatccaacta ttcagaactc cgcgaagata tccagacaaa ggggaaggaa gtcgagaatt 240

tcgagaagaa cctcgaggag tgcatcaccc gtatcacaaa cactgagaaa tgtctcaaag 300

aactcatgga acttaagaca aaagccaggg agcttcgaga ggagtgtcgg agtctgagat 360

ccaggtgtga ccagctcgag gagcgcgtga gcgcgatgga agacgagatg aacgagatga 420

aaagagaggg caaattcagg gagaagcgca ttaagaggaa cgaacagagt ctgcaggaga 480

tttgggatta cgtcaagagg cctaacctgc ggttgatcgg cgtccccgag agcgacgtag 540

aaaacgggac taaactggag aatacacttc aagacatcat tcaagaaaat tttccaaacc 600

tggctcggca agctaatgtg caaatccaag agatccaacg cacaccccag cggtatagct 660

ctcggcgtgc cacccctagg catattatcg tgcgctttac taaggtggag atgaaagaga 720

agatgctgcg agccgctcgg gaaaagggaa gggtgacttt gaagggcaaa cctattcggc 780

tgacggttga ccttagcgcc gagacactcc aggcacgccg ggaatggggc cccatcttta 840

atatcctgaa ggagaagaac ttccagccac gaatctctta ccctgcaaag ttgagtttta 900

tctccgaggg tgagattaag tatttcatcg ataaacagat gctgcgagac ttcgtgacaa 960

ctcgcccagc tctcaaggaa ctgctcaaag aggctcttaa tatggagcgc aataatagat 1020

atcaaccctt gcagaaccac gcaaagatgt gagacagccg tcagaccatc aagactagga 1080

agaaactgca tcaactaatg agcaaaatca ccagctaaca tcatagtata catgaccggc 1140

tctaactcac atatcaccat ccttacactt aacattaacg gcctcaactc agctatcaag 1200

cgccatcggc tggccagctg gatcaaatca caggatccaa gcgtttgttg catccaagag 1260

acccacctga cctgtagaga tactcaccgc ctcaagatca agggatggcg aaagatttat 1320

caggcgaacg gtaagcagaa gaaagccgga gtcgcaattc tggtctcaga caagacggat 1380

ttcaagccca ccaaaattaa gcgtgataag gaaggtcact atattatggt gaaaggcagc 1440

atacagcagg aagaacttac catattgaac atctacgcgc caaacaccgg cgcacctcgc 1500

tttatcaaac aggtcctgtc cgatctgcag cgagatctgg attctcatac gttgattatg 1560

ggtgatttca atacaccatt gagcaccctg gatcgcagca ccaggcaaaa ggtaaataaa 1620

gacacgcaag agctcaatag cgcactgcat caggcagatc tcattgatat ttatcgcact 1680

cttcatccta agagtaccga gtacacattc ttcagcgccc cacatcatac atactcaaag 1740

atcgatcata tcgtcggctc aaaggctctg ctgtcaaagt gcaagcgcac agagataatt 1800

acaaattacc tgtcagatca tagcgcgatc aagctcgagc tgagaatcaa gaacctgacc 1860

cagagccgga gtaccacttg gaagcttaat aacctgctgc tcaacgatta ttgggtccac 1920

aatgagatga aggcagagat taaaatgttc ttcgaaacaa atgagaataa ggatactacc 1980

tatcaaaacc tttgggatgc ctttaaggcc gtctgcagag gcaagttcat cgccctcaac 2040

gcctataaaa gaaaacaaga gagatctaag atcgatactc tcacctctca gctgaaggag 2100

ttggagaaac aggaacagac ccactccaag gcgtcaagac ggcaggagat cacaaagatt 2160

cgcgccgagt tgaaagagat cgaaacccaa aagactcttc agaaaattaa cgagtctcgt 2220

agttggttct tcgagcggat taataagata gacagacctc tggcacgact gattaagaag 2280

aagcgcgaaa agaaccagat tgataccatc aagaacgaca agggcgacat cactactgac 2340

ccgaccgaga tccagaccac tattcgggag tattataagc atttgtatgc taacaagctt 2400

gagaacctgg aagagatgga cacttttctg gatacctata ctctgccacg gcttaatcaa 2460

gaggaagtcg agtccctcaa ccgcccaatt acaggaagcg agattgtggc cataattaac 2520

tccctgccga caaagaaatc tcctggtccg gacgggttta cagctgagtt ttatcaacgg 2580

tatatggaag agcttgtacc gtttctgctc aagctctttc agtctataga aaaggaaggc 2640

atcttgccca attccttcta cgaagcttct ataatactta ttcccaaacc aggacgcgat 2700

accacaaaga aggaaaactt ccggcccatt agtctcatga atatcgacgc taaaatattg 2760

aacaagattc tcgccaacag aatccaacaa catattaaga aattgataca tcacgaccag 2820

gtggggttta tacctggcat gcagggctgg tttaacatcc ggaagagtat taacgtcatt 2880

caacacatta atagagctaa ggataagaat catatgatca tctctataga cgcggaaaag 2940

gcattcgata agattcagca gccatttatg ctcaagactc tgaacaaact cggcatcgac 3000

ggaacatatt ttaagattat tcgcgcaatt tacgataagc cgactgctaa cattatcctt 3060

aacggccaaa agctcgaggc ctttccgctc aagactggaa cccgccaagg ctgtcccctc 3120

tccccgcttt tgtttaatat tgtactcgag gtgctggcta gggctattcg tcaagagaaa 3180

gagattaaag ggatacagct cgggaaggaa gaggtcaagc tttccttgtt cgccgatgat 3240

atgattgtgt acctggagaa tcctattgtg tctgctcaga accttcttaa acttatttct 3300

aactttagca aggtcagcgg ctataagatt aacgtccaga aatctcaggc ctttctgtac 3360

acaaataatc gacagaccga atcccagata atgggtgagc ttccgtttgt catagccagc 3420

aaaaggataa agtatctcgg aatccagctg acacgagacg ttaaagattt gtttaaggaa 3480

aattacaagc ctctcctgaa agagattaag gaagatacta ataagtggaa gaatatcccc 3540

tgttcatggg ttggcagaat caacatagtg aagatggcaa tacttcctaa agtgatatat 3600

cgctttaacg ccatcccaat taaactgcct atgaccttct ttacggagct cgagaaaaca 3660

acccttaaat ttatatggaa tcaaaagaga gcaagaatag cgaagtccat cttgagccag 3720

aagaataagg ccggtgggat tactttgcct gattttaagt tgtattataa agccacagta 3780

actaagacag cctggtattg gtatcagaat agagacatcg accagtggaa tcggaccgaa 3840

ccatcagaga taatgcccca catctataat taccttatat tcgataagcc agaaaagaat 3900

aaacagtggg gcaaagacag cctcttcaac aagtggtgtt gggagaattg gctggccata 3960

tgccggaaac tcaagctcga ccccttttctt acaccctaca ctaaaatcaa cagtaggtgg 4020

atcaaggact tgaatgtcaa gccaaagact ataaagacac tggaagagaa tcttgggatc 4080

acaatacaag atataggcgt cggcaaagat tttatgtcaa agacgcccaa ggccatggcc 4140

actaaggata agattgataa gtgggacctt attaagctca aaagcttctg tactgccaag 4200

gagaccacga tcagagttaa taggcagccc actacatggg aaaagatttt cgccacttat 4260

tcatcagata aggggttgat aagcagaata tataacgagc tgaagcagat ctacaagaag 4320

aaaacgaata atcccatcaa gaagtgggca aaagatatga acaggcattt tagcaaagag 4380

gatatctacg ccgcgaagaa gcatatgaag aagtgtagtt caagcttggc cattcgtgag 4440

atgcagatta agacgaccat gcgataccac cttaccccag tgaggatggc aattatcaag 4500

aaatctggca ataatagatg ttggcggggc tgtggcgaga ttggcaccct gctccattgc 4560

tggtgggatt gcaagctggt gcagccgctt tggaaatcag tctggcgctt tctgagggac 4620

ctcgagcttg agattccctt cgatcccgca attcccttgc tcggaatcta tcctaacgaa 4680

tacaagagct gttgttacaa ggatacgtgt acccggatgt tcatcgcggc cttgtttacg 4740

atagctaaga cgtggaatca gcctaagtgc cccacaatga tcgattggat caagaaaatg 4800

tggcatattt ataccatgga gtattacgca gcaattaaga atgacgaatt tatttccttc 4860

gttgggacct ggatgaagct ggagactatt attctgagca agctgtctca ggagcaaaag 4920

acaaagcata gaatcttctc tctcattggt ggtaacgact acaaagacga tgacgacaag 4980

cccgccgcca agagggtgaa gctggactaa agcgcttcta gaagttgtct cctcctgcac 5040

tgactgactg atacaatcga tttctggatc cgcaggccta atcaacctct ggattacaaa 5100

atttgtgaaa gattgactgg tattcttaac tatgttgctc cttttacgct atgtggatac 5160
gctgctttaa tgcctttgta tcatgctatt gcttcccgta tggctttcat tttctcctcc 5220
ttgtataaat cctggttgct gtctctttat gaggagttgt ggcccgttgt caggcaacgt 5280
ggcgtggtgt gcactgtgtt tgctgacgca acccccactg gttggggcat tgccaccacc 5340
tgtcagctcc tttccgggac tttcgctttc cccctcccta ttgccacggc ggaactcatc 5400
gccgcctgcc ttgcccgctg ctggacaggg gctcggctgt tgggcactga caattccgtg 5460
gtgttgtcgg ggaagctgac gtcctttcca tggctgctcg cctgtgttgc cacctggatt 5520

ctgcgcggga cgtccttctg ctacgtccct tcggccctca atccagcgga ccttccttcc 5580
cgctgagaga cacaaaaaat tccaacacac tattgcaatg aaaataaatt tcctttatta 5640
gccagaagtc agatgctcaa ggggcttcat gatgtcccca taatttttgg cagagggaaa 5700
aagatctcag tggtatttgt gagccagggc attggccttc tgataggcag cctgcacctg 5760
aggagtgcgg ccgctttact tgtacagctc gtccatgccg agagtgatcc cggcggcggt 5820
cacgaactcc agcaggacca tgtgatcgcg cttctcgttg gggtctttgc tcagggcgga 5880
ctgggtgctc aggtagtggt tgtcgggcag cagcacgggg ccgtcgccga tgggggtgtt 5940

ctgctggtag tggtcggcga gctgcacgct gccgtcctcg atgttgtggc ggatcttgaa 6000
gttcaccttg atgccgttct tctgcttgtc ggccatgata tagacgttgt ggctgttgta 6060
gttgtactcc agcttgtgcc ccaggatgtt gccgtcctcc ttgaagtcga tgcccttcag 6120
ctcgatgcgg ttcaccaggg tgtcgccctc gaacttcacc tcggcgcggg tcttgtagtt 6180
gccgtcgtcc ttgaagaaga tggtgcgctc ctggacgtag ccttcgggca tggcggactt 6240
gaagaagtcg tgctgcttca tgtggtcggg gtagcggctg aagcactgca cgccgtaggt 6300
cagggtggtc acgagggtgg gccagggcac gggcagcttg ccggtggtgc agatgaactt 6360

cagggtcagc ttgccgtagg tggcatcgcc ctcgccctcg ccggacacgc tgaacttgtg 6420
gccgtttacg tcgccgtcca gctcgaccag gatgggcacc accccggtga acagctcctc 6480
gcccttgctc accatggtgg cgggatctga cggttcacta aaccagctct gcttatatag 6540
acctcccacc gtacacgcct accgcccatt tgcgtcaatg gggcggagtt gttacgacat 6600
tttggaaagt cccgttgatt ttggtgccaa aacaaactcc cattgacgtc aatggggtgg 6660
agacttggaa atccccgtga gtcaaaccgc tatccacgcc cattgatgta ctgccaaaac 6720
cgcatcacca tggtaatagc gatgactaat acgtagatgt actgccaagt aggaaagtcc 6780

cataaggtca tgtactgggc ataatgccag gcgggccatt taccgtcatt gacgtcaata 6840

gggggcgtac ttggcatatg atacacttga tgtactgcca agtgggcagt ttaccgtaaa 6900

tactccaccc attgacgtca atggaaagtc cctattggcg ttactatggg aacatacgtc 6960

attattgacg tcaatgggcg ggggtcgttg ggcggtcagc caggcgggcc atttaccgta 7020

agttatgtaa cgggcctgct gccggctctg cggcctcttc cgcgtcttcg ccttcgccct 7080

cagacgagtc ggatctccct ttgggccgcc tccccgcctg tctagcttga ctgactgaga 7140

tacagcgtac cttcagctca cagacatgat aagatacatt gatgagtttg gacaaaccac 7200

aactagaatg cagtgaaaaa aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt 7260

tgtaaccatt ataagctgca ataaacaagt t 7291

<210> 43

<211> 7376

<212> DNA

<213> Artificial Sequence

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

<400> 43

taatacgact cactataggg agaagtactg ccaccatggg caagaagcaa aatcgcaaga 60

cggggaattc caagacacaa tccgctagcc caccacctaa agagcgttct agctcccctg 120

ctactgagca gtcctggatg gaaaacgact tcgatgaact ccgggaagag ggatttaggc 180

gatccaacta ttcagaactc cgcgaagata tccagacaaa ggggaaggaa gtcgagaatt 240

tcgagaagaa cctcgaggag tgcatcaccc gtatcacaaa cactgagaaa tgtctcaaag 300

aactcatgga acttaagaca aaagccaggg agcttcgaga ggagtgtcgg agtctgagat 360

ccaggtgtga ccagctcgag gagcgcgtga gcgcgatgga agacgagatg aacgagatga 420

aaagagaggg caaattcagg gagaagcgca ttaagaggaa cgaacagagt ctgcaggaga 480

tttgggatta cgtcaagagg cctaacctgc ggttgatcgg cgtccccgag agcgacgtag 540

aaaacgggac taaactggag aatacacttc aagacatcat tcaagaaaat tttccaaacc 600

tggctcggca agctaatgtg caaatccaag agatccaacg cacaccccag cggtatagct 660

ctcggcgtgc cacccctagg catattatcg tgcgctttac taaggtggag atgaaagaga 720

agatgctgcg agccgctcgg gaaaagggaa gggtgacttt gaagggcaaa cctattcggc 780

tgacggttga ccttagcgcc gagacactcc aggcacgccg ggaatggggc cccatcttta 840

atatcctgaa ggagaagaac ttccagccac gaatctctta ccctgcaaag ttgagtttta 900

tctccgaggg tgagattaag tatttcatcg ataaacagat gctgcgagac ttcgtgacaa 960

ctcgcccagc tctcaaggaa ctgctcaaag aggctcttaa tatggagcgc aataatagat 1020

atcaaccctt gcagaaccac gcaaagatgt gagacagccg tcagaccatc aagactagga 1080

agaaactgca tcaactaatg agcaaaatca ccagctaaca tcatagtata catgaccggc 1140

tctaactcac atatcaccat ccttacactt aacattaacg gcctcaactc agctatcaag 1200

cgccatcggc tggccagctg gatcaaatca caggatccaa gcgtttgttg catccaagag 1260

acccacctga cctgtagaga tactcaccgc ctcaagatca agggatggcg aaagatttat 1320

caggcgaacg gtaagcagaa gaaagccgga gtcgcaattc tggtctcaga caagacggat 1380

ttcaagccca ccaaaattaa gcgtgataag gaaggtcact atattatggt gaaaggcagc 1440

atacagcagg aagaacttac catattgaac atctacgcgc caaacaccgg cgcacctcgc 1500

tttatcaaac aggtcctgtc cgatctgcag cgagatctgg attctcatac gttgattatg 1560

ggtgatttca atacaccatt gagcaccctg gatcgcagca ccaggcaaaa ggtaaataaa 1620

gacacgcaag agctcaatag cgcactgcat caggcagatc tcattgatat ttatcgcact 1680

cttcatccta agagtaccga gtacacattc ttcagcgccc cacatcatac atactcaaag 1740

atcgatcata tcgtcggctc aaaggctctg ctgtcaaagt gcaagcgcac agagataatt 1800

acaaattacc tgtcagatca tagcgcgatc aagctcgagc tgagaatcaa gaacctgacc 1860

cagagccgga gtaccacttg gaagcttaat aacctgctgc tcaacgatta ttgggtccac 1920

aatgagatga aggcagagat taaaatgttc ttcgaaacaa atgagaataa ggatactacc 1980

tatcaaaacc tttgggatgc ctttaaggcc gtctgcagag gcaagttcat cgccctcaac 2040

gcctataaaa gaaaacaaga gagatctaag atcgatactc tcacctctca gctgaaggag 2100

ttggagaaac aggaacagac ccactccaag gcgtcaagac ggcaggagat cacaaagatt 2160

cgcgccgagt tgaaagagat cgaaacccaa aagactcttc agaaaattaa cgagtctcgt 2220

agttggttct tcgagcggat taataagata gacagacctc tggcacgact gattaagaag 2280

aagcgcgaaa agaaccagat tgataccatc aagaacgaca agggcgacat cactactgac 2340

ccgaccgaga tccagaccac tattcgggag tattataagc atttgtatgc taacaagctt 2400

gagaacctgg aagagatgga cacttttctg gatacctata ctctgccacg gcttaatcaa 2460

gaggaagtcg agtccctcaa ccgcccaatt acaggaagcg agattgtggc cataattaac 2520

tccctgccga caaagaaatc tcctggtccg gacgggttta cagctgagtt ttatcaacgg 2580

tatatggaag agcttgtacc gtttctgctc aagctctttc agtctataga aaaggaaggc 2640

atcttgccca attccttcta cgaagcttct ataatactta ttcccaaacc aggacgcgat 2700

accacaaaga aggaaaactt ccggcccatt agtctcatga atatcgacgc taaaatattg 2760

aacaagattc tcgccaacag aatccaacaa catattaaga aattgataca tcacgaccag 2820

gtggggttta tacctggcat gcagggctgg tttaacatcc ggaagagtat taacgtcatt 2880

caacacatta atagagctaa ggataagaat catatgatca tctctataga cgcggaaaag 2940

gcattcgata agattcagca gccatttatg ctcaagactc tgaacaaact cggcatcgac 3000

ggaacatatt ttaagattat tcgcgcaatt tacgataagc cgactgctaa cattatcctt 3060

aacggccaaa agctcgaggc ctttccgctc aagactggaa cccgccaagg ctgtcccctc 3120

tccccgcttt tgtttaatat tgtactcgag gtgctggcta gggctattcg tcaagagaaa 3180

gagattaaag ggatacagct cgggaaggaa gaggtcaagc tttccttgtt cgccgatgat 3240

atgattgtgt acctggagaa tcctattgtg tctgctcaga accttcttaa acttatttct 3300

aactttagca aggtcagcgg ctataagatt aacgtccaga aatctcaggc ctttctgtac 3360

acaaataatc gacagaccga atcccagata atgggtgagc ttccgtttgt catagccagc 3420

aaaaggataa agtatctcgg aatccagctg acacgagacg ttaaagattt gtttaaggaa 3480

aattacaagc ctctcctgaa agagattaag gaagatacta ataagtggaa gaatatcccc 3540

tgttcatggg ttggcagaat caacatagtg aagatggcaa tacttcctaa agtgatatat 3600

cgctttaacg ccatcccaat taaactgcct atgaccttct ttacggagct cgagaaaaca 3660

acccttaaat ttatatggaa tcaaaagaga gcaagaatag cgaagtccat cttgagccag 3720

aagaataagg ccggtgggat tactttgcct gattttaagt tgtattataa agccacagta 3780

actaagacag cctggtattg gtatcagaat agagacatcg accagtggaa tcggaccgaa 3840

ccatcagaga taatgcccca catctataat taccttatat tcgataagcc agaaaagaat 3900

aaacagtggg gcaaagacag cctcttcaac aagtggtgtt gggagaattg gctggccata 3960

tgccggaaac tcaagctcga ccccttttctt acaccctaca ctaaaatcaa cagtaggtgg 4020

atcaaggact tgaatgtcaa gccaaagact ataaagacac tggaagagaa tcttgggatc 4080

acaatacaag atataggcgt cggcaaagat tttatgtcaa agacgcccaa ggccatggcc 4140

actaaggata agattgataa gtgggacctt attaagctca aaagcttctg tactgccaag 4200

gagaccacga tcagagttaa taggcagccc actacatggg aaaagatttt cgccacttat 4260

tcatcagata aggggttgat aagcagaata tataacgagc tgaagcagat ctacaagaag 4320

aaaacgaata atcccatcaa gaagtgggca aaagatatga acaggcattt tagcaaagag 4380

gatatctacg ccgcgaagaa gcatatgaag aagtgtagtt caagcttggc cattcgtgag 4440

atgcagatta agacgaccat gcgataccac cttaccccag tgaggatggc aattatcaag 4500

aaatctggca ataatagatg ttggcggggc tgtggcgaga ttggcaccct gctccattgc 4560

tggtgggatt gcaagctggt gcagccgctt tggaaatcag tctggcgctt tctgagggac 4620

ctcgagcttg agattccctt cgatcccgca attcccttgc tcggaatcta tcctaacgaa 4680

tacaagagct gttgttacaa ggatacgtgt acccggatgt tcatcgcggc cttgtttacg 4740

atagctaaga cgtggaatca gcctaagtgc cccacaatga tcgattggat caagaaaatg 4800

tggcatattt ataccatgga gtattacgca gcaattaaga atgacgaatt tatttccttc 4860

gttgggacct ggatgaagct ggagactatt attctgagca agctgtctca ggagcaaaag 4920

acaaagcata gaatcttctc tctcattggt ggtaacgact acaaagacga tgacgacaag 4980

taaagcggcc gggcgcggtg gctcacgcct gtaatcccag cactttggga ggccgaggcg 5040

ggaggatcgc agttcgagac cagcgcgaga ccccgtctct acaaaaatac aaaaattagc 5100

ttctagaagt tgtctcctcc tgcactgact gactgataca atcgatttct ggatccgcag 5160

gcctaatcaa cctctggatt acaaaatttg tgaaagattg actggtattc ttaactatgt 5220

tgctcctttt acgctatgtg gatacgctgc tttaatgcct ttgtatcatg ctattgcttc 5280

ccgtatggct ttcattttct cctccttgta taaatcctgg ttgctgtctc tttatgagga 5340

gttgtggccc gttgtcaggc aacgtggcgt ggtgtgcact gtgtttgctg acgcaacccc 5400

cactggttgg ggcattgcca ccacctgtca gctcctttcc gggactttcg ctttcccctt 5460

ccctattgcc acggcggaac tcatcgccgc ctgccttgcc cgctgctgga caggggctcg 5520

gctgttgggc actgacaatt ccgtggtgtt gtcggggaag ctgacgtcct ttccatggct 5580

gctcgcctgt gttgccacct ggattctgcg cgggacgtcc ttctgctacg tcccttcggc 5640

cctcaatcca gcggaccttc cttcccgctg agagacacaa aaaattccaa cacactattg 5700

caatgaaaat aaatttcctt tattagccag aagtcagatg ctcaaggggc ttcatgatgt 5760

ccccataatt tttggcagag ggaaaaagat ctcagtggta tttgtgagcc agggcattgg 5820

ccttctgata ggcagcctgc acctgaggag tgcggccgct ttacttgtac agctcgtcca 5880

tgccgagagt gatcccggcg gcggtcacga actccagcag gaccatgtga tcgcgcttct 5940

cgttggggtc tttgctcagg gcggactggg tgctcaggta gtggttgtcg ggcagcagca 6000

cggggccgtc gccgatgggg gtgttctgct ggtagtggtc ggcgagctgc acgctgccgt 6060

cctcgatgtt gtggcggatc ttgaagttca ccttgatgcc gttcttctgc ttgtcggcca 6120

tgatatagac gttgtggctg ttgtagttgt actccagctt gtgccccagg atgttgccgt 6180

cctccttgaa gtcgatgccc ttcagctcga tgcggttcac cagggtgtcg ccctcgaact 6240

tcacctcggc gcgggtcttg tagttgccgt cgtccttgaa gaagatggtg cgctcctgga 6300

cgtagccttc gggcatggcg gacttgaaga agtcgtgctg cttcatgtgg tcggggtagc 6360

ggctgaagca ctgcacgccg taggtcaggg tggtcacgag ggtgggccag ggcacgggca 6420

gcttgccggt ggtgcagatg aacttcaggg tcagcttgcc gtaggtggca tcgccctcgc 6480

cctcgccgga cacgctgaac ttgtggccgt ttacgtcgcc gtccagctcg accaggatgg 6540
gcaccacccc ggtgaacagc tcctcgccct tgctcaccat ggtggcggga tctgacggtt 6600
cactaaacca gctctgctta tatagacctc ccaccgtaca cgcctaccgc ccatttgcgt 6660
caatggggcg gagttgttac gacattttgg aaagtcccgt tgattttggt gccaaaacaa 6720
actcccattg acgtcaatgg ggtggagact tggaaatccc cgtgagtcaa accgctatcc 6780
acgcccattg atgtactgcc aaaaccgcat caccatggta atagcgatga ctaatacgta 6840
gatgtactgc caagtaggaa agtcccataa ggtcatgtac tgggcataat gccaggcggg 6900

ccatttaccg tcattgacgt caatggggcg cgtacttggc atatgataca cttgatgtac 6960
tgccaagtgg gcagtttacc gtaaatactc cacccattga cgtcaatgga aagtccctat 7020
tggcgttact atgggaacat acgtcattat tgacgtcaat gggcgggggt cgttgggcgg 7080
tcagccaggc gggccattta ccgtaagtta tgtaacgggc ctgctgccgg ctctgcggcc 7140
tcttccgcgt cttcgccttc gccctcagac gagtcggatc tccctttggg ccgcctcccc 7200
gcctgtctag cttgactgac tgagatacag cgtaccttca gctcacagac atgataagat 7260
acattgatga gtttggacaa accacaacta gaatgcagtg aaaaaaatgc tttatttgtg 7320

aaatttgtga tgctattgct ttatttgtaa ccattataag ctgcaataaa caagtt 7376

<210> 44
<211> 14122
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 44

taatacgact cactataggg agaagtactg ccaccatggg caagaagcaa aatcgcaaga 60
cggggaattc caagacacaa tccgctagcc caccacctaa agagcgttct agctcccctg 120
ctactgagca gtcctggatg gaaaacgact tcgatgaact ccgggaagag ggatttaggc 180
gatccaacta ttcagaactc cgcgaagata tccagacaaa ggggaaggaa gtcgagaatt 240

tcgagaagaa cctcgaggag tgcatcaccc gtatcacaaa cactgagaaa tgtctcaaag 300
aactcatgga acttaagaca aaagccaggg agcttcgaga ggagtgtcgg agtctgagat 360
ccaggtgtga ccagctcgag gagcgcgtga gcgcgatgga agacgagatg aacgagatga 420
aaagagaggg caaattcagg gagaagcgca ttaagaggaa cgaacagagt ctgcaggaga 480

tttgggatta cgtcaagagg cctaacctgc ggttgatcgg cgtccccgag agcgacgtag 540

aaaacgggac taaactggag aatacacttc aagacatcat tcaagaaaat tttccaaacc 600

tggctcggca agctaatgtg caaatccaag agatccaacg cacaccccag cggtatagct 660

ctcggcgtgc cacccctagg catattatcg tgcgctttac taaggtggag atgaaagaga 720

agatgctgcg agccgctcgg gaaaagggaa gggtgacttt gaagggcaaa cctattcggc 780

tgacggttga ccttagcgcc gagacactcc aggcacgccg ggaatggggc cccatcttta 840

atatcctgaa ggagaagaac ttccagccac gaatctctta ccctgcaaag ttgagtttta 900

tctccgaggg tgagattaag tatttcatcg ataaacagat gctgcgagac ttcgtgacaa 960

ctcgcccagc tctcaaggaa ctgctcaaag aggctcttaa tatggagcgc aataatagat 1020

atcaaccctt gcagaaccac gcaaagatgt gagacagtta aaacagcctg tgggttgatc 1080

ccacccacag gcccattggg cgctagcact ctggtatcac ggtacctttg tgcgcctgtt 1140

ttatacccccc tcccccaact gtaacttaga agtaacacac accgatcaac agtcagcgtg 1200

gcacaccagc cacgttttga tcaagcactt ctgttacccc ggactgagta tcaatagact 1260

gctcacgcgg ttgaaggaga aagcgttcgt tatccggcca actacttcga aaaacctagt 1320

aacaccgtgg aagttgcaga gtgtttcgct cagcactacc ccagtgtaga tcaggtcgat 1380

gagtcaccgc attccccacg ggcgaccgtg gcggtggctg cgttggcggc ctgcccatgg 1440

ggaaacccat gggacgctct aatacagaca tggtgcgaag agtctattga gctagttggt 1500

agtcctccgg cccctgaatg cggctaatcc taactgcgga gcacacaccc tcaagccaga 1560

gggcagtgtg tcgtaacggg caactctgca gcggaaccga ctactttggg tgtccgtgtt 1620

tcattttatt cctatactgg ctgcttatgg tgacaattga gagatcgtta ccatatagct 1680

attggattgg ccatccggtg actaatagag ctattatata tccctttgtt gggtttatac 1740

cacttagctt gaaagaggtt aaaacattac aattcattgt taagttgaat acagcaaata 1800

catgaccggc tctaactcac atatcaccat ccttacactt aacattaacg gcctcaactc 1860

agctatcaag cgccatcggc tggccagctg gatcaaatca caggatccaa gcgtttgttg 1920

catccaagag acccacctga cctgtagaga tactcaccgc ctcaagatca agggatggcg 1980

aaagatttat caggcgaacg gtaagcagaa gaaagccgga gtcgcaattc tggtctcaga 2040

caagacggat ttcaagccca ccaaaattaa gcgtgataag gaaggtcact atattatggt 2100

gaaaggcagc atacagcagg aagaacttac catattgaac atctacgcgc caaacaccgg 2160

cgcacctcgc tttatcaaac aggtcctgtc cgatctgcag cgagatctgg attctcatac 2220

gttgattatg ggtgatttca atacaccatt gagcaccctg gatcgcagca ccaggcaaaa 2280

ggtaaataaa gacacgcaag agctcaatag cgcactgcat caggcagatc tcattgatat 2340

ttatcgcact cttcatccta agagtaccga gtacacattc ttcagcgccc cacatcatac 2400
atactcaaag atcgatcata tcgtcggctc aaaggctctg ctgtcaaagt gcaagcgcac 2460
agagataatt acaaattacc tgtcagatca tagcgcgatc aagctcgagc tgagaatcaa 2520
gaacctgacc cagagccgga gtaccacttg gaagcttaat aacctgctgc tcaacgatta 2580
ttgggtccac aatgagatga aggcagagat taaaatgttc ttcgaaacaa atgagaataa 2640
ggatactacc tatcaaaacc tttgggatgc ctttaaggcc gtctgcagag gcaagttcat 2700
cgccctcaac gcctataaaa gaaaacaaga gagatctaag atcgatactc tcacctctca 2760

gctgaaggag ttggagaaac aggaacagac ccactccaag gcgtcaagac ggcaggagat 2820
cacaaagatt cgcgccgagt tgaaagagat cgaaacccaa aagactcttc agaaaattaa 2880
cgagtctcgt agttggttct tcgagcggat taataagata gacagacctc tggcacgact 2940
gattaagaag aagcgcgaaa agaaccagat tgataccatc aagaacgaca agggcgacat 3000
cactactgac ccgaccgaga tccagaccac tattcgggag tattataagc atttgtatgc 3060
taacaagctt gagaacctgg aagagatgga cacttttctg gatacctata ctctgccacg 3120
gcttaatcaa gaggaagtcg agtccctcaa ccgcccaatt acaggaagcg agattgtggc 3180

cataattaac tccctgccga caaagaaatc tcctggtccg gacgggttta cagctgagtt 3240
ttatcaacgg tatatggaag agcttgtacc gtttctgctc aagctctttc agtctataga 3300
aaaggaaggc atcttgccca attccttcta cgaagcttct ataatactta ttcccaaacc 3360
aggacgcgat accacaaaga aggaaaactt ccggcccatt agtctcatga atatcgacgc 3420
taaaatattg aacaagattc tcgccaacag aatccaacaa catattaaga aattgataca 3480
tcacgaccag gtggggttta tacctggcat gcagggctgg tttaacatcc ggaagagtat 3540
taacgtcatt caacacatta atagagctaa ggataagaat catatgatca tctctataga 3600

cgcggaaaag gcattcgata agattcagca gccatttatg ctcaagactc tgaacaaact 3660
cggcatcgac ggaacatatt ttaagattat tcgcgcaatt tacgataagc cgactgctaa 3720
cattatcctt aacggccaaa agctcgaggc ctttccgctc aagactggaa cccgccaagg 3780
ctgtcccctc tccccgcttt tgtttaatat tgtactcgag gtgctggcta gggctattcg 3840
tcaagagaaa gagattaaag ggatacagct cgggaaggaa gaggtcaagc tttccttgtt 3900
cgccgatgat atgattgtgt acctggagaa tcctattgtg tctgctcaga accttcttaa 3960
acttatttct aactttagca aggtcagcgg ctataagatt aacgtccaga aatctcaggc 4020

ctttctgtac acaaataatc gacagaccga atcccagata atgggtgagc ttccgtttgt 4080

catagccagc aaaaggataa agtatctcgg aatccagctg acacgagacg ttaaagattt 4140

gtttaaggaa aattacaagc ctctcctgaa agagattaag gaagatacta ataagtggaa 4200

gaatatcccc tgttcatggg ttggcagaat caacatagtg aagatggcaa tacttcctaa 4260

agtgatatat cgctttaacg ccatcccaat taaactgcct atgaccttct ttacggagct 4320

cgagaaaaca acccttaaat ttatatggaa tcaaaagaga gcaagaatag cgaagtccat 4380

cttgagccag aagaataagg ccggtgggat tactttgcct gattttaagt tgtattataa 4440

agccacagta actaagacag cctggtattg gtatcagaat agagacatcg accagtggaa 4500

tcggaccgaa ccatcagaga taatgcccca catctataat taccttatat tcgataagcc 4560

agaaaagaat aaacagtggg gcaaagacag cctcttcaac aagtggtgtt gggagaattg 4620

gctggccata tgccggaaac tcaagctcga ccccttttctt acaccctaca ctaaaatcaa 4680

cagtaggtgg atcaaggact tgaatgtcaa gccaaagact ataaagacac tggaagagaa 4740

tcttgggatc acaatacaag atataggcgt cggcaaagat tttatgtcaa agacgcccaa 4800

ggccatggcc actaaggata agattgataa gtgggacctt attaagctca aaagcttctg 4860

tactgccaag gagaccacga tcagagttaa taggcagccc actacatggg aaaagatttt 4920

cgccacttat tcatcagata aggggttgat aagcagaata tataacgagc tgaagcagat 4980

ctacaagaag aaaacgaata atcccatcaa gaagtgggca aaagatatga acaggcattt 5040

tagcaaagag gatatctacg ccgcgaagaa gcatatgaag aagtgtagtt caagcttggc 5100

cattcgtgag atgcagatta agacgaccat gcgataccac cttaccccag tgaggatggc 5160

aattatcaag aaatctggca ataatagatg ttggcggggc tgtggcgaga ttggcaccct 5220

gctccattgc tggtgggatt gcaagctggt gcagccgctt tggaaatcag tctggcgctt 5280

tctgagggac ctcgagcttg agattcccttcgatcccgca attcccttgc tcggaatcta 5340

tcctaacgaa tacaagagct gttgttacaa ggatacgtgt acccggatgt tcatcgcggc 5400

cttgtttacg atagctaaga cgtggaatca gcctaagtgc cccacaatga tcgattggat 5460

caagaaaatg tggcatattt ataccatgga gtattacgca gcaattaaga atgacgaatt 5520

tatttccttc gttgggacct ggatgaagct ggagactatt attctgagca agctgtctca 5580

ggagcaaaag acaaagcata gaatcttctc tctcattggt ggtaacgact acaaagacga 5640

tgacgacaag taaagcgctt ctagaagttg tctcctcctg cactgactga ctgatacaat 5700

cgatttctgg atccgcaggc ctaatcaacc tctggattac aaaatttgtg aaagattgac 5760

tggtattctt aactatgttg ctccttttac gctatgtgga tacgctgctt taatgccttt 5820

gtatcatgct attgcttccc gtatggcttt cattttctcc tccttgtata aatcctggtt 5880

gctgtctctt tatgaggagt tgtggcccgt tgtcaggcaa cgtggcgtgg tgtgcactgt 5940

gtttgctgac gcaaccccca ctggttgggg cattgccacc acctgtcagc tcctttccgg 6000

gactttcgct ttccccctcc ctattgccac ggcggaactc atcgccgcct gccttgcccg 6060

ctgctggaca ggggctcggc tgttgggcac tgacaattcc gtggtgttgt cggggaagct 6120

gacgtccttt ccatggctgc tcgcctgtgt tgccacctgg attctgcgcg ggacgtcctt 6180

ctgctacgtc ccttcggccc tcaatccagc ggaccttcct tcccgcgaac aaacgaccca 6240

acacccgtgc gttttattct gtctttttat tgccgatccc ctcagaagaa ctcgtcaaga 6300

aggcgataga aggcgatgcg ctgcgaatcg ggagcggcga taccgtaaag cacgaggaag 6360

cggtcagccc attcgccgcc aagctcttca gcaatatcac gggtagccaa cgctatgtcc 6420

tgatagcggt cggccgcttt acttgtacag ctcgtccatg ccgagagtga tcccggcggc 6480

ggtcacgaac tccagcagga ccatgtgatc gcgcttctcg ttggggtctt tgctcagggc 6540

ggactgggtg ctcaggtagt ggttgtcggg cagcagcacg gggccgtcgc cgatgggggt 6600

gttctgctgg tagtggtcgg ccaggtgagt ccaggagatg tttcagcact gttgccttta 6660

gtctcgaggc aacttagaca actgagtatt gatctgagca cagcagggtg tgagctgttt 6720

gaagatactg gggttggggg tgaagaaact gcagaggact aactgggctg agacccagtg 6780

gcaatgtttt agggcctaag gaatgcctct gaaaatctag atggacaact ttgactttga 6840

gaaaagagag gtggaaatga ggaaaatgac ttttctttat tagatttcgg tagaaagaac 6900

tttcatcttt cccctatttt tgttattcgt tttaaaacat ctatctggag gcaggacaag 6960

tatggtcatt aaaaagatgc aggcagaagg catatattgg ctcagtcaaa gtggggaact 7020

ttggtggcca aacatacatt gctaaggcta ttcctatatc agctggacac atataaaatg 7080

ctgctaatgc ttcattacaa acttatatcc tttaattcca gatgggggca aagtatgtcc 7140

aggggtgagg aacaattgaa acatttgggc tggagtagat tttgaaagtc agctctgtgt 7200

gtgtgtgtgt gtgtgtgtgt gtgagagcgt gtgtttcttt taacgttttc agcctacagc 7260

atacagggtt catggtggca agaagataac aagatttaaa ttatggccag tgactagtgc 7320

tgcaagaaga acaactacct gcatttaatg ggaaagcaaa atctcaggct ttgagggaag 7380

ttaacatagg cttgattctg ggtggaagct gggtgtgtag ttatctggag gccaggctgg 7440

agctctcagc tcactatggg ttcatcttta ttgtctcctt tcatctcaac agctgcacgc 7500

tgccgtcctc gatgttgtgg cggatcttga agttcacctt gatgccgttc ttctgcttgt 7560

cggccatgat atagacgttg tggctgttgt agttgtactc cagcttgtgc cccaggatgt 7620

tgccgtcctc cttgaagtcg atgcccttca gctcgatgcg gttcaccagg gtgtcgccct 7680

cgaacttcac ctcggcgcgg gtcttgtagt tgccgtcgtc cttgaagaag atggtgcgct 7740

cctggacgta gccttcgggc atggcggact tgaagaagtc gtgctgcttc atgtggtcgg 7800

ggtagcggct gaagcactgc acgccgtagg tcagggtggt cacgagggtg ggccagggca 7860
cgggcagctt gccggtggtg cagatgaact tcagggtcag cttgccgtag gtggcatcgc 7920
cctcgccctc gccggacacg ctgaacttgt ggccgtttac gtcgccgtcc agctcgacca 7980
ggatgggcac caccccggtg aacagctcct cgcccttgct caccatggtg gcgaattcga 8040
agcttgagca cgagatctga gtccggtagg cctagcggat ctgacggttc actaaaccag 8100
ctctgcttat atagacctcc caccgtacac gcctaccgcc catttgcgtc aatggggcgg 8160
agttgttacg acattttgga aagtcccgtt gattttggtg ccaaaacaaa ctcccattga 8220

cgtcaatggg gtggagactt ggaaatcccc gtgagtcaaa ccgctatcca cgcccattga 8280
tgtactgcca aaaccgcatc accatggtaa tagcgatgac taatacgtag atgtactgcc 8340
aagtaggaaa gtcccataag gtcatgtact gggcataatg ccaggcgggc catttaccgt 8400
cattgacgtc aataggggggc gtacttggca tatgatacac ttgatgtact gccaagtggg 8460
cagtttaccg taaatactcc acccattgac gtcaatggaa agtccctatt ggcgttacta 8520
tgggaacata cgtcattatt gacgtcaatg ggcgggggtc gttgggcggt cagccaggcg 8580
ggccatttac cgtaagttat gtaacgggcc tgctgccggc tctgcggcct cttccgcgtc 8640

ttcgccttcg ccctcagacg agtcggatct ccctttgggc cgcctccccg cctgtctagc 8700
ttgactgact gagatacagc gtaccttcag ctcacagaca tgataagata cattgatgag 8760
tttggacaaa ccacaactag aatgcagtga aaaaaatgct ttatttgtga aatttgtgat 8820
gctattgctt tatttgtaac cattataagc tgcaataaac aagttaacaa caacaattgc 8880
attcatttta tgtttcaggt tcaggggga gtgtgggagg ttttttaaag caagtaaaac 8940
ctctacaaat gtggtattgg cccatctcta tcggtatcgt agcataaccc cttggggcct 9000
ctaaacgggt cttgaggggt tttttgtgcc cctcgggccg gattgctatc taccggcatt 9060

ggcgcagaaa aaaatgcctg atgcgacgct gcgcgtctta tactcccaca tatgccagat 9120
tcagcaacgg atacggcttc cccaacttgc ccacttccat acgtgtcctc cttaccagaa 9180
atttatcctt aaggtcgtca gctatcctgc aggcgatctc tcgatttcga tcaagacatt 9240
cctttaatgg tcttttctgg acaccactag gggtcagaag tagttcatca aactttcttc 9300
cctccctaat ctcattggtt accttgggct atcgaaactt aattaagcga tctgcatctc 9360
aattagtcag caaccatagt cccgccccta actccgccca tcccgcccct aactccgccc 9420
agttccgccc attctccgcc ccatcgctga ctaatttttt ttatttatgc agaggccgag 9480

gccgcctcgg cctctgagct attccagaag tagtgaggag gctttttttgg aggcctaggc 9540

ttttgcaaag gaggtagcca acatgattga acaagatgga ttgcacgcag gttctcccgc 9600

cgcttgggtg gagaggctat tcggctatga ctgggcacaa cagacaatcg gctgctctga 9660

tgccgccgtg ttccggctgt cagcgcaggg gcgcccggtt ctttttgtca agaccgacct 9720

gtccggtgcc ctgaatgaac tccaggacga ggcagcgcgg ctatcgtggc tggccacgac 9780

gggcgttcct tgcgcagctg tgctcgacgt tgtcactgaa gcgggaaggg actggctgct 9840

attgggcgaa gtgccggggc aggatctcct gtcatctcac cttgctcctg ccgagaaagt 9900

atccatcatg gctgatgcaa tgcggcggct gcatacgctt gatccggcta cctgcccatt 9960

cgaccaccaa gcgaaacatc gcatcgagcg agcacgtact cggatggaag ccggtcttgt 10020

cgatcaggat gatctggacg aagagcatca ggggctcgcg ccagccgaac tgttcgccag 10080

gctcaaggcg cggatgcccg acggcgagga tctcgtcgtg acccacggcg atgcctgctt 10140

gccgaatatc atggtggaaa atggccgctt ttctggattc atcgactgtg gccggctggg 10200

tgtggcggac cgctatcagg acatagcgtt ggctacccgt gatattgctg aagagcttgg 10260

cggcgaatgg gctgaccgct tcctcgtgct ttacggtatc gccgctcccg attcgcagcg 10320

catcgccttc tatcgccttc ttgacgagtt cttctagtat gtaagccctg tgccttctag 10380

ttgccagcca tctgttgttt gcccctcccc cgtgccttcc ttgaccctgg aaggtgccac 10440

tcccactgtc ctttcctaat aaaatgagga aattgcatcg cattgtctga gtaggtgtca 10500

ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg aagacaatag 10560

caggcatgct ggggatgcgg tgggctctat ggttaattaa ccagtcaagt cagctacttg 10620

gcgagatcga cttgtctggg tttcgactac gctcagaatt gcgtcagtca agttcgatct 10680

ggtccttgct attgcacccg ttctccgatt acgagtttca tttaaatcat gtgagcaaaa 10740

ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc 10800

cgccccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca 10860

ggactataaa gataccaggc gtttcccccct ggaagctccc tcgtgcgctc tcctgttccg 10920

accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct 10980

catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt 11040

gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag 11100

tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc 11160

agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac 11220

actagaagaa cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga 11280

gttggtagct cttgatccgg caaacaaacc accgctggta gcggtggttt ttttgtttgc 11340

aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg 11400

gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca 11460

aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt 11520

atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca 11580

gcgatctgtc tatttcgttc atccatagtt gcatttaaat ttccgaactc tccaaggccc 11640

tcgtcggaaa atcttcaaac ctttcgtccg atccatcttg caggctacct ctcgaacgaa 11700

ctatcgcaag tctcttggcc ggccttgcgc cttggctatt gcttggcagc gcctatcgcc 11760

aggtattact ccaatcccga atatccgaga tcgggatcac ccgagagaag ttcaacctac 11820

atcctcaatc ccgatctatc cgagatccga ggaatatcga aatcggggcg cgcctggtgt 11880

accgagaacg atcctctcag tgcgagtctc gacgatccat atcgttgctt ggcagtcagc 11940

cagtcggaat ccagcttggg acccaggaag tccaatcgtc agatattgta ctcaagcctg 12000

gtcacggcag cgtaccgatc tgtttaaacc tagatattga tagtctgatc ggtcaacgta 12060

taatcgagtc ctagcttttg caaacatcta tcaagagaca ggatcagcag gaggctttcg 12120

catgagtatt caacatttcc gtgtcgccct tattcccttt tttgcggcat tttgccttcc 12180

tgtttttgct cacccagaaa cgctggtgaa agtaaaagat gctgaagatc agttgggtgc 12240

gcgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga gttttcgccc 12300

cgaagaacgc tttccaatga tgagcacttt taaagttctg ctatgtggcg cggtattatc 12360

ccgtattgac gccgggcaag agcaactcgg tcgccgcata cactattctc agaatgactt 12420

ggttgagtat tcaccagtca cagaaaagca tcttacggat ggcatgacag taagagaatt 12480

atgcagtgct gccataacca tgagtgataa cactgcggcc aacttacttc tgacaacgat 12540

tggaggaccg aaggagctaa ccgcttttt gcacaacatg ggggatcatg taactcgcct 12600

tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg acaccacgat 12660

gcctgtagca atggcaacaa ccttgcgtaa actattaact ggcgaactac ttactctagc 12720

ttcccggcaa cagttgatag actggatgga ggcggataaa gttgcaggac cacttctgcg 12780

ctcggccctt ccggctggct ggtttattgc tgataaatct ggagccggtg agcgtgggtc 12840

tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg tagttatcta 12900

cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg agataggtgc 12960

ctcactgatt aagcattggt aaccgattct aggtgcattg gcgcagaaaa aaatgcctga 13020

tgcgacgctg cgcgtcttat actcccacat atgccagatt cagcaacgga tacggcttcc 13080

ccaacttgcc cacttccata cgtgtcctcc ttaccagaaa tttatcctta agatcgttta 13140

aactcgactc tggctctatc gaatctccgt cgtttcgagc ttacgcgaac agccgtggcg 13200

ctcatttgct cgtcgggcat cgaatctcgt cagctatcgt cagcttacct ttttggcagc 13260

gatcgcggct cccgacatct tggaccatta gctccacagg tatcttcttc cctctagtgg 13320
tcataacagc agcttcagct acctctcaat tcaaaaaacc cctcaagacc cgtttagagg 13380
ccccaagggg ttatgctatc aatcgttgcg ttacacacac aaaaaaccaa cacacatcca 13440
tcttcgatgg atagcgattt tattatctaa ctgctgatcg agtgtagcca gatctagtaa 13500
tcaattacgg ggtcattagt tcatagccca tatatggagt tccgcgttac ataacttacg 13560
gtaaatggcc cgcctggctg accgcccaac gacccccgcc cattgacgtc aataatgacg 13620
tatgttccca tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta 13680

cggtaaactg cccacttggc agtacatcaa gtgtatcata tgccaagtac gccccctatt 13740
gacgtcaatg acggtaaatg gcccgcctgg cattatgccc agtacatgac cttatgggac 13800
tttcctactt ggcagtacat ctacgtatta gtcatcgcta ttaccatgct gatgcggttt 13860
tggcagtaca tcaatgggcg tggatagcgg tttgactcac ggggatttcc aagtctccac 13920
cccattgacg tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt 13980
cgtaacaact ccgccccatt gacgcaaatg ggcggtaggc gtgtacggtg ggaggtctat 14040
ataagcagag ctggtttagt gaaccgtcag atcagatctt tgtcgatcct accatccact 14100

cgacacaccc gccagcggcc gc 14122

<210> 45

<211> 14124

<212> DNA

<213> Artificial Sequence

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

<400> 45

taatacgact cactataggg agaagtactg ccaccatggg caagaagcaa aatcgcaaga 60
cggggaattc caagacacaa tccgctagcc caccacctaa agagcgttct agctcccctg 120
ctactgagca gtcctggatg gaaaacgact tcgatgaact ccgggaagag ggatttaggc 180
gatccaacta ttcagaactc cgcgaagata tccagacaaa ggggaaggaa gtcgagaatt 240

tcgagaagaa cctcgaggag tgcatcaccc gtatcacaaa cactgagaaa tgtctcaaag 300
aactcatgga acttaagaca aaagccaggg agcttcgaga ggagtgtcgg agtctgagat 360
ccaggtgtga ccagctcgag gagcgcgtga gcgcgatgga agacgagatg aacgagatga 420
aaagagaggg caaattcagg gagaagcgca ttaagaggaa cgaacagagt ctgcaggaga 480

tttgggatta cgtcaagagg cctaacctgc ggttgatcgg cgtccccgag agcgacgtag 540

aaaacgggac taaactggag aatacacttc aagacatcat tcaagaaaat tttccaaacc 600

tggctcggca agctaatgtg caaatccaag agatccaacg cacaccccag cggtatagct 660

ctcggcgtgc cacccctagg catattatcg tgcgctttac taaggtggag atgaaagaga 720

agatgctgcg agccgctcgg gaaaagggaa gggtgacttt gaagggcaaa cctattcggc 780

tgacggttga ccttagcgcc gagacactcc aggcacgccg ggaatggggc cccatcttta 840

atatcctgaa ggagaagaac ttccagccac gaatctctta ccctgcaaag ttgagtttta 900

tctccgaggg tgagattaag tatttcatcg ataaacagat gctgcgagac ttcgtgacaa 960

ctcgcccagc tctcaaggaa ctgctcaaag aggctcttaa tatggagcgc aataatagat 1020

atcaaccctt gcagaaccac gcaaagatgt gagacagtta aaacagctgt gggttgtcac 1080

ccacccacag ggtccactgg gcgctagtac actggtatct cggtaccttt gtacgcctgt 1140

tttatacccc ctccctgatt tgcaacttag aagcaacgca aaccagatca atagtaggtg 1200

tgacatacca gtcgcatctt gatcaagcac ttctgtatcc ccggaccgag tatcaataga 1260

ctgtgcacac ggttgaagga gaaaacgtcc gttacccggc taactacttc gagaagccta 1320

gtaacgccat tgaagttgca gagtgtttcg ctcagcactc cccccgtgta gatcaggtcg 1380

atgagtcacc gcattcccca cgggcgaccg tggcggtggc tgcgttggcg gcctgcctat 1440

ggggtaaccc ataggacgct ctaatacgga catggcgtga agagtctatt gagctagtta 1500

gtagtcctcc ggcccctgaa tgcggctaat cctaactgcg gagcacatac ccttaatcca 1560

aagggcagtg tgtcgtaacg ggcaactctg cagcggaacc gactactttg ggtgtccgtg 1620

tttcttttta ttcttgtatt ggctgcttat ggtgacaatt aaagaattgt taccatatag 1680

ctattggatt ggccatccag tgtcaaacag agctattgta tatctctttg ttggattcac 1740

acctctcact cttgaaacgt tacacaccct caattacatt atactgctga acacgaagcg 1800

tacatgaccg gctctaactc acatatcacc atccttacac ttaacattaa cggcctcaac 1860

tcagctatca agcgccatcg gctggccagc tggatcaaat cacaggatcc aagcgtttgt 1920

tgcatccaag agacccacct gacctgtaga gatactcacc gcctcaagat caagggatgg 1980

cgaaagattt atcaggcgaa cggtaagcag aagaaagccg gagtcgcaat tctggtctca 2040

gacaagacgg atttcaagcc caccaaaatt aagcgtgata aggaaggtca ctatattatg 2100

gtgaaaggca gcatacagca ggaagaactt accatattga acatctacgc gccaaacacc 2160

ggcgcacctc gctttatcaa acaggtcctg tccgatctgc agcgagatct ggattctcat 2220

acgttgatta tgggtgattt caatacacca ttgagcaccc tggatcgcag caccaggcaa 2280

aaggtaaata aagacacgca agagctcaat agcgcactgc atcaggcaga tctcattgat 2340

atttatcgca ctcttcatcc taagagtacc gagtacacat tcttcagcgc cccacatcat 2400
acatactcaa agatcgatca tatcgtcggc tcaaaggctc tgctgtcaaa gtgcaagcgc 2460
acagagataa ttacaaatta cctgtcagat catagcgcga tcaagctcga gctgagaatc 2520
aagaacctga cccagagccg gagtaccact tggaagctta ataacctgct gctcaacgat 2580
tattgggtcc acaatgagat gaaggcagag attaaaatgt tcttcgaaac aaatgagaat 2640
aaggatacta cctatcaaaa cctttgggat gcctttaagg ccgtctgcag aggcaagttc 2700
atcgccctca acgcctataa aagaaaacaa gagagatcta agatcgatac tctcacctct 2760

cagctgaagg agttggagaa acaggaacag acccactcca aggcgtcaag acggcaggag 2820
atcacaaaga ttcgcgccga gttgaaagag atcgaaaccc aaaagactct tcagaaaatt 2880
aacgagtctc gtagttggtt cttcgagcgg attaataaga tagacagacc tctggcacga 2940
ctgattaaga agaagcgcga aaagaaccag attgatacca tcaagaacga caagggcgac 3000
atcactactg acccgaccga gatccagacc actattcggg agtattataa gcatttgtat 3060
gctaacaagc ttgagaacct ggaagagatg gacacttttc tggataccta tactctgcca 3120
cggcttaatc aagaggaagt cgagtccctc aaccgcccaa ttacaggaag cgagattgtg 3180

gccataatta actccctgcc gacaaagaaa tctcctggtc cggacgggtt tacagctgag 3240
ttttatcaac ggtatatgga agagcttgta ccgtttctgc tcaagctctt tcagtctata 3300
gaaaaggaag gcatcttgcc caattccttc tacgaagctt ctataatact tattcccaaa 3360
ccaggacgcg ataccacaaa gaaggaaaac ttccggccca ttagtctcat gaatatcgac 3420
gctaaaatat tgaacaagat tctcgccaac agaatccaac aacatattaa gaaattgata 3480
catcacgacc aggtggggtt tatacctggc atgcagggct ggtttaacat ccggaagagt 3540
attaacgtca ttcaacacat taatagagct aaggataaga atcatatgat catctctata 3600

gacgcggaaa aggcattcga taagattcag cagccattta tgctcaagac tctgaacaaa 3660
ctcggcatcg acggaacata ttttaagatt attcgcgcaa tttacgataa gccgactgct 3720
aacattatcc ttaacggcca aaagctcgag gcctttccgc tcaagactgg aacccgccaa 3780
ggctgtcccc tctccccgct tttgtttaat attgtactcg aggtgctggc tagggctatt 3840
cgtcaagaga aagagattaa agggatacag ctcgggaagg aagaggtcaa gctttccttg 3900
ttcgccgatg atatgattgt gtacctggag aatcctattg tgtctgctca gaaccttctt 3960
aaacttattt ctaactttag caaggtcagc ggctataaga ttaacgtcca gaaatctcag 4020

gcctttctgt acacaaataa tcgacagacc gaatcccaga taatgggtga gcttccgttt 4080

gtcatagcca gcaaaaggat aaagtatctc ggaatccagc tgacacgaga cgttaaagat 4140

ttgtttaagg aaaattacaa gcctctcctg aaagagatta aggaagatac taataagtgg 4200

aagaatatcc cctgttcatg ggttggcaga atcaacatag tgaagatggc aatacttcct 4260

aaagtgatat atcgctttaa cgccatccca attaaactgc ctatgacctt ctttacggag 4320

ctcgagaaaa caacccttaa atttatatgg aatcaaaaga gagcaagaat agcgaagtcc 4380

atcttgagcc agaagaataa ggccggtggg attactttgc ctgattttaa gttgtattat 4440

aaagccacag taactaagac agcctggtat tggtatcaga atagagacat cgaccagtgg 4500

aatcggaccg aaccatcaga gataatgccc cacatctata attaccttat attcgataag 4560

ccagaaaaga ataaacagtg gggcaaagac agcctcttca acaagtggtg ttgggagaat 4620

tggctggcca tatgccggaa actcaagctc gacccctttc ttacacccta cactaaaatc 4680

aacagtaggt ggatcaagga cttgaatgtc aagccaaaga ctataaagac actggaagag 4740

aatcttggga tcacaataca agatataggc gtcggcaaag attttatgtc aaagacgccc 4800

aaggccatgg ccactaagga taagattgat aagtgggacc ttattaagct caaaagcttc 4860

tgtactgcca aggagaccac gatcagagtt aataggcagc ccactacatg ggaaaagatt 4920

ttcgccactt attcatcaga taaggggttg ataagcagaa tatataacga gctgaagcag 4980

atctacaaga agaaaacgaa taatcccatc aagaagtggg caaaagatat gaacaggcat 5040

tttagcaaag aggatatcta cgccgcgaag aagcatatga agaagtgtag ttcaagcttg 5100

gccattcgtg agatgcagat taagacgacc atgcgatacc accttacccc agtgaggatg 5160

gcaattatca agaaatctgg caataataga tgttggcggg gctgtggcga gattggcacc 5220

ctgctccatt gctggtggga ttgcaagctg gtgcagccgc tttggaaatc agtctggcgc 5280

tttctgaggg acctcgagct tgagattccc ttcgatcccg caattccctt gctcggaatc 5340

tatcctaacg aatacaagag ctgttgttac aaggatacgt gtacccggat gttcatcgcg 5400

gccttgttta cgatagctaa gacgtggaat cagcctaagt gccccacaat gatcgattgg 5460

atcaagaaaa tgtggcatat ttataccatg gagtattacg cagcaattaa gaatgacgaa 5520

tttatttcct tcgttgggac ctggatgaag ctggagacta ttattctgag caagctgtct 5580

caggagcaaa agacaaagca tagaatcttc tctctcattg gtggtaacga ctacaaagac 5640

gatgacgaca agtaaagcgc ttctagaagt tgtctcctcc tgcactgact gactgataca 5700

atcgatttct ggatccgcag gcctaatcaa cctctggatt acaaaatttg tgaaagattg 5760

actggtattc ttaactatgt tgctccttttt acgctatgtg gatacgctgc tttaatgcct 5820

ttgtatcatg ctattgcttc ccgtatggct ttcattttct cctccttgta taaatcctgg 5880

ttgctgtctc tttatgagga gttgtggccc gttgtcaggc aacgtggcgt ggtgtgcact 5940

gtgtttgctg acgcaacccc cactggttgg ggcattgcca ccacctgtca gctcctttcc 6000

gggactttcg ctttcccccc ccctattgcc acggcggaac tcatcgccgc ctgccttgcc 6060

cgctgctgga caggggctcg gctgttgggc actgacaatt ccgtggtgtt gtcggggaag 6120

ctgacgtcct ttccatggct gctcgcctgt gttgccacct ggattctgcg cgggacgtcc 6180

ttctgctacg tcccttcggc cctcaatcca gcggaccttc cttcccgcga acaaacgacc 6240

caacacccgt gcgttttatt ctgtcttttt attgccgatc ccctcagaag aactcgtcaa 6300

gaaggcgata gaaggcgatg cgctgcgaat cgggagcggc gataccgtaa agcacgagga 6360

agcggtcagc ccattcgccg ccaagctctt cagcaatatc acgggtagcc aacgctatgt 6420

cctgatagcg gtcggccgct ttacttgtac agctcgtcca tgccgagagt gatcccggcg 6480

gcggtcacga actccagcag gaccatgtga tcgcgcttct cgttggggtc tttgctcagg 6540

gcggactggg tgctcaggta gtggttgtcg ggcagcagca cggggccgtc gccgatgggg 6600

gtgttctgct ggtagtggtc ggccaggtga gtccaggaga tgtttcagca ctgttgcctt 6660

tagtctcgag gcaacttaga caactgagta ttgatctgag cacagcaggg tgtgagctgt 6720

ttgaagatac tggggttggg ggtgaagaaa ctgcagagga ctaactgggc tgagacccag 6780

tggcaatgtt ttagggccta aggaatgcct ctgaaaatct agatggacaa ctttgacttt 6840

gagaaaagag aggtggaaat gaggaaaatg acttttcttt attagatttc ggtagaaaga 6900

actttcatct ttcccctatt tttgttattc gttttaaaac atctatctgg aggcaggaca 6960

agtatggtca ttaaaaagat gcaggcagaa ggcatatatt ggctcagtca aagtggggaa 7020

ctttggtggc caaacataca ttgctaaggc tattcctata tcagctggac acatataaaa 7080

tgctgctaat gcttcattac aaacttatat cctttaattc cagatggggg caaagtatgt 7140

ccaggggtga ggaacaattg aaacatttgg gctggagtag attttgaaag tcagctctgt 7200

gtgtgtgtgt gtgtgtgtgt gtgtgagagc gtgtgtttct tttaacgttt tcagcctaca 7260

gcatacaggg ttcatggtgg caagaagata acaagattta aattatggcc agtgactagt 7320

gctgcaagaa gaacaactac ctgcatttaa tgggaaagca aaatctcagg ctttgaggga 7380

agttaacata ggcttgattc tgggtggaag ctgggtgtgt agttatctgg aggccaggct 7440

ggagctctca gctcactatg ggttcatctt tattgtctcc tttcatctca acagctgcac 7500

gctgccgtcc tcgatgttgt ggcggatctt gaagttcacc ttgatgccgt tcttctgctt 7560

gtcggccatg atatagacgt tgtggctgtt gtagttgtac tccagcttgt gccccaggat 7620

gttgccgtcc tccttgaagt cgatgcccct cagctcgatg cggttcacca gggtgtcgcc 7680

ctcgaacttc acctcggcgc gggtcttgta gttgccgtcg tccttgaaga agatggtgcg 7740

ctcctggacg tagccttcgg gcatggcgga cttgaagaag tcgtgctgct tcatgtggtc 7800

ggggtagcgg ctgaagcact gcacgccgta ggtcagggtg gtcacgaggg tgggccaggg 7860
cacgggcagc ttgccggtgg tgcagatgaa cttcagggtc agcttgccgt aggtggcatc 7920
gccctcgccc tcgccggaca cgctgaactt gtggccgttt acgtcgccgt ccagctcgac 7980
caggatgggc accaccccgg tgaacagctc ctcgcccttg ctcaccatgg tggcgaattc 8040
gaagcttgag cacgagatct gagtccggta ggcctagcgg atctgacggt tcactaaacc 8100
agctctgctt atatagacct cccaccgtac acgcctaccg cccatttgcg tcaatggggc 8160
ggagttgtta cgacattttg gaaagtcccg ttgattttgg tgccaaaaca aactcccatt 8220

gacgtcaatg gggtggagac ttggaaatcc ccgtgagtca aaccgctatc cacgcccatt 8280
gatgtactgc caaaaccgca tcaccatggt aatagcgatg actaatacgt agatgtactg 8340
ccaagtagga aagtcccata aggtcatgta ctgggcataa tgccaggcgg gccatttacc 8400
gtcattgacg tcaatagggg gcgtacttgg catatgatac acttgatgta ctgccaagtg 8460
ggcagtttac cgtaaatact ccacccattg acgtcaatgg aaagtcccta ttggcgttac 8520
tatgggaaca tacgtcatta ttgacgtcaa tgggcggggg tcgttgggcg gtcagccagg 8580
cgggccattt accgtaagtt atgtaacggg cctgctgccg gctctgcggc ctcttccgcg 8640

tcttcgcctt cgccctcaga cgagtcggat ctccctttgg gccgcctccc cgcctgtcta 8700
gcttgactga ctgagataca gcgtaccttc agctcacaga catgataaga tacattgatg 8760
agtttggaca aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg 8820
atgctattgc tttatttgta accattataa gctgcaataa acaagttaac aacaacaatt 8880
gcattcattt tatgtttcag gttcaggggg aggtgtggga ggtttttaa agcaagtaaa 8940
acctctacaa atgtggtatt ggcccatctc tatcggtatc gtagcataac cccttggggc 9000
ctctaaacgg gtcttgaggg gttttttgtg cccctcgggc cggattgcta tctaccggca 9060

ttggcgcaga aaaaaatgcc tgatgcgacg ctgcgcgtct tatactccca catatgccag 9120
attcagcaac ggatacggct tccccaactt gcccacttcc atacgtgtcc tccttaccag 9180
aaatttatcc ttaaggtcgt cagctatcct gcaggcgatc tctcgatttc gatcaagaca 9240
ttcctttaat ggtcttttct ggacaccact aggggtcaga agtagttcat caaactttct 9300
tccctcccta atctcattgg ttaccttggg ctatcgaaac ttaattaagc gatctgcatc 9360
tcaattagtc agcaaccata gtcccgcccc taactccgcc catcccgccc ctaactccgc 9420
ccagttccgc ccattctccg ccccatcgct gactaatttt ttttatttat gcagaggccg 9480

aggccgcctc ggcctctgag ctattccaga agtagtgagg aggctttttt ggaggcctag 9540

gcttttgcaa aggaggtagc caacatgatt gaacaagatg gattgcacgc aggttctccc 9600

gccgcttggg tggagaggct attcggctat gactgggcac aacagacaat cggctgctct 9660

gatgccgccg tgttccggct gtcagcgcag gggcgcccgg ttctttttgt caagaccgac 9720

ctgtccggtg ccctgaatga actccaggac gaggcagcgc ggctatcgtg gctggccacg 9780

acgggcgttc cttgcgcagc tgtgctcgac gttgtcactg aagcgggaag ggactggctg 9840

ctattgggcg aagtgccggg gcaggatctc ctgtcatctc accttgctcc tgccgagaaa 9900

gtatccatca tggctgatgc aatgcggcgg ctgcatacgc ttgatccggc tacctgccca 9960

ttcgaccacc aagcgaaaca tcgcatcgag cgagcacgta ctcggatgga agccggtctt 10020

gtcgatcagg atgatctgga cgaagagcat caggggctcg cgccagccga actgttcgcc 10080

aggctcaagg cgcggatgcc cgacggcgag gatctcgtcg tgacccacgg cgatgcctgc 10140

ttgccgaata tcatggtgga aaatggccgc ttttctggat tcatcgactg tggccggctg 10200

ggtgtggcgg accgctatca ggacatagcg ttggctaccc gtgatattgc tgaagagctt 10260

ggcggcgaat gggctgaccg cttcctcgtg ctttacggta tcgccgctcc cgattcgcag 10320

cgcatcgcct tctatcgcct tcttgacgag ttcttctagt atgtaagccc tgtgccttct 10380

agttgccagc catctgttgt ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc 10440

actcccactg tcctttccta ataaaatgag gaaattgcat cgcattgtct gagtaggtgt 10500

cattctattc tgggggggtgg ggtggggcag gacagcaagg gggaggattg ggaagacaat 10560

agcaggcatg ctggggatgc ggtgggctct atggttaatt aaccagtcaa gtcagctact 10620

tggcgagatc gacttgtctg ggtttcgact acgctcagaa ttgcgtcagt caagttcgat 10680

ctggtccttg ctattgcacc cgttctccga ttacgagttt catttaaatc atgtgagcaa 10740

aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc 10800

tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga 10860

caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc 10920

cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt 10980

ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc aagctgggct 11040

gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg 11100

agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta 11160

gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct 11220

acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa 11280

gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt 11340

gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta 11400

cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat 11460

caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa 11520

gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct 11580

cagcgatctg tctatttcgt tcatccatag ttgcatttaa atttccgaac tctccaaggc 11640

cctcgtcgga aaatcttcaa acctttcgtc cgatccatct tgcaggctac ctctcgaacg 11700

aactatcgca agtctcttgg ccggccttgc gccttggcta ttgcttggca gcgcctatcg 11760

ccaggtatta ctccaatccc gaatatccga gatcgggatc acccgagaga agttcaacct 11820

acatcctcaa tcccgatcta tccgagatcc gaggaatatc gaaatcgggg cgcgcctggt 11880

gtaccgagaa cgatcctctc agtgcgagtc tcgacgatcc atatcgttgc ttggcagtca 11940

gccagtcgga atccagcttg ggacccagga agtccaatcg tcagatattg tactcaagcc 12000

tggtcacggc agcgtaccga tctgtttaaa cctagatatt gatagtctga tcggtcaacg 12060

tataatcgag tcctagcttt tgcaaacatc tatcaagaga caggatcagc aggaggcttt 12120

cgcatgagta ttcaacattt ccgtgtcgcc cttattccct tttttgcggc attttgcctt 12180

cctgtttttg ctcacccaga aacgctggtg aaagtaaaag atgctgaaga tcagttgggt 12240

gcgcgagtgg gttacatcga actggatctc aacagcggta agatccttga gagttttcgc 12300

cccgaagaac gctttccaat gatgagcact tttaaagttc tgctatgtgg cgcggtatta 12360

tcccgtattg acgccgggca agagcaactc ggtcgccgca tacactattc tcagaatgac 12420

ttggttgagt attcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa 12480

ttatgcagtg ctgccataac catgagtgat aacactgcgg ccaacttact tctgacaacg 12540

attggaggac cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc 12600

cttgatcgtt gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg 12660

atgcctgtag caatggcaac aaccttgcgt aaactattaa ctggcgaact acttactcta 12720

gcttcccggc aacagttgat agactggatg gaggcggata aagttgcagg accacttctg 12780

cgctcggccc ttccggctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg 12840

tctcgcggta tcattgcagc actggggcca gatggtaagc cctcccgtat cgtagttatc 12900

tacacgacgg ggagtcaggc aactatggat gaacgaaata gacagatcgc tgagataggt 12960

gcctcactga ttaagcattg gtaaccgatt ctaggtgcat tggcgcagaa aaaaatgcct 13020

gatgcgacgc tgcgcgtctt atactcccac atatgccaga ttcagcaacg gatacggctt 13080

ccccaacttg cccacttcca tacgtgtcct ccttaccaga aatttatcct taagatcgtt 13140

taaactcgac tctggctcta tcgaatctcc gtcgtttcga gcttacgcga acagccgtgg 13200

cgctcatttg ctcgtcgggc atcgaatctc gtcagctatc gtcagcttac ctttttggca 13260

gcgatcgcgg ctcccgacat cttggaccat tagctccaca ggtatcttct tccctctagt 13320
ggtcataaca gcagcttcag ctacctctca attcaaaaaa ccccctcaaga cccgtttaga 13380
ggccccaagg ggttatgcta tcaatcgttg cgttacacac acaaaaaacc aacacacatc 13440
catcttcgat ggatagcgat tttattatct aactgctgat cgagtgtagc cagatctagt 13500
aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta 13560
cggtaaatgg cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga 13620
cgtatgttcc catagtaacg ccaatagggа ctttccattg acgtcaatgg gtggagtatt 13680

tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta 13740
ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg 13800
actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg ctgatgcggt 13860
tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc 13920
accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat 13980
gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct 14040
atataagcag agctggttta gtgaaccgtc agatcagatc tttgtcgatc ctaccatcca 14100

ctcgacacac ccgccagcgg ccgc 14124

<210> 46
<211> 13439
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 46

taatacgact cactataggg agaagtactg ccaccatggg caagaagcaa aatcgcaaga 60
cggggaattc caagacacaa tccgctagcc caccacctaa agagcgttct agctcccctg 120
ctactgagca gtcctggatg gaaaacgact tcgatgaact ccgggaagag ggatttaggc 180
gatccaacta ttcagaactc cgcgaagata tccagacaaa ggggaaggaa gtcgagaatt 240

tcgagaagaa cctcgaggag tgcatcaccc gtatcacaaa cactgagaaa tgtctcaaag 300
aactcatgga acttaagaca aaagccaggg agcttcgaga ggagtgtcgg agtctgagat 360
ccaggtgtga ccagctcgag gagcgcgtga gcgcgatgga agacgagatg aacgagatga 420
aaagagaggg caaattcagg gagaagcgca ttaagaggaa cgaacagagt ctgcaggaga 480

tttgggatta cgtcaagagg cctaacctgc ggttgatcgg cgtccccgag agcgacgtag 540

aaaacgggac taaactggag aatacacttc aagacatcat tcaagaaaat tttccaaacc 600

tggctcggca agctaatgtg caaatccaag agatccaacg cacaccccag cggtatagct 660

ctcggcgtgc cacccctagg catattatcg tgcgctttac taaggtggag atgaaagaga 720

agatgctgcg agccgctcgg gaaaagggaa gggtgacttt gaagggcaaa cctattcggc 780

tgacggttga ccttagcgcc gagacactcc aggcacgccg ggaatggggc cccatcttta 840

atatcctgaa ggagaagaac ttccagccac gaatctctta ccctgcaaag ttgagtttta 900

tctccgaggg tgagattaag tatttcatcg ataaacagat gctgcgagac ttcgtgacaa 960

ctcgcccagc tctcaaggaa ctgctcaaag aggctcttaa tatggagcgc aataatagat 1020

atcaaccctt gcagaaccac gcaaagatgg gaagcggaca gtgtactaat tatgctctct 1080

tgaaattggc tggagatgtt gagagcaacc ctggacctat gaccggctct aactcacata 1140

tcaccatcct tacacttaac attaacggcc tcaactcagc tatcaagcgc catcggctgg 1200

ccagctggat caaatcacag gatccaagcg tttgttgcat ccaagagacc cacctgacct 1260

gtagagatac tcaccgcctc aagatcaagg gatggcgaaa gatttatcag gcgaacggta 1320

agcagaagaa agccggagtc gcaattctgg tctcagacaa gacggatttc aagcccacca 1380

aaattaagcg tgataaggaa ggtcactata ttatggtgaa aggcagcata cagcaggaag 1440

aacttaccat attgaacatc tacgcgccaa acaccggcgc acctcgcttt atcaaacagg 1500

tcctgtccga tctgcagcga gatctggatt ctcatacgtt gattatgggt gatttcaata 1560

caccattgag caccctggat cgcagcacca ggcaaaaggt aaataaagac acgcaagagc 1620

tcaatagcgc actgcatcag gcagatctca ttgatattta tcgcactctt catcctaaga 1680

gtaccgagta cacattcttc agcgccccac atcatacata ctcaaagatc gatcatatcg 1740

tcggctcaaa ggctctgctg tcaaagtgca agcgcacaga gataattaca aattacctgt 1800

cagatcatag cgcgatcaag ctcgagctga gaatcaagaa cctgacccag agccggagta 1860

ccacttggaa gcttaataac ctgctgctca acgattattg ggtccacaat gagatgaagg 1920

cagagattaa aatgttcttc gaaacaaatg agaataagga tactacctat caaaaccttt 1980

gggatgcctt taaggccgtc tgcagaggca agttcatcgc cctcaacgcc tataaaagaa 2040

aacaagagag atctaagatc gatactctca cctctcagct gaaggagttg gagaaacagg 2100

aacagaccca ctccaaggcg tcaagacggc aggagatcac aaagattcgc gccgagttga 2160

aagagatcga aacccaaaag actcttcaga aaattaacga gtctcgtagt tggttcttcg 2220

agcggattaa taagatagac agacctctgg cacgactgat taagaagaag cgcgaaaaga 2280

accagattga taccatcaag aacgacaagg gcgacatcac tactgacccg accgagatcc 2340

agaccactat tcgggagtat tataagcatt tgtatgctaa caagcttgag aacctggaag 2400
agatggacac ttttctggat acctatactc tgccacggct taatcaagag gaagtcgagt 2460
ccctcaaccg cccaattaca ggaagcgaga ttgtggccat aattaactcc ctgccgacaa 2520
agaaatctcc tggtccggac gggtttacag ctgagtttta tcaacggtat atggaagagc 2580
ttgtaccgtt tctgctcaag ctctttcagt ctatagaaaa ggaaggcatc ttgcccaatt 2640
ccttctacga agcttctata atacttattc ccaaaccagg acgcgatacc acaaagaagg 2700
aaaacttccg gcccattagt ctcatgaata tcgacgctaa aatattgaac aagattctcg 2760

ccaacagaat ccaacaacat attaagaaat tgatacatca cgaccaggtg gggtttatac 2820
ctggcatgca gggctggttt aacatccgga agagtattaa cgtcattcaa cacattaata 2880
gagctaagga taagaatcat atgatcatct ctatagacgc ggaaaaggca ttcgataaga 2940
ttcagcagcc atttatgctc aagactctga acaaactcgg catcgacgga acatatttta 3000
agattattcg cgcaatttac gataagccga ctgctaacat tatccttaac ggccaaaagc 3060
tcgaggcctt tccgctcaag actggaaccc gccaaggctg tcccctctcc ccgcttttgt 3120
ttaatattgt actcgaggtg ctggctaggg ctattcgtca agagaaagag attaaaggga 3180

tacagctcgg gaaggaagag gtcaagcttt ccttgttcgc cgatgatatg attgtgtacc 3240
tggagaatcc tattgtgtct gctcagaacc ttcttaaact tatttctaac tttagcaagg 3300
tcagcggcta taagattaac gtccagaaat ctcaggcctt tctgtacaca aataatcgac 3360
agaccgaatc ccagataatg ggtgagcttc cgtttgtcat agccagcaaa aggataaagt 3420
atctcggaat ccagctgaca cgagacgtta aagatttgtt taaggaaaat tacaagcctc 3480
tcctgaaaga gattaaggaa gatactaata agtggaagaa tatcccctgt tcatgggttg 3540
gcagaatcaa catagtgaag atggcaatac ttcctaaagt gatatatcgc tttaacgcca 3600

tcccaattaa actgcctatg accttcttta cggagctcga gaaaacaacc cttaaattta 3660
tatggaatca aaagagagca agaatagcga agtccatctt gagccagaag aataaggccg 3720
gtgggattac tttgcctgat tttaagttgt attataaagc cacagtaact aagacagcct 3780
ggtattggta tcagaataga gacatcgacc agtggaatcg gaccgaacca tcagagataa 3840
tgccccacat ctataattac cttatattcg ataagccaga aaagaataaa cagtggggca 3900
aagacagcct cttcaacaag tggtgttggg agaattggct ggccatatgc cggaaactca 3960
agctcgaccc ctttcttaca ccctacacta aaatcaacag taggtggatc aaggacttga 4020

atgtcaagcc aaagactata aagacactgg aagagaatct tgggatcaca atacaagata 4080

taggcgtcgg caaagatttt atgtcaaaga cgcccaaggc catggccact aaggataaga 4140

ttgataagtg ggaccttatt aagctcaaaa gcttctgtac tgccaaggag accacgatca 4200

gagttaatag gcagcccact acatgggaaa agattttcgc cacttattca tcagataagg 4260

ggttgataag cagaatatat aacgagctga agcagatcta caagaagaaa acgaataatc 4320

ccatcaagaa gtgggcaaaa gatatgaaca ggcattttag caaagaggat atctacgccg 4380

cgaagaagca tatgaagaag tgtagttcaa gcttggccat tcgtgagatg cagattaaga 4440

cgaccatgcg ataccacctt accccagtga ggatggcaat tatcaagaaa tctggcaata 4500

atagatgttg gcggggctgt ggcgagattg gcaccctgct ccattgctgg tgggattgca 4560

agctggtgca gccgctttgg aaatcagtct ggcgctttct gagggacctc gagcttgaga 4620

ttcccttcga tcccgcaatt cccttgctcg gaatctatcc taacgaatac aagagctgtt 4680

gttacaagga tacgtgtacc cggatgttca tcgcggcctt gtttacgata gctaagacgt 4740

ggaatcagcc taagtgcccc acaatgatcg attggatcaa gaaaatgtgg catatttata 4800

ccatggagta ttacgcagca attaagaatg acgaatttat ttccttcgtt gggacctgga 4860

tgaagctgga gactattatt ctgagcaagc tgtctcagga gcaaaagaca aagcatagaa 4920

tcttctctct cattggtggt aacgactaca aagacgatga cgacaagtaa agcgcttcta 4980

gaagttgtct cctcctgcac tgactgactg atacaatcga tttctggatc cgcaggccta 5040

atcaacctct ggattacaaa atttgtgaaa gattgactgg tattcttaac tatgttgctc 5100

cttttacgct atgtggatac gctgctttaa tgcctttgta tcatgctatt gcttcccgta 5160

tggctttcat tttctcctcc ttgtataaat cctggttgct gtctctttat gaggagttgt 5220

ggcccgttgt caggcaacgt ggcgtggtgt gcactgtgtt tgctgacgca acccccactg 5280

gttggggcat tgccaccacc tgtcagctcc tttccgggac tttcgctttc cccctcccta 5340

ttgccacggc ggaactcatc gccgcctgcc ttgcccgctg ctggacaggg gctcggctgt 5400

tgggcactga caattccgtg gtgttgtcgg ggaagctgac gtcctttcca tggctgctcg 5460

cctgtgttgc cacctggatt ctgcgcggga cgtccttctg ctacgtccct tcggccctca 5520

atccagcgga ccttccttcc cgcgaacaaa cgacccaaca cccgtgcgtt ttattctgtc 5580

tttttattgc cgatcccctc agaagaactc gtcaagaagg cgatagaagg cgatgcgctg 5640

cgaatcggga gcggcgatac cgtaaagcac gaggaagcgg tcagcccatt cgccgccaag 5700

ctcttcagca atatcacggg tagccaacgc tatgtcctga tagcggtcgg ccgctttact 5760

tgtacagctc gtccatgccg agagtgatcc cggcggcggt cacgaactcc agcaggacca 5820

tgtgatcgcg cttctcgttg gggtctttgc tcagggcgga ctgggtgctc aggtagtggt 5880

tgtcgggcag cagcacgggg ccgtcgccga tgggggtgtt ctgctggtag tggtcggcca 5940

ggtgagtcca ggagatgttt cagcactgtt gcctttagtc tcgaggcaac ttagacaact 6000

gagtattgat ctgagcacag cagggtgtga gctgtttgaa gatactgggg ttgggggtga 6060

agaaactgca gaggactaac tgggctgaga cccagtggca atgttttagg gcctaaggaa 6120

tgcctctgaa aatctagatg gacaactttg actttgagaa aagagaggtg gaaatgagga 6180

aaatgacttt tctttattag atttcggtag aaagaacttt catctttccc ctatttttgt 6240

tattcgtttt aaaacatcta tctggaggca ggacaagtat ggtcattaaa aagatgcagg 6300

cagaaggcat atattggctc agtcaaagtg gggaactttg gtggccaaac atacattgct 6360

aaggctattc ctatatcagc tggacacata taaaatgctg ctaatgcttc attacaaact 6420

tatatccttt aattccagat gggggcaaag tatgtccagg ggtgaggaac aattgaaaca 6480

tttgggctgg agtagatttt gaaagtcagc tctgtgtgtg tgtgtgtgtg tgtgtgtgtg 6540

agagcgtgtg tttcttttaa cgttttcagc ctacagcata cagggttcat ggtggcaaga 6600

agataacaag atttaaatta tggccagtga ctagtgctgc aagaagaaca actacctgca 6660

tttaatggga aagcaaaatc tcaggctttg agggaagtta acataggctt gattctgggt 6720

ggaagctggg tgtgtagtta tctggaggcc aggctggagc tctcagctca ctatgggttc 6780

atctttattg tctcctttca tctcaacagc tgcacgctgc cgtcctcgat gttgtggcgg 6840

atcttgaagt tcaccttgat gccgttcttc tgcttgtcgg ccatgatata gacgttgtgg 6900

ctgttgtagt tgtactccag cttgtgcccc aggatgttgc cgtcctcctt gaagtcgatg 6960

cccttcagct cgatgcggtt caccagggtg tcgccctcga acttcacctc ggcgcgggtc 7020

ttgtagttgc cgtcgtcctt gaagaagatg gtgcgctcct ggacgtagcc ttcgggcatg 7080

gcggacttga agaagtcgtg ctgcttcatg tggtcggggt agcggctgaa gcactgcacg 7140

ccgtaggtca gggtggtcac gagggtgggc cagggcacgg gcagcttgcc ggtggtgcag 7200

atgaacttca gggtcagctt gccgtaggtg gcatcgccct cgccctcgcc ggacacgctg 7260

aacttgtggc cgtttacgtc gccgtccagc tcgaccagga tgggcaccac cccggtgaac 7320

agctcctcgc ccttgctcac catggtggcg aattcgaagc ttgagcacga gatctgagtc 7380

cggtaggcct agcggatctg acggttcact aaaccagctc tgcttatata gacctcccac 7440

cgtacacgcc taccgcccat ttgcgtcaat ggggcggagt tgttacgaca ttttggaaag 7500

tcccgttgat tttggtgcca aaacaaactc ccattgacgt caatggggtg gagacttgga 7560

aatccccgtg agtcaaaccg ctatccacgc ccattgatgt actgccaaaa ccgcatcacc 7620

atggtaatag cgatgactaa tacgtagatg tactgccaag taggaaagtc ccataaggtc 7680

atgtactggg cataatgcca ggcgggccat ttaccgtcat tgacgtcaat agggggcgta 7740

cttggcatat gatacacttg atgtactgcc aagtgggcag tttaccgtaa atactccacc 7800

cattgacgtc aatggaaagt ccctattggc gttactatgg gaacatacgt cattattgac 7860
gtcaatgggc gggggtcgtt gggcggtcag ccaggcgggc catttaccgt aagttatgta 7920
acgggcctgc tgccggctct gcggcctctt ccgcgtcttc gccttcgccc tcagacgagt 7980
cggatctccc tttgggccgc ctccccgcct gtctagcttg actgactgag atacagcgta 8040
ccttcagctc acagacatga taagatacat tgatgagttt ggacaaacca caactagaat 8100
gcagtgaaaa aaatgcttta tttgtgaaat ttgtgatgct attgctttat ttgtaaccat 8160
tataagctgc aataaacaag ttaacaacaa caattgcatt cattttatgt ttcaggttca 8220

gggggaggtg tgggaggttt tttaaagcaa gtaaaacctc tacaaatgtg gtattggccc 8280
atctctatcg gtatcgtagc ataacccctt ggggcctcta aacgggtctt gaggggtttt 8340
ttgtgcccct cgggccggat tgctatctac cggcattggc gcagaaaaaa atgcctgatg 8400
cgacgctgcg cgtcttatac tcccacatat gccagattca gcaacggata cggcttcccc 8460
aacttgccca cttccatacg tgtcctcctt accagaaatt tatccttaag gtcgtcagct 8520
atcctgcagg cgatctctcg atttcgatca agacattcct ttaatggtct tttctggaca 8580
ccactagggg tcagaagtag ttcatcaaac tttcttccct ccctaatctc attggttacc 8640

ttgggctatc gaaacttaat taagcgatct gcatctcaat tagtcagcaa ccatagtccc 8700
gcccctaact ccgcccatcc cgcccctaac tccgcccagt tccgcccatt ctccgcccca 8760
tcgctgacta attttttta tttatgcaga ggccgaggcc gcctcggcct ctgagctatt 8820
ccagaagtag tgaggaggct tttttggagg cctaggcttt tgcaaaggag gtagccaaca 8880
tgattgaaca agatggattg cacgcaggtt ctcccgccgc ttgggtggag aggctattcg 8940
gctatgactg ggcacaacag acaatcggct gctctgatgc cgccgtgttc cggctgtcag 9000
cgcaggggcg cccggttctt tttgtcaaga ccgacctgtc cggtgccctg aatgaactcc 9060

aggacgaggc agcgcggcta tcgtggctgg ccacgacggg cgttccttgc gcagctgtgc 9120
tcgacgttgt cactgaagcg ggaagggact ggctgctatt gggcgaagtg ccggggcagg 9180
atctcctgtc atctcacctt gctcctgccg agaaagtatc catcatggct gatgcaatgc 9240
ggcggctgca tacgcttgat ccggctacct gcccattcga ccaccaagcg aaacatcgca 9300
tcgagcgagc acgtactcgg atggaagccg gtcttgtcga tcaggatgat ctggacgaag 9360
agcatcaggg gctcgcgcca gccgaactgt tcgccaggct caaggcgcgg atgcccgacg 9420
gcgaggatct cgtcgtgacc cacggcgatg cctgcttgcc gaatatcatg gtggaaaatg 9480

gccgcttttc tggattcatc gactgtggcc ggctgggtgt ggcggaccgc tatcaggaca 9540

tagcgttggc tacccgtgat attgctgaag agcttggcgg cgaatgggct gaccgcttcc 9600

tcgtgcttta cggtatcgcc gctcccgatt cgcagcgcat cgccttctat cgccttcttg 9660

acgagttctt ctagtatgta agccctgtgc cttctagttg ccagccatct gttgtttgcc 9720

cctcccccgt gccttccttg accctggaag gtgccactcc cactgtcctt tcctaataaa 9780

atgaggaaat tgcatcgcat tgtctgagta ggtgtcattc tattctgggg ggtggggtgg 9840

ggcaggacag caagggggag gattgggaag acaatagcag gcatgctggg gatgcggtgg 9900

gctctatggt taattaacca gtcaagtcag ctacttggcg agatcgactt gtctgggttt 9960

cgactacgct cagaattgcg tcagtcaagt tcgatctggt ccttgctatt gcacccgttc 10020

tccgattacg agtttcattt aaatcatgtg agcaaaaggc cagcaaaagg ccaggaaccg 10080

taaaaaggcc gcgttgctgg cgtttttcca taggctccgc ccccctgacg agcatcacaa 10140

aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 10200

tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 10260

gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 10320

cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 10380

cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 10440

atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 10500

tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat 10560

ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 10620

acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 10680

aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 10740

aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct 10800

tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga 10860

cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc 10920

catagttgca tttaaatttc cgaactctcc aaggccctcg tcggaaaatc ttcaaacctt 10980

tcgtccgatc catcttgcag gctacctctc gaacgaacta tcgcaagtct cttggccggc 11040

cttgcgcctt ggctattgct tggcagcgcc tatcgccagg tattactcca atcccgaata 11100

tccgagatcg ggatcacccg agagaagttc aacctacatc ctcaatcccg atctatccga 11160

gatccgagga atatcgaaat cggggcgcgc ctggtgtacc gagaacgatc ctctcagtgc 11220

gagtctcgac gatccatatc gttgcttggc agtcagccag tcggaatcca gcttgggacc 11280

caggaagtcc aatcgtcaga tattgtactc aagcctggtc acggcagcgt accgatctgt 11340

ttaaacctag atattgatag tctgatcggt caacgtataa tcgagtccta gcttttgcaa 11400

acatctatca agagacagga tcagcaggag gctttcgcat gagtattcaa catttccgtg 11460

tcgcccttat tccctttttt gcggcatttt gccttcctgt ttttgctcac ccagaaacgc 11520

tggtgaaagt aaaagatgct gaagatcagt tgggtgcgcg agtgggttac atcgaactgg 11580

atctcaacag cggtaagatc cttgagagtt ttcgccccga agaacgcttt ccaatgatga 11640

gcacttttaa agttctgcta tgtggcgcgg tattatcccg tattgacgcc gggcaagagc 11700

aactcggtcg ccgcatacac tattctcaga atgacttggt tgagtattca ccagtcacag 11760

aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgctgcc ataaccatga 11820

gtgataacac tgcggccaac ttacttctga caacgattgg aggaccgaag gagctaaccg 11880

ctttttgca caacatgggg gatcatgtaa ctcgccttga tcgttgggaa ccggagctga 11940

atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacct 12000

tgcgtaaact attaactggc gaactactta ctctagcttc ccggcaacag ttgatagact 12060

ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg gctggctggt 12120

ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg 12180

ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta 12240

tggatgaacg aaatagacag atcgctgaga taggtgcctc actgattaag cattggtaac 12300

cgattctagg tgcattggcg cagaaaaaaa tgcctgatgc gacgctgcgc gtcttatact 12360

cccacatatg ccagattcag caacggatac ggcttcccca acttgcccac ttccatacgt 12420

gtcctcctta ccagaaattt atccttaaga tcgtttaaac tcgactctgg ctctatcgaa 12480

tctccgtcgt ttcgagctta cgcgaacagc cgtggcgctc atttgctcgt cgggcatcga 12540

atctcgtcag ctatcgtcag cttacctttt tggcagcgat cgcggctccc gacatcttgg 12600

accattagct ccacaggtat cttcttccct ctagtggtca taacagcagc ttcagctacc 12660

tctcaattca aaaaacccct caagacccgt ttagaggccc caaggggtta tgctatcaat 12720

cgttgcgtta cacacacaaa aaaccaacac acatccatct tcgatggata gcgattttat 12780

tatctaactg ctgatcgagt gtagccagat ctagtaatca attacggggt cattagttca 12840

tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 12900

gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 12960

agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 13020

acatcaagtg tatcatatgc caagtacgcc ccctattgac gtcaatgacg gtaaatggcc 13080

cgcctggcat tatgcccagt acatgacctt atgggacttt cctacttggc agtacatcta 13140

cgtattagtc atcgctatta ccatgctgat gcggttttgg cagtacatca atgggcgtgg 13200

atagcggttt gactcacggg gatttccaag tctccacccc attgacgtca atgggagttt 13260

gttttggcac caaaatcaac gggactttcc aaaatgtcgt aacaactccg ccccattgac 13320
gcaaatgggc ggtaggcgtg tacggtggga ggtctatata agcagagctg gtttagtgaa 13380
ccgtcagatc agatctttgt cgatcctacc atccactcga cacacccgcc agcggccgc 13439

<210> 47
<211> 13436
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic polynucleotide
<400> 47

taatacgact cactataggg agaagtactg ccaccatggg caagaagcaa aatcgcaaga 60
cggggaattc caagacacaa tccgctagcc caccacctaa agagcgttct agctcccctg 120

ctactgagca gtcctggatg gaaaacgact tcgatgaact ccgggaagag ggatttaggc 180
gatccaacta ttcagaactc cgcgaagata tccagacaaa ggggaaggaa gtcgagaatt 240
tcgagaagaa cctcgaggag tgcatcaccc gtatcacaaa cactgagaaa tgtctcaaag 300
aactcatgga acttaagaca aaagccaggg agcttcgaga ggagtgtcgg agtctgagat 360
ccaggtgtga ccagctcgag gagcgcgtga gcgcgatgga agacgagatg aacgagatga 420
aaagagaggg caaattcagg gagaagcgca ttaagaggaa cgaacagagt ctgcaggaga 480
tttgggatta cgtcaagagg cctaacctgc ggttgatcgg cgtccccgag agcgacgtag 540

aaaacgggac taaactggag aatacacttc aagacatcat tcaagaaaat tttccaaacc 600
tggctcggca agctaatgtg caaatccaag agatccaacg cacaccccag cggtatagct 660
ctcggcgtgc cacccctagg catattatcg tgcgctttac taaggtggag atgaaagaga 720
agatgctgcg agccgctcgg gaaaagggaa gggtgacttt gaagggcaaa cctattcggc 780
tgacggttga ccttagcgcc gagacactcc aggcacgccg ggaatggggc cccatcttta 840
atatcctgaa ggagaagaac ttccagccac gaatctctta ccctgcaaag ttgagtttta 900
tctccgaggg tgagattaag tatttcatcg ataaacagat gctgcgagac ttcgtgacaa 960

ctcgcccagc tctcaaggaa ctgctcaaag aggctcttaa tatggagcgc aataatagat 1020
atcaaccctt gcagaaccac gcaaagatgg gaagcggagc tactaacttc agcctgctga 1080
agcaggctgg agacgtggag gagaaccctg gacctatgac cggctctaac tcacatatca 1140
ccatccttac acttaacatt aacggcctca actcagctat caagcgccat cggctggcca 1200

gctggatcaa atcacaggat ccaagcgttt gttgcatcca agagacccac ctgacctgta 1260

gagatactca ccgcctcaag atcaagggat ggcgaaagat ttatcaggcg aacggtaagc 1320

agaagaaagc cggagtcgca attctggtct cagacaagac ggatttcaag cccaccaaaa 1380

ttaagcgtga taaggaaggt cactatatta tggtgaaagg cagcatacag caggaagaac 1440

ttaccatatt gaacatctac gcgccaaaca ccggcgcacc tcgctttatc aaacaggtcc 1500

tgtccgatct gcagcgagat ctggattctc atacgttgat tatgggtgat ttcaatacac 1560

cattgagcac cctggatcgc agcaccaggc aaaaggtaaa taaagacacg caagagctca 1620

atagcgcact gcatcaggca gatctcattg atatttatcg cactcttcat cctaagagta 1680

ccgagtacac attcttcagc gccccacatc atacatactc aaagatcgat catatcgtcg 1740

gctcaaaggc tctgctgtca aagtgcaagc gcacagagat aattacaaat tacctgtcag 1800

atcatagcgc gatcaagctc gagctgagaa tcaagaacct gacccagagc cggagtacca 1860

cttggaagct taataacctg ctgctcaacg attattgggt ccacaatgag atgaaggcag 1920

agattaaaat gttcttcgaa acaaatgaga ataaggatac tacctatcaa aacctttggg 1980

atgcctttaa ggccgtctgc agaggcaagt tcatcgccct caacgcctat aaaagaaaac 2040

aagagagatc taagatcgat actctcacct ctcagctgaa ggagttggag aaacaggaac 2100

agacccactc caaggcgtca agacggcagg agatcacaaa gattcgcgcc gagttgaaag 2160

agatcgaaac ccaaaagact cttcagaaaa ttaacgagtc tcgtagttgg ttcttcgagc 2220

ggattaataa gatagacaga cctctggcac gactgattaa gaagaagcgc gaaaagaacc 2280

agattgatac catcaagaac gacaagggcg acatcactac tgacccgacc gagatccaga 2340

ccactattcg ggagtattat aagcatttgt atgctaacaa gcttgagaac ctggaagaga 2400

tggacacttt tctggatacc tatactctgc cacggcttaa tcaagaggaa gtcgagtccc 2460

tcaaccgccc aattacagga agcgagattg tggccataat taactccctg ccgacaaaga 2520

aatctcctgg tccggacggg tttacagctg agttttatca acggtatatg gaagagcttg 2580

taccgtttct gctcaagctc tttcagtcta tagaaaagga aggcatcttg cccaattcct 2640

tctacgaagc ttctataata cttattccca aaccaggacg cgataccaca aagaaggaaa 2700

acttccggcc cattagtctc atgaatatcg acgctaaaat attgaacaag attctcgcca 2760

acagaatcca acaacatatt aagaaattga tacatcacga ccaggtgggg tttatacctg 2820

gcatgcaggg ctggtttaac atccggaaga gtattaacgt cattcaacac attaatagag 2880

ctaaggataa gaatcatatg atcatctcta tagacgcgga aaaggcattc gataagattc 2940

agcagccatt tatgctcaag actctgaaca aactcggcat cgacggaaca tattttaaga 3000

ttattcgcgc aatttacgat aagccgactg ctaacattat ccttaacggc caaaagctcg 3060

aggcctttcc gctcaagact ggaacccgcc aaggctgtcc cctctccccg cttttgttta 3120
atattgtact cgaggtgctg gctagggcta ttcgtcaaga gaaagagatt aaagggatac 3180
agctcgggaa ggaagaggtc aagctttcct tgttcgccga tgatatgatt gtgtacctgg 3240
agaatcctat tgtgtctgct cagaaccttc ttaaacttat ttctaacttt agcaaggtca 3300
gcggctataa gattaacgtc cagaaatctc aggcctttct gtacacaaat aatcgacaga 3360
ccgaatccca gataatgggt gagcttccgt ttgtcatagc cagcaaaagg ataaagtatc 3420
tcggaatcca gctgacacga gacgttaaag atttgtttaa ggaaaattac aagcctctcc 3480

tgaaagagat taaggaagat actaataagt ggaagaatat cccctgttca tgggttggca 3540
gaatcaacat agtgaagatg gcaatacttc ctaaagtgat atatcgcttt aacgccatcc 3600
caattaaact gcctatgacc ttctttacgg agctcgagaa aacaaccctt aaatttatat 3660
ggaatcaaaa gagagcaaga atagcgaagt ccatcttgag ccagaagaat aaggccggtg 3720
ggattacttt gcctgatttt aagttgtatt ataaagccac agtaactaag acagcctggt 3780
attggtatca gaatagagac atcgaccagt ggaatcggac cgaaccatca gagataatgc 3840
cccacatcta taattacctt atattcgata agccagaaaa gaataaacag tggggcaaag 3900

acagcctctt caacaagtgg tgttgggaga attggctggc catatgccgg aaactcaagc 3960
tcgacccctt tcttacaccc tacactaaaa tcaacagtag gtggatcaag gacttgaatg 4020
tcaagccaaa gactataaag acactggaag agaatcttgg gatcacaata caagatatag 4080
gcgtcggcaa agattttatg tcaaagacgc ccaaggccat ggccactaag gataagattg 4140
ataagtggga ccttattaag ctcaaaagct tctgtactgc caaggagacc acgatcagag 4200
ttaataggca gcccactaca tgggaaaaga ttttcgccac ttattcatca gataaggggt 4260
tgataagcag aatatataac gagctgaagc agatctacaa gaagaaaacg aataatccca 4320

tcaagaagtg ggcaaaagat atgaacaggc attttagcaa agaggatatc tacgccgcga 4380
agaagcatat gaagaagtgt agttcaagct tggccattcg tgagatgcag attaagacga 4440
ccatgcgata ccaccttacc ccagtgagga tggcaattat caagaaatct ggcaataata 4500
gatgttggcg gggctgtggc gagattggca ccctgctcca ttgctggtgg gattgcaagc 4560
tggtgcagcc gctttggaaa tcagtctggc gctttctgag ggacctcgag cttgagattc 4620
ccttcgatcc cgcaattccc ttgctcggaa tctatcctaa cgaatacaag agctgttgtt 4680
acaaggatac gtgtacccgg atgttcatcg cggccttgtt tacgatagct aagacgtgga 4740

atcagcctaa gtgccccaca atgatcgatt ggatcaagaa aatgtggcat atttatacca 4800

tggagtatta cgcagcaatt aagaatgacg aatttatttc cttcgttggg acctggatga 4860

agctggagac tattattctg agcaagctgt ctcaggagca aaagacaaag catagaatct 4920

tctctctcat tggtggtaac gactacaaag acgatgacga caagtaaagc gcttctagaa 4980

gttgtctcct cctgcactga ctgactgata caatcgattt ctggatccgc aggcctaatc 5040

aacctctgga ttacaaaatt tgtgaaagat tgactggtat tcttaactat gttgctcctt 5100

ttacgctatg tggatacgct gctttaatgc ctttgtatca tgctattgct tcccgtatgg 5160

ctttcatttt ctcctccttg tataaatcct ggttgctgtc tctttatgag gagttgtggc 5220

ccgttgtcag gcaacgtggc gtggtgtgca ctgtgtttgc tgacgcaacc cccactggtt 5280

ggggcattgc caccacctgt cagctccttt ccgggacttt cgctttcccc ctccctattg 5340

ccacggcgga actcatcgcc gcctgccttg cccgctgctg gacaggggct cggctgttgg 5400

gcactgacaa ttccgtggtg ttgtcgggga agctgacgtc ctttccatgg ctgctcgcct 5460

gtgttgccac ctggattctg cgcgggacgt ccttctgcta cgtcccttcg gccctcaatc 5520

cagcggacct tccttcccgc gaacaaacga cccaacaccc gtgcgtttta ttctgtcttt 5580

ttattgccga tcccctcaga agaactcgtc aagaaggcga tagaaggcga tgcgctgcga 5640

atcgggagcg gcgataccgt aaagcacgag gaagcggtca gcccattcgc cgccaagctc 5700

ttcagcaata tcacgggtag ccaacgctat gtcctgatag cggtcggccg ctttacttgt 5760

acagctcgtc catgccgaga gtgatcccgg cggcggtcac gaactccagc aggaccatgt 5820

gatcgcgctt ctcgttgggg tctttgctca gggcggactg ggtgctcagg tagtggttgt 5880

cgggcagcag cacggggccg tcgccgatgg gggtgttctg ctggtagtgg tcggccaggt 5940

gagtccagga gatgtttcag cactgttgcc tttagtctcg aggcaactta gacaactgag 6000

tattgatctg agcacagcag ggtgtgagct gtttgaagat actggggttg ggggtgaaga 6060

aactgcagag gactaactgg gctgagaccc agtggcaatg ttttagggcc taaggaatgc 6120

ctctgaaaat ctagatggac aactttgact ttgagaaaag agaggtggaa atgaggaaaa 6180

tgacttttct ttattagatt tcggtagaaa gaactttcat ctttccccta tttttgttat 6240

tcgttttaaa acatctatct ggaggcagga caagtatggt cattaaaaag atgcaggcag 6300

aaggcatata ttggctcagt caaagtgggg aactttggtg gccaaacata cattgctaag 6360

gctattccta tatcagctgg acacatataa aatgctgcta atgcttcatt acaaacttat 6420

atcctttaat tccagatggg ggcaaagtat gtccaggggt gaggaacaat tgaaacattt 6480

gggctggagt agattttgaa agtcagctct gtgtgtgtgt gtgtgtgtgt gtgtgtgaga 6540

gcgtgtgttt cttttaacgt tttcagccta cagcatacag ggttcatggt ggcaagaaga 6600

taacaagatt taaattatgg ccagtgacta gtgctgcaag aagaacaact acctgcattt 6660

aatgggaaag caaaatctca ggctttgagg gaagttaaca taggcttgat tctgggtgga 6720

agctgggtgt gtagttatct ggaggccagg ctggagctct cagctcacta tgggttcatc 6780

tttattgtct cctttcatct caacagctgc acgctgccgt cctcgatgtt gtggcggatc 6840

ttgaagttca ccttgatgcc gttcttctgc ttgtcggcca tgatatagac gttgtggctg 6900

ttgtagttgt actccagctt gtgccccagg atgttgccgt cctccttgaa gtcgatgccc 6960

ttcagctcga tgcggttcac cagggtgtcg ccctcgaact tcacctcggc gcgggtcttg 7020

tagttgccgt cgtccttgaa gaagatggtg cgctcctgga cgtagccttc gggcatggcg 7080

gacttgaaga agtcgtgctg cttcatgtgg tcggggtagc ggctgaagca ctgcacgccg 7140

taggtcaggg tggtcacgag ggtgggccag ggcacgggca gcttgccggt ggtgcagatg 7200

aacttcaggg tcagcttgcc gtaggtggca tcgccctcgc cctcgccgga cacgctgaac 7260

ttgtggccgt ttacgtcgcc gtccagctcg accaggatgg gcaccacccc ggtgaacagc 7320

tcctcgccct tgctcaccat ggtggcgaat tcgaagcttg agcacgagat ctgagtccgg 7380

taggcctagc ggatctgacg gttcactaaa ccagctctgc ttatatagac ctcccaccgt 7440

acacgcctac cgcccatttg cgtcaatggg gcggagttgt tacgacattt tggaaagtcc 7500

cgttgatttt ggtgccaaaa caaactccca ttgacgtcaa tggggtggag acttggaaat 7560

ccccgtgagt caaaccgcta tccacgccca ttgatgtact gccaaaaccg catcaccatg 7620

gtaatagcga tgactaatac gtagatgtac tgccaagtag gaaagtccca taaggtcatg 7680

tactgggcat aatgccaggc gggccattta ccgtcattga cgtcaatagg gggcgtactt 7740

ggcatatgat acacttgatg tactgccaag tgggcagttt accgtaaata ctccacccat 7800

tgacgtcaat ggaaagtccc tattggcgtt actatgggaa catacgtcat tattgacgtc 7860

aatgggcggg ggtcgttggg cggtcagcca ggcgggccat ttaccgtaag ttatgtaacg 7920

ggcctgctgc cggctctgcg gcctcttccg cgtcttcgcc ttcgccctca gacgagtcgg 7980

atctcccttt gggccgcctc cccgcctgtc tagcttgact gactgagata cagcgtacct 8040

tcagctcaca gacatgataa gatacattga tgagtttgga caaaccacaa ctagaatgca 8100

gtgaaaaaaa tgctttattt gtgaaatttg tgatgctatt gctttatttg taaccattat 8160

aagctgcaat aaacaagtta acaacaacaa ttgcattcat tttatgtttc aggttcaggg 8220

ggaggtgtgg gaggtttttt aaagcaagta aaacctctac aaatgtggta ttggcccatc 8280

tctatcggta tcgtagcata acccccttggg gcctctaaac gggtcttgag gggtttttg 8340

tgcccctcgg gccggattgc tatctaccgg cattggcgca gaaaaaaatg cctgatgcga 8400

cgctgcgcgt cttatactcc cacatatgcc agattcagca acggatacgg cttccccaac 8460

ttgcccactt ccatacgtgt cctccttacc agaaatttat ccttaaggtc gtcagctatc 8520

ctgcaggcga tctctcgatt tcgatcaaga cattccttta atggtctttt ctggacacca 8580

ctaggggtca gaagtagttc atcaaacttt cttccctccc taatctcatt ggttaccttg 8640

ggctatcgaa acttaattaa gcgatctgca tctcaattag tcagcaacca tagtcccgcc 8700

cctaactccg cccatcccgc ccctaactcc gcccagttcc gcccattctc cgccccatcg 8760

ctgactaatt ttttttattt atgcagaggc cgaggccgcc tcggcctctg agctattcca 8820

gaagtagtga ggaggctttt ttggaggcct aggcttttgc aaaggaggta gccaacatga 8880

ttgaacaaga tggattgcac gcaggttctc ccgccgcttg ggtggagagg ctattcggct 8940

atgactgggc acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc 9000

aggggcgccc ggttcttttt gtcaagaccg acctgtccgg tgccctgaat gaactccagg 9060

acgaggcagc gcggctatcg tggctggcca cgacgggcgt tccttgcgca gctgtgctcg 9120

acgttgtcac tgaagcggga agggactggc tgctattggg cgaagtgccg gggcaggatc 9180

tcctgtcatc tcaccttgct cctgccgaga aagtatccat catggctgat gcaatgcggc 9240

ggctgcatac gcttgatccg gctacctgcc cattcgacca ccaagcgaaa catcgcatcg 9300

agcgagcacg tactcggatg gaagccggtc ttgtcgatca ggatgatctg gacgaagagc 9360

atcaggggct cgcgccagcc gaactgttcg ccaggctcaa ggcgcggatg cccgacggcg 9420

aggatctcgt cgtgacccac ggcgatgcct gcttgccgaa tatcatggtg gaaaatggcc 9480

gcttttctgg attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag 9540

cgttggctac ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttcctcg 9600

tgctttacgg tatcgccgct cccgattcgc agcgcatcgc cttctatcgc cttcttgacg 9660

agttcttcta gtatgtaagc cctgtgcctt ctagttgcca gccatctgtt gtttgcccct 9720

cccccgtgcc ttccttgacc ctggaaggtg ccactcccac tgtcctttcc taataaaatg 9780

aggaaattgc atcgcattgt ctgagtaggt gtcattctat tctgggggggt ggggtggggc 9840

aggacagcaa gggggaggat tgggaagaca atagcaggca tgctggggat gcggtgggct 9900

ctatggttaa ttaaccagtc aagtcagcta cttggcgaga tcgacttgtc tgggtttcga 9960

ctacgctcag aattgcgtca gtcaagttcg atctggtcct tgctattgca cccgttctcc 10020

gattacgagt ttcatttaaa tcatgtgagc aaaaggccag caaaaggcca ggaaccgtaa 10080

aaaggccgcg ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa 10140

tcgacgctca agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc 10200

ccctggaagc tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc 10260

cgcctttctc ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag 10320

ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac gaacccccccg ttcagcccga 10380

ccgctgcgcc ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc 10440

gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac 10500

agagttcttg aagtggtggc ctaactacgg ctacactaga agaacagtat ttggtatctg 10560

cgctctgctg aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca 10620

aaccaccgct ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa 10680

aggatctcaa gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa 10740

ctcacgttaa gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt 10800

aaattaaaaa tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag 10860

ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat 10920

agttgcattt aaatttccga actctccaag gccctcgtcg gaaaatcttc aaacctttcg 10980

tccgatccat cttgcaggct acctctcgaa cgaactatcg caagtctctt ggccggcctt 11040

gcgccttggc tattgcttgg cagcgcctat cgccaggtat tactccaatc ccgaatatcc 11100

gagatcggga tcacccgaga gaagttcaac ctacatcctc aatcccgatc tatccgagat 11160

ccgaggaata tcgaaatcgg ggcgcgcctg gtgtaccgag aacgatcctc tcagtgcgag 11220

tctcgacgat ccatatcgtt gcttggcagt cagccagtcg gaatccagct tgggacccag 11280

gaagtccaat cgtcagatat tgtactcaag cctggtcacg gcagcgtacc gatctgttta 11340

aacctagata ttgatagtct gatcggtcaa cgtataatcg agtcctagct tttgcaaaca 11400

tctatcaaga gacaggatca gcaggaggct ttcgcatgag tattcaacat ttccgtgtcg 11460

cccttattcc ctttttgcg gcattttgcc ttcctgtttt tgctcaccca gaaacgctgg 11520

tgaaagtaaa agatgctgaa gatcagttgg gtgcgcgagt gggttacatc gaactggatc 11580

tcaacagcgg taagatcctt gagagttttc gccccgaaga acgctttcca atgatgagca 11640

cttttaaagt tctgctatgt ggcgcggtat tatcccgtat tgacgccggg caagagcaac 11700

tcggtcgccg catacactat tctcagaatg acttggttga gtattcacca gtcacagaaa 11760

agcatcttac ggatggcatg acagtaagag aattatgcag tgctgccata accatgagtg 11820

ataacactgc ggccaactta cttctgacaa cgattggagg accgaaggag ctaaccgctt 11880

ttttgcacaa catgggggat catgtaactc gccttgatcg ttgggaaccg gagctgaatg 11940

aagccatacc aaacgacgag cgtgacacca cgatgcctgt agcaatggca acaaccttgc 12000

gtaaactatt aactggcgaa ctacttactc tagcttcccg gcaacagttg atagactgga 12060

tggaggcgga taaagttgca ggaccacttc tgcgctcggc ccttccggct ggctggttta 12120

ttgctgataa atctggagcc ggtgagcgtg ggtctcgcgg tatcattgca gcactggggc 12180

cagatggtaa gccctcccgt atcgtagtta tctacacgac ggggagtcag gcaactatgg 12240

atgaacgaaa tagacagatc gctgagatag gtgcctcact gattaagcat tggtaaccga 12300

ttctaggtgc attggcgcag aaaaaaatgc ctgatgcgac gctgcgcgtc ttatactccc 12360

acatatgcca gattcagcaa cggatacggc ttccccaact tgcccacttc catacgtgtc 12420

ctccttacca gaaatttatc cttaagatcg tttaaactcg actctggctc tatcgaatct 12480

ccgtcgtttc gagcttacgc gaacagccgt ggcgctcatt tgctcgtcgg gcatcgaatc 12540

tcgtcagcta tcgtcagctt acctttttgg cagcgatcgc ggctcccgac atcttggacc 12600

attagctcca caggtatctt cttccctcta gtggtcataa cagcagcttc agctacctct 12660

caattcaaaa aacccctcaa gacccgttta gaggccccaa ggggttatgc tatcaatcgt 12720

tgcgttacac acacaaaaaa ccaacacaca tccatcttcg atggatagcg attttattat 12780

ctaactgctg atcgagtgta gccagatcta gtaatcaatt acggggtcat tagttcatag 12840

cccatatatg gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc 12900

caacgacccc cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg 12960

gactttccat tgacgtcaat gggtggagta tttacggtaa actgcccact tggcagtaca 13020

tcaagtgtat catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc 13080

ctggcattat gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt 13140

attagtcatc gctattacca tgctgatgcg gttttggcag tacatcaatg ggcgtggata 13200

gcggtttgac tcacggggat ttccaagtct ccaccccatt gacgtcaatg ggagtttgtt 13260

ttggcaccaa aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca 13320

aatgggcggt aggcgtgtac ggtgggaggt ctatataagc agagctggtt tagtgaaccg 13380

tcagatcaga tctttgtcga tcctaccatc cactcgacac acccgccagc ggccgc 13436

<210> 48

<211> 13433

<212> DNA

<213> Artificial Sequence

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

<400> 48

taatacgact cactataggg agaagtactg ccaccatggg caagaagcaa aatcgcaaga 60

cggggaattc caagacacaa tccgctagcc caccacctaa agagcgttct agctcccctg 120

ctactgagca gtcctggatg gaaaacgact tcgatgaact ccgggaagag ggatttaggc 180

gatccaacta ttcagaactc cgcgaagata tccagacaaa ggggaaggaa gtcgagaatt 240

tcgagaagaa cctcgaggag tgcatcaccc gtatcacaaa cactgagaaa tgtctcaaag 300

aactcatgga acttaagaca aaagccaggg agcttcgaga ggagtgtcgg agtctgagat 360

ccaggtgtga ccagctcgag gagcgcgtga gcgcgatgga agacgagatg aacgagatga 420

aaagagaggg caaattcagg gagaagcgca ttaagaggaa cgaacagagt ctgcaggaga 480

tttgggatta cgtcaagagg cctaacctgc ggttgatcgg cgtccccgag agcgacgtag 540

aaaacgggac taaactggag aatacacttc aagacatcat tcaagaaaat tttccaaacc 600

tggctcggca agctaatgtg caaatccaag agatccaacg cacaccccag cggtatagct 660

ctcggcgtgc caccctagg catattatcg tgcgctttac taaggtggag atgaaagaga 720

agatgctgcg agccgctcgg gaaaagggaa gggtgacttt gaagggcaaa cctattcggc 780

tgacggttga ccttagcgcc gagacactcc aggcacgccg ggaatggggc cccatcttta 840

atatcctgaa ggagaagaac ttccagccac gaatctctta ccctgcaaag ttgagtttta 900

tctccgaggg tgagattaag tatttcatcg ataaacagat gctgcgagac ttcgtgacaa 960

ctcgcccagc tctcaaggaa ctgctcaaag aggctcttaa tatggagcgc aataatagat 1020

atcaaccctt gcagaaccac gcaaagatgg gaagcggaga gggcagagga agtctgctaa 1080

catgcggtga cgtcgaggag aatcctggac ctatgaccgg ctctaactca catatcacca 1140

tccttacact taacattaac ggcctcaact cagctatcaa gcgccatcgg ctggccagct 1200

ggatcaaatc acaggatcca agcgtttgtt gcatccaaga gacccacctg acctgtagag 1260

atactcaccg cctcaagatc aagggatggc gaaagattta tcaggcgaac ggtaagcaga 1320

agaaagccgg agtcgcaatt ctggtctcag acaagacgga tttcaagccc accaaaatta 1380

agcgtgataa ggaaggtcac tatattatgg tgaaaggcag catacagcag gaagaactta 1440

ccatattgaa catctacgcg ccaaacaccg gcgcacctcg ctttatcaaa caggtcctgt 1500

ccgatctgca gcgagatctg gattctcata cgttgattat gggtgatttc aatacaccat 1560

tgagcaccct ggatcgcagc accaggcaaa aggtaaataa agacacgcaa gagctcaata 1620

gcgcactgca tcaggcagat ctcattgata tttatcgcac tcttcatcct aagagtaccg 1680

agtacacatt cttcagcgcc ccacatcata catactcaaa gatcgatcat atcgtcggct 1740

caaaggctct gctgtcaaag tgcaagcgca cagagataat tacaaattac ctgtcagatc 1800

atagcgcgat caagctcgag ctgagaatca agaacctgac ccagagccgg agtaccactt 1860

ggaagcttaa taacctgctg ctcaacgatt attgggtcca caatgagatg aaggcagaga 1920

ttaaaatgtt cttcgaaaca aatgagaata aggatactac ctatcaaaac ctttgggatg 1980

cctttaaggc cgtctgcaga ggcaagttca tcgccctcaa cgcctataaa agaaaacaag 2040

agagatctaa gatcgatact ctcacctctc agctgaagga gttggagaaa caggaacaga 2100

cccactccaa ggcgtcaaga cggcaggaga tcacaaagat tcgcgccgag ttgaaagaga 2160

tcgaaaccca aaagactctt cagaaaatta acgagtctcg tagttggttc ttcgagcgga 2220

ttaataagat agacagacct ctggcacgac tgattaagaa gaagcgcgaa aagaaccaga 2280

ttgataccat caagaacgac aagggcgaca tcactactga cccgaccgag atccagacca 2340

ctattcggga gtattataag catttgtatg ctaacaagct tgagaacctg gaagagatgg 2400

acacttttct ggatacctat actctgccac ggcttaatca agaggaagtc gagtccctca 2460

accgcccaat tacaggaagc gagattgtgg ccataattaa ctccctgccg acaaagaaat 2520

ctcctggtcc ggacgggttt acagctgagt tttatcaacg gtatatggaa gagcttgtac 2580

cgtttctgct caagctcttt cagtctatag aaaaggaagg catcttgccc aattccttct 2640

acgaagcttc tataatactt attcccaaac caggacgcga taccacaaag aaggaaaact 2700

tccggcccat tagtctcatg aatatcgacg ctaaaatatt gaacaagatt ctcgccaaca 2760

gaatccaaca acatattaag aaattgatac atcacgacca ggtggggttt atacctggca 2820

tgcagggctg gtttaacatc cggaagagta ttaacgtcat tcaacacatt aatagagcta 2880

aggataagaa tcatatgatc atctctatag acgcggaaaa ggcattcgat aagattcagc 2940

agccatttat gctcaagact ctgaacaaac tcggcatcga cggaacatat tttaagatta 3000

ttcgcgcaat ttacgataag ccgactgcta acattatcct taacggccaa aagctcgagg 3060

cctttccgct caagactgga acccgccaag gctgtcccct ctccccgctt ttgtttaata 3120

ttgtactcga ggtgctggct agggctattc gtcaagagaa agagattaaa gggatacagc 3180

tcgggaagga agaggtcaag ctttccttgt tcgccgatga tatgattgtg tacctggaga 3240

atcctattgt gtctgctcag aaccttctta aacttatttc taactttagc aaggtcagcg 3300

gctataagat taacgtccag aaatctcagg cctttctgta cacaaataat cgacagaccg 3360

aatcccagat aatgggtgag cttccgtttg tcatagccag caaaaggata aagtatctcg 3420

gaatccagct gacacgagac gttaaagatt tgtttaagga aaattacaag cctctcctga 3480

aagagattaa ggaagatact aataagtgga agaatatccc ctgttcatgg gttggcagaa 3540

tcaacatagt gaagatggca atacttccta aagtgatata tcgctttaac gccatcccaa 3600

ttaaactgcc tatgaccttc tttacggagc tcgagaaaac aacccttaaa tttatatgga 3660

atcaaaagag agcaagaata gcgaagtcca tcttgagcca gaagaataag gccggtggga 3720

ttactttgcc tgattttaag ttgtattata aagccacagt aactaagaca gcctggtatt 3780

ggtatcagaa tagagacatc gaccagtgga atcggaccga accatcagag ataatgcccc 3840

acatctataa ttaccttata ttcgataagc cagaaaagaa taaacagtgg ggcaaagaca 3900

gcctcttcaa caagtggtgt tgggagaatt ggctggccat atgccggaaa ctcaagctcg 3960

acccctttct tacaccctac actaaaatca acagtaggtg gatcaaggac ttgaatgtca 4020

agccaaagac tataaagaca ctggaagaga atcttgggat cacaatacaa gatataggcg 4080

tcggcaaaga ttttatgtca aagacgccca aggccatggc cactaaggat aagattgata 4140

agtgggacct tattaagctc aaaagcttct gtactgccaa ggagaccacg atcagagtta 4200

ataggcagcc cactacatgg gaaaagattt tcgccactta ttcatcagat aaggggttga 4260

taagcagaat atataacgag ctgaagcaga tctacaagaa gaaaacgaat aatcccatca 4320

agaagtgggc aaaagatatg aacaggcatt ttagcaaaga ggatatctac gccgcgaaga 4380

agcatatgaa gaagtgtagt tcaagcttgg ccattcgtga gatgcagatt aagacgacca 4440

tgcgatacca ccttacccca gtgaggatgg caattatcaa gaaatctggc aataatagat 4500

gttggcgggg ctgtggcgag attggcaccc tgctccattg ctggtgggat tgcaagctgg 4560

tgcagccgct ttggaaatca gtctggcgct ttctgaggga cctcgagctt gagattccct 4620

tcgatcccgc aattcccttg ctcggaatct atcctaacga atacaagagc tgttgttaca 4680

aggatacgtg tacccggatg ttcatcgcgg ccttgtttac gatagctaag acgtggaatc 4740

agcctaagtg ccccacaatg atcgattgga tcaagaaaat gtggcatatt tataccatgg 4800

agtattacgc agcaattaag aatgacgaat ttatttcctt cgttgggacc tggatgaagc 4860

tggagactat tattctgagc aagctgtctc aggagcaaaa gacaaagcat agaatcttct 4920

ctctcattgg tggtaacgac tacaaagacg atgacgacaa gtaaagcgct tctagaagtt 4980

gtctcctcct gcactgactg actgatacaa tcgatttctg gatccgcagg cctaatcaac 5040

ctctggatta caaaatttgt gaaagattga ctggtattct taactatgtt gctccttta 5100

cgctatgtgg atacgctgct ttaatgcctt tgtatcatgc tattgcttcc cgtatggctt 5160

tcattttctc ctccttgtat aaatcctggt tgctgtctct ttatgaggag ttgtggcccg 5220

ttgtcaggca acgtggcgtg gtgtgcactg tgtttgctga cgcaaccccc actggttggg 5280

gcattgccac cacctgtcag ctcctttccg ggactttcgc tttccccctc cctattgcca 5340

cggcggaact catcgccgcc tgccttgccc gctgctggac aggggctcgg ctgttgggca 5400

ctgacaattc cgtggtgttg tcggggaagc tgacgtcctt tccatggctg ctcgcctgtg 5460

ttgccacctg gattctgcgc gggacgtcct tctgctacgt cccttcggcc ctcaatccag 5520

cggaccttcc ttcccgcgaa caaacgaccc aacacccgtg cgttttattc tgtcttttta 5580

ttgccgatcc cctcagaaga actcgtcaag aaggcgatag aaggcgatgc gctgcgaatc 5640

gggagcggcg ataccgtaaa gcacgaggaa gcggtcagcc cattcgccgc caagctcttc 5700

agcaatatca cgggtagcca acgctatgtc ctgatagcgg tcggccgctt tacttgtaca 5760

gctcgtccat gccgagagtg atcccggcgg cggtcacgaa ctccagcagg accatgtgat 5820

cgcgcttctc gttggggtct ttgctcaggg cggactgggt gctcaggtag tggttgtcgg 5880

gcagcagcac ggggccgtcg ccgatggggg tgttctgctg gtagtggtcg gccaggtgag 5940

tccaggagat gtttcagcac tgttgccttt agtctcgagg caacttagac aactgagtat 6000

tgatctgagc acagcagggt gtgagctgtt tgaagatact ggggttgggg gtgaagaaac 6060

tgcagaggac taactgggct gagacccagt ggcaatgttt tagggcctaa ggaatgcctc 6120

tgaaaatcta gatggacaac tttgactttg agaaaagaga ggtggaaatg aggaaaatga 6180

cttttcttta ttagatttcg gtagaaagaa ctttcatctt tcccctattt ttgttattcg 6240

ttttaaaaca tctatctgga ggcaggacaa gtatggtcat taaaaagatg caggcagaag 6300

gcatatattg gctcagtcaa agtggggaac tttggtggcc aaacatacat tgctaaggct 6360

attcctatat cagctggaca catataaaat gctgctaatg cttcattaca aacttatatc 6420

ctttaattcc agatgggggc aaagtatgtc caggggtgag gaacaattga aacatttggg 6480

ctggagtaga ttttgaaagt cagctctgtg tgtgtgtgtg tgtgtgtgtg tgtgagagcg 6540

tgtgtttctt ttaacgtttt cagcctacag catacagggt tcatggtggc aagaagataa 6600

caagatttaa attatggcca gtgactagtg ctgcaagaag aacaactacc tgcatttaat 6660

gggaaagcaa aatctcaggc tttgagggaa gttaacatag gcttgattct gggtggaagc 6720

tgggtgtgta gttatctgga ggccaggctg gagctctcag ctcactatgg gttcatcttt 6780

attgtctcct ttcatctcaa cagctgcacg ctgccgtcct cgatgttgtg gcggatcttg 6840

aagttcacct tgatgccgtt cttctgcttg tcggccatga tatagacgtt gtggctgttg 6900

tagttgtact ccagcttgtg ccccaggatg ttgccgtcct ccttgaagtc gatgcccttc 6960

agctcgatgc ggttcaccag ggtgtcgccc tcgaacttca cctcggcgcg ggtcttgtag 7020

ttgccgtcgt ccttgaagaa gatggtgcgc tcctggacgt agccttcggg catggcggac 7080

ttgaagaagt cgtgctgctt catgtggtcg gggtagcggc tgaagcactg cacgccgtag 7140

gtcagggtgg tcacgagggt gggccagggc acgggcagct tgccggtggt gcagatgaac 7200

ttcagggtca gcttgccgta ggtggcatcg ccctcgccct cgccggacac gctgaacttg 7260

tggccgttta cgtcgccgtc cagctcgacc aggatgggca ccaccccggt gaacagctcc 7320

tcgcccttgc tcaccatggt ggcgaattcg aagcttgagc acgagatctg agtccggtag 7380

gcctagcgga tctgacggtt cactaaacca gctctgctta tatagacctc ccaccgtaca 7440

cgcctaccgc ccatttgcgt caatggggcg gagttgttac gacattttgg aaagtcccgt 7500

tgattttggt gccaaaacaa actcccattg acgtcaatgg ggtggagact tggaaatccc 7560

cgtgagtcaa accgctatcc acgcccattg atgtactgcc aaaaccgcat caccatggta 7620

atagcgatga ctaatacgta gatgtactgc caagtaggaa agtcccataa ggtcatgtac 7680

tgggcataat gccaggcggg ccatttaccg tcattgacgt caataggggg cgtacttggc 7740

atatgataca cttgatgtac tgccaagtgg gcagtttacc gtaaatactc cacccattga 7800

cgtcaatgga aagtccctat tggcgttact atgggaacat acgtcattat tgacgtcaat 7860

gggcgggggt cgttgggcgg tcagccaggc gggccattta ccgtaagtta tgtaacgggc 7920

ctgctgccgg ctctgcggcc tcttccgcgt cttcgccttc gccctcagac gagtcggatc 7980

tccctttggg ccgcctcccc gcctgtctag cttgactgac tgagatacag cgtaccttca 8040

gctcacagac atgataagat acattgatga gtttggacaa accacaacta gaatgcagtg 8100

aaaaaaatgc tttatttgtg aaatttgtga tgctattgct ttatttgtaa ccattataag 8160

ctgcaataaa caagttaaca acaacaattg cattcatttt atgtttcagg ttcaggggga 8220

ggtgtgggag gtttttttaaa gcaagtaaaa cctctacaaa tgtggtattg gcccatctct 8280

atcggtatcg tagcataacc ccttggggcc tctaaacggg tcttgagggg ttttttgtgc 8340

ccctcgggcc ggattgctat ctaccggcat tggcgcagaa aaaaatgcct gatgcgacgc 8400

tgcgcgtctt atactcccac atatgccaga ttcagcaacg gatacggctt ccccaacttg 8460

cccacttcca tacgtgtcct ccttaccaga aatttatcct taaggtcgtc agctatcctg 8520

caggcgatct ctcgatttcg atcaagacat tcctttaatg gtcttttctg gacaccacta 8580

ggggtcagaa gtagttcatc aaactttctt ccctccctaa tctcattggt taccttgggc 8640

tatcgaaact taattaagcg atctgcatct caattagtca gcaaccatag tcccgcccct 8700

aactccgccc atcccgcccc taactccgcc cagttccgcc cattctccgc cccatcgctg 8760

actaattttt tttatttatg cagaggccga ggccgcctcg gcctctgagc tattccagaa 8820

gtagtgagga ggcttttttg gaggcctagg cttttgcaaa ggaggtagcc aacatgattg 8880

aacaagatgg attgcacgca ggttctcccg ccgcttgggt ggagaggcta ttcggctatg 8940

actgggcaca acagacaatc ggctgctctg atgccgccgt gttccggctg tcagcgcagg 9000

ggcgcccggt tctttttgtc aagaccgacc tgtccggtgc cctgaatgaa ctccaggacg 9060

aggcagcgcg gctatcgtgg ctggccacga cgggcgttcc ttgcgcagct gtgctcgacg 9120

ttgtcactga agcgggaagg gactggctgc tattgggcga agtgccgggg caggatctcc 9180

tgtcatctca ccttgctcct gccgagaaag tatccatcat ggctgatgca atgcggcggc 9240

tgcatacgct tgatccggct acctgcccat tcgaccacca agcgaaacat cgcatcgagc 9300
gagcacgtac tcggatggaa gccggtcttg tcgatcagga tgatctggac gaagagcatc 9360
aggggctcgc gccagccgaa ctgttcgcca ggctcaaggc gcggatgccc gacggcgagg 9420
atctcgtcgt gacccacggc gatgcctgct tgccgaatat catggtggaa aatggccgct 9480
tttctggatt catcgactgt ggccggctgg gtgtggcgga ccgctatcag gacatagcgt 9540
tggctacccg tgatattgct gaagagcttg gcggcgaatg ggctgaccgc ttcctcgtgc 9600
tttacggtat cgccgctccc gattcgcagc gcatcgcctt ctatcgcctt cttgacgagt 9660

tcttctagta tgtaagccct gtgccttcta gttgccagcc atctgttgtt tgcccctccc 9720
ccgtgccttc cttgaccctg gaaggtgcca ctcccactgt cctttcctaa taaaatgagg 9780
aaattgcatc gcattgtctg agtaggtgtc attctattct ggggggtggg gtggggcagg 9840
acagcaaggg ggaggattgg gaagacaata gcaggcatgc tggggatgcg gtgggctcta 9900
tggttaatta accagtcaag tcagctactt ggcgagatcg acttgtctgg gtttcgacta 9960
cgctcagaat tgcgtcagtc aagttcgatc tggtccttgc tattgcaccc gttctccgat 10020
tacgagtttc atttaaatca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa 10080

ggccgcgttg ctggcgtttt tccataggct ccgccccccc gacgagcatc acaaaaatcg 10140
acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc 10200
tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc 10260
ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc 10320
ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg 10380
ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc 10440
actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga 10500

gttcttgaag tggtggccta actacggcta cactagaaga acagtatttg gtatctgcgc 10560
tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac 10620
caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg 10680
atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc 10740
acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa 10800
ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta 10860
ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt 10920

tgcatttaaa tttccgaact ctccaaggcc ctcgtcggaa aatcttcaaa cctttcgtcc 10980

gatccatctt gcaggctacc tctcgaacga actatcgcaa gtctcttggc cggccttgcg 11040

ccttggctat tgcttggcag cgcctatcgc caggtattac tccaatcccg aatatccgag 11100

atcgggatca cccgagagaa gttcaaccta catcctcaat cccgatctat ccgagatccg 11160

aggaatatcg aaatcggggc gcgcctggtg taccgagaac gatcctctca gtgcgagtct 11220

cgacgatcca tatcgttgct tggcagtcag ccagtcggaa tccagcttgg gacccaggaa 11280

gtccaatcgt cagatattgt actcaagcct ggtcacggca gcgtaccgat ctgtttaaac 11340

ctagatattg atagtctgat cggtcaacgt ataatcgagt cctagctttt gcaaacatct 11400

atcaagagac aggatcagca ggaggctttc gcatgagtat tcaacatttc cgtgtcgccc 11460

ttattccctt ttttgcggca ttttgccttc ctgtttttgc tcacccagaa acgctggtga 11520

aagtaaaaga tgctgaagat cagttgggtg cgcgagtggg ttacatcgaa ctggatctca 11580

acagcggtaa gatccttgag agttttcgcc ccgaagaacg ctttccaatg atgagcactt 11640

ttaaagttct gctatgtggc gcggtattat cccgtattga cgccgggcaa gagcaactcg 11700

gtcgccgcat acactattct cagaatgact tggttgagta ttcaccagtc acagaaaagc 11760

atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtgata 11820

acactgcggc caacttactt ctgacaacga ttggaggacc gaaggagcta accgcttttt 11880

tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag 11940

ccataccaaa cgacgagcgt gacaccacga tgcctgtagc aatggcaaca accttgcgta 12000

aactattaac tggcgaacta cttactctag cttcccggca acagttgata gactggatgg 12060

aggcggataa agttgcagga ccacttctgc gctcggccct tccggctggc tggtttattg 12120

ctgataaatc tggagccggt gagcgtgggt ctcgcggtat cattgcagca ctggggccag 12180

atggtaagcc ctcccgtatc gtagttatct acacgacggg gagtcaggca actatggatg 12240

aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg taaccgattc 12300

taggtgcatt ggcgcagaaa aaaatgcctg atgcgacgct gcgcgtctta tactcccaca 12360

tatgccagat tcagcaacgg atacggcttc cccaacttgc ccacttccat acgtgtcctc 12420

cttaccagaa atttatcctt aagatcgttt aaactcgact ctggctctat cgaatctccg 12480

tcgtttcgag cttacgcgaa cagccgtggc gctcatttgc tcgtcgggca tcgaatctcg 12540

tcagctatcg tcagcttacc tttttggcag cgatcgcggc tcccgacatc ttggaccatt 12600

agctccacag gtatcttctt ccctctagtg gtcataacag cagcttcagc tacctctcaa 12660

ttcaaaaaac ccctcaagac ccgtttagag gccccaaggg gttatgctat caatcgttgc 12720

gttacacaca caaaaaacca acacacatcc atcttcgatg gatagcgatt ttattatcta 12780

actgctgatc gagtgtagcc agatctagta atcaattacg gggtcattag ttcatagccc 12840

atatatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct gaccgcccaa 12900

cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac 12960

tttccattga cgtcaatggg tggagtattt acggtaaact gcccacttgg cagtacatca 13020

agtgtatcat atgccaagta cgcccccta tgacgtcaat gacggtaaat ggcccgcctg 13080

gcattatgcc cagtacatga ccttatggga ctttcctact tggcagtaca tctacgtatt 13140

agtcatcgct attaccatgc tgatgcggtt ttggcagtac atcaatgggc gtggatagcg 13200

gtttgactca cggggatttc caagtctcca ccccattgac gtcaatggga gtttgttttg 13260

gcaccaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat 13320

gggcggtagg cgtgtacggt gggaggtcta tataagcaga gctggtttag tgaaccgtca 13380

gatcagatct ttgtcgatcc taccatccac tcgacacacc cgccagcggc cgc 13433

<210> 49

<211> 9241

<212> DNA

<213> Artificial Sequence

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

<400> 49

taatacgact cactataggg agaagtactg ccaccatggg caagaagcaa aatcgcaaga 60

cggggaattc caagacacaa tccgctagcc caccacctaa agagcgttct agctcccctg 120

ctactgagca gtcctggatg gaaaacgact tcgatgaact ccgggaagag ggatttaggc 180

gatccaacta ttcagaactc cgcgaagata tccagacaaa ggggaaggaa gtcgagaatt 240

tcgagaagaa cctcgaggag tgcatcaccc gtatcacaaa cactgagaaa tgtctcaaag 300

aactcatgga acttaagaca aaagccaggg agcttcgaga ggagtgtcgg agtctgagat 360

ccaggtgtga ccagctcgag gagcgcgtga gcgcgatgga agacgagatg aacgagatga 420

aaagagaggg caaattcagg gagaagcgca ttaagaggaa cgaacagagt ctgcaggaga 480

tttgggatta cgtcaagagg cctaacctgc ggttgatcgg cgtccccgag agcgacgtag 540

aaaacgggac taaactggag aatacacttc aagacatcat tcaagaaaat tttccaaacc 600

tggctcggca agctaatgtg caaatccaag agatccaacg cacaccccag cggtatagct 660

ctcggcgtgc caccctagg catattatcg tgcgctttac taaggtggag atgaaagaga 720

agatgctgcg agccgctcgg gaaaagggaa gggtgacttt gaagggcaaa cctattcggc 780

tgacggttga ccttagcgcc gagacactcc aggcacgccg ggaatggggc cccatcttta 840

atatcctgaa ggagaagaac ttccagccac gaatctctta ccctgcaaag ttgagtttta 900

tctccgaggg tgagattaag tatttcatcg ataaacagat gctgcgagac ttcgtgacaa 960

ctcgcccagc tctcaaggaa ctgctcaaag aggctcttaa tatggagcgc aataatagat 1020

atcaaccctt gcagaaccac gcaaagatgt gagacagccg tcagaccatc aagactagga 1080

agaaactgca tcaactaatg agcaaaatca ccagctaaca tcatagtata catgaccggc 1140

tctaactcac atatcaccat ccttacactt aacattaacg gcctcaactc agctatcaag 1200

cgccatcggc tggccagctg gatcaaatca caggatccaa gcgtttgttg catccaagag 1260

acccacctga cctgtagaga tactcaccgc ctcaagatca agggatggcg aaagatttat 1320

caggcgaacg gtaagcagaa gaaagccgga gtcgcaattc tggtctcaga caagacggat 1380

ttcaagccca ccaaaattaa gcgtgataag gaaggtcact atattatggt gaaaggcagc 1440

atacagcagg aagaacttac catattgaac atctacgcgc caaacaccgg cgcacctcgc 1500

tttatcaaac aggtcctgtc cgatctgcag cgagatctgg attctcatac gttgattatg 1560

ggtgatttca atacaccatt gagcaccctg gatcgcagca ccaggcaaaa ggtaaataaa 1620

gacacgcaag agctcaatag cgcactgcat caggcagatc tcattgatat ttatcgcact 1680

cttcatccta agagtaccga gtacacattc ttcagcgccc cacatcatac atactcaaag 1740

atcgatcata tcgtcggctc aaaggctctg ctgtcaaagt gcaagcgcac agagataatt 1800

acaaattacc tgtcagatca tagcgcgatc aagctcgagc tgagaatcaa gaacctgacc 1860

cagagccgga gtaccacttg gaagcttaat aacctgctgc tcaacgatta ttgggtccac 1920

aatgagatga aggcagagat taaaatgttc ttcgaaacaa atgagaataa ggatactacc 1980

tatcaaaacc tttgggatgc ctttaaggcc gtctgcagag gcaagttcat cgccctcaac 2040

gcctataaaa gaaaacaaga gagatctaag atcgatactc tcacctctca gctgaaggag 2100

ttggagaaac aggaacagac ccactccaag gcgtcaagac ggcaggagat cacaaagatt 2160

cgcgccgagt tgaaagagat cgaaacccaa aagactcttc agaaaattaa cgagtctcgt 2220

agttggttct tcgagcggat taataagata gacagacctc tggcacgact gattaagaag 2280

aagcgcgaaa agaaccagat tgataccatc aagaacgaca agggcgacat cactactgac 2340

ccgaccgaga tccagaccac tattcgggag tattataagc atttgtatgc taacaagctt 2400

gagaacctgg aagagatgga cacttttctg gatacctata ctctgccacg gcttaatcaa 2460

gaggaagtcg agtccctcaa ccgcccaatt acaggaagcg agattgtggc cataattaac 2520

tccctgccga caaagaaatc tcctggtccg gacgggttta cagctgagtt ttatcaacgg 2580

tatatggaag agcttgtacc gtttctgctc aagctctttc agtctataga aaaggaaggc 2640

atcttgccca attccttcta cgaagcttct ataatactta ttcccaaacc aggacgcgat 2700

accacaaaga aggaaaactt ccggcccatt agtctcatga atatcgacgc taaaatattg 2760

aacaagattc tcgccaacag aatccaacaa catattaaga aattgataca tcacgaccag 2820

gtggggttta tacctggcat gcagggctgg tttaacatcc ggaagagtat taacgtcatt 2880

caacacatta atagagctaa ggataagaat catatgatca tctctataga cgcggaaaag 2940

gcattcgata agattcagca gccatttatg ctcaagactc tgaacaaact cggcatcgac 3000

ggaacatatt ttaagattat tcgcgcaatt tacgataagc cgactgctaa cattatcctt 3060

aacggccaaa agctcgaggc ctttccgctc aagactggaa cccgccaagg ctgtcccctc 3120

tccccgcttt tgtttaatat tgtactcgag gtgctggcta gggctattcg tcaagagaaa 3180

gagattaaag ggatacagct cgggaaggaa gaggtcaagc tttccttgtt cgccgatgat 3240

atgattgtgt acctggagaa tcctattgtg tctgctcaga accttcttaa acttatttct 3300

aactttagca aggtcagcgg ctataagatt aacgtccaga aatctcaggc ctttctgtac 3360

acaaataatc gacagaccga atcccagata atgggtgagc ttccgtttgt catagccagc 3420

aaaaggataa agtatctcgg aatccagctg acacgagacg ttaaagattt gtttaaggaa 3480

aattacaagc ctctcctgaa agagattaag gaagatacta ataagtggaa gaatatcccc 3540

tgttcatggg ttggcagaat caacatagtg aagatggcaa tacttcctaa agtgatatat 3600

cgctttaacg ccatcccaat taaactgcct atgaccttct ttacggagct cgagaaaaca 3660

acccttaaat ttatatggaa tcaaaagaga gcaagaatag cgaagtccat cttgagccag 3720

aagaataagg ccggtgggat tactttgcct gattttaagt tgtattataa agccacagta 3780

actaagacag cctggtattg gtatcagaat agagacatcg accagtggaa tcggaccgaa 3840

ccatcagaga taatgcccca catctataat taccttatat tcgataagcc agaaaagaat 3900

aaacagtggg gcaaagacag cctcttcaac aagtggtgtt gggagaattg gctggccata 3960

tgccggaaac tcaagctcga ccccttttctt acaccctaca ctaaaatcaa cagtaggtgg 4020

atcaaggact tgaatgtcaa gccaaagact ataaagacac tggaagagaa tcttgggatc 4080

acaatacaag atataggcgt cggcaaagat tttatgtcaa agacgcccaa ggccatggcc 4140

actaaggata agattgataa gtgggacctt attaagctca aaagcttctg tactgccaag 4200

gagaccacga tcagagttaa taggcagccc actacatggg aaaagatttt cgccacttat 4260

tcatcagata aggggttgat aagcagaata tataacgagc tgaagcagat ctacaagaag 4320

aaaacgaata atcccatcaa gaagtgggca aaagatatga acaggcattt tagcaaagag 4380

gatatctacg ccgcgaagaa gcatatgaag aagtgtagtt caagcttggc cattcgtgag 4440

atgcagatta agacgaccat gcgataccac cttaccccag tgaggatggc aattatcaag 4500

aaatctggca ataatagatg ttggcggggc tgtggcgaga ttggcaccct gctccattgc 4560
tggtgggatt gcaagctggt gcagccgctt tggaaatcag tctggcgctt tctgagggac 4620
ctcgagcttg agattccctt cgatcccgca attcccttgc tcggaatcta tcctaacgaa 4680
tacaagagct gttgttacaa ggatacgtgt acccggatgt tcatcgcggc cttgtttacg 4740
atagctaaga cgtggaatca gcctaagtgc cccacaatga tcgattggat caagaaaatg 4800
tggcatattt ataccatgga gtattacgca gcaattaaga atgacgaatt tatttccttc 4860
gttgggacct ggatgaagct ggagactatt attctgagca agctgtctca ggagcaaaag 4920

acaaagcata gaatcttctc tctcattggt ggtaacgctt ctaactttac tcagttcgtt 4980
ctcgtcgaca atggcggaac tggcgacgtg actgtcgccc caagcaactt cgctaacggg 5040
atcgctgaat ggatcagctc taactcgcgt tcacaggctt acaaagtaac ctgtagcgtt 5100
cgtcagagct ctgcgcagaa tcgcaaatac accatcaaag tcgaggtgcc taaaggcgcc 5160
tggcgttcgt acttaaatat ggaactaacc attccaattt tcgccacgaa ttccgactgc 5220
gagcttattg ttaaggcaat gcaaggtctc ctaaaagatg gaaacccgat tccctcagca 5280
atcgcagcaa actccggcat ctacgccatg gccagcaact tcacccagtt cgtgctggtg 5340

gacaacggcg gcaccggcga cgtgaccgtg gcccccagca acttcgccaa cggcatcgcc 5400
gagtggatca gcagcaacag cagaagccag gcctacaagg tgacctgcag cgtgagacag 5460
agcagcgccc agaacagaaa gtacaccatc aaggtggagg tgcccaaggg cgcctggaga 5520
agctacctga acatggagct gaccatcccc atcttcgcca ccaacagcga ctgcgagctg 5580
atcgtgaagg ccatgcaggg cctgctgaag gacggcaacc ccatccccag cgccatcgcc 5640
gccaacagcg gcatctacga ctacaaagac gatgacgaca agtaaagcaa cctacaaacg 5700
ggtggaggat caccccaccc gacacttcac aatcaagggg tacaatacac aagggtggag 5760

gaacacccca ccctccagac acattacaca gaaatccaat caaacagaag caccatcagg 5820
gcttctgcta ccaaatttat ctcaaaaaac tacaacaagg aatcaccatc agggattccc 5880
tgtgcaatat acgtcaaacg agggccacga cgggaggacg atcacgcctc ccgaatatcg 5940
gcatgtctgg ctttcgaatt cagtgcgtgg agcatcagcc cacgcagcca atcagagtcg 6000
aatacaagtc gactttcgcg aagagcatca gccttcgcgc cattcttaca caaaccacac 6060
tctcccctac aggaacagca tcagcgttcc tgcccagtac ccaactcaag aaaatttatg 6120
tccccatgca gcatcagcgc atgggcccca agaatacatc cccaacaaaa tcacatccga 6180

gcaccaacag ggctcggagt gttgtttctt gtccaactgg acaaaccctc catggaccat 6240

caggccatgg actctcacca acaagacaaa aactactctt ctcgaagcag catcagcgct 6300

tcgaaacact cgagcataca ttgtgcctat ttcttgggtg gacgatcacg ccacccatgc 6360

tctcacgaat ttcaaaacac ggacaaggac gagcaccacc agggctcgtc gttccacgtc 6420

caatacgatt acttaccttt cgggatcacg atcacggatc ccgcagctac atcacttcca 6480

ctcaggacat tcaagcatgc acgatcacgg catgctccac aagtctcaac cacagaaact 6540

accaaatggg ttcagcacca gcgaacccac tcctacctca aacctcttcc cacaaaactg 6600

gcaagcagga tcaccgcttg cccattccaa cataccaaat caaaaacaat tactggtaca 6660

gcatcagcgt accagcccac atctctcact actatcaaaa accaaaccgt tcagcaacag 6720

cgaacggtac acacggaaaa atcaactggt ttacaaatac gaaagacgat cacgctttcg 6780

tccagcgcaa actattacga aaaacatccg acgggaagag caacagcctt cccgcggcgg 6840

aaaacctcac aaaaacacga caaacggatg cacgaacacg gcatccgccg acaacccaca 6900

aacttacaac caggcaaacg gtgcaggatc accgcaccgt acatcaaaca cctcagatct 6960

catgcttcta gaagttgtct cctcctgcac tgactgactg atacaatcga tttctggatc 7020

cgcaggccta atcaacctct ggattacaaa atttgtgaaa gattgactgg tattcttaac 7080

tatgttgctc cttttacgct atgtggatac gctgctttaa tgcctttgta tcatgctatt 7140

gcttcccgta tggctttcat tttctcctcc ttgtataaat cctggttgct gtctctttat 7200

gaggagttgt ggcccgttgt caggcaacgt ggcgtggtgt gcactgtgtt tgctgacgca 7260

acccccactg gttggggcat tgccaccacc tgtcagctcc tttccgggac tttcgctttc 7320

cccctcccta ttgccacggc ggaactcatc gccgcctgcc ttgcccgctg ctggacaggg 7380

gctcggctgt tgggcactga caattccgtg gtgttgtcgg ggaagctgac gtcctttcca 7440

tggctgctcg cctgtgttgc cacctggatt ctgcgcggga cgtccttctg ctacgtccct 7500

tcggccctca atccagcgga ccttccttcc cgctgagaga cacaaaaaat tccaacacac 7560

tattgcaatg aaaataaatt tcctttatta gccagaagtc agatgctcaa ggggcttcat 7620

gatgtcccca taatttttgg cagagggaaa aagatctcag tggtatttgt gagccagggc 7680

attggccttc tgataggcag cctgcacctg aggagtgcgg ccgctttact tgtacagctc 7740

gtccatgccg agagtgatcc cggcggcggt cacgaactcc agcaggacca tgtgatcgcg 7800

cttctcgttg gggtctttgc tcagggcgga ctgggtgctc aggtagtggt tgtcgggcag 7860

cagcacgggg ccgtcgccga tgggggtgtt ctgctggtag tggtcggcga gctgcacgct 7920

gccgtcctcg atgttgtggc ggatcttgaa gttcaccttg atgccgttct tctgcttgtc 7980

ggccatgata tagacgttgt ggctgttgta gttgtactcc agcttgtgcc ccaggatgtt 8040

gccgtcctcc ttgaagtcga tgcccttcag ctcgatgcgg ttcaccaggg tgtcgccctc 8100

gaacttcacc tcggcgcggg tcttgtagtt gccgtcgtcc ttgaagaaga tggtgcgctc 8160

ctggacgtag ccttcgggca tggcggactt gaagaagtcg tgctgcttca tgtggtcggg 8220

gtagcggctg aagcactgca cgccgtaggt cagggtggtc acgagggtgg gccagggcac 8280

gggcagcttg ccggtggtgc agatgaactt cagggtcagc ttgccgtagg tggcatcgcc 8340

ctcgccctcg ccggacacgc tgaacttgtg gccgtttacg tcgccgtcca gctcgaccag 8400

gatgggcacc accccggtga acagctcctc gcccttgctc accatggtgg cgggatctga 8460

cggttcacta aaccagctct gcttatatag acctcccacc gtacacgcct accgcccatt 8520

tgcgtcaatg gggcggagtt gttacgacat tttggaaagt cccgttgatt ttggtgccaa 8580

aacaaactcc cattgacgtc aatggggtgg agacttggaa atccccgtga gtcaaaccgc 8640

tatccacgcc cattgatgta ctgccaaaac cgcatcacca tggtaatagc gatgactaat 8700

acgtagatgt actgccaagt aggaaagtcc cataaggtca tgtactgggc ataatgccag 8760

gcgggccatt taccgtcatt gacgtcaata gggggcgtac ttggcatatg atacacttga 8820

tgtactgcca agtgggcagt ttaccgtaaa tactccaccc attgacgtca atggaaagtc 8880

cctattggcg ttactatggg aacatacgtc attattgacg tcaatgggcg ggggtcgttg 8940

ggcggtcagc caggcgggcc atttaccgta agttatgtaa cgggcctgct gccggctctg 9000

cggcctcttc cgcgtcttcg ccttcgccct cagacgagtc ggatctccct ttgggccgcc 9060

tccccgcctg tctagcttga ctgactgaga tacagcgtac cttcagctca cagacatgat 9120

aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa aatgctttat 9180

ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca ataaacaagt 9240

t 9241

<210> 50

<211> 7309

<212> DNA

<213> Artificial Sequence

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

<400> 50

taatacgact cactataggg agaagtactg ccaccatggg caagaagcaa aatcgcaaga 60

cggggaattc caagacacaa tccgctagcc caccacctaa agagcgttct agctcccctg 120

ctactgagca gtcctggatg gaaaacgact tcgatgaact ccgggaagag ggatttaggc 180

gatccaacta ttcagaactc cgcgaagata tccagacaaa ggggaaggaa gtcgagaatt 240

tcgagaagaa cctcgaggag tgcatcaccc gtatcacaaa cactgagaaa tgtctcaaag 300

aactcatgga acttaagaca aaagccaggg agcttcgaga ggagtgtcgg agtctgagat 360

ccaggtgtga ccagctcgag gagcgcgtga gcgcgatgga agacgagatg aacgagatga 420

aaagagaggg caaattcagg gagaagcgca ttaagaggaa cgaacagagt ctgcaggaga 480

tttgggatta cgtcaagagg cctaacctgc ggttgatcgg cgtccccgag agcgacgtag 540

aaaacgggac taaactggag aatacacttc aagacatcat tcaagaaaat tttccaaacc 600

tggctcggca agctaatgtg caaatccaag agatccaacg cacaccccag cggtatagct 660

ctcggcgtgc cacccctagg catattatcg tgcgctttac taaggtggag atgaaagaga 720

agatgctgcg agccgctcgg gaaaagggaa gggtgacttt gaagggcaaa cctattcggc 780

tgacggttga ccttagcgcc gagacactcc aggcacgccg ggaatggggc cccatcttta 840

atatcctgaa ggagaagaac ttccagccac gaatctctta ccctgcaaag ttgagtttta 900

tctccgaggg tgagattaag tatttcatcg ataaacagat gctgcgagac ttcgtgacaa 960

ctcgcccagc tctcaaggaa ctgctcaaag aggctcttaa tatggagcgc aataatagat 1020

atcaaccctt gcagaaccac gcaaagatgt gagacagccg tcagaccatc aagactagga 1080

agaaactgca tcaactaatg agcaaaatca ccagctaaca tcatagtata catggtcata 1140

ggaacttaca tttcgattat taccttaaac gtgaatgggt taaatgcccc aaccaagaga 1200

catcggctgg ctgaatggat tcagaaacag gacccctata tttgctgtct gcaggagacc 1260

cacttccgtc ctcgcgacac atacagactg aaagtgaggg gctggaaaaa gatcttccat 1320

gccaatggaa atcaaaagaa agctggagtg gctattctca tctcagataa aattgacttc 1380

aaaataaaga atgttactcg agataaggag ggacactaca taatgatcca ggggtccatc 1440

caagaagagg atataactat tattaatatt tatgcaccca acattggcgc ccctcagtac 1500

atcaggcagc tgcttacagc tatcaaggag gaaatcgaca gtaacacgat tatcgtgggg 1560

gactttaaca ccagccttac tccgatggat agatcatcca aaatgaaaat aaataaggaa 1620

acagaggctc ttaatgacac cattgaccag atagatctga ttgatatata taggacattc 1680

catccaaaaa ctgccgatta cactttcttc agcagtgcgc atggaacctt ctccaggata 1740

gatcacatct tgggtcacaa aagtagcctc agtaagttta agaaaattga aatcattagc 1800

agcatctttt ctgaccataa cgctatgcgc ctggagatga atcacaggga gaagaacgta 1860

aagaagacaa acacctggag gctgaacaat acgctgctaa ataaccaaga gatcactgag 1920

gaaatcaaac aggaaataaa aaaatacttg gagacaaatg acaatgaaaa cacgaccacc 1980

cagaacttgt gggatgcagc taaagcggtt ctgagaggga agtttatagc tattcaagcc 2040

taccttaaga aacaggaaaa atctcaagtg aacaatttga ccttacacct aaagaaactg 2100

gagaaggagg agcagaccaa acccaaagtg agcaggagga aagaaatcat caagatcaga 2160

gccgaaatca atgaaataga aactaagaag acaattgcca agatcaataa aactaaatcc 2220

tggttctttg agaagatcaa caaaattgat aagccattag ccagactcat caagaaaaag 2280

agggagagga ctcagatcaa taagatcaga aatgagaaag gggaagttac aaccgacacc 2340

gcggagattc agaacatcct gagagactac tacaagcaac tttatgccaa taaaatggac 2400

aacctggaag aaatggacaa attcctggaa aggtataacc ttccccggct gaaccaggag 2460

gagactgaaa atatcaaccg cccaatcaca agtaatgaga ttgagactgt gattaagaat 2520

cttccaacta acaaaagtcc cggccccgat ggcttcacag gtgaattcta tcagaccttt 2580

cgggaggagt tgacacccat ccttctcaag ctcttccaaa aaattgcaga ggagggcaca 2640

ctcccgaact cattctatga ggccaccatc accctgatcc caaagcccga caaggacact 2700

acaaagaaag aaaattaccg accaatttcc ctgatgaata tcgatgccaa gatcctcaac 2760

aaaatcttgg caaacagaat ccagcagcac attaagagga tcatacacca cgatcaggtg 2820

ggctttatcc cggggatgca aggattcttc aatatccgca aatcaatcaa tgtgatccac 2880

catattaaca agttgaagaa gaagaaccat atgatcatct ccatcgatgc agagaaagct 2940

tttgacaaaa ttcaacaccc atttatgatc aaaactctcc agaaggtggg catcgagggg 3000

acctacctca acataattaa ggccatctat gataagccca cagccaacat cattctcaat 3060

ggtgaaaagc tgaaggcatt tcctctgcgg tccggaacga gacagggatg tcctctctct 3120

cctcttctgt tcaacatcgt tctggaagtc ctagccaccg ctatccgcga ggaaaaggaa 3180

attaaaggca tacagattgg aaaggaagag gtaaaactgt ctctgtttgc ggatgatatg 3240

atactgtaca tagagaatcc taaaactgcc acccggaagc tgttggagct aattaatgag 3300

tatggtaagg tcgccggtta caagattaat gctcagaagt ctcttgcttt cctgtacact 3360

aatgatgaaa agtctgaacg ggaaattatg gagacactcc cctttaccat tgcaaccaaa 3420

cgtattaaat accttggcat taacctgcct aaggagacaa aagacctgta tgctgaaaac 3480

tataagacac tgatgaaaga gattaaagat gataccaacc ggtggcggga tatcccatgt 3540

tcttggattg gcagaatcaa cattgtgaag atgagcatcc tgcccaaggc catctacaga 3600

ttcaatgcca tccctatcaa attacctatg gcatttttta cggagctgga acagatcatc 3660

ttaaaatttg tgtggcgcca caagcggccc cgaatcgcca aagcggtctt gaggcagaag 3720

aatggcgctg ggggaatccg actccctgac ttcagattgt actacaaagc taccgtcatc 3780

aagacaatct ggtactggca caagaacaga aacatcgatc agtggaacaa gatcgaaagc 3840

cctgagatta acccccgcac ctatggtcaa ctgatctatg acaaaggggg caaggatata 3900

caatggcgca aggacagcct cttcaataag tggtgctggg aaaactggac agccacctgc 3960
aagcgtatga agctggagta ctccctgaca ccatacacaa aaataaactc aaagtggatt 4020
cgagacctca atattcggct ggacactata aaactcctgg aggagaacat tgggcgtaca 4080
ctctttgaca ttaatcatag caagatcttt ttcgatcccc ctcctcgtgt aatggaaata 4140
aaaacaaaaa taaacaagtg ggatctgatg aaacttcaga gcttttgcac cgcaaaggag 4200
accataaaca agacgaagcg ccaaccctca gaatgggaga aaatatttgc gaatgagtct 4260
acggacaaag gcttaatctc caaaatatat aagcagctca ttcagctcaa tatcaaggaa 4320

acaaacaccc cgatccaaaa gtgggcagag gacctaaatc ggcatttctc caaggaagac 4380
atccagacgg ccacgaagca catgaagcga tgctcaactt ccctgattat tcgcgaaatg 4440
cagatcaaga ctactatgcg ctatcacctc actcctgttc ggatgggcat catccggaaa 4500
tctacaaaca acaagtgctg gagagggtgt ggcgaaaagg gaaccctctt gcattgttgg 4560
tgggagtgta agttgatcca gccactatgg cggaccatat ggaggttcct taaaaaactg 4620
aagattgagc tgccatatga cccagcaatc ccactgctgg gcatataccc ggagaaaacc 4680
gtgattcaga aagacacttg cacccgaatg ttcattgcag cattgtttac aatagccagg 4740

tcatggaagc agcctaagtg cccctcgaca gacgagtgga tcaagaagat gtggtacatt 4800
tatactatgg aatattacag cgccatcaaa cgcaacgaaa ttgggtcttt tctggagacg 4860
tggatggatc tagagactgt catccagagt gaggtaagtc agaaagagaa gaacaaatat 4920
cgtattttaa cgcatatttg tggaacctgg aagaatggta cagatgagcc ggtctgccga 4980
accgagattg agacccagat ggactacaaa gacgatgacg acaagtgaag cgcttctaga 5040
agttgtctcc tcctgcactg actgactgat acaatcgatt tctggatccg caggcctaat 5100
caacctctgg attacaaaat ttgtgaaaga ttgactggta ttcttaacta tgttgctcct 5160

tttacgctat gtggatacgc tgctttaatg cctttgtatc atgctattgc ttcccgtatg 5220
gctttcattt tctcctcctt gtataaatcc tggttgctgt ctctttatga ggagttgtgg 5280
cccgttgtca ggcaacgtgg cgtggtgtgc actgtgtttg ctgacgcaac ccccactggt 5340
tggggcattg ccaccacctg tcagctcctt tccgggactt tcgctttccc cctccctatt 5400
gccacggcgg aactcatcgc cgcctgcctt gcccgctgct ggacaggggc tcggctgttg 5460
ggcactgaca attccgtggt gttgtcgggg aagctgacgt cctttccatg gctgctcgcc 5520
tgtgttgcca cctggattct gcgcgggacg tccttctgct acgtcccttc ggccctcaat 5580

ccagcggacc ttccttcccg ctgagagaca caaaaaattc caacacacta ttgcaatgaa 5640

aataaatttc ctttattagc cagaagtcag atgctcaagg ggcttcatga tgtccccata 5700

atttttggca gagggaaaaa gatctcagtg gtatttgtga gccagggcat tggccttctg 5760

ataggcagcc tgcacctgag gagtgcggcc gctttacttg tacagctcgt ccatgccgag 5820

agtgatcccg gcggcggtca cgaactccag caggaccatg tgatcgcgct tctcgttggg 5880

gtctttgctc agggcggact gggtgctcag gtagtggttg tcgggcagca gcacggggcc 5940

gtcgccgatg ggggtgttct gctggtagtg gtcggcgagc tgcacgctgc cgtcctcgat 6000

gttgtggcgg atcttgaagt tcaccttgat gccgttcttc tgcttgtcgg ccatgatata 6060

gacgttgtgg ctgttgtagt tgtactccag cttgtgcccc aggatgttgc cgtcctcctt 6120

gaagtcgatg cccttcagct cgatgcggtt caccagggtg tcgccctcga acttcacctc 6180

ggcgcgggtc ttgtagttgc cgtcgtcctt gaagaagatg gtgcgctcct ggacgtagcc 6240

ttcgggcatg gcggacttga agaagtcgtg ctgcttcatg tggtcggggt agcggctgaa 6300

gcactgcacg ccgtaggtca gggtggtcac gagggtgggc cagggcacgg gcagcttgcc 6360

ggtggtgcag atgaacttca gggtcagctt gccgtaggtg gcatcgccct cgccctcgcc 6420

ggacacgctg aacttgtggc cgtttacgtc gccgtccagc tcgaccagga tgggcaccac 6480

cccggtgaac agctcctcgc ccttgctcac catggtggcg ggatctgacg gttcactaaa 6540

ccagctctgc ttatatagac ctcccaccgt acacgcctac cgcccatttg cgtcaatggg 6600

gcggagttgt tacgacattt tggaaagtcc cgttgatttt ggtgccaaaa caaactccca 6660

ttgacgtcaa tggggtggag acttggaaat ccccgtgagt caaaccgcta tccacgccca 6720

ttgatgtact gccaaaaccg catcaccatg gtaatagcga tgactaatac gtagatgtac 6780

tgccaagtag gaaagtccca taaggtcatg tactgggcat aatgccaggc gggccattta 6840

ccgtcattga cgtcaatagg gggcgtactt ggcatatgat acacttgatg tactgccaag 6900

tgggcagttt accgtaaata ctccacccat tgacgtcaat ggaaagtccc tattggcgtt 6960

actatgggaa catacgtcat tattgacgtc aatgggcggg ggtcgttggg cggtcagcca 7020

ggcgggccat ttaccgtaag ttatgtaacg ggcctgctgc cggctctgcg gcctcttccg 7080

cgtcttcgcc ttcgccctca gacgagtcgg atctcccttt gggccgcctc cccgcctgtc 7140

tagcttgact gactgagata cagcgtacct tcagctcaca gacatgataa gatacattga 7200

tgagtttgga caaaccacaa ctagaatgca gtgaaaaaaa tgctttattt gtgaaatttg 7260

tgatgctatt gctttatttg taaccattat aagctgcaat aaacaagtt 7309

<210> 51

<211> 61

<212> RNA

<213> Artificial Sequence

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

<400> 51

acuccucccc auccucuccc ucugucccuc ugucccucug acccugcacu gucccagcac 60

c 61

<210> 52

<211> 118

<212> RNA

<213> Artificial Sequence

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

<400> 52

caggacacag ccuuggauca ggacagagac uuggggggcca uccugccccu ccaacccgac 60

auguguaccu cagcuuuuuc ccucacuugc aucaauaaag cuucuguguu uggaacag 118

<210> 53

<211> 76

<212> RNA

<213> Artificial Sequence

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

<400> 53

ggaagguuuu ucuuuuccug aggcgaaagu cucagguuuu gcuuuuuggc cuuucuuaaa 60

aaaaaaaaaa gcaaaa 76

<210> 54

<211> 93

<212> RNA

<213
> Artificial Sequence

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

<400> 54

gaagguuuuu cuuuuccuga gaaaacaaca cguauuguuu ucucagguuu ugcuuuuugg 60

ccuuuuucua gcuuaaaaaa aaaaaaagca aaa 93

<210> 55

<211> 19

<212> DNA

<213> Unknown

<220><223> Description of Unknown:
Transposon end sequence

<400> 55

agatgtgtat aagagacag 19

<210> 56

<211> 19

<212> DNA

<213> Unknown

<220><223> Description of Unknown:
Transposon end sequence

<400> 56

ctgtctctta tacacatct 19

<210> 57

<211> 338

<212> PRT

<213> Artificial Sequence

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

<400> 57

Met Gly Lys Lys Gln Asn Arg Lys Thr Gly Asn Ser Lys Thr Gln Ser
1 5 10 15

Ala Ser Pro Pro Pro Lys Glu Arg Ser Ser Ser Pro Ala Thr Glu Gln

20 25 30

Ser Trp Met Glu Asn Asp Phe Asp Glu Leu Arg Glu Glu Gly Phe Arg
35 40 45

Arg Ser Asn Tyr Ser Glu Leu Arg Glu Asp Ile Gln Thr Lys Gly Lys
50 55 60

Glu Val Glu Asn Phe Glu Lys Asn Leu Glu Glu Cys Ile Thr Arg Ile

65 70 75 80

Thr Asn Thr Glu Lys Cys Leu Lys Glu Leu Met Glu Leu Lys Thr Lys

85 90 95

Ala Arg Glu Leu Arg Glu Glu Cys Arg Ser Leu Arg Ser Arg Cys Asp

100 105 110

Gln Leu Glu Glu Arg Val Ser Ala Met Glu Asp Glu Met Asn Glu Met

115 120 125

Lys Arg Glu Gly Lys Phe Arg Glu Lys Arg Ile Lys Arg Asn Glu Gln

130 135 140

Ser Leu Gln Glu Ile Trp Asp Tyr Val Lys Arg Pro Asn Leu Arg Leu

145 150 155 160

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

165 170 175

Thr Leu Gln Asp Ile Ile Gln Glu Asn Phe Pro Asn Leu Ala Arg Gln

180 185 190

Ala Asn Val Gln Ile Gln Glu Ile Gln Arg Thr Pro Gln Arg Tyr Ser

195 200 205

Ser Arg Arg Ala Thr Pro Arg His Ile Ile Val Arg Phe Thr Lys Val

210 215 220

Glu Met Lys Glu Lys Met Leu Arg Ala Ala Arg Glu Lys Gly Arg Val

225 230 235 240

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

245 250 255

Thr Leu Gln Ala Arg Arg Glu Trp Gly Pro Ile Phe Asn Ile Leu Lys

260 265 270

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

275 280 285

Ile Ser Glu Gly Glu Ile Lys Tyr Phe Ile Asp Lys Gln Met Leu Arg

290 295 300

Asp Phe Val Thr Thr Arg Pro Ala Leu Lys Glu Leu Leu Lys Glu Ala

305 310 315 320

Leu Asn Met Glu Arg Asn Asn Arg Tyr Gln Pro Leu Gln Asn His Ala
325 330 335

Lys Met

<210> 58

<211> 1017

<212

> DNA

<213> Artificial Sequence

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

<400> 58

atgggcaaga agcaaaatcg caagacgggg aattccaaga cacaatccgc tagcccacca 60

cctaaagagc gttctagctc ccctgctact gagcagtcct ggatggaaaa cgacttcgat 120

gaactccggg aagagggatt taggcgatcc aactattcag aactccgcga agatatccag 180

acaaagggga aggaagtcga gaatttcgag aagaacctcg aggagtgcat cacccgtatc 240

acaaacactg agaaatgtct caaagaactc atggaactta agacaaaagc cagggagctt 300

cgagaggagt gtcggagtct gagatccagg tgtgaccagc tcgaggagcg cgtgagcgcg 360

atggaagacg agatgaacga gatgaaaaga gagggcaaat tcagggagaa gcgcattaag 420

aggaacgaac agagtctgca ggagatttgg gattacgtca agaggcctaa cctgcggttg 480

atcggcgtcc ccgagagcga cgtagaaaac gggactaaac tggagaatac acttcaagac 540

atcattcaag aaaattttcc aaacctggct cggcaagcta atgtgcaaat ccaagagatc 600

caacgcacac cccagcggta tagctctcgg cgtgccaccc ctaggcatat tatcgtgcgc 660

tttactaagg tggagatgaa agagaagatg ctgcgagccg ctcgggaaaa gggaagggtg 720

actttgaagg gcaaacctat tcggctgacg gttgacctta gcgccgagac actccaggca 780

cgccgggaat ggggccccat ctttaatatc ctgaaggaga agaacttcca gccacgaatc 840

tcttaccctg caaagttgag ttttatctcc gagggtgaga ttaagtattt catcgataaa 900

cagatgctgc gagacttcgt gacaactcgc ccagctctca aggaactgct caaagaggct 960

cttaatatgg agcgcaataa tagatatcaa cccttgcaga accacgcaaa gatgtga 1017

<210> 59

<211> 1275

<212> PRT

<213> Artificial Sequence

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

<400> 59

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

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

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

Phe Phe Ser Ala Pro His His Thr Tyr Ser Lys Ile Asp His Ile Val
195 200 205
Gly Ser Lys Ala Leu Leu Ser Lys Cys Lys Arg Thr Glu Ile Ile Thr
210 215 220

Asn Tyr Leu Ser Asp His Ser Ala Ile Lys Leu Glu Leu Arg Ile Lys
225 230 235 240
Asn Leu Thr Gln Ser Arg Ser Thr Thr Trp Lys Leu Asn Asn Leu Leu
245 250 255

Leu Asn Asp Tyr Trp Val His Asn Glu Met Lys Ala Glu Ile Lys Met
260 265 270
Phe Phe Glu Thr Asn Glu Asn Lys Asp Thr Thr Tyr Gln Asn Leu Trp
275 280 285
Asp Ala Phe Lys Ala Val Cys Arg Gly Lys Phe Ile Ala Leu Asn Ala
290 295 300
Tyr Lys Arg Lys Gln Glu Arg Ser Lys Ile Asp Thr Leu Thr Ser Gln
305 310 315 320

Leu Lys Glu Leu Glu Lys Gln Glu Gln Thr His Ser Lys Ala Ser Arg
325 330 335
Arg Gln Glu Ile Thr Lys Ile Arg Ala Glu Leu Lys Glu Ile Glu Thr
340 345 350
Gln Lys Thr Leu Gln Lys Ile Asn Glu Ser Arg Ser Trp Phe Phe Glu
355 360 365
Arg Ile Asn Lys Ile Asp Arg Pro Leu Ala Arg Leu Ile Lys Lys Lys
370 375 380

Arg Glu Lys Asn Gln Ile Asp Thr Ile Lys Asn Asp Lys Gly Asp Ile
385 390 395 400
Thr Thr Asp Pro Thr Glu Ile Gln Thr Thr Ile Arg Glu Tyr Tyr Lys
405 410 415
His Leu Tyr Ala Asn Lys Leu Glu Asn Leu Glu Glu Met Asp Thr Phe
420 425 430
Leu Asp Thr Tyr Thr Leu Pro Arg Leu Asn Gln Glu Glu Val Glu Ser
435 440 445

Leu Asn Arg Pro Ile Thr Gly Ser Glu Ile Val Ala Ile Ile Asn Ser
450 455 460
Leu Pro Thr Lys Lys Ser Pro Gly Pro Asp Gly Phe Thr Ala Glu Phe

465 470 475 480

Tyr Gln Arg Tyr Met Glu Glu Leu Val Pro Phe Leu Leu Lys Leu Phe

485 490 495

Gln Ser Ile Glu Lys Glu Gly Ile Leu Pro Asn Ser Phe Tyr Glu Ala

500 505 510

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

515 520 525

Asn Phe Arg Pro Ile Ser Leu Met Asn Ile Asp Ala Lys Ile Leu Asn

530 535 540

Lys Ile Leu Ala Asn Arg Ile Gln Gln His Ile Lys Lys Leu Ile His

545 550 555 560

His Asp Gln Val Gly Phe Ile Pro Gly Met Gln Gly Trp Phe Asn Ile

565 570 575

Arg Lys Ser Ile Asn Val Ile Gln His Ile Asn Arg Ala Lys Asp Lys

580 585 590

Asn His Met Ile Ile Ser Ile Asp Ala Glu Lys Ala Phe Asp Lys Ile

595 600 605

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

610 615 620

Thr Tyr Phe Lys Ile Ile Arg Ala Ile Tyr Asp Lys Pro Thr Ala Asn

625 630 635 640

Ile Ile Leu Asn Gly Gln Lys Leu Glu Ala Phe Pro Leu Lys Thr Gly

645 650 655

Thr Arg Gln Gly Cys Pro Leu Ser Pro Leu Leu Phe Asn Ile Val Leu

660 665 670

Glu Val Leu Ala Arg Ala Ile Arg Gln Glu Lys Glu Ile Lys Gly Ile

675 680 685

Gln Leu Gly Lys Glu Glu Val Lys Leu Ser Leu Phe Ala Asp Asp Met

690 695 700

Ile Val Tyr Leu Glu Asn Pro Ile Val Ser Ala Gln Asn Leu Leu Lys

705 710 715 720

Leu Ile Ser Asn Phe Ser Lys Val Ser Gly Tyr Lys Ile Asn Val Gln
725 730 735
Lys Ser Gln Ala Phe Leu Tyr Thr Asn Asn Arg Gln Thr Glu Ser Gln
740 745 750
Ile Met Gly Glu Leu Pro Phe Val Ile Ala Ser Lys Arg Ile Lys Tyr
755 760 765

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

Pro Met Thr Phe Phe Thr Glu Leu Glu Lys Thr Thr Leu Lys Phe Ile
835 840 845
Trp Asn Gln Lys Arg Ala Arg Ile Ala Lys Ser Ile Leu Ser Gln Lys
850 855 860
Asn Lys Ala Gly Gly Ile Thr Leu Pro Asp Phe Lys Leu Tyr Tyr Lys
865 870 875 880
Ala Thr Val Thr Lys Thr Ala Trp Tyr Trp Tyr Gln Asn Arg Asp Ile
885 890 895

Asp Gln Trp Asn Arg Thr Glu Pro Ser Glu Ile Met Pro His Ile Tyr
900 905 910
Asn Tyr Leu Ile Phe Asp Lys Pro Glu Lys Asn Lys Gln Trp Gly Lys
915 920 925
Asp Ser Leu Phe Asn Lys Trp Cys Trp Glu Asn Trp Leu Ala Ile Cys
930 935 940
Arg Lys Leu Lys Leu Asp Pro Phe Leu Thr Pro Tyr Thr Lys Ile Asn
945 950 955 960

Ser Arg Trp Ile Lys Asp Leu Asn Val Lys Pro Lys Thr Ile Lys Thr

965 970 975

Leu Glu Glu Asn Leu Gly Ile Thr Ile Gln Asp Ile Gly Val Gly Lys

980 985 990

Asp Phe Met Ser Lys Thr Pro Lys Ala Met Ala Thr Lys Asp Lys Ile

995 1000 1005

Asp Lys Trp Asp Leu Ile Lys Leu Lys Ser Phe Cys Thr Ala Lys

1010 1015 1020

Glu Thr Thr Ile Arg Val Asn Arg Gln Pro Thr Thr Trp Glu Lys

1025 1030 1035

Ile Phe Ala Thr Tyr Ser Ser Asp Lys Gly Leu Ile Ser Arg Ile

1040 1045 1050

Tyr Asn Glu Leu Lys Gln Ile Tyr Lys Lys Lys Thr Asn Asn Pro

1055 1060 1065

Ile Lys Lys Trp Ala Lys Asp Met Asn Arg His Phe Ser Lys Glu

1070 1075 1080

Asp Ile Tyr Ala Ala Lys Lys His Met Lys Lys Cys Ser Ser Ser

1085 1090 1095

Leu Ala Ile Arg Glu Met Gln Ile Lys Thr Thr Met Arg Tyr His

1100 1105 1110

Leu Thr Pro Val Arg Met Ala Ile Ile Lys Lys Ser Gly Asn Asn

1115 1120 1125

Arg Cys Trp Arg Gly Cys Gly Glu Ile Gly Thr Leu Leu His Cys

1130 1135 1140

Trp Trp Asp Cys Lys Leu Val Gln Pro Leu Trp Lys Ser Val Trp

1145 1150 1155

Arg Phe Leu Arg Asp Leu Glu Leu Glu Ile Pro Phe Asp Pro Ala

1160 1165 1170

Ile Pro Leu Leu Gly Ile Tyr Pro Asn Glu Tyr Lys Ser Cys Cys

1175 1180 1185

Tyr Lys Asp Thr Cys Thr Arg Met Phe Ile Ala Ala Leu Phe Thr

1190 1195 1200

Ile Ala Lys Thr Trp Asn Gln Pro Lys Cys Pro Thr Met Ile Asp
1205 1210 1215

Trp Ile Lys Lys Met Trp His Ile Tyr Thr Met Glu Tyr Tyr Ala
1220 1225 1230

Ala Ile Lys Asn Asp Glu Phe Ile Ser Phe Val Gly Thr Trp Met
1235 1240 1245

Lys Leu Glu Thr Ile Ile Leu Ser Lys Leu Ser Gln Glu Gln Lys
1250 1255 1260

Thr Lys His Arg Ile Phe Ser Leu Ile Gly Gly Asn
1265 1270 1275

<210> 60

<211> 3828

<212> DNA

<213> Artificial Sequence

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

<400> 60

atgaccggct ctaactcaca tatcaccatc cttacactta acattaacgg cctcaactca 60

gctatcaagc gccatcggct ggccagctgg atcaaatcac aggatccaag cgtttgttgc 120

atccaagaga cccacctgac ctgtagagat actcaccgcc tcaagatcaa gggatggcga 180

aagatttatc aggcgaacgg taagcagaag aaagccggag tcgcaattct ggtctcagac 240

aagacggatt tcaagcccac caaaattaag cgtgataagg aaggtcacta tattatggtg 300

aaaggcagca tacagcagga agaacttacc atattgaaca tctacgcgcc aaacaccggc 360

gcacctcgct ttatcaaaca ggtcctgtcc gatctgcagc gagatctgga ttctcatacg 420

ttgattatgg gtgatttcaa tacaccattg agcaccctgg atcgcagcac caggcaaaag 480

gtaaataaag acacgcaaga gctcaatagc gcactgcatc aggcagatct cattgatatt 540

tatcgcactc ttcatcctaa gagtaccgag tacacattct tcagcgcccc acatcataca 600

tactcaaaga tcgatcatat cgtcggctca aaggctctgc tgtcaaagtg caagcgcaca 660

gagataatta caaattacct gtcagatcat agcgcgatca agctcgagct gagaatcaag 720

aacctgaccc agagccggag taccacttgg aagcttaata acctgctgct caacgattat 780

tgggtccaca atgagatgaa ggcagagatt aaaatgttct tcgaaacaaa tgagaataag 840

gatactacct atcaaaacct ttgggatgcc tttaaggccg tctgcagagg caagttcatc 900

gccctcaacg cctataaaag aaaacaagag agatctaaga tcgatactct cacctctcag 960

ctgaaggagt tggagaaaca ggaacagacc cactccaagg cgtcaagacg gcaggagatc 1020

acaaagattc gcgccgagtt gaaagagatc gaaacccaaa agactcttca gaaaattaac 1080

gagtctcgta gttggttctt cgagcggatt aataagatag acagacctct ggcacgactg 1140

attaagaaga agcgcgaaaa gaaccagatt gataccatca agaacgacaa gggcgacatc 1200

actactgacc cgaccgagat ccagaccact attcgggagt attataagca tttgtatgct 1260

aacaagcttg agaacctgga agagatggac acttttctgg atacctatac tctgccacgg 1320

cttaatcaag aggaagtcga gtccctcaac cgcccaatta caggaagcga gattgtggcc 1380

ataattaact ccctgccgac aaagaaatct cctggtccgg acgggtttac agctgagttt 1440

tatcaacggt atatggaaga gcttgtaccg tttctgctca agctctttca gtctatagaa 1500

aaggaaggca tcttgcccaa ttccttctac gaagcttcta taatacttat tcccaaacca 1560

ggacgcgata ccacaaagaa ggaaaacttc cggcccatta gtctcatgaa tatcgacgct 1620

aaaatattga acaagattct cgccaacaga atccaacaac atattaagaa attgatacat 1680

cacgaccagg tggggtttat acctggcatg cagggctggt ttaacatccg gaagagtatt 1740

aacgtcattc aacacattaa tagagctaag gataagaatc atatgatcat ctctatagac 1800

gcggaaaagg cattcgataa gattcagcag ccatttatgc tcaagactct gaacaaactc 1860

ggcatcgacg gaacatattt taagattatt cgcgcaattt acgataagcc gactgctaac 1920

attatcctta acggccaaaa gctcgaggcc tttccgctca agactggaac ccgccaaggc 1980

tgtcccctct ccccgctttt gtttaatatt gtactcgagg tgctggctag ggctattcgt 2040

caagagaaag agattaaagg gatacagctc gggaaggaag aggtcaagct ttccttgttc 2100

gccgatgata tgattgtgta cctggagaat cctattgtgt ctgctcagaa ccttcttaaa 2160

cttatttcta actttagcaa ggtcagcggc tataagatta acgtccagaa atctcaggcc 2220

tttctgtaca caaataatcg acagaccgaa tcccagataa tgggtgagct tccgtttgtc 2280

atagccagca aaaggataaa gtatctcgga atccagctga cacgagacgt taaagatttg 2340

tttaaggaaa attacaagcc tctcctgaaa gagattaagg aagatactaa taagtggaag 2400

aatatcccct gttcatgggt tggcagaatc aacatagtga agatggcaat acttcctaaa 2460

gtgatatatc gctttaacgc catcccaatt aaactgccta tgaccttctt tacggagctc 2520

gagaaaacaa cccttaaatt tatatggaat caaaagagag caagaatagc gaagtccatc 2580

ttgagccaga agaataaggc cggtgggatt actttgcctg attttaagtt gtattataaa 2640

gccacagtaa ctaagacagc ctggtattgg tatcagaata gagacatcga ccagtggaat 2700

cggaccgaac catcagagat aatgccccac atctataatt accttatatt cgataagcca 2760

gaaaagaata aacagtgggg caaagacagc ctcttcaaca agtggtgttg ggagaattgg 2820
ctggccatat gccggaaact caagctcgac ccctttctta caccctacac taaaatcaac 2880
agtaggtgga tcaaggactt gaatgtcaag ccaaagacta taaagacact ggaagagaat 2940
cttgggatca caatacaaga tataggcgtc ggcaaagatt ttatgtcaaa gacgcccaag 3000
gccatggcca ctaaggataa gattgataag tgggacctta ttaagctcaa aagcttctgt 3060
actgccaagg agaccacgat cagagttaat aggcagccca ctacatggga aaagattttc 3120
gccacttatt catcagataa ggggttgata agcagaatat ataacgagct gaagcagatc 3180

tacaagaaga aaacgaataa tcccatcaag aagtgggcaa aagatatgaa caggcatttt 3240
agcaaagagg atatctacgc cgcgaagaag catatgaaga agtgtagttc aagcttggcc 3300
attcgtgaga tgcagattaa gacgaccatg cgataccacc ttaccccagt gaggatggca 3360
attatcaaga aatctggcaa taatagatgt tggcggggct gtggcgagat tggcaccctg 3420
ctccattgct ggtgggattg caagctggtg cagccgcttt ggaaatcagt ctggcgcttt 3480
ctgagggacc tcgagcttga gattcccttc gatcccgcaa ttcccttgct cggaatctat 3540
cctaacgaat acaagagctg ttgttacaag gatacgtgta cccggatgtt catcgcggcc 3600

ttgtttacga tagctaagac gtggaatcag cctaagtgcc ccacaatgat cgattggatc 3660
aagaaaatgt ggcatattta taccatggag tattacgcag caattaagaa tgacgaattt 3720
atttccttcg ttgggacctg gatgaagctg gagactatta ttctgagcaa gctgtctcag 3780
gagcaaaaga caaagcatag aatcttctct ctcattggtg gtaactaa 3828

<210> 61
<211> 1290
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 61

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

35 40 45

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

50 55 60

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

65 70 75 80

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

85 90 95

Ile Met Ile Gln Gly Ser Ile Gln Glu Glu Asp Ile Thr Ile Ile Asn

100 105 110

Ile Tyr Ala Pro Asn Ile Gly Ala Pro Gln Tyr Ile Arg Gln Leu Leu

115 120 125

Thr Ala Ile Lys Glu Glu Ile Asp Ser Asn Thr Ile Ile Val Gly Asp

130 135 140

Phe Asn Thr Ser Leu Thr Pro Met Asp Arg Ser Ser Lys Met Lys Ile

145 150 155 160

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

165 170 175

Ile Asp Ile Tyr Arg Thr Phe His Pro Lys Thr Ala Asp Tyr Thr Phe

180 185 190

Phe Ser Ser Ala His Gly Thr Phe Ser Arg Ile Asp His Ile Leu Gly

195 200 205

His Lys Ser Ser Leu Ser Lys Phe Lys Lys Ile Glu Ile Ile Ser Ser

210 215 220

Ile Phe Ser Asp His Asn Ala Met Arg Leu Glu Met Asn His Arg Glu

225 230 235 240

Lys Asn Val Lys Lys Thr Asn Thr Trp Arg Leu Asn Asn Thr Leu Leu

245 250 255

Asn Asn Gln Glu Ile Thr Glu Glu Ile Lys Gln Glu Ile Lys Lys Tyr

260 265 270

Leu Glu Thr Asn Asp Asn Glu Asn Thr Thr Thr Gln Asn Leu Trp Asp

275 280 285

Ala Ala Lys Ala Val Leu Arg Gly Lys Phe Ile Ala Ile Gln Ala Tyr
290 295 300
Leu Lys Lys Gln Glu Lys Ser Gln Val Asn Asn Leu Thr Leu His Leu
305 310 315 320
Lys Lys Leu Glu Lys Glu Glu Gln Thr Lys Pro Lys Val Ser Arg Arg

325 330 335
Lys Glu Ile Ile Lys Ile Arg Ala Glu Ile Asn Glu Ile Glu Thr Lys
340 345 350
Lys Thr Ile Ala Lys Ile Asn Lys Thr Lys Ser Trp Phe Phe Glu Lys
355 360 365
Ile Asn Lys Ile Asp Lys Pro Leu Ala Arg Leu Ile Lys Lys Lys Arg
370 375 380
Glu Arg Thr Gln Ile Asn Lys Ile Arg Asn Glu Lys Gly Glu Val Thr

385 390 395 400
Thr Asp Thr Ala Glu Ile Gln Asn Ile Leu Arg Asp Tyr Tyr Lys Gln
405 410 415
Leu Tyr Ala Asn Lys Met Asp Asn Leu Glu Glu Met Asp Lys Phe Leu
420 425 430
Glu Arg Tyr Asn Leu Pro Arg Leu Asn Gln Glu Glu Thr Glu Asn Ile
435 440 445
Asn Arg Pro Ile Thr Ser Asn Glu Ile Glu Thr Val Ile Lys Asn Leu

450 455 460
Pro Thr Asn Lys Ser Pro Gly Pro Asp Gly Phe Thr Gly Glu Phe Tyr
465 470 475 480
Gln Thr Phe Arg Glu Glu Leu Thr Pro Ile Leu Leu Lys Leu Phe Gln
485 490 495
Lys Ile Ala Glu Glu Gly Thr Leu Pro Asn Ser Phe Tyr Glu Ala Thr
500 505 510
Ile Thr Leu Ile Pro Lys Pro Asp Lys Asp Thr Thr Lys Lys Glu Asn

515 520 525
Tyr Arg Pro Ile Ser Leu Met Asn Ile Asp Ala Lys Ile Leu Asn Lys

530 535 540

Ile Leu Ala Asn Arg Ile Gln Gln His Ile Lys Arg Ile Ile His His

545 550 555 560

Asp Gln Val Gly Phe Ile Pro Gly Met Gln Gly Phe Phe Asn Ile Arg

565 570 575

Lys Ser Ile Asn Val Ile His His Ile Asn Lys Leu Lys Lys Lys Asn

580 585 590

His Met Ile Ile Ser Ile Asp Ala Glu Lys Ala Phe Asp Lys Ile Gln

595 600 605

His Pro Phe Met Ile Lys Thr Leu Gln Lys Val Gly Ile Glu Gly Thr

610 615 620

Tyr Leu Asn Ile Ile Lys Ala Ile Tyr Asp Lys Pro Thr Ala Asn Ile

625 630 635 640

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

645 650 655

Arg Gln Gly Cys Pro Leu Ser Pro Leu Leu Phe Asn Ile Val Leu Glu

660 665 670

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

675 680 685

Ile Gly Lys Glu Glu Val Lys Leu Ser Leu Phe Ala Asp Asp Met Ile

690 695 700

Leu Tyr Ile Glu Asn Pro Lys Thr Ala Thr Arg Lys Leu Leu Glu Leu

705 710 715 720

Ile Asn Glu Tyr Gly Lys Val Ala Gly Tyr Lys Ile Asn Ala Gln Lys

725 730 735

Ser Leu Ala Phe Leu Tyr Thr Asn Asp Glu Lys Ser Glu Arg Glu Ile

740 745 750

Met Glu Thr Leu Pro Phe Thr Ile Ala Thr Lys Arg Ile Lys Tyr Leu

755 760 765

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

770 775 780

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

835 840 845
Arg His Lys Arg Pro Arg Ile Ala Lys Ala Val Leu Arg Gln Lys Asn
850 855 860
Gly Ala Gly Gly Ile Arg Leu Pro Asp Phe Arg Leu Tyr Tyr Lys Ala
865 870 875 880
Thr Val Ile Lys Thr Ile Trp Tyr Trp His Lys Asn Arg Asn Ile Asp
885 890 895
Gln Trp Asn Lys Ile Glu Ser Pro Glu Ile Asn Pro Arg Thr Tyr Gly

900 905 910
Gln Leu Ile Tyr Asp Lys Gly Gly Lys Asp Ile Gln Trp Arg Lys Asp
915 920 925
Ser Leu Phe Asn Lys Trp Cys Trp Glu Asn Trp Thr Ala Thr Cys Lys
930 935 940
Arg Met Lys Leu Glu Tyr Ser Leu Thr Pro Tyr Thr Lys Ile Asn Ser
945 950 955 960
Lys Trp Ile Arg Asp Leu Asn Ile Arg Leu Asp Thr Ile Lys Leu Leu

965 970 975
Glu Glu Asn Ile Gly Arg Thr Leu Phe Asp Ile Asn His Ser Lys Ile
980 985 990
Phe Phe Asp Pro Pro Pro Arg Val Met Glu Ile Lys Thr Lys Ile Asn
995 1000 1005
Lys Trp Asp Leu Met Lys Leu Gln Ser Phe Cys Thr Ala Lys Glu
1010 1015 1020
Thr Ile Asn Lys Thr Lys Arg Gln Pro Ser Glu Trp Glu Lys Ile

1025 1030 1035

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

1040 1045 1050

Lys Gln Leu Ile Gln Leu Asn Ile Lys Glu Thr Asn Thr Pro Ile

1055 1060 1065

Gln Lys Trp Ala Glu Asp Leu Asn Arg His Phe Ser Lys Glu Asp

1070 1075 1080

Ile Gln Thr Ala Thr Lys His Met Lys Arg Cys Ser Thr Ser Leu

1085 1090 1095

Ile Ile Arg Glu Met Gln Ile Lys Thr Thr Met Arg Tyr His Leu

1100 1105 1110

Thr Pro Val Arg Met Gly Ile Ile Arg Lys Ser Thr Asn Asn Lys

1115 1120 1125

Cys Trp Arg Gly Cys Gly Glu Lys Gly Thr Leu Leu His Cys Trp

1130 1135 1140

Trp Glu Cys Lys Leu Ile Gln Pro Leu Trp Arg Thr Ile Trp Arg

1145 1150 1155

Phe Leu Lys Lys Leu Lys Ile Glu Leu Pro Tyr Asp Pro Ala Ile

1160 1165 1170

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

1175 1180 1185

Thr Cys Thr Arg Met Phe Ile Ala Ala Leu Phe Thr Ile Ala Arg

1190 1195 1200

Ser Trp Lys Gln Pro Lys Cys Pro Ser Thr Asp Glu Trp Ile Lys

1205 1210 1215

Lys Met Trp Tyr Ile Tyr Thr Met Glu Tyr Tyr Ser Ala Ile Lys

1220 1225 1230

Arg Asn Glu Ile Gly Ser Phe Leu Glu Thr Trp Met Asp Leu Glu

1235 1240 1245

Thr Val Ile Gln Ser Glu Val Ser Gln Lys Glu Lys Asn Lys Tyr

1250 1255 1260

Arg Ile Leu Thr His Ile Cys Gly Thr Trp Lys Asn Gly Thr Asp

1265 1270 1275

Glu Pro Val Cys Arg Thr Glu Ile Glu Thr Gln Met

1280 1285 1290

<210> 62

<211> 3873

<212> DNA

<213> Artificial Sequence

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

<400> 62

atggtcatag gaacatacat atcgataatt accttaaacg tgaatggatt aaatgcccca 60

accaaaagac atagactggc tgaatggata caaaaacaag acccatatat atgctgtcta 120

caagagaccc acttcagacc tagggacaca tacagactga aagtgagggg atggaaaaag 180

atattccatg caaatggaaa tcaaaagaaa gctggagtag ctatactcat atcagataaa 240

atagacttta aaataaagaa tgttacaaga gacaaggaag gacactacat aatgatccag 300

ggatcaatcc aagaagaaga tataacaatt ataaatatat atgcacccaa cataggagca 360

cctcaataca taaggcaact gctaacagct ataaaagagg aaatcgacag taacacaata 420

atagtggggg actttaacac ctcacttaca ccaatggaca gatcatccaa aatgaaaata 480

aataaggaaa cagaagcttt aaatgacaca atagaccaga tagatttaat tgatatatat 540

aggacattcc atccaaaaac agcagattac acgttcttct caagtgcgca cggaacattc 600

tccaggatag atcacatctt gggtcacaaa tcaagcctca gtaaatttaa gaaaattgaa 660

atcatatcaa gcatcttttc tgaccacaac gctatgagat tagaaatgaa tcacagggaa 720

aaaaacgtaa aaaagacaaa cacatggagg ctaaacaata cgttactaaa taaccaagag 780

atcactgaag aaatcaaaca ggaaataaaa aaatacctag agacaaatga caatgaaaac 840

acgacgaccc aaaacctatg ggatgcagca aaagcggttc taagagggaa gtttatagct 900

atacaagcct acctaaagaa acaagaaaaa tctcaagtaa acaatctaac cttacaccta 960

aagaaactag agaaagaaga acaaacaaaa cccaaagtta gcagaaggaa agaaatcata 1020

aagatcagag cagaaataaa tgaaatagaa acaaagaaaa caatagcaaa gatcaataaa 1080

actaaaagtt ggttctttga gaagataaac aaaattgata agccattagc cagactcatc 1140

aagaaaaaga gggagaggac tcaaatcaat aaaatcagaa atgaaaaagg agaagttaca 1200

acagacaccg cagaaataca aaacatccta agagactact acaagcaact ttatgccaat 1260

aaaatggaca acctggaaga aatggacaaa ttcttagaaa ggtataacct tccaagactg 1320

aaccaggaag aaacagaaaa tatcaacaga ccaatcacaa gtaatgaaat tgaaactgtg 1380

attaaaaatc ttccaacaaa caaaagtcca ggaccagatg gcttcacagg tgaattctat 1440

caaacattta gagaagagct aacacccatc cttctcaaac tcttccaaaa aattgcagaa 1500

gaaggaacac tcccaaactc attctatgag gccaccatca ccctgatacc aaaaccagac 1560

aaagacacta caaaaaaaga aaattacaga ccaatatcac tgatgaatat agatgcaaaa 1620

atcctcaaca aaatactagc aaacagaatc caacaacaca ttaaaaggat catacaccac 1680

gatcaagtgg gatttatccc agggatgcaa ggattcttca atatacgcaa atcaatcaat 1740

gtgatacacc atattaacaa attgaagaag aaaaaccata tgatcatctc aatagatgca 1800

gaaaaagctt ttgacaaaat tcaacaccca tttatgataa aaactctcca gaaagtgggc 1860

atagagggaa cctacctcaa cataataaag gccatatatg acaaacccac agcaaacatc 1920

attctcaatg gtgaaaaact gaaagcattt cctctaagat caggaacgag acaaggatgt 1980

ccactctcac cactattatt caacatagtt ctggaagtcc tagccacggc aatcagagaa 2040

gaaaaagaaa taaaaggaat acaaattgga aaagaagaag taaaactgtc actgtttgcg 2100

gatgacatga tactatacat agagaatcct aaaactgcca ccagaaaact gctagagcta 2160

attaatgaat atggtaaagt tgcaggttac aaaattaatg cacagaaatc tcttgcattc 2220

ctatacacta atgatgaaaa atctgaaaga gaaattatgg aaacactccc atttaccatt 2280

gcaacaaaaa gaataaaata cctaggaata aacctaccta aggagacaaa agacctgtat 2340

gcagaaaact ataagacact gatgaaagaa attaaagatg ataccaacag atggagagat 2400

ataccatgtt cttggattgg aagaatcaac attgtgaaaa tgagtatact acccaaagca 2460

atctacagat tcaatgcaat ccctatcaaa ttaccaatgg catttttac ggagctagaa 2520

caaatcatct taaaatttgt atggagacac aaaagacccc gaatagccaa agcagtcttg 2580

aggcaaaaaa atggagctgg aggaatcaga ctccctgact tcagactata ctacaaagct 2640

acagtaatca agacaatatg gtactggcac aaaaacagaa acatagatca atggaacaag 2700

atagaaagcc cagagattaa cccacgcacc tatggtcaac taatctatga caaaggaggc 2760

aaagatatac aatggagaaa agacagtctc ttcaataagt ggtgctggga aaactggaca 2820

gccacatgta aaagaatgaa attagaatac tccctaacac catacacaaa aataaactca 2880

aaatggatta gagacctaaa tataagactg gacactataa aactcttaga ggaaaacata 2940

ggaagaacac tctttgacat aaatcacagc aagatctttt tcgatccacc tcctagagta 3000

atggaaataa aaacaaaaat aaacaagtgg gacctaatga aacttcaaag cttttgcaca 3060

gcaaaggaaa ccataaacaa gacgaaaaga caaccctcag aatgggagaa aatatttgca 3120

aatgaatcaa cggacaaagg attaatctcc aaaatatata aacagctcat tcagctcaat 3180

atcaaagaaa caaacacccc aatccaaaaa tgggcagaag acctaaatag acatttctcc 3240

aaagaagaca tacagacggc cacgaagcac atgaaaagat gctcaacatc actaattatt 3300

agagaaatgc aaatcaaaac tacaatgagg tatcacctca ctcctgttag aatgggcatc 3360

atcagaaaat ctacaaacaa caaatgctgg agagggtgtg gagaaaaggg aaccctcttg 3420

cactgttggt gggaatgtaa attgatacag ccactatgga gaacaatatg gaggttcctt 3480

aaaaaactaa aaatagaatt accatatgac ccagcaatcc cactactggg catataccca 3540

gagaaaaccg taattcaaaa agacacatgc acccgaatgt tcattgcagc actatttaca 3600

atagccaggt catggaagca acctaaatgc ccatcgacag acgaatggat aaagaagatg 3660

tggtacatat atacaatgga atattactca gccataaaaa ggaacgaaat tgggtcattt 3720

ttagagacgt ggatggatct agagactgtc atacagagtg aagtaagtca gaaagagaaa 3780

aacaaatatc gtatattaac gcatatatgt ggaacctgga aaaatggtac agatgaaccg 3840

gtctgcagga cagaaattga gacacaaatg taa 3873

<210> 63

<211> 9

<212> PRT

<213> Artificial Sequence

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

<400> 63

Pro Ala Ala Lys Arg Val Lys Leu Asp

1 5

<210> 64

<211> 8

<212> PRT

<213> Artificial Sequence

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

<400> 64

Asp Tyr Lys Asp Asp Asp Asp Lys

1 5

<210> 65

<211> 234

<212> PRT

<213> Artificial Sequence

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

<400> 65

Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr Gly
1 5 10 15

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

Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser Val
35 40 45

Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu Val
50 55 60

Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile Pro
65 70 75 80

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

Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala Asn
100 105 110

Ser Gly Ile Tyr Ala Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val
115 120 125

Asp Asn Gly Gly Thr Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala
130 135 140

Asn Gly Ile Ala Glu Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr
145 150 155 160

Lys Val Thr Cys Ser Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr
165 170 175

Thr Ile Lys Val Glu Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn
180 185 190

Met Glu Leu Thr Ile Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu
195 200 205

Ile Val Lys Ala Met Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro
210 215 220

Ser Ala Ile Ala Ala Asn Ser Gly Ile Tyr
225 230

<210> 66

<211> 9

<212> PRT

<213> Unknown

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

<400> 66

Leu Ala Gly Leu Ile Asp Ala Asp Gly

1 5

<210> 67

<211> 45

<212> PRT

<213> Artificial Sequence

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

<400> 67

Leu Tyr Cys Arg Arg Leu Lys Ile Gln Val Arg Lys Ala Ala Ile Thr
1 5 10 15

Ser Tyr Glu Lys Ser Asp Gly Val Tyr Thr Gly Leu Ser Thr Arg Asn
20 25 30

Gln Glu Thr Tyr Glu Thr Leu Lys His Glu Lys Pro Pro
35 40 45

<210> 68

<211> 61

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 68

Lys Val Ala Lys Lys Pro Thr Asn Lys Ala Pro His Pro Lys Gln Glu
1 5 10 15

Pro Gln Glu Ile Asn Phe Pro Asp Asp Leu Pro Gly Ser Asn Thr Ala
20 25 30

Ala Pro Val Gln Glu Thr Leu His Gly Cys Gln Pro Val Thr Gln Glu
35 40 45

Asp Gly Lys Glu Ser Arg Ile Ser Val Gln Glu Arg Gln
50 55 60

<210> 69

<211> 415

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 69

Glu Ile Gln Leu Val Gln Ser Gly Gly Gly Leu Val Lys Pro Gly Gly
1 5 10 15

Ser Val Arg Ile Ser Cys Ala Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Met
35 40 45

Gly Trp Ile Asn Thr His Thr Gly Glu Pro Thr Tyr Ala Asp Ser Phe
50 55 60

Lys Gly Arg Phe Thr Phe Ser Leu Asp Asp Ser Lys Asn Thr Ala Tyr
65 70 75 80

Leu Gln Ile Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Phe Cys
85 90 95

Thr Arg Arg Gly Tyr Asp Trp Tyr Phe Asp Val Trp Gly Gln Gly Thr
100 105 110

Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

115 120 125

Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu

130 135 140

Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln

145 150 155 160

Asp Ile Asn Ser Tyr Leu Ser Trp Phe Gln Gln Lys Pro Gly Lys Ala

165 170 175

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

180 185 190

Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Thr Leu Thr Ile

195 200 205

Ser Ser Leu Gln Tyr Glu Asp Phe Gly Ile Tyr Tyr Cys Gln Gln Tyr

210 215 220

Asp Glu Ser Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys

225 230 235 240

Ser Gly Gly Gly Gly Ser Gly Ala Leu Ser Asn Ser Ile Met Tyr Phe

245 250 255

Ser His Phe Val Pro Val Phe Leu Pro Ala Lys Pro Thr Thr Thr Pro

260 265 270

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

275 280 285

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

290 295 300

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

305 310 315 320

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

325 330 335

Lys Ile Gln Val Arg Lys Ala Ala Ile Thr Ser Tyr Glu Lys Ser Asp

340 345 350

Gly Val Tyr Thr Gly Leu Ser Thr Arg Asn Gln Glu Thr Tyr Glu Thr

355 360 365

Leu Lys His Glu Lys Pro Pro Gln Gly Ser Gly Ser Tyr Glu Asp Met
370 375 380
Arg Gly Ile Leu Tyr Ala Ala Pro Gln Leu Arg Ser Ile Arg Gly Gln
385 390 395 400
Pro Gly Pro Asn His Glu Glu Asp Ala Asp Ser Tyr Glu Asn Met
405 410 415

<210> 70
<211> 437
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic polypeptide
<400> 70
Glu Ile Gln Leu Val Gln Ser Gly Gly Gly Leu Val Lys Pro Gly Gly
1 5 10 15
Ser Val Arg Ile Ser Cys Ala Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30
Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Met
35 40 45

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

Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
115 120 125
Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu
130 135 140
Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln

145 150 155 160

Asp Ile Asn Ser Tyr Leu Ser Trp Phe Gln Gln Lys Pro Gly Lys Ala

165 170 175

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

180 185 190

Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Thr Leu Thr Ile

195 200 205

Ser Ser Leu Gln Tyr Glu Asp Phe Gly Ile Tyr Tyr Cys Gln Gln Tyr

210 215 220

Asp Glu Ser Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys

225 230 235 240

Ser Gly Gly Gly Gly Ser Gly Ala Leu Ser Asn Ser Ile Met Tyr Phe

245 250 255

Ser His Phe Val Pro Val Phe Leu Pro Ala Lys Pro Thr Thr Thr Pro

260 265 270

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

275 280 285

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

290 295 300

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

305 310 315 320

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

325 330 335

Ile Gln Val Arg Lys Ala Ala Ile Thr Ser Tyr Glu Lys Ser Asp Gly

340 345 350

Val Tyr Thr Gly Leu Ser Thr Arg Asn Gln Glu Thr Tyr Glu Thr Leu

355 360 365

Lys His Glu Lys Pro Pro Gln Lys Lys Val Ala Lys Lys Pro Thr Asn

370 375 380

Lys Ala Pro His Pro Lys Gln Glu Pro Gln Glu Ile Asn Phe Pro Asp

385 390 395 400

Asp Leu Pro Gly Ser Asn Thr Ala Ala Pro Val Gln Glu Thr Leu His
405 410 415
Gly Cys Gln Pro Val Thr Gln Glu Asp Gly Lys Glu Ser Arg Ile Ser
420 425 430

Val Gln Glu Arg Gln
435

<210> 71
<211> 438
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic polypeptide
<400> 71
Glu Ile Gln Leu Val Gln Ser Gly Gly Gly Leu Val Lys Pro Gly Gly
1 5 10 15
Ser Val Arg Ile Ser Cys Ala Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30
Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Met

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

100 105 110
Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
115 120 125
Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu
130 135 140
Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln

145 150 155 160

Asp Ile Asn Ser Tyr Leu Ser Trp Phe Gln Gln Lys Pro Gly Lys Ala

165 170 175

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

180 185 190

Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Thr Leu Thr Ile

195 200 205

Ser Ser Leu Gln Tyr Glu Asp Phe Gly Ile Tyr Tyr Cys Gln Gln Tyr

210 215 220

Asp Glu Ser Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys

225 230 235 240

Ser Gly Gly Gly Gly Ser Gly Ala Leu Ser Asn Ser Ile Met Tyr Phe

245 250 255

Ser His Phe Val Pro Val Phe Leu Pro Ala Lys Pro Thr Thr Thr Pro

260 265 270

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

275 280 285

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

290 295 300

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

305 310 315 320

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

325 330 335

Lys Ile Gln Val Arg Lys Ala Ala Ile Thr Ser Tyr Glu Lys Ser Asp

340 345 350

Gly Val Tyr Thr Gly Leu Ser Thr Arg Asn Gln Glu Thr Tyr Glu Thr

355 360 365

Leu Lys His Glu Lys Pro Pro Gln Lys Lys Val Ala Lys Lys Pro Thr

370 375 380

Asn Lys Ala Pro His Pro Lys Gln Glu Pro Gln Glu Ile Asn Phe Pro

385 390 395 400

Asp Asp Leu Pro Gly Ser Asn Thr Ala Ala Pro Val Gln Glu Thr Leu
405 410 415
His Gly Cys Gln Pro Val Thr Gln Glu Asp Gly Lys Glu Ser Arg Ile

420 425 430
Ser Val Gln Glu Arg Gln
435

<210> 72
<211> 44
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic polypeptide
<400> 72
Met Arg Asn Lys Lys Ile Leu Lys Glu Asp Glu Leu Leu Ser Glu Thr
1 5 10 15
Gln Gln Ala Ala Phe His Gln Ile Ala Met Glu Pro Phe Glu Ile Asn
20 25 30

Val Pro Lys Pro Lys Arg Arg Asn Gly Val Asn Phe
35 40

<210> 73
<211> 50
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic polypeptide
<400> 73
Met Glu Gln Trp Asp His Phe His Asn Gln Gln Glu Asp Thr Asp Ser
1 5 10 15
Cys Ser Glu Ser Val Lys Phe Asp Ala Arg Ser Met Thr Ala Leu Leu
20 25 30

Pro Pro Asn Pro Lys Asn Ser Pro Ser Leu Gln Glu Lys Leu Lys Ser
35 40 45

Phe Lys

50

<210> 74

<211> 203

<212> RNA

<213> Artificial Sequence

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

<220><221> modified_base

<222> (66)..(185)

<223> a, c, u or g

<220>

<223> See specification as filed for detailed description of substitutions and preferred embodiments

<400> 74

gcuggguuuu uccuuguucg caccggacac cuccagugac cagacggcaa gguuuuuauc 60

ccagunnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 120

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 180

nnnnnaaaaa aaaaaaaaaa aaa 203

<210> 75

<211> 33

<212> PRT

<213> Artificial Sequence

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

<400> 75

Gly Ala Ala Pro Ala Ala Ala Pro Ala Lys Gln Glu Ala Ala Ala Pro

1 5 10 15

Ala Pro Ala Ala Lys Ala Glu Ala Pro Ala Ala Ala Pro Ala Ala Lys

20 25 30

Ala

<210> 76

<211> 20

<212> PRT

<213> Artificial Sequence

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

<220><221> SITE

<222> (1)..(20)

<223> This sequence may encompass 1-4 "Gly Gly Gly Gly Ser" repeating units

<220>

<223> See specification as filed for detailed description of substitutions and preferred embodiments

<400> 76

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1 5 10 15

Gly Gly Gly Ser
20

<210> 77

<211> 6

<212> PRT

<213> Artificial Sequence

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

<400> 77

Gly Gly Gly Gly Gly Gly
1 5

<210> 78

<211> 8

<212> PRT

<213> Artificial Sequence

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

<400> 78

Gly Gly Gly Gly Gly Gly Gly Gly

1 5

<210> 79

<211> 20

<212> PRT

<213> Artificial Sequence

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

<220><221> SITE

<222> (1)..(20)

<223> This sequence may encompass 1-4 "Glu Ala Ala Ala Lys" repeating units

<220>

<223> See specification as filed for detailed description of substitutions and preferred embodiments

<400> 79

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu

1 5 10 15

Ala Ala Ala Lys

20

<210> 80

<211> 7

<212> PRT

<213> Simian virus 40

<400> 80

Pro Lys Lys Lys Arg Lys Val

1 5

<210> 81

<211> 21

<212> DNA

<213> Simian virus 40

<400> 81

ccaaagaaga agcggaaggt c 21

<210> 82

<211> 11

<212> PRT

<213> Artificial Sequence

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

<400> 82

Pro Lys Lys Lys Arg Lys Val Gly Gly Gly Ser

1 5 10

<210> 83

<211> 33

<212> DNA

<213> Artificial Sequence

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

<400> 83

ccaaagaaga agcggaaggt cggcggcggc agc 33

<210> 84

<211> 16

<212> PRT

<213> Unknown

<220><223> Description of Unknown: Nucleoplasmin bipartite NLS sequence

<400> 84

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys

1 5 10 15

<210> 85

<211> 48

<212> DNA

<213> Unknown

<220><223> Description of Unknown: Nucleoplasmin bipartite NLS sequence

<400> 85

aaaaggccgg cggccacgaa aaaggccggc caggcaaaaa agaaaaag 48

<210> 86

<211> 20

<212> PRT

<213> Artificial Sequence

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

<400> 86

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

Gly Gly Gly Ser
20

<210> 87

<211> 60

<212> DNA

<213> Artificial Sequence

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

<400> 87

aaaaggccgg cggccacgaa aaaggccggc caggcaaaaa agaaaaaggg cggcggcagc 60

<210> 88

<211> 20

<212> PRT

<213> Artificial Sequence

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

<400> 88

Gly Gly Gly Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala
1 5 10 15

Lys Lys Lys Lys
20

<210> 89

<211> 60

<212> DNA

<213> Artificial Sequence

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

<400> 89

ggcggcggca gcaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag 60

<210> 90

<211> 11

<212> PRT

<213> Artificial Sequence

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

<400> 90

Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys Val
1 5 10

<210
> 91

<211> 33

<212> DNA

<213> Artificial Sequence

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

<400> 91

ggcggcggca gcccaaagaa gaagcggaag gtc 33

<210> 92

<211> 12542

<212> DNA

<213> Artificial Sequence

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

<400> 92

ctttcgtccg atccatcttg caggctacct ctcgaacgaa ctatcgcaag tctcttggcc 60

ggccttgcgc cttggctatt gcttggcagc gcctatcgcc aggtattact ccaatcccga 120

atatccgaga tcgggatcac ccgagagaag ttcaacctac atcctcaatc ccgatctatc 180

cgagatccga ggaatatcga aatcggggcg cgcctggtgt accgagaacg atcctctcag 240

tgcgagtctc gacgatccat atcgttgctt ggcagtcagc cagtcggaat ccagcttggg 300

acccaggaag tccaatcgtc agatattgta ctcaagcctg gtcacggcag cgtaccgatc 360

tgtttaaacc tagatattga tagtctgatc ggtcaacgta taatcgagtc ctagcttttg 420

caaacatcta tcaagagaca ggatcagcag gaggctttcg catgagtatt caacatttcc 480

gtgtcgccct tattcccttt tttgcggcat tttgccttcc tgtttttgct cacccagaaa 540

cgctggtgaa agtaaaagat gctgaagatc agttgggtgc gcgagtgggt tacatcgaac 600

tggatctcaa cagcggtaag atccttgaga gttttcgccc cgaagaacgc tttccaatga 660

tgagcacttt taaagttctg ctatgtggcg cggtattatc ccgtattgac gccgggcaag 720

agcaactcgg tcgccgcata cactattctc agaatgactt ggttgagtat tcaccagtca 780

cagaaaagca tcttacggat ggcatgacag taagagaatt atgcagtgct gccataacca 840

tgagtgataa cactgcggcc aacttacttc tgacaacgat tggaggaccg aaggagctaa 900

ccgctttttt gcacaacatg ggggatcatg taactcgcct tgatcgttgg gaaccggagc 960

tgaatgaagc cataccaaac gacgagcgtg acaccacgat gcctgtagca atggcaacaa 1020

ccttgcgtaa actattaact ggcgaactac ttactctagc ttcccggcaa cagttgatag 1080

actggatgga ggcggataaa gttgcaggac cacttctgcg ctcggccctt ccggctggct 1140

ggtttattgc tgataaatct ggagccggtg agcgtgggtc tcgcggtatc attgcagcac 1200

tggggccaga tggtaagccc tcccgtatcg tagttatcta cacgacgggg agtcaggcaa 1260

ctatggatga acgaaataga cagatcgctg agataggtgc ctcactgatt aagcattggt 1320

aaccgattct aggtgcattg gcgcagaaaa aaatgcctga tgcgacgctg cgcgtcttat 1380

actcccacat atgccagatt cagcaacgga tacggcttcc ccaacttgcc cacttccata 1440

cgtgtcctcc ttaccagaaa tttatcctta agatcgttta aactcgactc tggctctatc 1500

gaatctccgt cgtttcgagc ttacgcgaac agccgtggcg ctcatttgct cgtcgggcat 1560

cgaatctcgt cagctatcgt cagcttacct ttttggcagc gatcgcggct cccgacatct 1620

tggaccatta gctccacagg tatcttcttc cctctagtgg tcataacagc agcttcagct 1680

acctctcaat tcaaaaaacc cctcaagacc cgtttagagg ccccaagggg ttatgctatc 1740

aatcgttgcg ttacacacac aaaaaaccaa cacacatcca tcttcgatgg atagcgattt 1800

tattatctaa ctgctgatcg agtgtagcca gatctagtaa tcaattacgg ggtcattagt 1860

tcatagccca tatatggagt tccgcgttac ataacttacg gtaaatggcc cgcctggctg 1920

accgcccaac gacccccgcc cattgacgtc aataatgacg tatgttccca tagtaacgcc 1980

aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaactg cccacttggc 2040

agtacatcaa gtgtatcata tgccaagtac gccccctatt gacgtcaatg acggtaaatg 2100

gcccgcctgg cattatgccc agtacatgac cttatgggac tttcctactt ggcagtacat 2160

ctacgtatta gtcatcgcta ttaccatgct gatgcggttt tggcagtaca tcaatgggcg 2220

tggatagcgg tttgactcac ggggatttcc aagtctccac cccattgacg tcaatgggag 2280

tttgttttgg caccaaaatc aacgggactt tccaaaatgt cgtaacaact ccgccccatt 2340

gacgcaaatg ggcggtaggc gtgtacggtg ggaggtctat ataagcagag ctggtttagt 2400

gaaccgtcag atcagatctt tgtcgatcct accatccact cgacacaccc gccagcggcc 2460

gctaatacga ctcactatag ggagaagtac tgccaccatg ggcaagaagc aaaatcgcaa 2520

gacggggaat tccaagacac aatccgctag cccaccacct aaagagcgtt ctagctcccc 2580

tgctactgag cagtcctgga tggaaaacga cttcgatgaa ctccgggaag agggatttag 2640

gcgatccaac tattcagaac tccgcgaaga tatccagaca aaggggaagg aagtcgagaa 2700

tttcgagaag aacctcgagg agtgcatcac ccgtatcaca aacactgaga aatgtctcaa 2760

agaactcatg gaacttaaga caaaagccag ggagcttcga gaggagtgtc ggagtctgag 2820

atccaggtgt gaccagctcg aggagcgcgt gagcgcgatg gaagacgaga tgaacgagat 2880

gaaaagagag ggcaaattca gggagaagcg cattaagagg aacgaacaga gtctgcagga 2940

gatttgggat tacgtcaaga ggcctaacct gcggttgatc ggcgtccccg agagcgacgt 3000

agaaaacggg actaaactgg agaatacact tcaagacatc attcaagaaa attttccaaa 3060

cctggctcgg caagctaatg tgcaaatcca agagatccaa cgcacacccc agcggtatag 3120

ctctcggcgt gccaccccta ggcatattat cgtgcgcttt actaaggtgg agatgaaaga 3180

gaagatgctg cgagccgctc gggaaaaggg aagggtgact ttgaagggca aacctattcg 3240

gctgacggtt gaccttagcg ccgagacact ccaggcacgc cgggaatggg gccccatctt 3300

taatatcctg aaggagaaga acttccagcc acgaatctct taccctgcaa agttgagttt 3360

tatctccgag ggtgagatta agtatttcat cgataaacag atgctgcgag acttcgtgac 3420

aactcgccca gctctcaagg aactgctcaa agaggctctt aatatggagc gcaataatag 3480

atatcaaccc ttgcagaacc acgcaaagat gccaaagaag aagcggaagg tctgagacag 3540

ccgtcagacc atcaagacta ggaagaaact gcatcaacta atgagcaaaa tcaccagcta 3600

acatcatagt atacatgacc ggctctaact cacatatcac catccttaca cttaacatta 3660

acggcctcaa ctcagctatc aagcgccatc ggctggccag ctggatcaaa tcacaggatc 3720

caagcgtttg ttgcatccaa gagacccacc tgacctgtag agatactcac cgcctcaaga 3780

tcaagggatg gcgaaagatt tatcaggcga acggtaagca gaagaaagcc ggagtcgcaa 3840

ttctggtctc agacaagacg gatttcaagc ccaccaaaat taagcgtgat aaggaaggtc 3900

actatattat ggtgaaaggc agcatacagc aggaagaact taccatattg aacatctacg 3960
cgccaaacac cggcgcacct cgctttatca aacaggtcct gtccgatctg cagcgagatc 4020
tggattctca tacgttgatt atgggtgatt tcaatacacc attgagcacc ctggatcgca 4080
gcaccaggca aaaggtaaat aaagacacgc aagagctcaa tagcgcactg catcaggcag 4140
atctcattga tatttatcgc actcttcatc ctaagagtac cgagtacaca ttcttcagcg 4200
ccccacatca tacatactca aagatcgatc atatcgtcgg ctcaaaggct ctgctgtcaa 4260
agtgcaagcg cacagagata attacaaatt acctgtcaga tcatagcgcg atcaagctcg 4320

agctgagaat caagaacctg acccagagcc ggagtaccac ttggaagctt aataacctgc 4380
tgctcaacga ttattgggtc cacaatgaga tgaaggcaga gattaaaatg ttcttcgaaa 4440
caaatgagaa taaggatact acctatcaaa acctttggga tgcctttaag gccgtctgca 4500
gaggcaagtt catcgccctc aacgcctata aaagaaaaca agagagatct aagatcgata 4560
ctctcacctc tcagctgaag gagttggaga aacaggaaca gacccactcc aaggcgtcaa 4620
gacggcagga gatcacaaag attcgcgccg agttgaaaga gatcgaaacc caaaagactc 4680
ttcagaaaat taacgagtct cgtagttggt tcttcgagcg gattaataag atagacagac 4740

ctctggcacg actgattaag aagaagcgcg aaaagaacca gattgatacc atcaagaacg 4800
acaagggcga catcactact gacccgaccg agatccagac cactattcgg gagtattata 4860
agcatttgta tgctaacaag cttgagaacc tggaagagat ggacactttt ctggatacct 4920
atactctgcc acggcttaat caagaggaag tcgagtccct caaccgccca attacaggaa 4980
gcgagattgt ggccataatt aactccctgc cgacaaagaa atctcctggt ccggacgggt 5040
ttacagctga gttttatcaa cggtatatgg aagagcttgt accgtttctg ctcaagctct 5100
ttcagtctat agaaaaggaa ggcatcttgc ccaattcctt ctacgaagct tctataatac 5160

ttattcccaa accaggacgc gataccacaa agaaggaaaa cttccggccc attagtctca 5220
tgaatatcga cgctaaaata ttgaacaaga ttctcgccaa cagaatccaa caacatatta 5280
agaaattgat acatcacgac caggtggggt ttatacctgg catgcagggc tggtttaaca 5340
tccggaagag tattaacgtc attcaacaca ttaatagagc taaggataag aatcatatga 5400
tcatctctat agacgcggaa aaggcattcg ataagattca gcagccattt atgctcaaga 5460
ctctgaacaa actcggcatc gacggaacat attttaagat tattcgcgca atttacgata 5520
agccgactgc taacattatc cttaacggcc aaaagctcga ggcctttccg ctcaagactg 5580

gaacccgcca aggctgtccc ctctccccgc ttttgtttaa tattgtactc gaggtgctgg 5640

ctagggctat tcgtcaagag aaagagatta aagggataca gctcgggaag gaagaggtca 5700

agctttcctt gttcgccgat gatatgattg tgtacctgga gaatcctatt gtgtctgctc 5760

agaaccttct taaacttatt tctaacttta gcaaggtcag cggctataag attaacgtcc 5820

agaaatctca ggcctttctg tacacaaata atcgacagac cgaatcccag ataatgggtg 5880

agcttccgtt tgtcatagcc agcaaaagga taaagtatct cggaatccag ctgacacgag 5940

acgttaaaga tttgtttaag gaaaattaca agcctctcct gaaagagatt aaggaagata 6000

ctaataagtg gaagaatatc ccctgttcat gggttggcag aatcaacata gtgaagatgg 6060

caatacttcc taaagtgata tatcgcttta acgccatccc aattaaactg cctatgacct 6120

tctttacgga gctcgagaaa acaaccctta aatttatatg gaatcaaaag agagcaagaa 6180

tagcgaagtc catcttgagc cagaagaata aggccggtgg gattactttg cctgatttta 6240

agttgtatta taaagccaca gtaactaaga cagcctggta ttggtatcag aatagagaca 6300

tcgaccagtg gaatcggacc gaaccatcag agataatgcc ccacatctat aattacctta 6360

tattcgataa gccagaaaag aataaacagt ggggcaaaga cagcctcttc aacaagtggt 6420

gttgggagaa ttggctggcc atatgccgga aactcaagct cgaccccttt cttacaccct 6480

acactaaaat caacagtagg tggatcaagg acttgaatgt caagccaaag actataaaga 6540

cactggaaga gaatcttggg atcacaatac aagatatagg cgtcggcaaa gattttatgt 6600

caaagacgcc caaggccatg gccactaagg ataagattga taagtgggac cttattaagc 6660

tcaaaagctt ctgtactgcc aaggagacca cgatcagagt taataggcag cccactacat 6720

gggaaaagat tttcgccact tattcatcag ataaggggtt gataagcaga atatataacg 6780

agctgaagca gatctacaag aagaaaacga ataatcccat caagaagtgg gcaaaagata 6840

tgaacaggca ttttagcaaa gaggatatct acgccgcgaa gaagcatatg aagaagtgta 6900

gttcaagctt ggccattcgt gagatgcaga ttaagacgac catgcgatac caccttaccc 6960

cagtgaggat ggcaattatc aagaaatctg gcaataatag atgttggcgg ggctgtggcg 7020

agattggcac cctgctccat tgctggtggg attgcaagct ggtgcagccg ctttggaaat 7080

cagtctggcg ctttctgagg gacctcgagc ttgagattcc cttcgatccc gcaattccct 7140

tgctcggaat ctatcctaac gaatacaaga gctgttgtta caaggatacg tgtacccgga 7200

tgttcatcgc ggccttgttt acgatagcta agacgtggaa tcagcctaag tgccccacaa 7260

tgatcgattg gatcaagaaa atgtggcata tttataccat ggagtattac gcagcaatta 7320

agaatgacga atttatttcc ttcgttggga cctggatgaa gctggagact attattctga 7380

gcaagctgtc tcaggagcaa aagacaaagc atagaatctt ctctctcatt ggtggtaacg 7440

actacaaaga cgatgacgac aagtaaagcg cttctagaag ttgtctcctc ctgcactgac 7500

tgactgatac aatcgatttc tggatccgca ggcctaatca acctctggat tacaaaattt 7560

gtgaaagatt gactggtatt cttaactatg ttgctccttt tacgctatgt ggatacgctg 7620

ctttaatgcc tttgtatcat gctattgctt cccgtatggc tttcattttc tcctccttgt 7680

ataaatcctg gttgctgtct ctttatgagg agttgtggcc cgttgtcagg caacgtggcg 7740

tggtgtgcac tgtgtttgct gacgcaaccc ccactggttg gggcattgcc accacctgtc 7800

agctcctttc cgggactttc gctttccccc tccctattgc cacggcggaa ctcatcgccg 7860

cctgccttgc ccgctgctgg acaggggctc ggctgttggg cactgacaat tccgtggtgt 7920

tgtcggggaa gctgacgtcc tttccatggc tgctcgcctg tgttgccacc tggattctgc 7980

gcgggacgtc cttctgctac gtcccttcgg ccctcaatcc agcggacctt ccttcccgct 8040

gagagacaca aaaaattcca acacactatt gcaatgaaaa taaatttcct ttattagcca 8100

gaagtcagat gctcaagggg cttcatgatg tccccataat ttttggcaga gggaaaaaga 8160

tctcagtggt atttgtgagc cagggcattg gccttctgat aggcagcctg cacctgagga 8220

gtgcggccgc tttacttgta cagctcgtcc atgccgagag tgatcccggc ggcggtcacg 8280

aactccagca ggaccatgtg atcgcgcttc tcgttggggt ctttgctcag ggcggactgg 8340

gtgctcaggt agtggttgtc gggcagcagc acggggccgt cgccgatggg ggtgttctgc 8400

tggtagtggt cggcgagctg cacgctgccg tcctcgatgt tgtggcggat cttgaagttc 8460

accttgatgc cgttcttctg cttgtcggcc atgatataga cgttgtggct gttgtagttg 8520

tactccagct tgtgccccag gatgttgccg tcctccttga agtcgatgcc cttcagctcg 8580

atgcggttca ccagggtgtc gccctcgaac ttcacctcgg cgcgggtctt gtagttgccg 8640

tcgtccttga agaagatggt gcgctcctgg acgtagcctt cgggcatggc ggacttgaag 8700

aagtcgtgct gcttcatgtg gtcggggtag cggctgaagc actgcacgcc gtaggtcagg 8760

gtggtcacga gggtgggcca gggcacgggc agcttgccgg tggtgcagat gaacttcagg 8820

gtcagcttgc cgtaggtggc atcgccctcg ccctcgccgg acacgctgaa cttgtggccg 8880

tttacgtcgc cgtccagctc gaccaggatg ggcaccaccc cggtgaacag ctcctcgccc 8940

ttgctcacca tggtggcggg atctgacggt tcactaaacc agctctgctt atatagacct 9000

cccaccgtac acgcctaccg cccatttgcg tcaatggggc ggagttgtta cgacattttg 9060

gaaagtcccg ttgattttgg tgccaaaaca aactcccatt gacgtcaatg gggtggagac 9120

ttggaaatcc ccgtgagtca aaccgctatc cacgcccatt gatgtactgc caaaaccgca 9180

tcaccatggt aatagcgatg actaatacgt agatgtactg ccaagtagga aagtcccata 9240

aggtcatgta ctgggcataa tgccaggcgg gccatttacc gtcattgacg tcaatagggg 9300

gcgtacttgg catatgatac acttgatgta ctgccaagtg ggcagtttac cgtaaatact 9360

ccacccattg acgtcaatgg aaagtcccta ttggcgttac tatgggaaca tacgtcatta 9420
ttgacgtcaa tgggcggggg tcgttgggcg gtcagccagg cgggccattt accgtaagtt 9480
atgtaacggg cctgctgccg gctctgcggc ctcttccgcg tcttcgcctt cgccctcaga 9540
cgagtcggat ctccctttgg gccgcctccc cgcctgtcta gcttgactga ctgagataca 9600
gcgtaccttc agctcacaga catgataaga tacattgatg agtttggaca aaccacaact 9660
agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc tttatttgta 9720
accattataa gctgcaataa acaagttaac aacaacaatt gcattcattt tatgtttcag 9780

gttcaggggg aggtgtggga ggtttttaa agcaagtaaa acctctacaa atgtggtatt 9840
ggcccatctc tatcggtatc gtagcataac cccttggggc ctctaaacgg gtcttgaggg 9900
gttttttgtg cccctcgggc cggattgcta tctaccggca ttggcgcaga aaaaaatgcc 9960
tgatgcgacg ctgcgcgtct tatactccca catatgccag attcagcaac ggatacggct 10020
tccccaactt gcccacttcc atacgtgtcc tccttaccag aaatttatcc ttaaggtcgt 10080
cagctatcct gcaggcgatc tctcgatttc gatcaagaca ttcctttaat ggtcttttct 10140
ggacaccact aggggtcaga agtagttcat caaactttct tccctcccta atctcattgg 10200

ttaccttggg ctatcgaaac ttaattaagc gatctgcatc tcaattagtc agcaaccata 10260
gtcccgcccc taactccgcc catcccgccc ctaactccgc ccagttccgc ccattctccg 10320
ccccatcgct gactaatttt ttttatttat gcagaggccg aggccgcctc ggcctctgag 10380
ctattccaga agtagtgagg aggctttttt ggaggcctag gcttttgcaa aggaggtagc 10440
caacatgatt gaacaagatg gattgcacgc aggttctccc gccgcttggg tggagaggct 10500
attcggctat gactgggcac aacagacaat cggctgctct gatgccgccg tgttccggct 10560
gtcagcgcag gggcgcccgg ttctttttgt caagaccgac ctgtccggtg ccctgaatga 10620

actccaggac gaggcagcgc ggctatcgtg gctggccacg acgggcgttc cttgcgcagc 10680
tgtgctcgac gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg 10740
gcaggatctc ctgtcatctc accttgctcc tgccgagaaa gtatccatca tggctgatgc 10800
aatgcggcgg ctgcatacgc ttgatccggc tacctgccca ttcgaccacc aagcgaaaca 10860
tcgcatcgag cgagcacgta ctcggatgga agccggtctt gtcgatcagg atgatctgga 10920
cgaagagcat caggggctcg cgccagccga actgttcgcc aggctcaagg cgcggatgcc 10980
cgacggcgag gatctcgtcg tgacccacgg cgatgcctgc ttgccgaata tcatggtgga 11040

aaatggccgc ttttctggat tcatcgactg tggccggctg ggtgtggcgg accgctatca 11100

ggacatagcg ttggctaccc gtgatattgc tgaagagctt ggcggcgaat gggctgaccg 11160

cttcctcgtg ctttacggta tcgccgctcc cgattcgcag cgcatcgcct tctatcgcct 11220

tcttgacgag ttcttctagt atgtaagccc tgtgccttct agttgccagc catctgttgt 11280

ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta 11340

ataaaatgag gaaattgcat cgcattgtct gagtaggtgt cattctattc tgggggggtgg 11400

ggtggggcag gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc 11460

ggtgggctct atggttaatt aaccagtcaa gtcagctact tggcgagatc gacttgtctg 11520

ggtttcgact acgctcagaa ttgcgtcagt caagttcgat ctggtccttg ctattgcacc 11580

cgttctccga ttacgagttt catttaaatc atgtgagcaa aaggccagca aaaggccagg 11640

aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc tccgcccccc tgacgagcat 11700

cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga caggactata aagataccag 11760

gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga 11820

tacctgtccg cctttctccc ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg 11880

tatctcagtt cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga acccccgtt 11940

cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc ggtaagacac 12000

gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc 12060

ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag aacagtattt 12120

ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc 12180

ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt gcaagcagca gattacgcgc 12240

agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga cgctcagtgg 12300

aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat cttcacctag 12360

atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga gtaaacttgg 12420

tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg tctatttcgt 12480

tcatccatag ttgcatttaa atttccgaac tctccaaggc cctcgtcgga aaatcttcaa 12540

ac 12542

<210> 93

<211> 12554

<212> DNA

<213> Artificial Sequence

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

polynucleotide

<400> 93

cctctcagtg cgagtctcga cgatccatat cgttgcttgg cagtcagcca gtcggaatcc 60

agcttgggac ccaggaagtc caatcgtcag atattgtact caagcctggt cacggcagcg 120

taccgatctg tttaaaccta gatattgata gtctgatcgg tcaacgtata atcgagtcct 180

agcttttgca aacatctatc aagagacagg atcagcagga ggctttcgca tgagtattca 240

acatttccgt gtcgccctta ttccctttt tgcggcattt tgccttcctg tttttgctca 300

cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcgc gagtgggtta 360

catcgaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg aagaacgctt 420

tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc gtattgacgc 480

cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg ttgagtattc 540

accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat gcagtgctgc 600

cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgattg gaggaccgaa 660

ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg atcgttggga 720

accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc ctgtagcaat 780

ggcaacaacc ttgcgtaaac tattaactgg cgaactactt actctagctt cccggcaaca 840

gttgatagac tggatggagg cggataaagt tgcaggacca cttctgcgct cggcccttcc 900

ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc gcggtatcat 960

tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca cgacggggag 1020

tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct cactgattaa 1080

gcattggtaa ccgattctag gtgcattggc gcagaaaaaa atgcctgatg cgacgctgcg 1140

cgtcttatac tcccacatat gccagattca gcaacggata cggcttcccc aacttgccca 1200

cttccatacg tgtcctcctt accagaaatt tatccttaag atcgtttaaa ctcgactctg 1260

gctctatcga atctccgtcg tttcgagctt acgcgaacag ccgtggcgct catttgctcg 1320

tcgggcatcg aatctcgtca gctatcgtca gcttaccttt ttggcagcga tcgcggctcc 1380

cgacatcttg gaccattagc tccacaggta tcttcttccc tctagtggtc ataacagcag 1440

cttcagctac ctctcaattc aaaaaacccc tcaagacccg tttagaggcc ccaaggggtt 1500

atgctatcaa tcgttgcgtt acacacacaa aaaaccaaca cacatccatc ttcgatggat 1560

agcgatttta ttatctaact gctgatcgag tgtagccaga tctagtaatc aattacgggg 1620

tcattagttc atagcccata tatggagttc cgcgttacat aacttacggt aaatggcccg 1680

cctggctgac cgcccaacga cccccgccca ttgacgtcaa taatgacgta tgttcccata 1740
gtaacgccaa tagggacttt ccattgacgt caatgggtgg agtatttacg gtaaactgcc 1800
cacttggcag tacatcaagt gtatcatatg ccaagtacgc cccctattga cgtcaatgac 1860
ggtaaatggc ccgcctggca ttatgcccag tacatgacct tatgggactt tcctacttgg 1920
cagtacatct acgtattagt catcgctatt accatgctga tgcggttttg gcagtacatc 1980
aatgggcgtg gatagcggtt tgactcacgg ggatttccaa gtctccaccc cattgacgtc 2040
aatgggagtt tgttttggca ccaaaatcaa cgggactttc caaaatgtcg taacaactcc 2100

gccccattga cgcaaatggg cggtaggcgt gtacggtggg aggtctatat aagcagagct 2160
ggtttagtga accgtcagat cagatctttg tcgatcctac catccactcg acacacccgc 2220
cagcggccgc taatacgact cactataggg agaagtactg ccaccatggg caagaagcaa 2280
aatcgcaaga cggggaattc caagacacaa tccgctagcc caccacctaa agagcgttct 2340
agctcccctg ctactgagca gtcctggatg gaaaacgact tcgatgaact ccgggaagag 2400
ggatttaggc gatccaacta ttcagaactc cgcgaagata tccagacaaa ggggaaggaa 2460
gtcgagaatt tcgagaagaa cctcgaggag tgcatcaccc gtatcacaaa cactgagaaa 2520

tgtctcaaag aactcatgga acttaagaca aaagccaggg agcttcgaga ggagtgtcgg 2580
agtctgagat ccaggtgtga ccagctcgag gagcgcgtga gcgcgatgga agacgagatg 2640
aacgagatga aaagagaggg caaattcagg gagaagcgca ttaagaggaa cgaacagagt 2700
ctgcaggaga tttgggatta cgtcaagagg cctaacctgc ggttgatcgg cgtccccgag 2760
agcgacgtag aaaacgggac taaactggag aatacacttc aagacatcat tcaagaaaat 2820
tttccaaacc tggctcggca agctaatgtg caaatccaag agatccaacg cacaccccag 2880
cggtatagct ctcggcgtgc cacccctagg catattatcg tgcgctttac taaggtggag 2940

atgaaagaga agatgctgcg agccgctcgg gaaaagggaa gggtgacttt gaagggcaaa 3000
cctattcggc tgacggttga ccttagcgcc gagacactcc aggcacgccg ggaatggggc 3060
cccatcttta atatcctgaa ggagaagaac ttccagccac gaatctctta ccctgcaaag 3120
ttgagtttta tctccgaggg tgagattaag tatttcatcg ataaacagat gctgcgagac 3180
ttcgtgacaa ctcgcccagc tctcaaggaa ctgctcaaag aggctcttaa tatggagcgc 3240
aataatagat atcaaccctt gcagaaccac gcaaagatgg gcggcggcag cccaaagaag 3300
aagcggaagg tctgagacag ccgtcagacc atcaagacta ggaagaaact gcatcaacta 3360

atgagcaaaa tcaccagcta acatcatagt atacatgacc ggctctaact cacatatcac 3420

catccttaca cttaacatta acggcctcaa ctcagctatc aagcgccatc ggctggccag 3480

ctggatcaaa tcacaggatc caagcgtttg ttgcatccaa gagacccacc tgacctgtag 3540

agatactcac cgcctcaaga tcaagggatg gcgaaagatt tatcaggcga acggtaagca 3600

gaagaaagcc ggagtcgcaa ttctggtctc agacaagacg gatttcaagc ccaccaaaat 3660

taagcgtgat aaggaaggtc actatattat ggtgaaaggc agcatacagc aggaagaact 3720

taccatattg aacatctacg cgccaaacac cggcgcacct cgctttatca aacaggtcct 3780

gtccgatctg cagcgagatc tggattctca tacgttgatt atgggtgatt tcaatacacc 3840

attgagcacc ctggatcgca gcaccaggca aaaggtaaat aaagacacgc aagagctcaa 3900

tagcgcactg catcaggcag atctcattga tatttatcgc actcttcatc ctaagagtac 3960

cgagtacaca ttcttcagcg ccccacatca tacatactca aagatcgatc atatcgtcgg 4020

ctcaaaggct ctgctgtcaa agtgcaagcg cacagagata attacaaatt acctgtcaga 4080

tcatagcgcg atcaagctcg agctgagaat caagaacctg acccagagcc ggagtaccac 4140

ttggaagctt aataacctgc tgctcaacga ttattgggtc cacaatgaga tgaaggcaga 4200

gattaaaatg ttcttcgaaa caaatgagaa taaggatact acctatcaaa acctttggga 4260

tgcctttaag gccgtctgca gaggcaagtt catcgccctc aacgcctata aaagaaaaca 4320

agagagatct aagatcgata ctctcacctc tcagctgaag gagttggaga aacaggaaca 4380

gacccactcc aaggcgtcaa gacggcagga gatcacaaag attcgcgccg agttgaaaga 4440

gatcgaaacc caaaagactc ttcagaaaat taacgagtct cgtagttggt tcttcgagcg 4500

gattaataag atagacagac ctctggcacg actgattaag aagaagcgcg aaaagaacca 4560

gattgatacc atcaagaacg acaagggcga catcactact gacccgaccg agatccagac 4620

cactattcgg gagtattata agcatttgta tgctaacaag cttgagaacc tggaagagat 4680

ggacactttt ctggatacct atactctgcc acggcttaat caagaggaag tcgagtccct 4740

caaccgccca attacaggaa gcgagattgt ggccataatt aactccctgc cgacaaagaa 4800

atctcctggt ccggacgggt ttacagctga gttttatcaa cggtatatgg aagagcttgt 4860

accgtttctg ctcaagctct ttcagtctat agaaaaggaa ggcatcttgc ccaattcctt 4920

ctacgaagct tctataatac ttattcccaa accaggacgc gataccacaa agaaggaaaa 4980

cttccggccc attagtctca tgaatatcga cgctaaaata ttgaacaaga ttctcgccaa 5040

cagaatccaa caacatatta agaaattgat acatcacgac caggtggggt ttatacctgg 5100

catgcagggc tggtttaaca tccggaagag tattaacgtc attcaacaca ttaatagagc 5160

taaggataag aatcatatga tcatctctat agacgcggaa aaggcattcg ataagattca 5220

gcagccattt atgctcaaga ctctgaacaa actcggcatc gacggaacat attttaagat 5280

tattcgcgca atttacgata agccgactgc taacattatc cttaacggcc aaaagctcga 5340

ggcctttccg ctcaagactg gaacccgcca aggctgtccc ctctccccgc ttttgtttaa 5400

tattgtactc gaggtgctgg ctagggctat tcgtcaagag aaagagatta aagggataca 5460

gctcgggaag gaagaggtca agctttcctt gttcgccgat gatatgattg tgtacctgga 5520

gaatcctatt gtgtctgctc agaaccttct taaacttatt tctaacttta gcaaggtcag 5580

cggctataag attaacgtcc agaaatctca ggcctttctg tacacaaata atcgacagac 5640

cgaatcccag ataatgggtg agcttccgtt tgtcatagcc agcaaaagga taaagtatct 5700

cggaatccag ctgacacgag acgttaaaga tttgtttaag gaaaattaca agcctctcct 5760

gaaagagatt aaggaagata ctaataagtg gaagaatatc ccctgttcat gggttggcag 5820

aatcaacata gtgaagatgg caatacttcc taaagtgata tatcgcttta acgccatccc 5880

aattaaactg cctatgacct tctttacgga gctcgagaaa acaaccctta aatttatatg 5940

gaatcaaaag agagcaagaa tagcgaagtc catcttgagc cagaagaata aggccggtgg 6000

gattactttg cctgatttta agttgtatta taaagccaca gtaactaaga cagcctggta 6060

ttggtatcag aatagagaca tcgaccagtg gaatcggacc gaaccatcag agataatgcc 6120

ccacatctat aattacctta tattcgataa gccagaaaag aataaacagt ggggcaaaga 6180

cagcctcttc aacaagtggt gttgggagaa ttggctggcc atatgccgga aactcaagct 6240

cgaccccttt cttacaccct acactaaaat caacagtagg tggatcaagg acttgaatgt 6300

caagccaaag actataaaga cactggaaga gaatcttggg atcacaatac aagatatagg 6360

cgtcggcaaa gattttatgt caaagacgcc caaggccatg gccactaagg ataagattga 6420

taagtgggac cttattaagc tcaaaagctt ctgtactgcc aaggagacca cgatcagagt 6480

taataggcag cccactacat gggaaaagat tttcgccact tattcatcag ataaggggtt 6540

gataagcaga atatataacg agctgaagca gatctacaag aagaaaacga ataatcccat 6600

caagaagtgg gcaaaagata tgaacaggca ttttagcaaa gaggatatct acgccgcgaa 6660

gaagcatatg aagaagtgta gttcaagctt ggccattcgt gagatgcaga ttaagacgac 6720

catgcgatac caccttaccc cagtgaggat ggcaattatc aagaaatctg gcaataatag 6780

atgttggcgg ggctgtggcg agattggcac cctgctccat tgctggtggg attgcaagct 6840

ggtgcagccg ctttggaaat cagtctggcg ctttctgagg gacctcgagc ttgagattcc 6900

cttcgatccc gcaattccct tgctcggaat ctatcctaac gaatacaaga gctgttgtta 6960

caaggatacg tgtacccgga tgttcatcgc ggccttgttt acgatagcta agacgtggaa 7020

tcagcctaag tgccccacaa tgatcgattg gatcaagaaa atgtggcata tttataccat 7080

ggagtattac gcagcaatta agaatgacga atttatttcc ttcgttggga cctggatgaa 7140

gctggagact attattctga gcaagctgtc tcaggagcaa aagacaaagc atagaatctt 7200
ctctctcatt ggtggtaacg actacaaaga cgatgacgac aagtaaagcg cttctagaag 7260
ttgtctcctc ctgcactgac tgactgatac aatcgatttc tggatccgca ggcctaatca 7320
acctctggat tacaaaattt gtgaaagatt gactggtatt cttaactatg ttgctccttt 7380
tacgctatgt ggatacgctg ctttaatgcc tttgtatcat gctattgctt cccgtatggc 7440
tttcattttc tcctccttgt ataaatcctg gttgctgtct ctttatgagg agttgtggcc 7500
cgttgtcagg caacgtggcg tggtgtgcac tgtgtttgct gacgcaaccc ccactggttg 7560

gggcattgcc accacctgtc agctcctttc cgggactttc gctttccccc tccctattgc 7620
cacggcggaa ctcatcgccg cctgccttgc ccgctgctgg acaggggctc ggctgttggg 7680
cactgacaat tccgtggtgt tgtcggggaa gctgacgtcc tttccatggc tgctcgcctg 7740
tgttgccacc tggattctgc gcgggacgtc cttctgctac gtcccttcgg ccctcaatcc 7800
agcggacctt ccttcccgct gagagacaca aaaaattcca acacactatt gcaatgaaaa 7860
taaatttcct ttattagcca gaagtcagat gctcaagggg cttcatgatg tccccataat 7920
ttttggcaga gggaaaaaga tctcagtggt atttgtgagc cagggcattg gccttctgat 7980

aggcagcctg cacctgagga gtgcggccgc tttacttgta cagctcgtcc atgccgagag 8040
tgatcccggc ggcggtcacg aactccagca ggaccatgtg atcgcgcttc tcgttggggt 8100
ctttgctcag ggcggactgg gtgctcaggt agtggttgtc gggcagcagc acggggccgt 8160
cgccgatggg ggtgttctgc tggtagtggt cggcgagctg cacgctgccg tcctcgatgt 8220
tgtggcggat cttgaagttc accttgatgc cgttcttctg cttgtcggcc atgatataga 8280
cgttgtggct gttgtagttg tactccagct tgtgccccag gatgttgccg tcctccttga 8340
agtcgatgcc cttcagctcg atgcggttca ccagggtgtc gccctcgaac ttcacctcgg 8400

cgcgggtctt gtagttgccg tcgtccttga agaagatggt gcgctcctgg acgtagcctt 8460
cgggcatggc ggacttgaag aagtcgtgct gcttcatgtg gtcggggtag cggctgaagc 8520
actgcacgcc gtaggtcagg gtggtcacga gggtgggcca gggcacgggc agcttgccgg 8580
tggtgcagat gaacttcagg gtcagcttgc cgtaggtggc atcgccctcg ccctcgccgg 8640
acacgctgaa cttgtggccg tttacgtcgc cgtccagctc gaccaggatg ggcaccaccc 8700
cggtgaacag ctcctcgccc ttgctcacca tggtggcggg atctgacggt tcactaaacc 8760
agctctgctt atatagacct cccaccgtac acgcctaccg cccatttgcg tcaatggggc 8820

ggagttgtta cgacattttg gaaagtcccg ttgattttgg tgccaaaaca aactcccatt 8880

gacgtcaatg gggtggagac ttggaaatcc ccgtgagtca aaccgctatc cacgcccatt 8940

gatgtactgc caaaaccgca tcaccatggt aatagcgatg actaatacgt agatgtactg 9000

ccaagtagga aagtcccata aggtcatgta ctgggcataa tgccaggcgg gccatttacc 9060

gtcattgacg tcaatagggg gcgtacttgg catatgatac acttgatgta ctgccaagtg 9120

ggcagtttac cgtaaatact ccacccattg acgtcaatgg aaagtcccta ttggcgttac 9180

tatgggaaca tacgtcatta ttgacgtcaa tgggcggggg tcgttgggcg gtcagccagg 9240

cgggccattt accgtaagtt atgtaacggg cctgctgccg gctctgcggc ctcttccgcg 9300

tcttcgcctt cgccctcaga cgagtcggat ctccctttgg gccgcctccc cgcctgtcta 9360

gcttgactga ctgagataca gcgtaccttc agctcacaga catgataaga tacattgatg 9420

agtttggaca aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg 9480

atgctattgc tttatttgta accattataa gctgcaataa acaagttaac aacaacaatt 9540

gcattcattt tatgtttcag gttcaggggg aggtgtggga ggtttttaa agcaagtaaa 9600

acctctacaa atgtggtatt ggcccatctc tatcggtatc gtagcataac cccttggggc 9660

ctctaaacgg gtcttgaggg gttttttgtg cccctcgggc cggattgcta tctaccggca 9720

ttggcgcaga aaaaaatgcc tgatgcgacg ctgcgcgtct tatactccca catatgccag 9780

attcagcaac ggatacggct tccccaactt gcccacttcc atacgtgtcc tccttaccag 9840

aaatttatcc ttaaggtcgt cagctatcct gcaggcgatc tctcgatttc gatcaagaca 9900

ttcctttaat ggtcttttct ggacaccact aggggtcaga agtagttcat caaactttct 9960

tccctcccta atctcattgg ttaccttggg ctatcgaaac ttaattaagc gatctgcatc 10020

tcaattagtc agcaaccata gtcccgcccc taactccgcc catcccgccc ctaactccgc 10080

ccagttccgc ccattctccg ccccatcgct gactaatttt ttttatttat gcagaggccg 10140

aggccgcctc ggcctctgag ctattccaga agtagtgagg aggctttttt ggaggcctag 10200

gcttttgcaa aggaggtagc caacatgatt gaacaagatg gattgcacgc aggttctccc 10260

gccgcttggg tggagaggct attcggctat gactgggcac aacagacaat cggctgctct 10320

gatgccgccg tgttccggct gtcagcgcag gggcgcccgg ttctttttgt caagaccgac 10380

ctgtccggtg ccctgaatga actccaggac gaggcagcgc ggctatcgtg gctggccacg 10440

acgggcgttc cttgcgcagc tgtgctcgac gttgtcactg aagcgggaag ggactggctg 10500

ctattgggcg aagtgccggg gcaggatctc ctgtcatctc accttgctcc tgccgagaaa 10560

gtatccatca tggctgatgc aatgcggcgg ctgcatacgc ttgatccggc tacctgccca 10620

ttcgaccacc aagcgaaaca tcgcatcgag cgagcacgta ctcggatgga agccggtctt 10680

gtcgatcagg atgatctgga cgaagagcat caggggctcg cgccagccga actgttcgcc 10740

aggctcaagg cgcggatgcc cgacggcgag gatctcgtcg tgacccacgg cgatgcctgc 10800

ttgccgaata tcatggtgga aaatggccgc ttttctggat tcatcgactg tggccggctg 10860

ggtgtggcgg accgctatca ggacatagcg ttggctaccc gtgatattgc tgaagagctt 10920

ggcggcgaat gggctgaccg cttcctcgtg ctttacggta tcgccgctcc cgattcgcag 10980

cgcatcgcct tctatcgcct tcttgacgag ttcttctagt atgtaagccc tgtgccttct 11040

agttgccagc catctgttgt ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc 11100

actcccactg tcctttccta ataaaatgag gaaattgcat cgcattgtct gagtaggtgt 11160

cattctattc tgggggggtgg ggtggggcag gacagcaagg gggaggattg ggaagacaat 11220

agcaggcatg ctggggatgc ggtgggctct atggttaatt aaccagtcaa gtcagctact 11280

tggcgagatc gacttgtctg ggtttcgact acgctcagaa ttgcgtcagt caagttcgat 11340

ctggtccttg ctattgcacc cgttctccga ttacgagttt catttaaatc atgtgagcaa 11400

aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc 11460

tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga 11520

caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc 11580

cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt 11640

ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc aagctgggct 11700

gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg 11760

agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta 11820

gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct 11880

acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa 11940

gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt 12000

gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta 12060

cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat 12120

caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa 12180

gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct 12240

cagcgatctg tctatttcgt tcatccatag ttgcatttaa atttccgaac tctccaaggc 12300

cctcgtcgga aaatcttcaa acctttcgtc cgatccatct tgcaggctac ctctcgaacg 12360

aactatcgca agtctcttgg ccggccttgc gccttggcta ttgcttggca gcgcctatcg 12420

ccaggtatta ctccaatccc gaatatccga gatcgggatc acccgagaga agttcaacct 12480

acatcctcaa tcccgatcta tccgagatcc gaggaatatc gaaatcgggg cgcgcctggt 12540

gtaccgagaa cgat 12554

<210> 94

<211> 12569

<212> DNA

<213> Artificial Sequence

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

<400> 94

tcgaaatcgg ggcgcgcctg gtgtaccgag aacgatcctc tcagtgcgag tctcgacgat 60

ccatatcgtt gcttggcagt cagccagtcg gaatccagct tgggacccag gaagtccaat 120

cgtcagatat tgtactcaag cctggtcacg gcagcgtacc gatctgttta aacctagata 180

ttgatagtct gatcggtcaa cgtataatcg agtcctagct tttgcaaaca tctatcaaga 240

gacaggatca gcaggaggct ttcgcatgag tattcaacat ttccgtgtcg cccttattcc 300

ctttttgcg gcattttgcc ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa 360

agatgctgaa gatcagttgg gtgcgcgagt gggttacatc gaactggatc tcaacagcgg 420

taagatcctt gagagttttc gccccgaaga acgctttcca atgatgagca cttttaaagt 480

tctgctatgt ggcgcggtat tatcccgtat tgacgccggg caagagcaac tcggtcgccg 540

catacactat tctcagaatg acttggttga gtattcacca gtcacagaaa agcatcttac 600

ggatggcatg acagtaagag aattatgcag tgctgccata accatgagtg ataacactgc 660

ggccaactta cttctgacaa cgattggagg accgaaggag ctaaccgctt ttttgcacaa 720

catgggggat catgtaactc gccttgatcg ttgggaaccg gagctgaatg aagccatacc 780

aaacgacgag cgtgacacca cgatgcctgt agcaatggca acaaccttgc gtaaactatt 840

aactggcgaa ctacttactc tagcttcccg gcaacagttg atagactgga tggaggcgga 900

taaagttgca ggaccacttc tgcgctcggc ccttccggct ggctggttta ttgctgataa 960

atctggagcc ggtgagcgtg ggtctcgcgg tatcattgca gcactggggc cagatggtaa 1020

gccctcccgt atcgtagtta tctacacgac ggggagtcag gcaactatgg atgaacgaaa 1080

tagacagatc gctgagatag gtgcctcact gattaagcat tggtaaccga ttctaggtgc 1140

attggcgcag aaaaaaatgc ctgatgcgac gctgcgcgtc ttatactccc acatatgcca 1200

gattcagcaa cggatacggc ttccccaact tgcccacttc catacgtgtc ctccttacca 1260

gaaatttatc cttaagatcg tttaaactcg actctggctc tatcgaatct ccgtcgtttc 1320

gagcttacgc gaacagccgt ggcgctcatt tgctcgtcgg gcatcgaatc tcgtcagcta 1380

tcgtcagctt acctttttgg cagcgatcgc ggctcccgac atcttggacc attagctcca 1440

caggtatctt cttccctcta gtggtcataa cagcagcttc agctacctct caattcaaaa 1500

aacccctcaa gacccgttta gaggccccaa ggggttatgc tatcaatcgt tgcgttacac 1560

acacaaaaaa ccaacacaca tccatcttcg atggatagcg attttattat ctaactgctg 1620

atcgagtgta gccagatcta gtaatcaatt acggggtcat tagttcatag cccatatatg 1680

gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 1740

cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 1800

tgacgtcaat gggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 1860

catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 1920

gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcatc 1980

gctattacca tgctgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 2040

tcacggggat ttccaagtct ccaccccatt gacgtcaatg ggagtttgtt ttggcaccaa 2100

aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt 2160

aggcgtgtac ggtgggaggt ctatataagc agagctggtt tagtgaaccg tcagatcaga 2220

tctttgtcga tcctaccatc cactcgacac acccgccagc ggccgctaat acgactcact 2280

atagggagaa gtactgccac catgggcaag aagcaaaatc gcaagacggg gaattccaag 2340

acacaatccg ctagcccacc acctaaagag cgttctagct cccctgctac tgagcagtcc 2400

tggatggaaa acgacttcga tgaactccgg gaagagggat ttaggcgatc caactattca 2460

gaactccgcg aagatatcca gacaaagggg aaggaagtcg agaatttcga gaagaacctc 2520

gaggagtgca tcacccgtat cacaaacact gagaaatgtc tcaaagaact catggaactt 2580

aagacaaaag ccagggagct tcgagaggag tgtcggagtc tgagatccag gtgtgaccag 2640

ctcgaggagc gcgtgagcgc gatggaagac gagatgaacg agatgaaaag agagggcaaa 2700

ttcagggaga agcgcattaa gaggaacgaa cagagtctgc aggagatttg ggattacgtc 2760

aagaggccta acctgcggtt gatcggcgtc cccgagagcg acgtagaaaa cgggactaaa 2820

ctggagaata cacttcaaga catcattcaa gaaaattttc caaacctggc tcggcaagct 2880

aatgtgcaaa tccaagagat ccaacgcaca ccccagcggt atagctctcg gcgtgccacc 2940

cctaggcata ttatcgtgcg ctttactaag gtggagatga aagagaagat gctgcgagcc 3000

gctcgggaaa agggaagggt gactttgaag ggcaaaccta ttcggctgac ggttgacctt 3060

agcgccgaga cactccaggc acgccgggaa tggggcccca tctttaatat cctgaaggag 3120

aagaacttcc agccacgaat ctcttaccct gcaaagttga gttttatctc cgagggtgag 3180

attaagtatt tcatcgataa acagatgctg cgagacttcg tgacaactcg cccagctctc 3240

aaggaactgc tcaaagaggc tcttaatatg gagcgcaata atagatatca acccttgcag 3300
aaccacgcaa agatgaaaag gccggcggcc acgaaaaagg ccggccaggc aaaaaagaaa 3360
aagtgagaca gccgtcagac catcaagact aggaagaaac tgcatcaact aatgagcaaa 3420
atcaccagct aacatcatag tatacatgac cggctctaac tcacatatca ccatccttac 3480
acttaacatt aacggcctca actcagctat caagcgccat cggctggcca gctggatcaa 3540
atcacaggat ccaagcgttt gttgcatcca agagacccac ctgacctgta gagatactca 3600
ccgcctcaag atcaagggat ggcgaaagat ttatcaggcg aacggtaagc agaagaaagc 3660

cggagtcgca attctggtct cagacaagac ggatttcaag cccaccaaaa ttaagcgtga 3720
taaggaaggt cactatatta tggtgaaagg cagcatacag caggaagaac ttaccatatt 3780
gaacatctac gcgccaaaca ccggcgcacc tcgctttatc aaacaggtcc tgtccgatct 3840
gcagcgagat ctggattctc atacgttgat tatgggtgat ttcaatacac cattgagcac 3900
cctggatcgc agcaccaggc aaaaggtaaa taaagacacg caagagctca atagcgcact 3960
gcatcaggca gatctcattg atatttatcg cactcttcat cctaagagta ccgagtacac 4020
attcttcagc gccccacatc atacatactc aaagatcgat catatcgtcg gctcaaaggc 4080

tctgctgtca aagtgcaagc gcacagagat aattacaaat tacctgtcag atcatagcgc 4140
gatcaagctc gagctgagaa tcaagaacct gacccagagc cggagtacca cttggaagct 4200
taataacctg ctgctcaacg attattgggt ccacaatgag atgaaggcag agattaaaat 4260
gttcttcgaa acaaatgaga ataaggatac tacctatcaa aacctttggg atgcctttaa 4320
ggccgtctgc agaggcaagt tcatcgccct caacgcctat aaaagaaaac aagagagatc 4380
taagatcgat actctcacct ctcagctgaa ggagttggag aaacaggaac agacccactc 4440
caaggcgtca agacggcagg agatcacaaa gattcgcgcc gagttgaaag agatcgaaac 4500

ccaaaagact cttcagaaaa ttaacgagtc tcgtagttgg ttcttcgagc ggattaataa 4560
gatagacaga cctctggcac gactgattaa gaagaagcgc gaaaagaacc agattgatac 4620
catcaagaac gacaagggcg acatcactac tgacccgacc gagatccaga ccactattcg 4680
ggagtattat aagcatttgt atgctaacaa gcttgagaac ctggaagaga tggacacttt 4740
tctggatacc tatactctgc cacggcttaa tcaagaggaa gtcgagtccc tcaaccgccc 4800
aattacagga agcgagattg tggccataat taactccctg ccgacaaaga aatctcctgg 4860
tccggacggg tttacagctg agttttatca acggtatatg gaagagcttg taccgtttct 4920

gctcaagctc tttcagtcta tagaaaagga aggcatcttg cccaattcct tctacgaagc 4980

ttctataata cttattccca aaccaggacg cgataccaca aagaaggaaa acttccggcc 5040

cattagtctc atgaatatcg acgctaaaat attgaacaag attctcgcca acagaatcca 5100

acaacatatt aagaaattga tacatcacga ccaggtgggg tttatacctg gcatgcaggg 5160

ctggtttaac atccggaaga gtattaacgt cattcaacac attaatagag ctaaggataa 5220

gaatcatatg atcatctcta tagacgcgga aaaggcattc gataagattc agcagccatt 5280

tatgctcaag actctgaaca aactcggcat cgacggaaca tattttaaga ttattcgcgc 5340

aatttacgat aagccgactg ctaacattat ccttaacggc caaaagctcg aggcctttcc 5400

gctcaagact ggaacccgcc aaggctgtcc cctctccccg cttttgttta atattgtact 5460

cgaggtgctg gctagggcta ttcgtcaaga gaaagagatt aaagggatac agctcgggaa 5520

ggaagaggtc aagctttcct tgttcgccga tgatatgatt gtgtacctgg agaatcctat 5580

tgtgtctgct cagaaccttc ttaaacttat ttctaacttt agcaaggtca gcggctataa 5640

gattaacgtc cagaaatctc aggcctttct gtacacaaat aatcgacaga ccgaatccca 5700

gataatgggt gagcttccgt ttgtcatagc cagcaaaagg ataaagtatc tcggaatcca 5760

gctgacacga gacgttaaag atttgtttaa ggaaaattac aagcctctcc tgaaagagat 5820

taaggaagat actaataagt ggaagaatat cccctgttca tgggttggca gaatcaacat 5880

agtgaagatg gcaatacttc ctaaagtgat atatcgcttt aacgccatcc caattaaact 5940

gcctatgacc ttctttacgg agctcgagaa aacaaccctt aaatttatat ggaatcaaaa 6000

gagagcaaga atagcgaagt ccatcttgag ccagaagaat aaggccggtg ggattacttt 6060

gcctgatttt aagttgtatt ataaagccac agtaactaag acagcctggt attggtatca 6120

gaatagagac atcgaccagt ggaatcggac cgaaccatca gagataatgc cccacatcta 6180

taattacctt atattcgata agccagaaaa gaataaacag tggggcaaag acagcctctt 6240

caacaagtgg tgttgggaga attggctggc catatgccgg aaactcaagc tcgacccctt 6300

tcttacaccc tacactaaaa tcaacagtag gtggatcaag gacttgaatg tcaagccaaa 6360

gactataaag acactggaag agaatcttgg gatcacaata caagatatag gcgtcggcaa 6420

agattttatg tcaaagacgc ccaaggccat ggccactaag gataagattg ataagtggga 6480

ccttattaag ctcaaaagct tctgtactgc caaggagacc acgatcagag ttaataggca 6540

gcccactaca tgggaaaaga ttttcgccac ttattcatca gataaggggt tgataagcag 6600

aatatataac gagctgaagc agatctacaa gaagaaaacg aataatccca tcaagaagtg 6660

ggcaaaagat atgaacaggc attttagcaa agaggatatc tacgccgcga agaagcatat 6720

gaagaagtgt agttcaagct tggccattcg tgagatgcag attaagacga ccatgcgata 6780

ccaccttacc ccagtgagga tggcaattat caagaaatct ggcaataata gatgttggcg 6840

gggctgtggc gagattggca ccctgctcca ttgctggtgg gattgcaagc tggtgcagcc 6900

gctttggaaa tcagtctggc gctttctgag ggacctcgag cttgagattc ccttcgatcc 6960

cgcaattccc ttgctcggaa tctatcctaa cgaatacaag agctgttgtt acaaggatac 7020

gtgtacccgg atgttcatcg cggccttgtt tacgatagct aagacgtgga atcagcctaa 7080

gtgccccaca atgatcgatt ggatcaagaa aatgtggcat atttatacca tggagtatta 7140

cgcagcaatt aagaatgacg aatttatttc cttcgttggg acctggatga agctggagac 7200

tattattctg agcaagctgt ctcaggagca aaagacaaag catagaatct tctctctcat 7260

tggtggtaac gactacaaag acgatgacga caagtaaagc gcttctagaa gttgtctcct 7320

cctgcactga ctgactgata caatcgattt ctggatccgc aggcctaatc aacctctgga 7380

ttacaaaatt tgtgaaagat tgactggtat tcttaactat gttgctcctt ttacgctatg 7440

tggatacgct gctttaatgc ctttgtatca tgctattgct tcccgtatgg ctttcatttt 7500

ctcctccttg tataaatcct ggttgctgtc tctttatgag gagttgtggc ccgttgtcag 7560

gcaacgtggc gtggtgtgca ctgtgtttgc tgacgcaacc cccactggtt ggggcattgc 7620

caccacctgt cagctccttt ccgggacttt cgctttcccc ctccctattg ccacggcgga 7680

actcatcgcc gcctgccttg cccgctgctg gacaggggct cggctgttgg gcactgacaa 7740

ttccgtggtg ttgtcgggga agctgacgtc ctttccatgg ctgctcgcct gtgttgccac 7800

ctggattctg cgcgggacgt ccttctgcta cgtcccttcg gccctcaatc cagcggacct 7860

tccttcccgc tgagagacac aaaaaattcc aacacactat tgcaatgaaa ataaatttcc 7920

tttattagcc agaagtcaga tgctcaaggg gcttcatgat gtccccataa tttttggcag 7980

agggaaaaag atctcagtgg tatttgtgag ccagggcatt ggccttctga taggcagcct 8040

gcacctgagg agtgcggccg ctttacttgt acagctcgtc catgccgaga gtgatcccgg 8100

cggcggtcac gaactccagc aggaccatgt gatcgcgctt ctcgttgggg tctttgctca 8160

gggcggactg ggtgctcagg tagtggttgt cgggcagcag cacggggccg tcgccgatgg 8220

gggtgttctg ctggtagtgg tcggcgagct gcacgctgcc gtcctcgatg ttgtggcgga 8280

tcttgaagtt caccttgatg ccgttcttct gcttgtcggc catgatatag acgttgtggc 8340

tgttgtagtt gtactccagc ttgtgcccca ggatgttgcc gtcctccttg aagtcgatgc 8400

ccttcagctc gatgcggttc accagggtgt cgccctcgaa cttcacctcg gcgcgggtct 8460

tgtagttgcc gtcgtccttg aagaagatgg tgcgctcctg gacgtagcct tcgggcatgg 8520

cggacttgaa gaagtcgtgc tgcttcatgt ggtcggggta gcggctgaag cactgcacgc 8580

cgtaggtcag ggtggtcacg agggtgggcc agggcacggg cagcttgccg gtggtgcaga 8640

tgaacttcag ggtcagcttg ccgtaggtgg catcgccctc gccctcgccg gacacgctga 8700

acttgtggcc gtttacgtcg ccgtccagct cgaccaggat gggcaccacc ccggtgaaca 8760
gctcctcgcc cttgctcacc atggtggcgg gatctgacgg ttcactaaac cagctctgct 8820
tatatagacc tcccaccgta cacgcctacc gcccatttgc gtcaatgggg cggagttgtt 8880
acgacatttt ggaaagtccc gttgattttg gtgccaaaac aaactcccat tgacgtcaat 8940
ggggtggaga cttggaaatc cccgtgagtc aaaccgctat ccacgcccat tgatgtactg 9000
ccaaaaccgc atcaccatgg taatagcgat gactaatacg tagatgtact gccaagtagg 9060
aaagtcccat aaggtcatgt actgggcata atgccaggcg ggccatttac cgtcattgac 9120

gtcaataggg ggcgtacttg gcatatgata cacttgatgt actgccaagt gggcagttta 9180
ccgtaaatac tccacccatt gacgtcaatg gaaagtccct attggcgtta ctatgggaac 9240
atacgtcatt attgacgtca atgggcgggg gtcgttgggc ggtcagccag gcgggccatt 9300
taccgtaagt tatgtaacgg gcctgctgcc ggctctgcgg cctcttccgc gtcttcgcct 9360
tcgccctcag acgagtcgga tctccctttg ggccgcctcc ccgcctgtct agcttgactg 9420
actgagatac agcgtacctt cagctcacag acatgataag atacattgat gagtttggac 9480
aaaccacaac tagaatgcag tgaaaaaaat gctttatttg tgaaatttgt gatgctattg 9540

ctttatttgt aaccattata agctgcaata aacaagttaa caacaacaat tgcattcatt 9600
ttatgtttca ggttcagggg gaggtgtggg aggtttttta aagcaagtaa aacctctaca 9660
aatgtggtat tggcccatct ctatcggtat cgtagcataa ccccttgggg cctctaaacg 9720
ggtcttgagg ggtttttttgt gcccctcggg ccggattgct atctaccggc attggcgcag 9780
aaaaaaatgc ctgatgcgac gctgcgcgtc ttatactccc acatatgcca gattcagcaa 9840
cggatacggc ttccccaact tgcccacttc catacgtgtc ctccttacca gaaatttatc 9900
cttaaggtcg tcagctatcc tgcaggcgat ctctcgattt cgatcaagac attcctttaa 9960

tggtcttttc tggacaccac taggggtcag aagtagttca tcaaactttc ttccctccct 10020
aatctcattg gttaccttgg gctatcgaaa cttaattaag cgatctgcat ctcaattagt 10080
cagcaaccat agtcccgccc ctaactccgc ccatcccgcc cctaactccg cccagttccg 10140
cccattctcc gccccatcgc tgactaattt tttttattta tgcagaggcc gaggccgcct 10200
cggcctctga gctattccag aagtagtgag gaggcttttt tggaggccta ggcttttgca 10260
aaggaggtag ccaacatgat tgaacaagat ggattgcacg caggttctcc cgccgcttgg 10320
gtggagaggc tattcggcta tgactgggca caacagacaa tcggctgctc tgatgccgcc 10380

gtgttccggc tgtcagcgca ggggcgcccg gttctttttg tcaagaccga cctgtccggt 10440

gccctgaatg aactccagga cgaggcagcg cggctatcgt ggctggccac gacgggcgtt 10500

ccttgcgcag ctgtgctcga cgttgtcact gaagcgggaa gggactggct gctattgggc 10560

gaagtgccgg ggcaggatct cctgtcatct caccttgctc ctgccgagaa agtatccatc 10620

atggctgatg caatgcggcg gctgcatacg cttgatccgg ctacctgccc attcgaccac 10680

caagcgaaac atcgcatcga gcgagcacgt actcggatgg aagccggtct tgtcgatcag 10740

gatgatctgg acgaagagca tcaggggctc gcgccagccg aactgttcgc caggctcaag 10800

gcgcggatgc ccgacggcga ggatctcgtc gtgacccacg gcgatgcctg cttgccgaat 10860

atcatggtgg aaaatggccg cttttctgga ttcatcgact gtggccggct gggtgtggcg 10920

gaccgctatc aggacatagc gttggctacc cgtgatattg ctgaagagct tggcggcgaa 10980

tgggctgacc gcttcctcgt gctttacggt atcgccgctc ccgattcgca gcgcatcgcc 11040

ttctatcgcc ttcttgacga gttcttctag tatgtaagcc ctgtgccttc tagttgccag 11100

ccatctgttg tttgcccctc ccccgtgcct tccttgaccc tggaaggtgc cactcccact 11160

gtcctttcct aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt 11220

ctgggggggtg gggtggggca ggacagcaag ggggaggatt gggaagacaa tagcaggcat 11280

gctggggatg cggtgggctc tatggttaat taaccagtca agtcagctac ttggcgagat 11340

cgacttgtct gggtttcgac tacgctcaga attgcgtcag tcaagttcga tctggtcctt 11400

gctattgcac ccgttctccg attacgagtt tcatttaaat catgtgagca aaaggccagc 11460

aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgccccc 11520

ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat 11580

aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc 11640

cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct 11700

cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg 11760

aacccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc 11820

cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga 11880

ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa 11940

gaacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta 12000

gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttttgtt tgcaagcagc 12060

agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg 12120

acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga 12180

tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg 12240

agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct 12300

gtctatttcg ttcatccata gttgcattta aatttccgaa ctctccaagg ccctcgtcgg 12360

aaaatcttca aacctttcgt ccgatccatc ttgcaggcta cctctcgaac gaactatcgc 12420

aagtctcttg gccggccttg cgccttggct attgcttggc agcgcctatc gccaggtatt 12480

actccaatcc cgaatatccg agatcgggat cacccgagag aagttcaacc tacatcctca 12540

atcccgatct atccgagatc cgaggaata 12569

<210> 95

<211> 12581

<212> DNA

<213> Artificial Sequence

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

<400> 95

tatcgaaatc ggggcgcgcc tggtgtaccg agaacgatcc tctcagtgcg agtctcgacg 60

atccatatcg ttgcttggca gtcagccagt cggaatccag cttgggaccc aggaagtcca 120

atcgtcagat attgtactca agcctggtca cggcagcgta ccgatctgtt taaacctaga 180

tattgatagt ctgatcggtc aacgtataat cgagtcctag cttttgcaaa catctatcaa 240

gagacaggat cagcaggagg ctttcgcatg agtattcaac atttccgtgt cgcccttatt 300

ccctttttg cggcattttg ccttcctgtt tttgctcacc cagaaacgct ggtgaaagta 360

aaagatgctg aagatcagtt gggtgcgcga gtgggttaca tcgaactgga tctcaacagc 420

ggtaagatcc ttgagagttt tcgccccgaa gaacgctttc caatgatgag cacttttaaa 480

gttctgctat gtggcgcggt attatcccgt attgacgccg ggcaagagca actcggtcgc 540

cgcatacact attctcagaa tgacttggtt gagtattcac cagtcacaga aaagcatctt 600

acggatggca tgacagtaag agaattatgc agtgctgcca taaccatgag tgataacact 660

gcggccaact tacttctgac aacgattgga ggaccgaagg agctaaccgc ttttttgcac 720

aacatggggg atcatgtaac tcgccttgat cgttgggaac cggagctgaa tgaagccata 780

ccaaacgacg agcgtgacac cacgatgcct gtagcaatgg caacaacctt gcgtaaacta 840

ttaactggcg aactacttac tctagcttcc cggcaacagt tgatagactg gatggaggcg 900

gataaagttg caggaccact tctgcgctcg gcccttccgg ctggctggtt tattgctgat 960

aaatctggag ccggtgagcg tgggtctcgc ggtatcattg cagcactggg gccagatggt 1020

aagccctccc gtatcgtagt tatctacacg acggggagtc aggcaactat ggatgaacga 1080

aatagacaga tcgctgagat aggtgcctca ctgattaagc attggtaacc gattctaggt 1140

gcattggcgc agaaaaaaat gcctgatgcg acgctgcgcg tcttatactc ccacatatgc 1200

cagattcagc aacggatacg gcttccccaa cttgcccact tccatacgtg tcctccttac 1260

cagaaattta tccttaagat cgtttaaact cgactctggc tctatcgaat ctccgtcgtt 1320

tcgagcttac gcgaacagcc gtggcgctca tttgctcgtc gggcatcgaa tctcgtcagc 1380

tatcgtcagc ttacctttt ggcagcgatc gcggctcccg acatcttgga ccattagctc 1440

cacaggtatc ttcttccctc tagtggtcat aacagcagct tcagctacct ctcaattcaa 1500

aaaacccctc aagacccgtt tagaggcccc aaggggttat gctatcaatc gttgcgttac 1560

acacacaaaa aaccaacaca catccatctt cgatggatag cgattttatt atctaactgc 1620

tgatcgagtg tagccagatc tagtaatcaa ttacggggtc attagttcat agcccatata 1680

tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 1740

cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 1800

attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt 1860

atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 1920

atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 1980

tcgctattac catgctgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 2040

actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 2100

aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 2160

gtaggcgtgt acggtgggag gtctatataa gcagagctgg tttagtgaac cgtcagatca 2220

gatctttgtc gatcctacca tccactcgac acacccgcca gcggccgcta atacgactca 2280

ctatagggag aagtactgcc accatgggca agaagcaaaa tcgcaagacg gggaattcca 2340

agacacaatc cgctagccca ccacctaaag agcgttctag ctcccctgct actgagcagt 2400

cctggatgga aaacgacttc gatgaactcc gggaagaggg atttaggcga tccaactatt 2460

cagaactccg cgaagatatc cagacaaagg ggaaggaagt cgagaatttc gagaagaacc 2520

tcgaggagtg catcacccgt atcacaaaca ctgagaaatg tctcaaagaa ctcatggaac 2580

ttaagacaaa agccagggag cttcgagagg agtgtcggag tctgagatcc aggtgtgacc 2640

agctcgagga gcgcgtgagc gcgatggaag acgagatgaa cgagatgaaa agagagggca 2700

aattcaggga gaagcgcatt aagaggaacg aacagagtct gcaggagatt tgggattacg 2760

tcaagaggcc taacctgcgg ttgatcggcg tccccgagag cgacgtagaa aacgggacta 2820

aactggagaa tacacttcaa gacatcattc aagaaaattt tccaaacctg gctcggcaag 2880

ctaatgtgca aatccaagag atccaacgca caccccagcg gtatagctct cggcgtgcca 2940

ccctaggca tattatcgtg cgctttacta aggtggagat gaaagagaag atgctgcgag 3000

ccgctcggga aaagggaagg gtgactttga agggcaaacc tattcggctg acggttgacc 3060

ttagcgccga gacactccag gcacgccggg aatggggccc catctttaat atcctgaagg 3120

agaagaactt ccagccacga atctcttacc ctgcaaagtt gagttttatc tccgagggtg 3180

agattaagta tttcatcgat aaacagatgc tgcgagactt cgtgacaact cgcccagctc 3240

tcaaggaact gctcaaagag gctcttaata tggagcgcaa taatagatat caacccttgc 3300

agaaccacgc aaagatgggc ggcggcagca aaaggccggc ggccacgaaa aaggccggcc 3360

aggcaaaaaa gaaaaagtga gacagccgtc agaccatcaa gactaggaag aaactgcatc 3420

aactaatgag caaaatcacc agctaacatc atagtataca tgaccggctc taactcacat 3480

atcaccatcc ttacacttaa cattaacggc ctcaactcag ctatcaagcg ccatcggctg 3540

gccagctgga tcaaatcaca ggatccaagc gtttgttgca tccaagagac ccacctgacc 3600

tgtagagata ctcaccgcct caagatcaag ggatggcgaa agatttatca ggcgaacggt 3660

aagcagaaga aagccggagt cgcaattctg gtctcagaca agacggattt caagcccacc 3720

aaaattaagc gtgataagga aggtcactat attatggtga aaggcagcat acagcaggaa 3780

gaacttacca tattgaacat ctacgcgcca aacaccggcg cacctcgctt tatcaaacag 3840

gtcctgtccg atctgcagcg agatctggat tctcatacgt tgattatggg tgatttcaat 3900

acaccattga gcaccctgga tcgcagcacc aggcaaaagg taaataaaga cacgcaagag 3960

ctcaatagcg cactgcatca ggcagatctc attgatattt atcgcactct tcatcctaag 4020

agtaccgagt acacattctt cagcgcccca catcatacat actcaaagat cgatcatatc 4080

gtcggctcaa aggctctgct gtcaaagtgc aagcgcacag agataattac aaattacctg 4140

tcagatcata gcgcgatcaa gctcgagctg agaatcaaga acctgaccca gagccggagt 4200

accacttgga agcttaataa cctgctgctc aacgattatt gggtccacaa tgagatgaag 4260

gcagagatta aaatgttctt cgaaacaaat gagaataagg atactaccta tcaaaacctt 4320

tgggatgcct ttaaggccgt ctgcagaggc aagttcatcg ccctcaacgc ctataaaaga 4380

aaacaagaga gatctaagat cgatactctc acctctcagc tgaaggagtt ggagaaacag 4440

gaacagaccc actccaaggc gtcaagacgg caggagatca caaagattcg cgccgagttg 4500

aaagagatcg aaacccaaaa gactcttcag aaaattaacg agtctcgtag ttggttcttc 4560

gagcggatta ataagataga cagacctctg gcacgactga ttaagaagaa gcgcgaaaag 4620

aaccagattg ataccatcaa gaacgacaag ggcgacatca ctactgaccc gaccgagatc 4680

cagaccacta ttcgggagta ttataagcat ttgtatgcta acaagcttga gaacctggaa 4740

gagatggaca cttttctgga tacctatact ctgccacggc ttaatcaaga ggaagtcgag 4800

tccctcaacc gcccaattac aggaagcgag attgtggcca taattaactc cctgccgaca 4860
aagaaatctc ctggtccgga cgggtttaca gctgagtttt atcaacggta tatggaagag 4920
cttgtaccgt ttctgctcaa gctctttcag tctatagaaa aggaaggcat cttgcccaat 4980
tccttctacg aagcttctat aatacttatt cccaaaccag gacgcgatac cacaaagaag 5040
gaaaacttcc ggcccattag tctcatgaat atcgacgcta aaatattgaa caagattctc 5100
gccaacagaa tccaacaaca tattaagaaa ttgatacatc acgaccaggt ggggtttata 5160
cctggcatgc agggctggtt taacatccgg aagagtatta acgtcattca acacattaat 5220

agagctaagg ataagaatca tatgatcatc tctatagacg cggaaaaggc attcgataag 5280
attcagcagc catttatgct caagactctg aacaaactcg gcatcgacgg aacatatttt 5340
aagattattc gcgcaattta cgataagccg actgctaaca ttatccttaa cggccaaaag 5400
ctcgaggcct ttccgctcaa gactggaacc cgccaaggct gtcccctctc cccgcttttg 5460
tttaatattg tactcgaggt gctggctagg gctattcgtc aagagaaaga gattaaaggg 5520
atacagctcg ggaaggaaga ggtcaagctt tccttgttcg ccgatgatat gattgtgtac 5580
ctggagaatc ctattgtgtc tgctcagaac cttcttaaac ttatttctaa ctttagcaag 5640

gtcagcggct ataagattaa cgtccagaaa tctcaggcct ttctgtacac aaataatcga 5700
cagaccgaat cccagataat gggtgagctt ccgtttgtca tagccagcaa aaggataaag 5760
tatctcggaa tccagctgac acgagacgtt aaagatttgt ttaaggaaaa ttacaagcct 5820
ctcctgaaag agattaagga agatactaat aagtggaaga atatcccctg ttcatgggtt 5880
ggcagaatca acatagtgaa gatggcaata cttcctaaag tgatatatcg ctttaacgcc 5940
atcccaatta aactgcctat gaccttcttt acggagctcg agaaaacaac ccttaaattt 6000
atatggaatc aaaagagagc aagaatagcg aagtccatct tgagccagaa gaataaggcc 6060

ggtgggatta ctttgcctga ttttaagttg tattataaag ccacagtaac taagacagcc 6120
tggtattggt atcagaatag agacatcgac cagtggaatc ggaccgaacc atcagagata 6180
atgccccaca tctataatta ccttatattc gataagccag aaaagaataa acagtggggc 6240
aaagacagcc tcttcaacaa gtggtgttgg gagaattggc tggccatatg ccggaaactc 6300
aagctcgacc cctttcttac accctacact aaaatcaaca gtaggtggat caaggacttg 6360
aatgtcaagc caaagactat aaagacactg gaagagaatc ttgggatcac aatacaagat 6420
ataggcgtcg gcaaagattt tatgtcaaag acgcccaagg ccatggccac taaggataag 6480

attgataagt gggaccttat taagctcaaa agcttctgta ctgccaagga gaccacgatc 6540

agagttaata ggcagcccac tacatgggaa aagattttcg ccacttattc atcagataag 6600

gggttgataa gcagaatata taacgagctg aagcagatct acaagaagaa aacgaataat 6660

cccatcaaga agtgggcaaa agatatgaac aggcatttta gcaaagagga tatctacgcc 6720

gcgaagaagc atatgaagaa gtgtagttca agcttggcca ttcgtgagat gcagattaag 6780

acgaccatgc gataccacct taccccagtg aggatggcaa ttatcaagaa atctggcaat 6840

aatagatgtt ggcggggctg tggcgagatt ggcaccctgc tccattgctg gtgggattgc 6900

aagctggtgc agccgctttg gaaatcagtc tggcgctttc tgagggacct cgagcttgag 6960

attcccttcg atcccgcaat tcccttgctc ggaatctatc ctaacgaata caagagctgt 7020

tgttacaagg atacgtgtac ccggatgttc atcgcggcct tgtttacgat agctaagacg 7080

tggaatcagc ctaagtgccc cacaatgatc gattggatca agaaaatgtg gcatatttat 7140

accatggagt attacgcagc aattaagaat gacgaattta tttccttcgt tgggacctgg 7200

atgaagctgg agactattat tctgagcaag ctgtctcagg agcaaaagac aaagcataga 7260

atcttctctc tcattggtgg taacgactac aaagacgatg acgacaagta aagcgcttct 7320

agaagttgtc tcctcctgca ctgactgact gatacaatcg atttctggat ccgcaggcct 7380

aatcaacctc tggattacaa aatttgtgaa agattgactg gtattcttaa ctatgttgct 7440

ccttttacgc tatgtggata cgctgcttta atgcctttgt atcatgctat tgcttcccgt 7500

atggctttca ttttctcctc cttgtataaa tcctggttgc tgtctcttta tgaggagttg 7560

tggcccgttg tcaggcaacg tggcgtggtg tgcactgtgt ttgctgacgc aacccccact 7620

ggttggggca ttgccaccac ctgtcagctc ctttccggga ctttcgcttt ccccctccct 7680

attgccacgg cggaactcat cgccgcctgc cttgcccgct gctggacagg ggctcggctg 7740

ttgggcactg acaattccgt ggtgttgtcg gggaagctga cgtcctttcc atggctgctc 7800

gcctgtgttg ccacctggat tctgcgcggg acgtccttct gctacgtccc ttcggccctc 7860

aatccagcgg accttccttc ccgctgagag acacaaaaaa ttccaacaca ctattgcaat 7920

gaaaataaat ttcctttatt agccagaagt cagatgctca aggggcttca tgatgtcccc 7980

ataatttttg gcagagggaa aaagatctca gtggtatttg tgagccaggg cattggcctt 8040

ctgataggca gcctgcacct gaggagtgcg gccgctttac ttgtacagct cgtccatgcc 8100

gagagtgatc ccggcggcgg tcacgaactc cagcaggacc atgtgatcgc gcttctcgtt 8160

ggggtctttg ctcagggcgg actgggtgct caggtagtgg ttgtcgggca gcagcacggg 8220

gccgtcgccg atgggggtgt tctgctggta gtggtcggcg agctgcacgc tgccgtcctc 8280

gatgttgtgg cggatcttga agttcacctt gatgccgttc ttctgcttgt cggccatgat 8340

atagacgttg tggctgttgt agttgtactc cagcttgtgc cccaggatgt tgccgtcctc 8400

cttgaagtcg atgcccttca gctcgatgcg gttcaccagg gtgtcgccct cgaacttcac 8460

ctcggcgcgg gtcttgtagt tgccgtcgtc cttgaagaag atggtgcgct cctggacgta 8520

gccttcgggc atggcggact tgaagaagtc gtgctgcttc atgtggtcgg ggtagcggct 8580

gaagcactgc acgccgtagg tcagggtggt cacgagggtg ggccagggca cgggcagctt 8640

gccggtggtg cagatgaact tcagggtcag cttgccgtag gtggcatcgc cctcgccctc 8700

gccggacacg ctgaacttgt ggccgtttac gtcgccgtcc agctcgacca ggatgggcac 8760

caccccggtg aacagctcct cgcccttgct caccatggtg gcgggatctg acggttcact 8820

aaaccagctc tgcttatata gacctcccac cgtacacgcc taccgcccat ttgcgtcaat 8880

ggggcggagt tgttacgaca ttttggaaag tcccgttgat tttggtgcca aaacaaactc 8940

ccattgacgt caatggggtg gagacttgga aatccccgtg agtcaaaccg ctatccacgc 9000

ccattgatgt actgccaaaa ccgcatcacc atggtaatag cgatgactaa tacgtagatg 9060

tactgccaag taggaaagtc ccataaggtc atgtactggg cataatgcca ggcgggccat 9120

ttaccgtcat tgacgtcaat aggggggcgta cttggcatat gatacacttg atgtactgcc 9180

aagtgggcag tttaccgtaa atactccacc cattgacgtc aatggaaagt ccctattggc 9240

gttactatgg gaacatacgt cattattgac gtcaatgggc gggggtcgtt gggcggtcag 9300

ccaggcgggc catttaccgt aagttatgta acgggcctgc tgccggctct gcggcctctt 9360

ccgcgtcttc gccttcgccc tcagacgagt cggatctccc tttgggccgc ctccccgcct 9420

gtctagcttg actgactgag atacagcgta ccttcagctc acagacatga taagatacat 9480

tgatgagttt ggacaaacca caactagaat gcagtgaaaa aaatgcttta tttgtgaaat 9540

ttgtgatgct attgctttat ttgtaaccat tataagctgc aataaacaag ttaacaacaa 9600

caattgcatt cattttatgt ttcaggttca gggggaggtg tgggaggttt tttaaagcaa 9660

gtaaaacctc tacaaatgtg gtattggccc atctctatcg gtatcgtagc ataacccctt 9720

ggggcctcta aacgggtctt gaggggtttt ttgtgcccct cgggccggat tgctatctac 9780

cggcattggc gcagaaaaaa atgcctgatg cgacgctgcg cgtcttatac tcccacatat 9840

gccagattca gcaacggata cggcttcccc aacttgccca cttccatacg tgtcctcctt 9900

accagaaatt tatccttaag gtcgtcagct atcctgcagg cgatctctcg atttcgatca 9960

agacattcct ttaatggtct tttctggaca ccactagggg tcagaagtag ttcatcaaac 10020

tttcttccct ccctaatctc attggttacc ttgggctatc gaaacttaat taagcgatct 10080

gcatctcaat tagtcagcaa ccatagtccc gcccctaact ccgcccatcc cgcccctaac 10140

tccgcccagt tccgcccatt ctccgcccca tcgctgacta atttttttta tttatgcaga 10200

ggccgaggcc gcctcggcct ctgagctatt ccagaagtag tgaggaggct tttttggagg 10260

cctaggcttt tgcaaaggag gtagccaaca tgattgaaca agatggattg cacgcaggtt 10320

ctcccgccgc ttgggtggag aggctattcg gctatgactg ggcacaacag acaatcggct 10380

gctctgatgc cgccgtgttc cggctgtcag cgcaggggcg cccggttctt tttgtcaaga 10440

ccgacctgtc cggtgccctg aatgaactcc aggacgaggc agcgcggcta tcgtggctgg 10500

ccacgacggg cgttccttgc gcagctgtgc tcgacgttgt cactgaagcg ggaagggact 10560

ggctgctatt gggcgaagtg ccggggcagg atctcctgtc atctcacctt gctcctgccg 10620

agaaagtatc catcatggct gatgcaatgc ggcggctgca tacgcttgat ccggctacct 10680

gcccattcga ccaccaagcg aaacatcgca tcgagcgagc acgtactcgg atggaagccg 10740

gtcttgtcga tcaggatgat ctggacgaag agcatcaggg gctcgcgcca gccgaactgt 10800

tcgccaggct caaggcgcgg atgcccgacg gcgaggatct cgtcgtgacc cacggcgatg 10860

cctgcttgcc gaatatcatg gtggaaaatg gccgcttttc tggattcatc gactgtggcc 10920

ggctgggtgt ggcggaccgc tatcaggaca tagcgttggc tacccgtgat attgctgaag 10980

agcttggcgg cgaatgggct gaccgcttcc tcgtgcttta cggtatcgcc gctcccgatt 11040

cgcagcgcat cgccttctat cgccttcttg acgagttctt ctagtatgta agccctgtgc 11100

cttctagttg ccagccatct gttgtttgcc cctcccccgt gccttccttg accctggaag 11160

gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcatcgcat tgtctgagta 11220

ggtgtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag gattgggaag 11280

acaatagcag gcatgctggg gatgcggtgg gctctatggt taattaacca gtcaagtcag 11340

ctacttggcg agatcgactt gtctgggttt cgactacgct cagaattgcg tcagtcaagt 11400

tcgatctggt ccttgctatt gcacccgttc tccgattacg agtttcattt aaatcatgtg 11460

agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca 11520

taggctccgc ccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 11580

cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc 11640

tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc 11700

gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct 11760

gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg 11820

tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag 11880

gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta 11940

cggctacact agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg 12000

aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt 12060

tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt 12120

ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag 12180

attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat 12240

ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc 12300

tatctcagcg atctgtctat ttcgttcatc catagttgca tttaaatttc cgaactctcc 12360

aaggccctcg tcggaaaatc ttcaaacctt tcgtccgatc catcttgcag gctacctctc 12420

gaacgaacta tcgcaagtct cttggccggc cttgcgcctt ggctattgct tggcagcgcc 12480

tatcgccagg tattactcca atcccgaata tccgagatcg ggatcacccg agagaagttc 12540

aacctacatc ctcaatcccg atctatccga gatccgagga a 12581

<210> 96

<211> 12542

<212> DNA

<213> Artificial Sequence

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

<400> 96

gtcggaaaat cttcaaacct ttcgtccgat ccatcttgca ggctacctct cgaacgaact 60

atcgcaagtc tcttggccgg ccttgcgcct tggctattgc ttggcagcgc ctatcgccag 120

gtattactcc aatcccgaat atccgagatc gggatcaccc gagagaagtt caacctacat 180

cctcaatccc gatctatccg agatccgagg aatatcgaaa tcggggcgcg cctggtgtac 240

cgagaacgat cctctcagtg cgagtctcga cgatccatat cgttgcttgg cagtcagcca 300

gtcggaatcc agcttgggac ccaggaagtc caatcgtcag atattgtact caagcctggt 360

cacggcagcg taccgatctg tttaaaccta gatattgata gtctgatcgg tcaacgtata 420

atcgagtcct agcttttgca aacatctatc aagagacagg atcagcagga ggctttcgca 480

tgagtattca acatttccgt gtcgccctta ttccctttt tgcggcattt tgccttcctg 540

tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcgc 600

gagtgggtta catcgaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg 660

aagaacgctt tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc 720

gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg 780

ttgagtattc accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat 840

gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgattg 900

gaggaccgaa ggagctaacc gctttttttgc acaacatggg ggatcatgta actcgccttg 960
atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc 1020
ctgtagcaat ggcaacaacc ttgcgtaaac tattaactgg cgaactactt actctagctt 1080
cccggcaaca gttgatagac tggatggagg cggataaagt tgcaggacca cttctgcgct 1140
cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc 1200
gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca 1260
cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct 1320

cactgattaa gcattggtaa ccgattctag gtgcattggc gcagaaaaaa atgcctgatg 1380
cgacgctgcg cgtcttatac tcccacatat gccagattca gcaacggata cggcttcccc 1440
aacttgccca cttccatacg tgtcctcctt accagaaatt tatccttaag atcgtttaaa 1500
ctcgactctg gctctatcga atctccgtcg tttcgagctt acgcgaacag ccgtggcgct 1560
catttgctcg tcgggcatcg aatctcgtca gctatcgtca gcttaccttt ttggcagcga 1620
tcgcggctcc cgacatcttg gaccattagc tccacaggta tcttcttccc tctagtggtc 1680
ataacagcag cttcagctac ctctcaattc aaaaaacccc tcaagacccg tttagaggcc 1740

ccaaggggtt atgctatcaa tcgttgcgtt acacacacaa aaaaccaaca cacatccatc 1800
ttcgatggat agcgatttta ttatctaact gctgatcgag tgtagccaga tctagtaatc 1860
aattacgggg tcattagttc atagcccata tatggagttc cgcgttacat aacttacggt 1920
aaatggcccg cctggctgac cgcccaacga cccccgccca ttgacgtcaa taatgacgta 1980
tgttcccata gtaacgccaa tagggacttt ccattgacgt caatgggtgg agtatttacg 2040
gtaaactgcc cacttggcag tacatcaagt gtatcatatg ccaagtacgc cccctattga 2100
cgtcaatgac ggtaaatggc ccgcctggca ttatgcccag tacatgacct tatgggactt 2160

tcctacttgg cagtacatct acgtattagt catcgctatt accatgctga tgcggttttg 2220
gcagtacatc aatgggcgtg gatagcggtt tgactcacgg ggatttccaa gtctccaccc 2280
cattgacgtc aatgggagtt tgttttggca ccaaaatcaa cgggactttc caaaatgtcg 2340
taacaactcc gccccattga cgcaaatggg cggtaggcgt gtacggtggg aggtctatat 2400
aagcagagct ggtttagtga accgtcagat cagatctttg tcgatcctac catccactcg 2460
acacacccgc cagcggccgc taatacgact cactataggg agaagtactg ccaccatgcc 2520
aaagaagaag cggaaggtcg gcaagaagca aaatcgcaag acggggaatt ccaagacaca 2580

atccgctagc ccaccaccta aagagcgttc tagctcccct gctactgagc agtcctggat 2640

ggaaaacgac ttcgatgaac tccgggaaga gggatttagg cgatccaact attcagaact 2700

ccgcgaagat atccagacaa aggggaagga agtcgagaat ttcgagaaga acctcgagga 2760

gtgcatcacc cgtatcacaa acactgagaa atgtctcaaa gaactcatgg aacttaagac 2820

aaaagccagg gagcttcgag aggagtgtcg gagtctgaga tccaggtgtg accagctcga 2880

ggagcgcgtg agcgcgatgg aagacgagat gaacgagatg aaaagagagg gcaaattcag 2940

ggagaagcgc attaagagga acgaacagag tctgcaggag atttgggatt acgtcaagag 3000

gcctaacctg cggttgatcg gcgtccccga gagcgacgta gaaaacggga ctaaactgga 3060

gaatacactt caagacatca ttcaagaaaa ttttccaaac ctggctcggc aagctaatgt 3120

gcaaatccaa gagatccaac gcacacccca gcggtatagc tctcggcgtg ccacccctag 3180

gcatattatc gtgcgcttta ctaaggtgga gatgaaagag aagatgctgc gagccgctcg 3240

ggaaaaggga agggtgactt tgaagggcaa acctattcgg ctgacggttg accttagcgc 3300

cgagacactc caggcacgcc gggaatgggg ccccatcttt aatatcctga aggagaagaa 3360

cttccagcca cgaatctctt accctgcaaa gttgagtttt atctccgagg gtgagattaa 3420

gtatttcatc gataaacaga tgctgcgaga cttcgtgaca actcgcccag ctctcaagga 3480

actgctcaaa gaggctctta atatggagcg caataataga tatcaaccct tgcagaacca 3540

cgcaaagatg tgagacagcc gtcagaccat caagactagg aagaaactgc atcaactaat 3600

gagcaaaatc accagctaac atcatagtat acatgaccgg ctctaactca catatcacca 3660

tccttacact taacattaac ggcctcaact cagctatcaa gcgccatcgg ctggccagct 3720

ggatcaaatc acaggatcca agcgtttgtt gcatccaaga gacccacctg acctgtagag 3780

atactcaccg cctcaagatc aagggatggc gaaagattta tcaggcgaac ggtaagcaga 3840

agaaagccgg agtcgcaatt ctggtctcag acaagacgga tttcaagccc accaaaatta 3900

agcgtgataa ggaaggtcac tatattatgg tgaaaggcag catacagcag gaagaactta 3960

ccatattgaa catctacgcg ccaaacaccg gcgcacctcg ctttatcaaa caggtcctgt 4020

ccgatctgca gcgagatctg gattctcata cgttgattat gggtgatttc aatacaccat 4080

tgagcaccct ggatcgcagc accaggcaaa aggtaaataa agacacgcaa gagctcaata 4140

gcgcactgca tcaggcagat ctcattgata tttatcgcac tcttcatcct aagagtaccg 4200

agtacacatt cttcagcgcc ccacatcata catactcaaa gatcgatcat atcgtcggct 4260

caaaggctct gctgtcaaag tgcaagcgca cagagataat tacaaattac ctgtcagatc 4320

atagcgcgat caagctcgag ctgagaatca agaacctgac ccagagccgg agtaccactt 4380

ggaagcttaa taacctgctg ctcaacgatt attgggtcca caatgagatg aaggcagaga 4440

ttaaaatgtt cttcgaaaca aatgagaata aggatactac ctatcaaaac ctttgggatg 4500

cctttaaggc cgtctgcaga ggcaagttca tcgccctcaa cgcctataaa agaaaacaag 4560

agagatctaa gatcgatact ctcacctctc agctgaagga gttggagaaa caggaacaga 4620

cccactccaa ggcgtcaaga cggcaggaga tcacaaagat tcgcgccgag ttgaaagaga 4680

tcgaaaccca aaagactctt cagaaaatta acgagtctcg tagttggttc ttcgagcgga 4740

ttaataagat agacagacct ctggcacgac tgattaagaa gaagcgcgaa aagaaccaga 4800

ttgataccat caagaacgac aagggcgaca tcactactga cccgaccgag atccagacca 4860

ctattcggga gtattataag catttgtatg ctaacaagct tgagaacctg gaagagatgg 4920

acacttttct ggatacctat actctgccac ggcttaatca agaggaagtc gagtccctca 4980

accgcccaat tacaggaagc gagattgtgg ccataattaa ctccctgccg acaaagaaat 5040

ctcctggtcc ggacgggttt acagctgagt tttatcaacg gtatatggaa gagcttgtac 5100

cgtttctgct caagctcttt cagtctatag aaaaggaagg catcttgccc aattccttct 5160

acgaagcttc tataatactt attcccaaac caggacgcga taccacaaag aaggaaaact 5220

tccggcccat tagtctcatg aatatcgacg ctaaaatatt gaacaagatt ctcgccaaca 5280

gaatccaaca acatattaag aaattgatac atcacgacca ggtggggttt atacctggca 5340

tgcagggctg gtttaacatc cggaagagta ttaacgtcat tcaacacatt aatagagcta 5400

aggataagaa tcatatgatc atctctatag acgcggaaaa ggcattcgat aagattcagc 5460

agccatttat gctcaagact ctgaacaaac tcggcatcga cggaacatat tttaagatta 5520

ttcgcgcaat ttacgataag ccgactgcta acattatcct taacggccaa aagctcgagg 5580

cctttccgct caagactgga acccgccaag gctgtcccct ctccccgctt ttgtttaata 5640

ttgtactcga ggtgctggct agggctattc gtcaagagaa agagattaaa gggatacagc 5700

tcgggaagga agaggtcaag ctttccttgt tcgccgatga tatgattgtg tacctggaga 5760

atcctattgt gtctgctcag aaccttctta aacttatttc taactttagc aaggtcagcg 5820

gctataagat taacgtccag aaatctcagg cctttctgta cacaaataat cgacagaccg 5880

aatcccagat aatgggtgag cttccgtttg tcatagccag caaaaggata aagtatctcg 5940

gaatccagct gacacgagac gttaaagatt tgtttaagga aaattacaag cctctcctga 6000

aagagattaa ggaagatact aataagtgga agaatatccc ctgttcatgg gttggcagaa 6060

tcaacatagt gaagatggca atacttccta aagtgatata tcgctttaac gccatcccaa 6120

ttaaactgcc tatgaccttc tttacggagc tcgagaaaac aacccttaaa tttatatgga 6180

atcaaaagag agcaagaata gcgaagtcca tcttgagcca gaagaataag gccggtggga 6240

ttactttgcc tgattttaag ttgtattata aagccacagt aactaagaca gcctggtatt 6300

ggtatcagaa tagagacatc gaccagtgga atcggaccga accatcagag ataatgcccc 6360

acatctataa ttaccttata ttcgataagc cagaaaagaa taaacagtgg ggcaaagaca 6420
gcctcttcaa caagtggtgt tgggagaatt ggctggccat atgccggaaa ctcaagctcg 6480
acccctttct tacaccctac actaaaatca acagtaggtg gatcaaggac ttgaatgtca 6540
agccaaagac tataaagaca ctggaagaga atcttgggat cacaatacaa gatataggcg 6600
tcggcaaaga ttttatgtca aagacgccca aggccatggc cactaaggat aagattgata 6660
agtgggacct tattaagctc aaaagcttct gtactgccaa ggagaccacg atcagagtta 6720
ataggcagcc cactacatgg gaaaagattt tcgccactta ttcatcagat aaggggttga 6780

taagcagaat atataacgag ctgaagcaga tctacaagaa gaaaacgaat aatcccatca 6840
agaagtgggc aaaagatatg aacaggcatt ttagcaaaga ggatatctac gccgcgaaga 6900
agcatatgaa gaagtgtagt tcaagcttgg ccattcgtga gatgcagatt aagacgacca 6960
tgcgatacca ccttacccca gtgaggatgg caattatcaa gaaatctggc aataatagat 7020
gttggcgggg ctgtggcgag attggcaccc tgctccattg ctggtgggat tgcaagctgg 7080
tgcagccgct ttggaaatca gtctggcgct ttctgaggga cctcgagctt gagattccct 7140
tcgatcccgc aattcccttg ctcggaatct atcctaacga atacaagagc tgttgttaca 7200

aggatacgtg tacccggatg ttcatcgcgg ccttgtttac gatagctaag acgtggaatc 7260
agcctaagtg ccccacaatg atcgattgga tcaagaaaat gtggcatatt tataccatgg 7320
agtattacgc agcaattaag aatgacgaat ttatttcctt cgttgggacc tggatgaagc 7380
tggagactat tattctgagc aagctgtctc aggagcaaaa gacaaagcat agaatcttct 7440
ctctcattgg tggtaacgac tacaaagacg atgacgacaa gtaaagcgct tctagaagtt 7500
gtctcctcct gcactgactg actgatacaa tcgatttctg gatccgcagg cctaatcaac 7560
ctctggatta caaaatttgt gaaagattga ctggtattct taactatgtt gctcctttta 7620

cgctatgtgg atacgctgct ttaatgcctt tgtatcatgc tattgcttcc cgtatggctt 7680
tcattttctc ctccttgtat aaatcctggt tgctgtctct ttatgaggag ttgtggcccg 7740
ttgtcaggca acgtggcgtg gtgtgcactg tgtttgctga cgcaaccccc actggttggg 7800
gcattgccac cacctgtcag ctcctttccg ggactttcgc tttccccctc cctattgcca 7860
cggcggaact catcgccgcc tgccttgccc gctgctggac aggggctcgg ctgttgggca 7920
ctgacaattc cgtggtgttg tcggggaagc tgacgtcctt tccatggctg ctcgcctgtg 7980
ttgccacctg gattctgcgc gggacgtcct tctgctacgt cccttcggcc ctcaatccag 8040

cggaccttcc ttcccgctga gagacacaaa aaattccaac acactattgc aatgaaaata 8100

aatttccttt attagccaga agtcagatgc tcaaggggct tcatgatgtc cccataattt 8160

ttggcagagg gaaaaagatc tcagtggtat ttgtgagcca gggcattggc cttctgatag 8220

gcagcctgca cctgaggagt gcggccgctt tacttgtaca gctcgtccat gccgagagtg 8280

atcccggcgg cggtcacgaa ctccagcagg accatgtgat cgcgcttctc gttggggtct 8340

ttgctcaggg cggactgggt gctcaggtag tggttgtcgg gcagcagcac ggggccgtcg 8400

ccgatggggg tgttctgctg gtagtggtcg gcgagctgca cgctgccgtc ctcgatgttg 8460

tggcggatct tgaagttcac cttgatgccg ttcttctgct tgtcggccat gatatagacg 8520

ttgtggctgt tgtagttgta ctccagcttg tgccccagga tgttgccgtc ctccttgaag 8580

tcgatgccct tcagctcgat gcggttcacc agggtgtcgc cctcgaactt cacctcggcg 8640

cgggtcttgt agttgccgtc gtccttgaag aagatggtgc gctcctggac gtagccttcg 8700

ggcatggcgg acttgaagaa gtcgtgctgc ttcatgtggt cggggtagcg gctgaagcac 8760

tgcacgccgt aggtcagggt ggtcacgagg gtgggccagg gcacgggcag cttgccggtg 8820

gtgcagatga acttcagggt cagcttgccg taggtggcat cgccctcgcc ctcgccggac 8880

acgctgaact tgtggccgtt tacgtcgccg tccagctcga ccaggatggg caccaccccg 8940

gtgaacagct cctcgccctt gctcaccatg gtggcgggat ctgacggttc actaaaccag 9000

ctctgcttat atagacctcc caccgtacac gcctaccgcc catttgcgtc aatggggcgg 9060

agttgttacg acattttgga aagtcccgtt gattttggtg ccaaaacaaa ctcccattga 9120

cgtcaatggg gtggagactt ggaaatcccc gtgagtcaaa ccgctatcca cgcccattga 9180

tgtactgcca aaaccgcatc accatggtaa tagcgatgac taatacgtag atgtactgcc 9240

aagtaggaaa gtcccataag gtcatgtact gggcataatg ccaggcgggc catttaccgt 9300

cattgacgtc aataggggggc gtacttggca tatgatacac ttgatgtact gccaagtggg 9360

cagtttaccg taaatactcc acccattgac gtcaatggaa agtccctatt ggcgttacta 9420

tgggaacata cgtcattatt gacgtcaatg ggcgggggtc gttgggcggt cagccaggcg 9480

ggccatttac cgtaagttat gtaacgggcc tgctgccggc tctgcggcct cttccgcgtc 9540

ttcgccttcg ccctcagacg agtcggatct ccctttgggc cgcctccccg cctgtctagc 9600

ttgactgact gagatacagc gtaccttcag ctcacagaca tgataagata cattgatgag 9660

tttggacaaa ccacaactag aatgcagtga aaaaaatgct ttatttgtga aatttgtgat 9720

gctattgctt tatttgtaac cattataagc tgcaataaac aagttaacaa caacaattgc 9780

attcatttta tgtttcaggt tcagggggag gtgtgggagg ttttttaaag caagtaaaac 9840

ctctacaaat gtggtattgg cccatctcta tcggtatcgt agcataaccc cttggggcct 9900

ctaaacgggt cttgaggggt tttttgtgcc cctcgggccg gattgctatc taccggcatt 9960

ggcgcagaaa aaaatgcctg atgcgacgct gcgcgtctta tactcccaca tatgccagat 10020

tcagcaacgg atacggcttc cccaacttgc ccacttccat acgtgtcctc cttaccagaa 10080

atttatcctt aaggtcgtca gctatcctgc aggcgatctc tcgatttcga tcaagacatt 10140

cctttaatgg tcttttctgg acaccactag gggtcagaag tagttcatca aactttcttc 10200

cctccctaat ctcattggtt accttgggct atcgaaactt aattaagcga tctgcatctc 10260

aattagtcag caaccatagt cccgccccta actccgccca tcccgcccct aactccgccc 10320

agttccgccc attctccgcc ccatcgctga ctaatttttt ttatttatgc agaggccgag 10380

gccgcctcgg cctctgagct attccagaag tagtgaggag gctttttttgg aggcctaggc 10440

ttttgcaaag gaggtagcca acatgattga acaagatgga ttgcacgcag gttctcccgc 10500

cgcttgggtg gagaggctat tcggctatga ctgggcacaa cagacaatcg gctgctctga 10560

tgccgccgtg ttccggctgt cagcgcaggg gcgcccggtt ctttttgtca agaccgacct 10620

gtccggtgcc ctgaatgaac tccaggacga ggcagcgcgg ctatcgtggc tggccacgac 10680

gggcgttcct tgcgcagctg tgctcgacgt tgtcactgaa gcgggaaggg actggctgct 10740

attgggcgaa gtgccggggc aggatctcct gtcatctcac cttgctcctg ccgagaaagt 10800

atccatcatg gctgatgcaa tgcggcggct gcatacgctt gatccggcta cctgcccatt 10860

cgaccaccaa gcgaaacatc gcatcgagcg agcacgtact cggatggaag ccggtcttgt 10920

cgatcaggat gatctggacg aagagcatca ggggctcgcg ccagccgaac tgttcgccag 10980

gctcaaggcg cggatgcccg acggcgagga tctcgtcgtg acccacggcg atgcctgctt 11040

gccgaatatc atggtggaaa atggccgctt ttctggattc atcgactgtg gccggctggg 11100

tgtggcggac cgctatcagg acatagcgtt ggctacccgt gatattgctg aagagcttgg 11160

cggcgaatgg gctgaccgct tcctcgtgct ttacggtatc gccgctcccg attcgcagcg 11220

catcgccttc tatcgccttc ttgacgagtt cttctagtat gtaagccctg tgccttctag 11280

ttgccagcca tctgttgttt gcccctcccc cgtgccttcc ttgaccctgg aaggtgccac 11340

tcccactgtc ctttcctaat aaaatgagga aattgcatcg cattgtctga gtaggtgtca 11400

ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg aagacaatag 11460

caggcatgct ggggatgcgg tgggctctat ggttaattaa ccagtcaagt cagctacttg 11520

gcgagatcga cttgtctggg tttcgactac gctcagaatt gcgtcagtca agttcgatct 11580

ggtccttgct attgcacccg ttctccgatt acgagtttca tttaaatcat gtgagcaaaa 11640

ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc 11700

cgccccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca 11760

ggactataaa gataccaggc gtttcccccc ggaagctccc tcgtgcgctc tcctgttccg 11820

accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct 11880
catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt 11940
gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag 12000
tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc 12060
agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac 12120
actagaagaa cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga 12180
gttggtagct cttgatccgg caaacaaacc accgctggta gcggtggttt ttttgtttgc 12240

aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg 12300
gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca 12360
aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt 12420
atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca 12480
gcgatctgtc tatttcgttc atccatagtt gcatttaaat ttccgaactc tccaaggccc 12540
tc 12542

<210> 97
<211> 12554
<212> DNA
<213> Artificial Sequence

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

<400> 97
ggaaaatctt caaacctttc gtccgatcca tcttgcaggc tacctctcga acgaactatc 60
gcaagtctct tggccggcct tgcgccttgg ctattgcttg gcagcgccta tcgccaggta 120
ttactccaat cccgaatatc cgagatcggg atcacccgag agaagttcaa cctacatcct 180
caatcccgat ctatccgaga tccgaggaat atcgaaatcg gggcgcgcct ggtgtaccga 240
gaacgatcct ctcagtgcga gtctcgacga tccatatcgt tgcttggcag tcagccagtc 300
ggaatccagc ttgggaccca ggaagtccaa tcgtcagata ttgtactcaa gcctggtcac 360

ggcagcgtac cgatctgttt aaacctagat attgatagtc tgatcggtca acgtataatc 420
gagtcctagc ttttgcaaac atctatcaag agacaggatc agcaggaggc tttcgcatga 480
gtattcaaca tttccgtgtc gcccttattc cctttttgc ggcattttgc cttcctgttt 540
ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcgcgag 600

tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag 660

aacgctttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgta 720

ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg 780

agtattcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca 840

gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca acgattggag 900

gaccgaagga gctaaccgct tttttgcaca acatgggggga tcatgtaact cgccttgatc 960

gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg 1020

tagcaatggc aacaaccttg cgtaaactat taactggcga actacttact ctagcttccc 1080

ggcaacagtt gatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg 1140

cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt gggtctcgcg 1200

gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga 1260

cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 1320

tgattaagca ttggtaaccg attctaggtg cattggcgca gaaaaaaatg cctgatgcga 1380

cgctgcgcgt cttatactcc cacatatgcc agattcagca acggatacgg cttccccaac 1440

ttgcccactt ccatacgtgt cctccttacc agaaatttat ccttaagatc gtttaaactc 1500

gactctggct ctatcgaatc tccgtcgttt cgagcttacg cgaacagccg tggcgctcat 1560

ttgctcgtcg ggcatcgaat ctcgtcagct atcgtcagct taccttttg gcagcgatcg 1620

cggctcccga catcttggac cattagctcc acaggtatct tcttccctct agtggtcata 1680

acagcagctt cagctacctc tcaattcaaa aaacccctca agacccgttt agaggcccca 1740

aggggttatg ctatcaatcg ttgcgttaca cacacaaaaa accaacacac atccatcttc 1800

gatggatagc gattttatta tctaactgct gatcgagtgt agccagatct agtaatcaat 1860

tacggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa 1920

tggcccgcct ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt 1980

tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta 2040

aactgcccac ttggcagtac atcaagtgta tcatatgcca agtacgcccc ctattgacgt 2100

caatgacggt aaatggcccg cctggcatta tgcccagtac atgaccttat gggactttcc 2160

tacttggcag tacatctacg tattagtcat cgctattacc atgctgatgc ggttttggca 2220

gtacatcaat gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat 2280

tgacgtcaat gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa 2340

caactccgcc ccattgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag 2400

cagagctggt ttagtgaacc gtcagatcag atctttgtcg atcctaccat ccactcgaca 2460

cacccgccag cggccgctaa tacgactcac tatagggaga agtactgcca ccatgccaaa 2520
gaagaagcgg aaggtcggcg gcggcagcgg caagaagcaa aatcgcaaga cggggaattc 2580
caagacacaa tccgctagcc caccacctaa agagcgttct agctcccctg ctactgagca 2640
gtcctggatg gaaaacgact tcgatgaact ccgggaagag ggatttaggc gatccaacta 2700
ttcagaactc cgcgaagata tccagacaaa ggggaaggaa gtcgagaatt tcgagaagaa 2760
cctcgaggag tgcatcaccc gtatcacaaa cactgagaaa tgtctcaaag aactcatgga 2820
acttaagaca aaagccaggg agcttcgaga ggagtgtcgg agtctgagat ccaggtgtga 2880

ccagctcgag gagcgcgtga gcgcgatgga agacgagatg aacgagatga aaagagaggg 2940
caaattcagg gagaagcgca ttaagaggaa cgaacagagt ctgcaggaga tttgggatta 3000
cgtcaagagg cctaacctgc ggttgatcgg cgtccccgag agcgacgtag aaaacgggac 3060
taaactggag aatacacttc aagacatcat tcaagaaaat tttccaaacc tggctcggca 3120
agctaatgtg caaatccaag agatccaacg cacaccccag cggtatagct ctcggcgtgc 3180
caccctagg catattatcg tgcgctttac taaggtggag atgaaagaga agatgctgcg 3240
agccgctcgg gaaaagggaa gggtgacttt gaagggcaaa cctattcggc tgacggttga 3300

ccttagcgcc gagacactcc aggcacgccg ggaatggggc cccatcttta atatcctgaa 3360
ggagaagaac ttccagccac gaatctctta ccctgcaaag ttgagtttta tctccgaggg 3420
tgagattaag tatttcatcg ataaacagat gctgcgagac ttcgtgacaa ctcgcccagc 3480
tctcaaggaa ctgctcaaag aggctcttaa tatggagcgc aataatagat atcaaccctt 3540
gcagaaccac gcaaagatgt gagacagccg tcagaccatc aagactagga agaaactgca 3600
tcaactaatg agcaaaatca ccagctaaca tcatagtata catgaccggc tctaactcac 3660
atatcaccat ccttacactt aacattaacg gcctcaactc agctatcaag cgccatcggc 3720

tggccagctg gatcaaatca caggatccaa gcgtttgttg catccaagag acccacctga 3780
cctgtagaga tactcaccgc ctcaagatca agggatggcg aaagatttat caggcgaacg 3840
gtaagcagaa gaaagccgga gtcgcaattc tggtctcaga caagacggat ttcaagccca 3900
ccaaaattaa gcgtgataag gaaggtcact atattatggt gaaaggcagc atacagcagg 3960
aagaacttac catattgaac atctacgcgc caaacaccgg cgcacctcgc tttatcaaac 4020
aggtcctgtc cgatctgcag cgagatctgg attctcatac gttgattatg ggtgatttca 4080
atacaccatt gagcaccctg gatcgcagca ccaggcaaaa ggtaaataaa gacacgcaag 4140

agctcaatag cgcactgcat caggcagatc tcattgatat ttatcgcact cttcatccta 4200

agagtaccga gtacacattc ttcagcgccc cacatcatac atactcaaag atcgatcata 4260

tcgtcggctc aaaggctctg ctgtcaaagt gcaagcgcac agagataatt acaaattacc 4320

tgtcagatca tagcgcgatc aagctcgagc tgagaatcaa gaacctgacc cagagccgga 4380

gtaccacttg gaagcttaat aacctgctgc tcaacgatta ttgggtccac aatgagatga 4440

aggcagagat taaaatgttc ttcgaaacaa atgagaataa ggatactacc tatcaaaacc 4500

tttgggatgc ctttaaggcc gtctgcagag gcaagttcat cgccctcaac gcctataaaa 4560

gaaaacaaga gagatctaag atcgatactc tcacctctca gctgaaggag ttggagaaac 4620

aggaacagac ccactccaag gcgtcaagac ggcaggagat cacaaagatt cgcgccgagt 4680

tgaaagagat cgaaacccaa aagactcttc agaaaattaa cgagtctcgt agttggttct 4740

tcgagcggat taataagata gacagacctc tggcacgact gattaagaag aagcgcgaaa 4800

agaaccagat tgataccatc aagaacgaca agggcgacat cactactgac ccgaccgaga 4860

tccagaccac tattcgggag tattataagc atttgtatgc taacaagctt gagaacctgg 4920

aagagatgga cacttttctg gatacctata ctctgccacg gcttaatcaa gaggaagtcg 4980

agtccctcaa ccgcccaatt acaggaagcg agattgtggc cataattaac tccctgccga 5040

caaagaaatc tcctggtccg gacgggttta cagctgagtt ttatcaacgg tatatggaag 5100

agcttgtacc gtttctgctc aagctctttc agtctataga aaaggaaggc atcttgccca 5160

attccttcta cgaagcttct ataatactta ttcccaaacc aggacgcgat accacaaaga 5220

aggaaaactt ccggcccatt agtctcatga atatcgacgc taaaatattg aacaagattc 5280

tcgccaacag aatccaacaa catattaaga aattgataca tcacgaccag gtggggttta 5340

tacctggcat gcagggctgg tttaacatcc ggaagagtat taacgtcatt caacacatta 5400

atagagctaa ggataagaat catatgatca tctctataga cgcggaaaag gcattcgata 5460

agattcagca gccatttatg ctcaagactc tgaacaaact cggcatcgac ggaacatatt 5520

ttaagattat tcgcgcaatt tacgataagc cgactgctaa cattatcctt aacggccaaa 5580

agctcgaggc ctttccgctc aagactggaa cccgccaagg ctgtcccctc tccccgcttt 5640

tgtttaatat tgtactcgag gtgctggcta gggctattcg tcaagagaaa gagattaaag 5700

ggatacagct cgggaaggaa gaggtcaagc tttccttgtt cgccgatgat atgattgtgt 5760

acctggagaa tcctattgtg tctgctcaga accttcttaa acttatttct aactttagca 5820

aggtcagcgg ctataagatt aacgtccaga aatctcaggc ctttctgtac acaaataatc 5880

gacagaccga atcccagata atgggtgagc ttccgtttgt catagccagc aaaaggataa 5940

agtatctcgg aatccagctg acacgagacg ttaaagattt gtttaaggaa aattacaagc 6000

ctctcctgaa agagattaag gaagatacta ataagtggaa gaatatcccc tgttcatggg 6060

ttggcagaat caacatagtg aagatggcaa tacttcctaa agtgatatat cgctttaacg 6120

ccatcccaat taaactgcct atgaccttct ttacggagct cgagaaaaca acccttaaat 6180

ttatatggaa tcaaaagaga gcaagaatag cgaagtccat cttgagccag aagaataagg 6240

ccggtgggat tactttgcct gattttaagt tgtattataa agccacagta actaagacag 6300

cctggtattg gtatcagaat agagacatcg accagtggaa tcggaccgaa ccatcagaga 6360

taatgcccca catctataat taccttatat tcgataagcc agaaaagaat aaacagtggg 6420

gcaaagacag cctcttcaac aagtggtgtt gggagaattg gctggccata tgccggaaac 6480

tcaagctcga ccccttttctt acaccctaca ctaaaatcaa cagtaggtgg atcaaggact 6540

tgaatgtcaa gccaaagact ataaagacac tggaagagaa tcttgggatc acaatacaag 6600

atataggcgt cggcaaagat tttatgtcaa agacgcccaa ggccatggcc actaaggata 6660

agattgataa gtgggacctt attaagctca aaagcttctg tactgccaag gagaccacga 6720

tcagagttaa taggcagccc actacatggg aaaagatttt cgccacttat tcatcagata 6780

aggggttgat aagcagaata tataacgagc tgaagcagat ctacaagaag aaaacgaata 6840

atcccatcaa gaagtgggca aaagatatga acaggcattt tagcaaagag gatatctacg 6900

ccgcgaagaa gcatatgaag aagtgtagtt caagcttggc cattcgtgag atgcagatta 6960

agacgaccat gcgataccac cttaccccag tgaggatggc aattatcaag aaatctggca 7020

ataatagatg ttggcggggc tgtggcgaga ttggcaccct gctccattgc tggtgggatt 7080

gcaagctggt gcagccgctt tggaaatcag tctggcgctt tctgagggac ctcgagcttg 7140

agattccctt cgatcccgca attcccttgc tcggaatcta tcctaacgaa tacaagagct 7200

gttgttacaa ggatacgtgt acccggatgt tcatcgcggc cttgtttacg atagctaaga 7260

cgtggaatca gcctaagtgc cccacaatga tcgattggat caagaaaatg tggcatattt 7320

ataccatgga gtattacgca gcaattaaga atgacgaatt tatttccttc gttgggacct 7380

ggatgaagct ggagactatt attctgagca agctgtctca ggagcaaaag acaaagcata 7440

gaatcttctc tctcattggt ggtaacgact acaaagacga tgacgacaag taaagcgctt 7500

ctagaagttg tctcctcctg cactgactga ctgatacaat cgatttctgg atccgcaggc 7560

ctaatcaacc tctggattac aaaatttgtg aaagattgac tggtattctt aactatgttg 7620

ctccttttac gctatgtgga tacgctgctt taatgccttt gtatcatgct attgcttccc 7680

gtatggcttt cattttctcc tccttgtata aatcctggtt gctgtctctt tatgaggagt 7740

tgtggcccgt tgtcaggcaa cgtggcgtgg tgtgcactgt gtttgctgac gcaacccccа 7800

ctggttgggg cattgccacc acctgtcagc tcctttccgg gactttcgct ttcccctcc 7860

ctattgccac ggcggaactc atcgccgcct gccttgcccg ctgctggaca ggggctcggc 7920

tgttgggcac tgacaattcc gtggtgttgt cggggaagct gacgtccttt ccatggctgc 7980
tcgcctgtgt tgccacctgg attctgcgcg ggacgtcctt ctgctacgtc ccttcggccc 8040
tcaatccagc ggaccttcct tcccgctgag agacacaaaa aattccaaca cactattgca 8100
atgaaaataa atttccttta ttagccagaa gtcagatgct caaggggctt catgatgtcc 8160
ccataatttt tggcagaggg aaaaagatct cagtggtatt tgtgagccag ggcattggcc 8220
ttctgatagg cagcctgcac ctgaggagtg cggccgcttt acttgtacag ctcgtccatg 8280
ccgagagtga tcccggcggc ggtcacgaac tccagcagga ccatgtgatc gcgcttctcg 8340

ttggggtctt tgctcagggc ggactgggtg ctcaggtagt ggttgtcggg cagcagcacg 8400
gggccgtcgc cgatgggggt gttctgctgg tagtggtcgg cgagctgcac gctgccgtcc 8460
tcgatgttgt ggcggatctt gaagttcacc ttgatgccgt tcttctgctt gtcggccatg 8520
atatagacgt tgtggctgtt gtagttgtac tccagcttgt gccccaggat gttgccgtcc 8580
tccttgaagt cgatgcccct cagctcgatg cggttcacca gggtgtcgcc ctcgaacttc 8640
acctcggcgc gggtcttgta gttgccgtcg tccttgaaga agatggtgcg ctcctggacg 8700
tagccttcgg gcatggcgga cttgaagaag tcgtgctgct tcatgtggtc ggggtagcgg 8760

ctgaagcact gcacgccgta ggtcagggtg gtcacgaggg tgggccaggg cacgggcagc 8820
ttgccggtgg tgcagatgaa cttcagggtc agcttgccgt aggtggcatc gccctcgccc 8880
tcgccggaca cgctgaactt gtggccgttt acgtcgccgt ccagctcgac caggatgggc 8940
accaccccgg tgaacagctc ctcgcccttg ctcaccatgg tggcgggatc tgacggttca 9000
ctaaaccagc tctgcttata tagacctccc accgtacacg cctaccgccc atttgcgtca 9060
atggggcgga gttgttacga cattttggaa agtcccgttg attttggtgc caaaacaaac 9120
tcccattgac gtcaatgggg tggagacttg gaaatccccg tgagtcaaac cgctatccac 9180

gcccattgat gtactgccaa aaccgcatca ccatggtaat agcgatgact aatacgtaga 9240
tgtactgcca agtaggaaag tcccataagg tcatgtactg ggcataatgc caggcgggcc 9300
atttaccgtc attgacgtca ataggggggcg tacttggcat atgatacact tgatgtactg 9360
ccaagtgggc agtttaccgt aaatactcca cccattgacg tcaatggaaa gtccctattg 9420
gcgttactat gggaacatac gtcattattg acgtcaatgg gcgggggtcg ttgggcggtc 9480
agccaggcgg gccatttacc gtaagttatg taacgggcct gctgccggct ctgcggcctc 9540
ttccgcgtct tcgccttcgc cctcagacga gtcggatctc cctttgggcc gcctccccgc 9600

ctgtctagct tgactgactg agatacagcg taccttcagc tcacagacat gataagatac 9660

attgatgagt ttggacaaac cacaactaga atgcagtgaa aaaaatgctt tatttgtgaa 9720

atttgtgatg ctattgcttt atttgtaacc attataagct gcaataaaca agttaacaac 9780

aacaattgca ttcattttat gtttcaggtt caggggggagg tgtgggaggt tttttaaagc 9840

aagtaaaacc tctacaaatg tggtattggc ccatctctat cggtatcgta gcataacccc 9900

ttggggcctc taaacgggtc ttgaggggtt ttttgtgccc ctcgggccgg attgctatct 9960

accggcattg gcgcagaaaa aaatgcctga tgcgacgctg cgcgtcttat actcccacat 10020

atgccagatt cagcaacgga tacggcttcc ccaacttgcc cacttccata cgtgtcctcc 10080

ttaccagaaa tttatcctta aggtcgtcag ctatcctgca ggcgatctct cgatttcgat 10140

caagacattc ctttaatggt cttttctgga caccactagg ggtcagaagt agttcatcaa 10200

actttcttcc ctccctaatc tcattggtta ccttgggcta tcgaaactta attaagcgat 10260

ctgcatctca attagtcagc aaccatagtc ccgcccctaa ctccgcccat cccgccccta 10320

actccgccca gttccgccca ttctccgccc catcgctgac taatttttt tatttatgca 10380

gaggccgagg ccgcctcggc ctctgagcta ttccagaagt agtgaggagg ctttttgga 10440

ggcctaggct tttgcaaagg aggtagccaa catgattgaa caagatggat tgcacgcagg 10500

ttctcccgcc gcttgggtgg agaggctatt cggctatgac tgggcacaac agacaatcgg 10560

ctgctctgat gccgccgtgt tccggctgtc agcgcagggg cgcccggttc tttttgtcaa 10620

gaccgacctg tccggtgccc tgaatgaact ccaggacgag gcagcgcggc tatcgtggct 10680

ggccacgacg ggcgttcctt gcgcagctgt gctcgacgtt gtcactgaag cgggaaggga 10740

ctggctgcta ttgggcgaag tgccggggca ggatctcctg tcatctcacc ttgctcctgc 10800

cgagaaagta tccatcatgg ctgatgcaat gcggcggctg catacgcttg atccggctac 10860

ctgcccattc gaccaccaag cgaaacatcg catcgagcga gcacgtactc ggatggaagc 10920

cggtcttgtc gatcaggatg atctggacga agagcatcag gggctcgcgc cagccgaact 10980

gttcgccagg ctcaaggcgc ggatgcccga cggcgaggat ctcgtcgtga cccacggcga 11040

tgcctgcttg ccgaatatca tggtggaaaa tggccgcttt tctggattca tcgactgtgg 11100

ccggctgggt gtggcggacc gctatcagga catagcgttg gctacccgtg atattgctga 11160

agagcttggc ggcgaatggg ctgaccgctt cctcgtgctt tacggtatcg ccgctcccga 11220

ttcgcagcgc atcgccttct atcgccttct tgacgagttc ttctagtatg taagccctgt 11280

gccttctagt tgccagccat ctgttgtttg cccctccccc gtgccttcct tgaccctgga 11340

aggtgccact cccactgtcc tttcctaata aaatgaggaa attgcatcgc attgtctgag 11400

taggtgtcat tctattctgg ggggtggggt ggggcaggac agcaaggggg aggattggga 11460

agacaatagc aggcatgctg gggatgcggt gggctctatg gttaattaac cagtcaagtc 11520

agctacttgg cgagatcgac ttgtctgggt ttcgactacg ctcagaattg cgtcagtcaa 11580

gttcgatctg gtccttgcta ttgcacccgt tctccgatta cgagtttcat ttaaatcatg 11640

tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 11700

cataggctcc gccccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 11760

aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 11820

cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg 11880

gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag 11940

ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat 12000

cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac 12060

aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac 12120

tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc agttaccttc 12180

ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt 12240

tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc 12300

ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg 12360

agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca 12420

atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca 12480

cctatctcag cgatctgtct atttcgttca tccatagttg catttaaatt tccgaactct 12540

ccaaggccct cgtc 12554

<210> 98

<211> 12569

<212> DNA

<213> Artificial Sequence

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

<400> 98

ggaaaatctt caaacctttc gtccgatcca tcttgcaggc tacctctcga acgaactatc 60

gcaagtctct tggccggcct tgcgccttgg ctattgcttg gcagcgccta tcgccaggta 120

ttactccaat cccgaatatc cgagatcggg atcacccgag agaagttcaa cctacatcct 180

caatcccgat ctatccgaga tccgaggaat atcgaaatcg gggcgcgcct ggtgtaccga 240

gaacgatcct ctcagtgcga gtctcgacga tccatatcgt tgcttggcag tcagccagtc 300

ggaatccagc ttgggaccca ggaagtccaa tcgtcagata ttgtactcaa gcctggtcac 360

ggcagcgtac cgatctgttt aaacctagat attgatagtc tgatcggtca acgtataatc 420

gagtcctagc ttttgcaaac atctatcaag agacaggatc agcaggaggc tttcgcatga 480

gtattcaaca tttccgtgtc gcccttattc cctttttgc ggcattttgc cttcctgttt 540

ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcgcgag 600

tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag 660

aacgctttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgta 720

ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg 780

agtattcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca 840

gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca acgattggag 900

gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc 960

gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg 1020

tagcaatggc aacaaccttg cgtaaactat taactggcga actacttact ctagcttccc 1080

ggcaacagtt gatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg 1140

cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt gggtctcgcg 1200

gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga 1260

cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 1320

tgattaagca ttggtaaccg attctaggtg cattggcgca gaaaaaaatg cctgatgcga 1380

cgctgcgcgt cttatactcc cacatatgcc agattcagca acggatacgg cttccccaac 1440

ttgcccactt ccatacgtgt cctccttacc agaaatttat ccttaagatc gtttaaactc 1500

gactctggct ctatcgaatc tccgtcgttt cgagcttacg cgaacagccg tggcgctcat 1560

ttgctcgtcg ggcatcgaat ctcgtcagct atcgtcagct taccttttg gcagcgatcg 1620

cggctcccga catcttggac cattagctcc acaggtatct tcttccctct agtggtcata 1680

acagcagctt cagctacctc tcaattcaaa aaacccctca agacccgttt agaggcccca 1740

aggggttatg ctatcaatcg ttgcgttaca cacacaaaaa accaacacac atccatcttc 1800

gatggatagc gattttatta tctaactgct gatcgagtgt agccagatct agtaatcaat 1860

tacggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa 1920

tggcccgcct ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt 1980

tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta 2040

aactgcccac ttggcagtac atcaagtgta tcatatgcca agtacgcccc ctattgacgt 2100

caatgacggt aaatggcccg cctggcatta tgcccagtac atgaccttat gggactttcc 2160

tacttggcag tacatctacg tattagtcat cgctattacc atgctgatgc ggttttggca 2220

gtacatcaat gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat 2280

tgacgtcaat gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa 2340

caactccgcc ccattgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag 2400

cagagctggt ttagtgaacc gtcagatcag atctttgtcg atcctaccat ccactcgaca 2460

cacccgccag cggccgctaa tacgactcac tatagggaga agtactgcca ccatgaaaag 2520

gccggcggcc acgaaaaagg ccggccaggc aaaaaagaaa aagggcaaga agcaaaatcg 2580

caagacgggg aattccaaga cacaatccgc tagcccacca cctaaagagc gttctagctc 2640

ccctgctact gagcagtcct ggatggaaaa cgacttcgat gaactccggg aagagggatt 2700

taggcgatcc aactattcag aactccgcga agatatccag acaaagggga aggaagtcga 2760

gaatttcgag aagaacctcg aggagtgcat cacccgtatc acaaacactg agaaatgtct 2820

caaagaactc atggaactta agacaaaagc cagggagctt cgagaggagt gtcggagtct 2880

gagatccagg tgtgaccagc tcgaggagcg cgtgagcgcg atggaagacg agatgaacga 2940

gatgaaaaga gagggcaaat tcagggagaa gcgcattaag aggaacgaac agagtctgca 3000

ggagatttgg gattacgtca agaggcctaa cctgcggttg atcggcgtcc ccgagagcga 3060

cgtagaaaac gggactaaac tggagaatac acttcaagac atcattcaag aaaattttcc 3120

aaacctggct cggcaagcta atgtgcaaat ccaagagatc caacgcacac cccagcggta 3180

tagctctcgg cgtgccaccc ctaggcatat tatcgtgcgc tttactaagg tggagatgaa 3240

agagaagatg ctgcgagccg ctcgggaaaa gggaagggtg actttgaagg gcaaacctat 3300

tcggctgacg gttgacctta gcgccgagac actccaggca cgccgggaat ggggccccat 3360

ctttaatatc ctgaaggaga agaacttcca gccacgaatc tcttaccctg caaagttgag 3420

ttttatctcc gagggtgaga ttaagtattt catcgataaa cagatgctgc gagacttcgt 3480

gacaactcgc ccagctctca aggaactgct caaagaggct cttaatatgg agcgcaataa 3540

tagatatcaa cccttgcaga accacgcaaa gatgtgagac agccgtcaga ccatcaagac 3600

taggaagaaa ctgcatcaac taatgagcaa aatcaccagc taacatcata gtatacatga 3660

ccggctctaa ctcacatatc accatcctta cacttaacat taacggcctc aactcagcta 3720

tcaagcgcca tcggctggcc agctggatca aatcacagga tccaagcgtt tgttgcatcc 3780

aagagaccca cctgacctgt agagatactc accgcctcaa gatcaaggga tggcgaaaga 3840

tttatcaggc gaacggtaag cagaagaaag ccggagtcgc aattctggtc tcagacaaga 3900

cggatttcaa gcccaccaaa attaagcgtg ataaggaagg tcactatatt atggtgaaag 3960

gcagcataca gcaggaagaa cttaccatat tgaacatcta cgcgccaaac accggcgcac 4020

ctcgctttat caaacaggtc ctgtccgatc tgcagcgaga tctggattct catacgttga 4080
ttatgggtga tttcaataca ccattgagca ccctggatcg cagcaccagg caaaaggtaa 4140
ataaagacac gcaagagctc aatagcgcac tgcatcaggc agatctcatt gatatttatc 4200
gcactcttca tcctaagagt accgagtaca cattcttcag cgccccacat catacatact 4260
caaagatcga tcatatcgtc ggctcaaagg ctctgctgtc aaagtgcaag cgcacagaga 4320
taattacaaa ttacctgtca gatcatagcg cgatcaagct cgagctgaga atcaagaacc 4380
tgacccagag ccggagtacc acttggaagc ttaataacct gctgctcaac gattattggg 4440

tccacaatga gatgaaggca gagattaaaa tgttcttcga aacaaatgag aataaggata 4500
ctacctatca aaacctttgg gatgccttta aggccgtctg cagaggcaag ttcatcgccc 4560
tcaacgccta taaaagaaaa caagagagat ctaagatcga tactctcacc tctcagctga 4620
aggagttgga gaaacaggaa cagacccact ccaaggcgtc aagacggcag gagatcacaa 4680
agattcgcgc cgagttgaaa gagatcgaaa cccaaaagac tcttcagaaa attaacgagt 4740
ctcgtagttg gttcttcgag cggattaata agatagacag acctctggca cgactgatta 4800
agaagaagcg cgaaaagaac cagattgata ccatcaagaa cgacaagggc gacatcacta 4860

ctgacccgac cgagatccag accactattc gggagtatta taagcatttg tatgctaaca 4920
agcttgagaa cctggaagag atggacactt ttctggatac ctatactctg ccacggctta 4980
atcaagagga agtcgagtcc ctcaaccgcc caattacagg aagcgagatt gtggccataa 5040
ttaactccct gccgacaaag aaatctcctg gtccggacgg gtttacagct gagttttatc 5100
aacggtatat ggaagagctt gtaccgtttc tgctcaagct ctttcagtct atagaaaagg 5160
aaggcatctt gcccaattcc ttctacgaag cttctataat acttattccc aaaccaggac 5220
gcgataccac aaagaaggaa aacttccggc ccattagtct catgaatatc gacgctaaaa 5280

tattgaacaa gattctcgcc aacagaatcc aacaacatat taagaaattg atacatcacg 5340
accaggtggg gtttatacct ggcatgcagg gctggtttaa catccggaag agtattaacg 5400
tcattcaaca cattaataga gctaaggata agaatcatat gatcatctct atagacgcgg 5460
aaaaggcatt cgataagatt cagcagccat ttatgctcaa gactctgaac aaactcggca 5520
tcgacggaac atattttaag attattcgcg caatttacga taagccgact gctaacatta 5580
tccttaacgg ccaaaagctc gaggcctttc cgctcaagac tggaacccgc caaggctgtc 5640
ccctctcccc gcttttgttt aatattgtac tcgaggtgct ggctagggct attcgtcaag 5700

agaaagagat taaagggata cagctcggga aggaagaggt caagctttcc ttgttcgccg 5760

atgatatgat tgtgtacctg gagaatccta ttgtgtctgc tcagaacctt cttaaactta 5820

tttctaactt tagcaaggtc agcggctata agattaacgt ccagaaatct caggcctttc 5880

tgtacacaaa taatcgacag accgaatccc agataatggg tgagcttccg tttgtcatag 5940

ccagcaaaag gataaagtat ctcggaatcc agctgacacg agacgttaaa gatttgttta 6000

aggaaaatta caagcctctc ctgaaagaga ttaaggaaga tactaataag tggaagaata 6060

tcccctgttc atgggttggc agaatcaaca tagtgaagat ggcaatactt cctaaagtga 6120

tatatcgctt taacgccatc ccaattaaac tgcctatgac cttctttacg gagctcgaga 6180

aaacaaccct taaatttata tggaatcaaa agagagcaag aatagcgaag tccatcttga 6240

gccagaagaa taaggccggt gggattactt tgcctgattt taagttgtat tataaagcca 6300

cagtaactaa gacagcctgg tattggtatc agaatagaga catcgaccag tggaatcgga 6360

ccgaaccatc agagataatg ccccacatct ataattacct tatattcgat aagccagaaa 6420

agaataaaca gtggggcaaa gacagcctct tcaacaagtg gtgttgggag aattggctgg 6480

ccatatgccg gaaactcaag ctcgacccct ttcttacacc ctacactaaa atcaacagta 6540

ggtggatcaa ggacttgaat gtcaagccaa agactataaa gacactggaa gagaatcttg 6600

ggatcacaat acaagatata ggcgtcggca aagattttat gtcaaagacg cccaaggcca 6660

tggccactaa ggataagatt gataagtggg accttattaa gctcaaaagc ttctgtactg 6720

ccaaggagac cacgatcaga gttaataggc agcccactac atgggaaaag attttcgcca 6780

cttattcatc agataagggg ttgataagca gaatatataa cgagctgaag cagatctaca 6840

agaagaaaac gaataatccc atcaagaagt gggcaaaaga tatgaacagg cattttagca 6900

aagaggatat ctacgccgcg aagaagcata tgaagaagtg tagttcaagc ttggccattc 6960

gtgagatgca gattaagacg accatgcgat accaccttac cccagtgagg atggcaatta 7020

tcaagaaatc tggcaataat agatgttggc ggggctgtgg cgagattggc accctgctcc 7080

attgctggtg ggattgcaag ctggtgcagc cgctttggaa atcagtctgg cgctttctga 7140

gggacctcga gcttgagatt cccttcgatc ccgcaattcc cttgctcgga atctatccta 7200

acgaatacaa gagctgttgt tacaaggata cgtgtacccg gatgttcatc gcggccttgt 7260

ttacgatagc taagacgtgg aatcagccta agtgccccac aatgatcgat tggatcaaga 7320

aaatgtggca tatttatacc atggagtatt acgcagcaat taagaatgac gaatttattt 7380

ccttcgttgg gacctggatg aagctggaga ctattattct gagcaagctg tctcaggagc 7440

aaaagacaaa gcatagaatc ttctctctca ttggtggtaa cgactacaaa gacgatgacg 7500

acaagtaaag cgcttctaga agttgtctcc tcctgcactg actgactgat acaatcgatt 7560

tctggatccg caggcctaat caacctctgg attacaaaat ttgtgaaaga ttgactggta 7620

ttcttaacta tgttgctcct tttacgctat gtggatacgc tgctttaatg cctttgtatc 7680

atgctattgc ttcccgtatg gctttcattt tctcctcctt gtataaatcc tggttgctgt 7740

ctctttatga ggagttgtgg cccgttgtca ggcaacgtgg cgtggtgtgc actgtgtttg 7800

ctgacgcaac ccccactggt tggggcattg ccaccacctg tcagctcctt tccgggactt 7860

tcgctttccc cctccctatt gccacggcgg aactcatcgc cgcctgcctt gcccgctgct 7920

ggacaggggc tcggctgttg ggcactgaca attccgtggt gttgtcgggg aagctgacgt 7980

cctttccatg gctgctcgcc tgtgttgcca cctggattct gcgcgggacg tccttctgct 8040

acgtcccttc ggccctcaat ccagcggacc ttccttcccg ctgagagaca caaaaaattc 8100

caacacacta ttgcaatgaa aataaatttc ctttattagc cagaagtcag atgctcaagg 8160

ggcttcatga tgtccccata atttttggca gagggaaaaa gatctcagtg gtatttgtga 8220

gccagggcat tggccttctg ataggcagcc tgcacctgag gagtgcggcc gctttacttg 8280

tacagctcgt ccatgccgag agtgatcccg gcggcggtca cgaactccag caggaccatg 8340

tgatcgcgct tctcgttggg gtctttgctc agggcggact gggtgctcag gtagtggttg 8400

tcgggcagca gcacggggcc gtcgccgatg ggggtgttct gctggtagtg gtcggcgagc 8460

tgcacgctgc cgtcctcgat gttgtggcgg atcttgaagt tcaccttgat gccgttcttc 8520

tgcttgtcgg ccatgatata gacgttgtgg ctgttgtagt tgtactccag cttgtgcccc 8580

aggatgttgc cgtcctcctt gaagtcgatg cccttcagct cgatgcggtt caccagggtg 8640

tcgccctcga acttcacctc ggcgcgggtc ttgtagttgc cgtcgtcctt gaagaagatg 8700

gtgcgctcct ggacgtagcc ttcgggcatg gcggacttga agaagtcgtg ctgcttcatg 8760

tggtcggggt agcggctgaa gcactgcacg ccgtaggtca gggtggtcac gagggtgggc 8820

cagggcacgg gcagcttgcc ggtggtgcag atgaacttca gggtcagctt gccgtaggtg 8880

gcatcgccct cgccctcgcc ggacacgctg aacttgtggc cgtttacgtc gccgtccagc 8940

tcgaccagga tgggcaccac cccggtgaac agctcctcgc ccttgctcac catggtggcg 9000

ggatctgacg gttcactaaa ccagctctgc ttatatagac ctcccaccgt acacgcctac 9060

cgcccatttg cgtcaatggg gcggagttgt tacgacattt tggaaagtcc cgttgatttt 9120

ggtgccaaaa caaactccca ttgacgtcaa tggggtggag acttggaaat ccccgtgagt 9180

caaaccgcta tccacgccca ttgatgtact gccaaaaccg catcaccatg gtaatagcga 9240

tgactaatac gtagatgtac tgccaagtag gaaagtccca taaggtcatg tactgggcat 9300

aatgccaggc gggccattta ccgtcattga cgtcaatagg gggcgtactt ggcatatgat 9360

acacttgatg tactgccaag tgggcagttt accgtaaata ctccacccat tgacgtcaat 9420

ggaaagtccc tattggcgtt actatgggaa catacgtcat tattgacgtc aatgggcggg 9480

ggtcgttggg cggtcagcca ggcgggccat ttaccgtaag ttatgtaacg ggcctgctgc 9540
cggctctgcg gcctcttccg cgtcttcgcc ttcgccctca gacgagtcgg atctcccttt 9600
gggccgcctc cccgcctgtc tagcttgact gactgagata cagcgtacct tcagctcaca 9660
gacatgataa gatacattga tgagtttgga caaaccacaa ctagaatgca gtgaaaaaaa 9720
tgctttattt gtgaaatttg tgatgctatt gctttatttg taaccattat aagctgcaat 9780
aaacaagtta acaacaacaa ttgcattcat tttatgtttc aggttcaggg ggaggtgtgg 9840
gaggtttttt aaagcaagta aaacctctac aaatgtggta ttggcccatc tctatcggta 9900

tcgtagcata accccttggg gcctctaaac gggtcttgag gggtttttg tgcccctcgg 9960
gccggattgc tatctaccgg cattggcgca gaaaaaaatg cctgatgcga cgctgcgcgt 10020
cttatactcc cacatatgcc agattcagca acggatacgg cttccccaac ttgcccactt 10080
ccatacgtgt cctccttacc agaaatttat ccttaaggtc gtcagctatc ctgcaggcga 10140
tctctcgatt tcgatcaaga cattccttta atggtctttt ctggacacca ctaggggtca 10200
gaagtagttc atcaaacttt cttccctccc taatctcatt ggttaccttg ggctatcgaa 10260
acttaattaa gcgatctgca tctcaattag tcagcaacca tagtcccgcc cctaactccg 10320

cccatcccgc ccctaactcc gcccagttcc gcccattctc cgccccatcg ctgactaatt 10380
ttttttattt atgcagaggc cgaggccgcc tcggcctctg agctattcca gaagtagtga 10440
ggaggctttt ttggaggcct aggcttttgc aaaggaggta gccaacatga ttgaacaaga 10500
tggattgcac gcaggttctc ccgccgcttg ggtggagagg ctattcggct atgactgggc 10560
acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc 10620
ggttcttttt gtcaagaccg acctgtccgg tgccctgaat gaactccagg acgaggcagc 10680
gcggctatcg tggctggcca cgacgggcgt tccttgcgca gctgtgctcg acgttgtcac 10740

tgaagcggga agggactggc tgctattggg cgaagtgccg gggcaggatc tcctgtcatc 10800
tcaccttgct cctgccgaga aagtatccat catggctgat gcaatgcggc ggctgcatac 10860
gcttgatccg gctacctgcc cattcgacca ccaagcgaaa catcgcatcg agcgagcacg 10920
tactcggatg gaagccggtc ttgtcgatca ggatgatctg gacgaagagc atcaggggct 10980
cgcgccagcc gaactgttcg ccaggctcaa ggcgcggatg cccgacggcg aggatctcgt 11040
cgtgacccac ggcgatgcct gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg 11100
attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag cgttggctac 11160

ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttcctcg tgctttacgg 11220

tatcgccgct cccgattcgc agcgcatcgc cttctatcgc cttcttgacg agttcttcta 11280

gtatgtaagc cctgtgcctt ctagttgcca gccatctgtt gtttgcccct cccccgtgcc 11340

ttccttgacc ctggaaggtg ccactcccac tgtcctttcc taataaaatg aggaaattgc 11400

atcgcattgt ctgagtaggt gtcattctat tctgggggggt ggggtggggc aggacagcaa 11460

gggggaggat tgggaagaca atagcaggca tgctggggat gcggtgggct ctatggttaa 11520

ttaaccagtc aagtcagcta cttggcgaga tcgacttgtc tgggtttcga ctacgctcag 11580

aattgcgtca gtcaagttcg atctggtcct tgctattgca cccgttctcc gattacgagt 11640

ttcatttaaa tcatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg 11700

ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca 11760

agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc 11820

tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc cgcctttctc 11880

ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag 11940

gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc 12000

ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca 12060

gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg 12120

aagtggtggc ctaactacgg ctacactaga agaacagtat ttggtatctg cgctctgctg 12180

aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct 12240

ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 12300

gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa 12360

gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa 12420

tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag ttaccaatgc 12480

ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat agttgcattt 12540

aaatttccga actctccaag gccctcgtc 12569

<210> 99

<211> 12581

<212> DNA

<213> Artificial Sequence

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

<400> 99

caaacctttc gtccgatcca tcttgcaggc tacctctcga acgaactatc gcaagtctct 60

tggccggcct tgcgccttgg ctattgcttg gcagcgccta tcgccaggta ttactccaat 120

cccgaatatc cgagatcggg atcacccgag agaagttcaa cctacatcct caatcccgat 180
ctatccgaga tccgaggaat atcgaaatcg gggcgcgcct ggtgtaccga gaacgatcct 240
ctcagtgcga gtctcgacga tccatatcgt tgcttggcag tcagccagtc ggaatccagc 300
ttgggaccca ggaagtccaa tcgtcagata ttgtactcaa gcctggtcac ggcagcgtac 360
cgatctgttt aaacctagat attgatagtc tgatcggtca acgtataatc gagtcctagc 420
ttttgcaaac atctatcaag agacaggatc agcaggaggc tttcgcatga gtattcaaca 480
tttccgtgtc gcccttattc cctttttgc ggcattttgc cttcctgttt ttgctcaccc 540

agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcgcgag tgggttacat 600
cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag aacgctttcc 660
aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgta ttgacgccgg 720
gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg agtattcacc 780
agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca gtgctgccat 840
aaccatgagt gataacactg cggccaactt acttctgaca acgattggag gaccgaagga 900
gctaaccgct tttttgcaca acatgggggga tcatgtaact cgccttgatc gttgggaacc 960

ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg tagcaatggc 1020
aacaaccttg cgtaaactat taactggcga actacttact ctagcttccc ggcaacagtt 1080
gatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg cccttccggc 1140
tggctggttt attgctgata aatctggagc cggtgagcgt gggtctcgcg gtatcattgc 1200
agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga cggggagtca 1260
ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac tgattaagca 1320
ttggtaaccg attctaggtg cattggcgca gaaaaaaatg cctgatgcga cgctgcgcgt 1380

cttatactcc cacatatgcc agattcagca acggatacgg cttccccaac ttgcccactt 1440
ccatacgtgt cctccttacc agaaatttat ccttaagatc gtttaaactc gactctggct 1500
ctatcgaatc tccgtcgttt cgagcttacg cgaacagccg tggcgctcat ttgctcgtcg 1560
ggcatcgaat ctcgtcagct atcgtcagct tacctttttg gcagcgatcg cggctcccga 1620
catcttggac cattagctcc acaggtatct tcttccctct agtggtcata acagcagctt 1680
cagctacctc tcaattcaaa aaacccctca agacccgttt agaggcccca aggggttatg 1740
ctatcaatcg ttgcgttaca cacacaaaaa accaacacac atccatcttc gatggatagc 1800

gattttatta tctaactgct gatcgagtgt agccagatct agtaatcaat tacggggtca 1860

ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa tggcccgcct 1920

ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt tcccatagta 1980

acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac 2040

ttggcagtac atcaagtgta tcatatgcca agtacgcccc ctattgacgt caatgacggt 2100

aaatggcccg cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag 2160

tacatctacg tattagtcat cgctattacc atgctgatgc ggttttggca gtacatcaat 2220

gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat 2280

gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc 2340

ccattgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag cagagctggt 2400

ttagtgaacc gtcagatcag atctttgtcg atcctaccat ccactcgaca cacccgccag 2460

cggccgctaa tacgactcac tatagggaga agtactgcca ccatgaaaag gccggcggcc 2520

acgaaaaagg ccggccaggc aaaaaagaaa aagggcggcg gcagcggcaa gaagcaaaat 2580

cgcaagacgg ggaattccaa gacacaatcc gctagcccac cacctaaaga gcgttctagc 2640

tcccctgcta ctgagcagtc ctggatggaa aacgacttcg atgaactccg ggaagaggga 2700

tttaggcgat ccaactattc agaactccgc gaagatatcc agacaaaggg gaaggaagtc 2760

gagaatttcg agaagaacct cgaggagtgc atcacccgta tcacaaacac tgagaaatgt 2820

ctcaaagaac tcatggaact taagacaaaa gccagggagc ttcgagagga gtgtcggagt 2880

ctgagatcca ggtgtgacca gctcgaggag cgcgtgagcg cgatggaaga cgagatgaac 2940

gagatgaaaa gagagggcaa attcagggag aagcgcatta agaggaacga acagagtctg 3000

caggagattt gggattacgt caagaggcct aacctgcggt tgatcggcgt ccccgagagc 3060

gacgtagaaa acgggactaa actggagaat acacttcaag acatcattca agaaaatttt 3120

ccaaacctgg ctcggcaagc taatgtgcaa atccaagaga tccaacgcac accccagcgg 3180

tatagctctc ggcgtgccac ccctaggcat attatcgtgc gctttactaa ggtggagatg 3240

aaagagaaga tgctgcgagc cgctcgggaa aagggaaggg tgactttgaa gggcaaacct 3300

attcggctga cggttgacct tagcgccgag acactccagg cacgccggga atggggcccc 3360

atctttaata tcctgaagga gaagaacttc cagccacgaa tctcttaccc tgcaaagttg 3420

agttttatct ccgagggtga gattaagtat ttcatcgata aacagatgct gcgagacttc 3480

gtgacaactc gcccagctct caaggaactg ctcaaagagg ctcttaatat ggagcgcaat 3540

aatagatatc aacccttgca gaaccacgca aagatgtgag acagccgtca gaccatcaag 3600

actaggaaga aactgcatca actaatgagc aaaatcacca gctaacatca tagtatacat 3660

gaccggctct aactcacata tcaccatcct tacacttaac attaacggcc tcaactcagc 3720

tatcaagcgc catcggctgg ccagctggat caaatcacag gatccaagcg tttgttgcat 3780

ccaagagacc cacctgacct gtagagatac tcaccgcctc aagatcaagg gatggcgaaa 3840

gatttatcag gcgaacggta agcagaagaa agccggagtc gcaattctgg tctcagacaa 3900

gacggatttc aagcccacca aaattaagcg tgataaggaa ggtcactata ttatggtgaa 3960

aggcagcata cagcaggaag aacttaccat attgaacatc tacgcgccaa acaccggcgc 4020

acctcgcttt atcaaacagg tcctgtccga tctgcagcga gatctggatt ctcatacgtt 4080

gattatgggt gatttcaata caccattgag caccctggat cgcagcacca ggcaaaaggt 4140

aaataaagac acgcaagagc tcaatagcgc actgcatcag gcagatctca ttgatattta 4200

tcgcactctt catcctaaga gtaccgagta cacattcttc agcgccccac atcatacata 4260

ctcaaagatc gatcatatcg tcggctcaaa ggctctgctg tcaaagtgca agcgcacaga 4320

gataattaca aattacctgt cagatcatag cgcgatcaag ctcgagctga gaatcaagaa 4380

cctgacccag agccggagta ccacttggaa gcttaataac ctgctgctca acgattattg 4440

ggtccacaat gagatgaagg cagagattaa aatgttcttc gaaacaaatg agaataagga 4500

tactacctat caaaaccttt gggatgcctt taaggccgtc tgcagaggca agttcatcgc 4560

cctcaacgcc tataaaagaa aacaagagag atctaagatc gatactctca cctctcagct 4620

gaaggagttg gagaaacagg aacagaccca ctccaaggcg tcaagacggc aggagatcac 4680

aaagattcgc gccgagttga aagagatcga aacccaaaag actcttcaga aaattaacga 4740

gtctcgtagt tggttcttcg agcggattaa taagatagac agacctctgg cacgactgat 4800

taagaagaag cgcgaaaaga accagattga taccatcaag aacgacaagg gcgacatcac 4860

tactgacccg accgagatcc agaccactat tcgggagtat tataagcatt tgtatgctaa 4920

caagcttgag aacctggaag agatggacac ttttctggat acctatactc tgccacggct 4980

taatcaagag gaagtcgagt ccctcaaccg cccaattaca ggaagcgaga ttgtggccat 5040

aattaactcc ctgccgacaa agaaatctcc tggtccggac gggtttacag ctgagtttta 5100

tcaacggtat atggaagagc ttgtaccgtt tctgctcaag ctctttcagt ctatagaaaa 5160

ggaaggcatc ttgcccaatt ccttctacga agcttctata atacttattc ccaaaccagg 5220

acgcgatacc acaaagaagg aaaacttccg gcccattagt ctcatgaata tcgacgctaa 5280

aatattgaac aagattctcg ccaacagaat ccaacaacat attaagaaat tgatacatca 5340

cgaccaggtg gggtttatac ctggcatgca gggctggttt aacatccgga agagtattaa 5400

cgtcattcaa cacattaata gagctaagga taagaatcat atgatcatct ctatagacgc 5460

ggaaaaggca ttcgataaga ttcagcagcc atttatgctc aagactctga acaaactcgg 5520

catcgacgga acatatttta agattattcg cgcaatttac gataagccga ctgctaacat 5580

tatccttaac ggccaaaagc tcgaggcctt tccgctcaag actggaaccc gccaaggctg 5640
tcccctctcc ccgcttttgt ttaatattgt actcgaggtg ctggctaggg ctattcgtca 5700
agagaaagag attaaaggga tacagctcgg gaaggaagag gtcaagcttt ccttgttcgc 5760
cgatgatatg attgtgtacc tggagaatcc tattgtgtct gctcagaacc ttcttaaact 5820
tatttctaac tttagcaagg tcagcggcta taagattaac gtccagaaat ctcaggcctt 5880
tctgtacaca aataatcgac agaccgaatc ccagataatg ggtgagcttc cgtttgtcat 5940
agccagcaaa aggataaagt atctcggaat ccagctgaca cgagacgtta aagatttgtt 6000

taaggaaaat tacaagcctc tcctgaaaga gattaaggaa gatactaata agtggaagaa 6060
tatcccctgt tcatgggttg gcagaatcaa catagtgaag atggcaatac ttcctaaagt 6120
gatatatcgc tttaacgcca tcccaattaa actgcctatg accttcttta cggagctcga 6180
gaaaacaacc cttaaattta tatggaatca aaagagagca agaatagcga agtccatctt 6240
gagccagaag aataaggccg gtgggattac tttgcctgat tttaagttgt attataaagc 6300
cacagtaact aagacagcct ggtattggta tcagaataga gacatcgacc agtggaatcg 6360
gaccgaacca tcagagataa tgccccacat ctataattac cttatattcg ataagccaga 6420

aaagaataaa cagtggggca aagacagcct cttcaacaag tggtgttggg agaattggct 6480
ggccatatgc cggaaactca agctcgaccc ctttcttaca ccctacacta aaatcaacag 6540
taggtggatc aaggacttga atgtcaagcc aaagactata aagacactgg aagagaatct 6600
tgggatcaca atacaagata taggcgtcgg caaagatttt atgtcaaaga cgcccaaggc 6660
catggccact aaggataaga ttgataagtg ggaccttatt aagctcaaaa gcttctgtac 6720
tgccaaggag accacgatca gagttaatag gcagcccact acatgggaaa agattttcgc 6780
cacttattca tcagataagg ggttgataag cagaatatat aacgagctga agcagatcta 6840

caagaagaaa acgaataatc ccatcaagaa gtgggcaaaa gatatgaaca ggcattttag 6900
caaagaggat atctacgccg cgaagaagca tatgaagaag tgtagttcaa gcttggccat 6960
tcgtgagatg cagattaaga cgaccatgcg ataccacctt accccagtga ggatggcaat 7020
tatcaagaaa tctggcaata atagatgttg gcggggctgt ggcgagattg gcaccctgct 7080
ccattgctgg tgggattgca agctggtgca gccgctttgg aaatcagtct ggcgctttct 7140
gagggacctc gagcttgaga ttcccttcga tcccgcaatt cccttgctcg gaatctatcc 7200
taacgaatac aagagctgtt gttacaagga tacgtgtacc cggatgttca tcgcggcctt 7260

gtttacgata gctaagacgt ggaatcagcc taagtgcccc acaatgatcg attggatcaa 7320

gaaaatgtgg catatttata ccatggagta ttacgcagca attaagaatg acgaatttat 7380

ttccttcgtt gggacctgga tgaagctgga gactattatt ctgagcaagc tgtctcagga 7440

gcaaaagaca aagcatagaa tcttctctct cattggtggt aacgactaca aagacgatga 7500

cgacaagtaa agcgcttcta gaagttgtct cctcctgcac tgactgactg atacaatcga 7560

tttctggatc cgcaggccta atcaacctct ggattacaaa atttgtgaaa gattgactgg 7620

tattcttaac tatgttgctc cttttacgct atgtggatac gctgctttaa tgcctttgta 7680

tcatgctatt gcttcccgta tggctttcat tttctcctcc ttgtataaat cctggttgct 7740

gtctctttat gaggagttgt ggcccgttgt caggcaacgt ggcgtggtgt gcactgtgtt 7800

tgctgacgca acccccactg gttggggcat tgccaccacc tgtcagctcc tttccgggac 7860

tttcgctttc cccctcccta ttgccacggc ggaactcatc gccgcctgcc ttgcccgctg 7920

ctggacaggg gctcggctgt tgggcactga caattccgtg gtgttgtcgg ggaagctgac 7980

gtcctttcca tggctgctcg cctgtgttgc cacctggatt ctgcgcggga cgtccttctg 8040

ctacgtccct tcggccctca atccagcgga ccttccttcc cgctgagaga cacaaaaaat 8100

tccaacacac tattgcaatg aaaataaatt tcctttatta gccagaagtc agatgctcaa 8160

ggggcttcat gatgtcccca taatttttgg cagagggaaa aagatctcag tggtatttgt 8220

gagccagggc attggccttc tgataggcag cctgcacctg aggagtgcgg ccgctttact 8280

tgtacagctc gtccatgccg agagtgatcc cggcggcggt cacgaactcc agcaggacca 8340

tgtgatcgcg cttctcgttg gggtctttgc tcagggcgga ctgggtgctc aggtagtggt 8400

tgtcgggcag cagcacgggg ccgtcgccga tgggggtgtt ctgctggtag tggtcggcga 8460

gctgcacgct gccgtcctcg atgttgtggc ggatcttgaa gttcaccttg atgccgttct 8520

tctgcttgtc ggccatgata tagacgttgt ggctgttgta gttgtactcc agcttgtgcc 8580

ccaggatgtt gccgtcctcc ttgaagtcga tgcccttcag ctcgatgcgg ttcaccaggg 8640

tgtcgccctc gaacttcacc tcggcgcggg tcttgtagtt gccgtcgtcc ttgaagaaga 8700

tggtgcgctc ctggacgtag ccttcgggca tggcggactt gaagaagtcg tgctgcttca 8760

tgtggtcggg gtagcggctg aagcactgca cgccgtaggt cagggtggtc acgagggtgg 8820

gccagggcac gggcagcttg ccggtggtgc agatgaactt cagggtcagc ttgccgtagg 8880

tggcatcgcc ctcgccctcg ccggacacgc tgaacttgtg gccgtttacg tcgccgtcca 8940

gctcgaccag gatgggcacc accccggtga acagctcctc gcccttgctc accatggtgg 9000

cgggatctga cggttcacta aaccagctct gcttatatag acctcccacc gtacacgcct 9060

accgcccatt tgcgtcaatg gggcggagtt gttacgacat tttggaaagt cccgttgatt 9120

ttggtgccaa aacaaactcc cattgacgtc aatggggtgg agacttggaa atccccgtga 9180

gtcaaaccgc tatccacgcc cattgatgta ctgccaaaac cgcatcacca tggtaatagc 9240

gatgactaat acgtagatgt actgccaagt aggaaagtcc cataaggtca tgtactgggc 9300

ataatgccag gcgggccatt taccgtcatt gacgtcaata gggggcgtac ttggcatatg 9360

atacacttga tgtactgcca agtgggcagt ttaccgtaaa tactccaccc attgacgtca 9420

atggaaagtc cctattggcg ttactatggg aacatacgtc attattgacg tcaatgggcg 9480

ggggtcgttg ggcggtcagc caggcgggcc atttaccgta agttatgtaa cgggcctgct 9540

gccggctctg cggcctcttc cgcgtcttcg ccttcgccct cagacgagtc ggatctccct 9600

ttgggccgcc tccccgcctg tctagcttga ctgactgaga tacagcgtac cttcagctca 9660

cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa 9720

aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca 9780

ataaacaagt taacaacaac aattgcattc attttatgtt tcaggttcag ggggaggtgt 9840

gggaggtttt ttaaagcaag taaaacctct acaaatgtgg tattggccca tctctatcgg 9900

tatcgtagca taacccctg gggcctctaa acgggtcttg aggggttttt tgtgcccctc 9960

gggccggatt gctatctacc ggcattggcg cagaaaaaaa tgcctgatgc gacgctgcgc 10020

gtcttatact cccacatatg ccagattcag caacggatac ggcttcccca acttgcccac 10080

ttccatacgt gtcctcctta ccagaaattt atccttaagg tcgtcagcta tcctgcaggc 10140

gatctctcga tttcgatcaa gacattcctt taatggtctt ttctggacac cactaggggt 10200

cagaagtagt tcatcaaact ttcttccctc cctaatctca ttggttacct tgggctatcg 10260

aaacttaatt aagcgatctg catctcaatt agtcagcaac catagtcccg cccctaactc 10320

cgcccatccc gcccctaact ccgcccagtt ccgcccattc tccgccccat cgctgactaa 10380

tttttttttat ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt 10440

gaggaggctt ttttggaggc ctaggctttt gcaaaggagg tagccaacat gattgaacaa 10500

gatggattgc acgcaggttc tcccgccgct tgggtggaga ggctattcgg ctatgactgg 10560

gcacaacaga caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc 10620

ccggttcttt ttgtcaagac cgacctgtcc ggtgccctga atgaactcca ggacgaggca 10680

gcgcggctat cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc 10740

actgaagcgg gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca 10800

tctcaccttg ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat 10860

acgcttgatc cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca 10920

cgtactcgga tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg 10980

ctcgcgccag ccgaactgtt cgccaggctc aaggcgcgga tgcccgacgg cgaggatctc 11040

gtcgtgaccc acggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct 11100

ggattcatcg actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct 11160

acccgtgata ttgctgaaga gcttggcggc gaatgggctg accgcttcct cgtgctttac 11220

ggtatcgccg ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagttcttc 11280

tagtatgtaa gccctgtgcc ttctagttgc cagccatctg ttgtttgccc ctccccgtg 11340

ccttccttga ccctggaagg tgccactccc actgtccttt cctaataaaa tgaggaaatt 11400

gcatcgcatt gtctgagtag gtgtcattct attctggggg gtggggtggg gcaggacagc 11460

aaggggagg attgggaaga caatagcagg catgctgggg atgcggtggg ctctatggtt 11520

aattaaccag tcaagtcagc tacttggcga gatcgacttg tctgggtttc gactacgctc 11580

agaattgcgt cagtcaagtt cgatctggtc cttgctattg cacccgttct ccgattacga 11640

gtttcattta aatcatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 11700

cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct 11760

caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccccctggaa 11820

gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc 11880

tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt 11940

aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 12000

ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta tcgccactgg 12060

cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct 12120

tgaagtggtg gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc 12180

tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 12240

ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 12300

aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 12360

aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa 12420

aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 12480

gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcat 12540

ttaaatttcc gaactctcca aggccctcgt cggaaaatct t 12581

<210> 100

<211> 12542

<212> DNA

<213> Artificial Sequence

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

<400> 100

ggggcgcgcc tggtgtaccg agaacgatcc tctcagtgcg agtctcgacg atccatatcg 60

ttgcttggca gtcagccagt cggaatccag cttgggaccc aggaagtcca atcgtcagat 120

attgtactca agcctggtca cggcagcgta ccgatctgtt taaacctaga tattgatagt 180

ctgatcggtc aacgtataat cgagtcctag cttttgcaaa catctatcaa gagacaggat 240

cagcaggagg ctttcgcatg agtattcaac atttccgtgt cgcccttatt ccctttttg 300

cggcattttg ccttcctgtt tttgctcacc cagaaacgct ggtgaaagta aaagatgctg 360

aagatcagtt gggtgcgcga gtgggttaca tcgaactgga tctcaacagc ggtaagatcc 420

ttgagagttt tcgccccgaa gaacgctttc caatgatgag cacttttaaa gttctgctat 480

gtggcgcggt attatcccgt attgacgccg ggcaagagca actcggtcgc cgcatacact 540

attctcagaa tgacttggtt gagtattcac cagtcacaga aaagcatctt acggatggca 600

tgacagtaag agaattatgc agtgctgcca taaccatgag tgataacact gcggccaact 660

tacttctgac aacgattgga ggaccgaagg agctaaccgc ttttttgcac aacatggggg 720

atcatgtaac tcgccttgat cgttgggaac cggagctgaa tgaagccata ccaaacgacg 780

agcgtgacac cacgatgcct gtagcaatgg caacaacctt gcgtaaacta ttaactggcg 840

aactacttac tctagcttcc cggcaacagt tgatagactg gatggaggcg gataaagttg 900

caggaccact tctgcgctcg gcccttccgg ctggctggtt tattgctgat aaatctggag 960

ccggtgagcg tgggtctcgc ggtatcattg cagcactggg gccagatggt aagccctccc 1020

gtatcgtagt tatctacacg acggggagtc aggcaactat ggatgaacga aatagacaga 1080

tcgctgagat aggtgcctca ctgattaagc attggtaacc gattctaggt gcattggcgc 1140

agaaaaaaat gcctgatgcg acgctgcgcg tcttatactc ccacatatgc cagattcagc 1200

aacggatacg gcttccccaa cttgcccact tccatacgtg tcctccttac cagaaattta 1260

tccttaagat cgtttaaact cgactctggc tctatcgaat ctccgtcgtt tcgagcttac 1320

gcgaacagcc gtggcgctca tttgctcgtc gggcatcgaa tctcgtcagc tatcgtcagc 1380

ttacctttt ggcagcgatc gcggctcccg acatcttgga ccattagctc cacaggtatc 1440

ttcttccctc tagtggtcat aacagcagct tcagctacct ctcaattcaa aaaacccctc 1500

aagacccgtt tagaggcccc aaggggttat gctatcaatc gttgcgttac acacacaaaa 1560

aaccaacaca catccatctt cgatggatag cgattttatt atctaactgc tgatcgagtg 1620

tagccagatc tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg 1680

cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 1740
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 1800
atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 1860
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 1920
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 1980
catgctgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 2040
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 2100

ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 2160
acggtgggag gtctatataa gcagagctgg tttagtgaac cgtcagatca gatctttgtc 2220
gatcctacca tccactcgac acacccgcca gcggccgcta atacgactca ctatagggag 2280
aagtactgcc accatgggca agaagcaaaa tcgcaagacg gggaattcca agacacaatc 2340
cgctagccca ccacctaaag agcgttctag ctcccctgct actgagcagt cctggatgga 2400
aaacgacttc gatgaactcc gggaagaggg atttaggcga tccaactatt cagaactccg 2460
cgaagatatc cagacaaagg ggaaggaagt cgagaatttc gagaagaacc tcgaggagtg 2520

catcacccgt atcacaaaca ctgagaaatg tctcaaagaa ctcatggaac ttaagacaaa 2580
agccagggag cttcgagagg agtgtcggag tctgagatcc aggtgtgacc agctcgagga 2640
gcgcgtgagc gcgatggaag acgagatgaa cgagatgaaa agagagggca aattcaggga 2700
gaagcgcatt aagaggaacg aacagagtct gcaggagatt tgggattacg tcaagaggcc 2760
taacctgcgg ttgatcggcg tccccgagag cgacgtagaa aacgggacta aactggagaa 2820
tacacttcaa gacatcattc aagaaaattt tccaaacctg gctcggcaag ctaatgtgca 2880
aatccaagag atccaacgca caccccagcg gtatagctct cggcgtgcca cccctaggca 2940

tattatcgtg cgctttacta aggtggagat gaaagagaag atgctgcgag ccgctcggga 3000
aaagggaagg gtgactttga agggcaaacc tattcggctg acggttgacc ttagcgccga 3060
gacactccag gcacgccggg aatggggccc catctttaat atcctgaagg agaagaactt 3120
ccagccacga atctcttacc ctgcaaagtt gagttttatc tccgagggtg agattaagta 3180
tttcatcgat aaacagatgc tgcgagactt cgtgacaact cgcccagctc tcaaggaact 3240
gctcaaagag gctcttaata tggagcgcaa taatagatat caacccttgc agaaccacgc 3300
aaagatgtga gacagccgtc agaccatcaa gactaggaag aaactgcatc aactaatgag 3360

caaaatcacc agctaacatc atagtataca tgccaaagaa gaagcggaag gtcaccggct 3420

ctaactcaca tatcaccatc cttacactta acattaacgg cctcaactca gctatcaagc 3480

gccatcggct ggccagctgg atcaaatcac aggatccaag cgtttgttgc atccaagaga 3540

cccacctgac ctgtagagat actcaccgcc tcaagatcaa gggatggcga aagatttatc 3600

aggcgaacgg taagcagaag aaagccggag tcgcaattct ggtctcagac aagacggatt 3660

tcaagcccac caaaattaag cgtgataagg aaggtcacta tattatggtg aaaggcagca 3720

tacagcagga agaacttacc atattgaaca tctacgcgcc aaacaccggc gcacctcgct 3780

ttatcaaaca ggtcctgtcc gatctgcagc gagatctgga ttctcatacg ttgattatgg 3840

gtgatttcaa tacaccattg agcaccctgg atcgcagcac caggcaaaag gtaaataaag 3900

acacgcaaga gctcaatagc gcactgcatc aggcagatct cattgatatt tatcgcactc 3960

ttcatcctaa gagtaccgag tacacattct tcagcgcccc acatcataca tactcaaaga 4020

tcgatcatat cgtcggctca aaggctctgc tgtcaaagtg caagcgcaca gagataatta 4080

caaattacct gtcagatcat agcgcgatca agctcgagct gagaatcaag aacctgaccc 4140

agagccggag taccacttgg aagcttaata acctgctgct caacgattat tgggtccaca 4200

atgagatgaa ggcagagatt aaaatgttct tcgaaacaaa tgagaataag gatactacct 4260

atcaaaacct ttgggatgcc tttaaggccg tctgcagagg caagttcatc gccctcaacg 4320

cctataaaag aaaacaagag agatctaaga tcgatactct cacctctcag ctgaaggagt 4380

tggagaaaca ggaacagacc cactccaagg cgtcaagacg gcaggagatc acaaagattc 4440

gcgccgagtt gaaagagatc gaaacccaaa agactcttca gaaaattaac gagtctcgta 4500

gttggttctt cgagcggatt aataagatag acagacctct ggcacgactg attaagaaga 4560

agcgcgaaaa gaaccagatt gataccatca agaacgacaa gggcgacatc actactgacc 4620

cgaccgagat ccagaccact attcgggagt attataagca tttgtatgct aacaagcttg 4680

agaacctgga agagatggac acttttctgg atacctatac tctgccacgg cttaatcaag 4740

aggaagtcga gtccctcaac cgcccaatta caggaagcga gattgtggcc ataattaact 4800

ccctgccgac aaagaaatct cctggtccgg acgggtttac agctgagttt tatcaacggt 4860

atatggaaga gcttgtaccg tttctgctca agctctttca gtctatagaa aaggaaggca 4920

tcttgcccaa ttccttctac gaagcttcta taatacttat tcccaaacca ggacgcgata 4980

ccacaaagaa ggaaaacttc cggcccatta gtctcatgaa tatcgacgct aaaatattga 5040

acaagattct cgccaacaga atccaacaac atattaagaa attgatacat cacgaccagg 5100

tggggtttat acctggcatg cagggctggt ttaacatccg gaagagtatt aacgtcattc 5160

aacacattaa tagagctaag gataagaatc atatgatcat ctctatagac gcggaaaagg 5220

cattcgataa gattcagcag ccatttatgc tcaagactct gaacaaactc ggcatcgacg 5280

gaacatattt taagattatt cgcgcaattt acgataagcc gactgctaac attatcctta 5340

acggccaaaa gctcgaggcc tttccgctca agactggaac ccgccaaggc tgtcccctct 5400

ccccgctttt gtttaatatt gtactcgagg tgctggctag ggctattcgt caagagaaag 5460

agattaaagg gatacagctc gggaaggaag aggtcaagct ttccttgttc gccgatgata 5520

tgattgtgta cctggagaat cctattgtgt ctgctcagaa ccttcttaaa cttatttcta 5580

actttagcaa ggtcagcggc tataagatta acgtccagaa atctcaggcc tttctgtaca 5640

caaataatcg acagaccgaa tcccagataa tgggtgagct tccgtttgtc atagccagca 5700

aaaggataaa gtatctcgga atccagctga cacgagacgt taaagatttg tttaaggaaa 5760

attacaagcc tctcctgaaa gagattaagg aagatactaa taagtggaag aatatcccct 5820

gttcatgggt tggcagaatc aacatagtga agatggcaat acttcctaaa gtgatatatc 5880

gctttaacgc catcccaatt aaactgccta tgaccttctt tacggagctc gagaaaacaa 5940

cccttaaatt tatatggaat caaaagagag caagaatagc gaagtccatc ttgagccaga 6000

agaataaggc cggtgggatt actttgcctg attttaagtt gtattataaa gccacagtaa 6060

ctaagacagc ctggtattgg tatcagaata gagacatcga ccagtggaat cggaccgaac 6120

catcagagat aatgccccac atctataatt accttatatt cgataagcca gaaaagaata 6180

aacagtgggg caaagacagc ctcttcaaca agtggtgttg ggagaattgg ctggccatat 6240

gccggaaact caagctcgac ccctttctta caccctacac taaaatcaac agtaggtgga 6300

tcaaggactt gaatgtcaag ccaaagacta taaagacact ggaagagaat cttgggatca 6360

caatacaaga tataggcgtc ggcaaagatt ttatgtcaaa gacgcccaag gccatggcca 6420

ctaaggataa gattgataag tgggacctta ttaagctcaa aagcttctgt actgccaagg 6480

agaccacgat cagagttaat aggcagccca ctacatggga aaagattttc gccacttatt 6540

catcagataa ggggttgata agcagaatat ataacgagct gaagcagatc tacaagaaga 6600

aaacgaataa tcccatcaag aagtgggcaa aagatatgaa caggcatttt agcaaagagg 6660

atatctacgc cgcgaagaag catatgaaga agtgtagttc aagcttggcc attcgtgaga 6720

tgcagattaa gacgaccatg cgataccacc ttaccccagt gaggatggca attatcaaga 6780

aatctggcaa taatagatgt tggcggggct gtggcgagat tggcaccctg ctccattgct 6840

ggtgggattg caagctggtg cagccgcttt ggaaatcagt ctggcgcttt ctgagggacc 6900

tcgagcttga gattcccttc gatcccgcaa ttcccttgct cggaatctat cctaacgaat 6960

acaagagctg ttgttacaag gatacgtgta cccggatgtt catcgcggcc ttgtttacga 7020

tagctaagac gtggaatcag cctaagtgcc ccacaatgat cgattggatc aagaaaatgt 7080

ggcatattta taccatggag tattacgcag caattaagaa tgacgaattt atttccttcg 7140

ttgggacctg gatgaagctg gagactatta ttctgagcaa gctgtctcag gagcaaaaga 7200
caaagcatag aatcttctct ctcattggtg gtaacgacta caaagacgat gacgacaagt 7260
aaagcgcttc tagaagttgt ctcctcctgc actgactgac tgatacaatc gatttctgga 7320
tccgcaggcc taatcaacct ctggattaca aaatttgtga aagattgact ggtattctta 7380
actatgttgc tccttttacg ctatgtggat acgctgcttt aatgcctttg tatcatgcta 7440
ttgcttcccg tatggctttc attttctcct ccttgtataa atcctggttg ctgtctcttt 7500
atgaggagtt gtggcccgtt gtcaggcaac gtggcgtggt gtgcactgtg tttgctgacg 7560

caacccccac tggttggggc attgccacca cctgtcagct cctttccggg actttcgctt 7620
tccccctccc tattgccacg gcggaactca tcgccgcctg ccttgcccgc tgctggacag 7680
gggctcggct gttgggcact gacaattccg tggtgttgtc ggggaagctg acgtcctttc 7740
catggctgct cgcctgtgtt gccacctgga ttctgcgcgg gacgtccttc tgctacgtcc 7800
cttcggccct caatccagcg gaccttcctt cccgctgaga gacacaaaaa attccaacac 7860
actattgcaa tgaaaataaa tttcctttat tagccagaag tcagatgctc aaggggcttc 7920
atgatgtccc cataattttt ggcagaggga aaaagatctc agtggtattt gtgagccagg 7980

gcattggcct tctgataggc agcctgcacc tgaggagtgc ggccgcttta cttgtacagc 8040
tcgtccatgc cgagagtgat cccggcggcg gtcacgaact ccagcaggac catgtgatcg 8100
cgcttctcgt tggggtcttt gctcagggcg gactgggtgc tcaggtagtg gttgtcgggc 8160
agcagcacgg ggccgtcgcc gatgggggtg ttctgctggt agtggtcggc gagctgcacg 8220
ctgccgtcct cgatgttgtg gcggatcttg aagttcacct tgatgccgtt cttctgcttg 8280
tcggccatga tatagacgtt gtggctgttg tagttgtact ccagcttgtg ccccaggatg 8340
ttgccgtcct ccttgaagtc gatgcccttc agctcgatgc ggttcaccag ggtgtcgccc 8400

tcgaacttca cctcggcgcg ggtcttgtag ttgccgtcgt ccttgaagaa gatggtgcgc 8460
tcctggacgt agccttcggg catggcggac ttgaagaagt cgtgctgctt catgtggtcg 8520
gggtagcggc tgaagcactg cacgccgtag gtcagggtgg tcacgagggt gggccagggc 8580
acgggcagct tgccggtggt gcagatgaac ttcagggtca gcttgccgta ggtggcatcg 8640
ccctcgccct cgccggacac gctgaacttg tggccgttta cgtcgccgtc cagctcgacc 8700
aggatgggca ccaccccggt gaacagctcc tcgcccttgc tcaccatggt ggcgggatct 8760
gacggttcac taaaccagct ctgcttatat agacctccca ccgtacacgc ctaccgccca 8820

tttgcgtcaa tggggcggag ttgttacgac attttggaaa gtcccgttga ttttggtgcc 8880

aaaacaaact cccattgacg tcaatggggt ggagacttgg aaatccccgt gagtcaaacc 8940

gctatccacg cccattgatg tactgccaaa accgcatcac catggtaata gcgatgacta 9000

atacgtagat gtactgccaa gtaggaaagt cccataaggt catgtactgg gcataatgcc 9060

aggcgggcca tttaccgtca ttgacgtcaa taggggggcgt acttggcata tgatacactt 9120

gatgtactgc caagtgggca gtttaccgta aatactccac ccattgacgt caatggaaag 9180

tccctattgg cgttactatg ggaacatacg tcattattga cgtcaatggg cgggggtcgt 9240

tgggcggtca gccaggcggg ccatttaccg taagttatgt aacgggcctg ctgccggctc 9300

tgcggcctct tccgcgtctt cgccttcgcc ctcagacgag tcggatctcc ctttgggccg 9360

cctccccgcc tgtctagctt gactgactga gatacagcgt accttcagct cacagacatg 9420

ataagataca ttgatgagtt tggacaaacc acaactagaa tgcagtgaaa aaaatgcttt 9480

atttgtgaaa tttgtgatgc tattgcttta tttgtaacca ttataagctg caataaacaa 9540

gttaacaaca acaattgcat tcattttatg tttcaggttc agggggaggt gtgggaggtt 9600

ttttaaagca agtaaaacct ctacaaatgt ggtattggcc catctctatc ggtatcgtag 9660

cataacccct tggggcctct aaacgggtct tgaggggttt tttgtgcccc tcgggccgga 9720

ttgctatcta ccggcattgg cgcagaaaaa aatgcctgat gcgacgctgc gcgtcttata 9780

ctcccacata tgccagattc agcaacggat acggcttccc caacttgccc acttccatac 9840

gtgtcctcct taccagaaat ttatccttaa ggtcgtcagc tatcctgcag gcgatctctc 9900

gatttcgatc aagacattcc tttaatggtc ttttctggac accactaggg gtcagaagta 9960

gttcatcaaa ctttcttccc tccctaatct cattggttac cttgggctat cgaaacttaa 10020

ttaagcgatc tgcatctcaa ttagtcagca accatagtcc cgcccctaac tccgcccatc 10080

ccgcccctaa ctccgcccag ttccgcccat tctccgcccc atcgctgact aatttttttt 10140

atttatgcag aggccgaggc cgcctcggcc tctgagctat tccagaagta gtgaggaggc 10200

ttttttggag gcctaggctt ttgcaaagga ggtagccaac atgattgaac aagatggatt 10260

gcacgcaggt tctcccgccg cttgggtgga gaggctattc ggctatgact gggcacaaca 10320

gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca gcgcaggggc gcccggttct 10380

ttttgtcaag accgacctgt ccggtgccct gaatgaactc caggacgagg cagcgcggct 10440

atcgtggctg gccacgacgg gcgttccttg cgcagctgtg ctcgacgttg tcactgaagc 10500

gggaagggac tggctgctat tgggcgaagt gccggggcag gatctcctgt catctcacct 10560

tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg cggcggctgc atacgcttga 10620

tccggctacc tgcccattcg accaccaagc gaaacatcgc atcgagcgag cacgtactcg 10680

gatggaagcc ggtcttgtcg atcaggatga tctggacgaa gagcatcagg ggctcgcgcc 10740

agccgaactg ttcgccaggc tcaaggcgcg gatgcccgac ggcgaggatc tcgtcgtgac 10800

ccacggcgat gcctgcttgc cgaatatcat ggtggaaaat ggccgctttt ctggattcat 10860

cgactgtggc cggctgggtg tggcggaccg ctatcaggac atagcgttgg ctacccgtga 10920

tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc ctcgtgcttt acggtatcgc 10980

cgctcccgat tcgcagcgca tcgccttcta tcgccttctt gacgagttct tctagtatgt 11040

aagccctgtg ccttctagtt gccagccatc tgttgtttgc ccctcccccg tgccttcctt 11100

gaccctggaa ggtgccactc ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca 11160

ttgtctgagt aggtgtcatt ctattctggg gggtggggtg gggcaggaca gcaaggggga 11220

ggattgggaa gacaatagca ggcatgctgg ggatgcggtg ggctctatgg ttaattaacc 11280

agtcaagtca gctacttggc gagatcgact tgtctgggtt tcgactacgc tcagaattgc 11340

gtcagtcaag ttcgatctgg tccttgctat tgcacccgtt ctccgattac gagtttcatt 11400

taaatcatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg 11460

gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag 11520

aggtggcgaa acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc 11580

gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg 11640

ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt 11700

cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc 11760

ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc 11820

actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg 11880

tggcctaact acggctacac tagaagaaca gtatttggta tctgcgctct gctgaagcca 11940

gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc 12000

ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat 12060

cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt 12120

ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt 12180

tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc 12240

agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc atttaaattt 12300

ccgaactctc caaggccctc gtcggaaaat cttcaaacct ttcgtccgat ccatcttgca 12360

ggctacctct cgaacgaact atcgcaagtc tcttggccgg ccttgcgcct tggctattgc 12420

ttggcagcgc ctatcgccag gtattactcc aatcccgaat atccgagatc gggatcaccc 12480

gagagaagtt caacctacat cctcaatccc gatctatccg agatccgagg aatatcgaaa 12540

tc 12542

<210> 101

<211> 12554

<212> DNA

<213> Artificial Sequence

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

<400> 101

tggccggcct tgcgccttgg ctattgcttg gcagcgccta tcgccaggta ttactccaat 60

cccgaatatc cgagatcggg atcacccgag agaagttcaa cctacatcct caatcccgat 120

ctatccgaga tccgaggaat atcgaaatcg gggcgcgcct ggtgtaccga gaacgatcct 180

ctcagtgcga gtctcgacga tccatatcgt tgcttggcag tcagccagtc ggaatccagc 240

ttgggaccca ggaagtccaa tcgtcagata ttgtactcaa gcctggtcac ggcagcgtac 300

cgatctgttt aaacctagat attgatagtc tgatcggtca acgtataatc gagtcctagc 360

ttttgcaaac atctatcaag agacaggatc agcaggaggc tttcgcatga gtattcaaca 420

tttccgtgtc gcccttattc cctttttgc ggcattttgc cttcctgttt ttgctcaccc 480

agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcgcgag tgggttacat 540

cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag aacgctttcc 600

aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgta ttgacgccgg 660

gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg agtattcacc 720

agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca gtgctgccat 780

aaccatgagt gataacactg cggccaactt acttctgaca acgattggag gaccgaagga 840

gctaaccgct tttttgcaca acatgggggga tcatgtaact cgccttgatc gttgggaacc 900

ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg tagcaatggc 960

aacaaccttg cgtaaactat taactggcga actacttact ctagcttccc ggcaacagtt 1020

gatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg cccttccggc 1080

tggctggttt attgctgata aatctggagc cggtgagcgt gggtctcgcg gtatcattgc 1140

agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga cggggagtca 1200

ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac tgattaagca 1260

ttggtaaccg attctaggtg cattggcgca gaaaaaaatg cctgatgcga cgctgcgcgt 1320

cttatactcc cacatatgcc agattcagca acggatacgg cttccccaac ttgcccactt 1380

ccatacgtgt cctccttacc agaaatttat ccttaagatc gtttaaactc gactctggct 1440

ctatcgaatc tccgtcgttt cgagcttacg cgaacagccg tggcgctcat ttgctcgtcg 1500

ggcatcgaat ctcgtcagct atcgtcagct tacctttttg gcagcgatcg cggctcccga 1560

catcttggac cattagctcc acaggtatct tcttccctct agtggtcata acagcagctt 1620

cagctacctc tcaattcaaa aaacccctca agacccgttt agaggcccca aggggttatg 1680

ctatcaatcg ttgcgttaca cacacaaaaa accaacacac atccatcttc gatggatagc 1740

gattttatta tctaactgct gatcgagtgt agccagatct agtaatcaat tacggggtca 1800

ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa tggcccgcct 1860

ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt tcccatagta 1920

acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac 1980

ttggcagtac atcaagtgta tcatatgcca agtacgcccc ctattgacgt caatgacggt 2040

aaatggcccg cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag 2100

tacatctacg tattagtcat cgctattacc atgctgatgc ggttttggca gtacatcaat 2160

gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat 2220

gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc 2280

ccattgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag cagagctggt 2340

ttagtgaacc gtcagatcag atctttgtcg atcctaccat ccactcgaca cacccgccag 2400

cggccgctaa tacgactcac tatagggaga agtactgcca ccatgggcaa gaagcaaaat 2460

cgcaagacgg ggaattccaa gacacaatcc gctagcccac cacctaaaga gcgttctagc 2520

tcccctgcta ctgagcagtc ctggatggaa aacgacttcg atgaactccg ggaagaggga 2580

tttaggcgat ccaactattc agaactccgc gaagatatcc agacaaaggg gaaggaagtc 2640

gagaatttcg agaagaacct cgaggagtgc atcacccgta tcacaaacac tgagaaatgt 2700

ctcaaagaac tcatggaact taagacaaaa gccagggagc ttcgagagga gtgtcggagt 2760

ctgagatcca ggtgtgacca gctcgaggag cgcgtgagcg cgatggaaga cgagatgaac 2820

gagatgaaaa gagagggcaa attcagggag aagcgcatta agaggaacga acagagtctg 2880

caggagattt gggattacgt caagaggcct aacctgcggt tgatcggcgt ccccgagagc 2940

gacgtagaaa acgggactaa actggagaat acacttcaag acatcattca agaaaatttt 3000

ccaaacctgg ctcggcaagc taatgtgcaa atccaagaga tccaacgcac accccagcgg 3060

tatagctctc ggcgtgccac ccctaggcat attatcgtgc gctttactaa ggtggagatg 3120

aaagagaaga tgctgcgagc cgctcgggaa aagggaaggg tgactttgaa gggcaaacct 3180

attcggctga cggttgacct tagcgccgag acactccagg cacgccggga atggggcccc 3240

atctttaata tcctgaagga gaagaacttc cagccacgaa tctcttaccc tgcaaagttg 3300
agttttatct ccgagggtga gattaagtat ttcatcgata aacagatgct gcgagacttc 3360
gtgacaactc gcccagctct caaggaactg ctcaaagagg ctcttaatat ggagcgcaat 3420
aatagatatc aacccttgca gaaccacgca aagatgtgag acagccgtca gaccatcaag 3480
actaggaaga aactgcatca actaatgagc aaaatcacca gctaacatca tagtatacat 3540
gccaaagaag aagcggaagg tcggcggcgg cagcaccggc tctaactcac atatcaccat 3600
ccttacactt aacattaacg gcctcaactc agctatcaag cgccatcggc tggccagctg 3660

gatcaaatca caggatccaa gcgtttgttg catccaagag acccacctga cctgtagaga 3720
tactcaccgc ctcaagatca agggatggcg aaagatttat caggcgaacg gtaagcagaa 3780
gaaagccgga gtcgcaattc tggtctcaga caagacggat ttcaagccca ccaaaattaa 3840
gcgtgataag gaaggtcact atattatggt gaaaggcagc atacagcagg aagaacttac 3900
catattgaac atctacgcgc caaacaccgg cgcacctcgc tttatcaaac aggtcctgtc 3960
cgatctgcag cgagatctgg attctcatac gttgattatg ggtgatttca atacaccatt 4020
gagcaccctg gatcgcagca ccaggcaaaa ggtaaataaa gacacgcaag agctcaatag 4080

cgcactgcat caggcagatc tcattgatat ttatcgcact cttcatccta agagtaccga 4140
gtacacattc ttcagcgccc cacatcatac atactcaaag atcgatcata tcgtcggctc 4200
aaaggctctg ctgtcaaagt gcaagcgcac agagataatt acaaattacc tgtcagatca 4260
tagcgcgatc aagctcgagc tgagaatcaa gaacctgacc cagagccgga gtaccacttg 4320
gaagcttaat aacctgctgc tcaacgatta ttgggtccac aatgagatga aggcagagat 4380
taaaatgttc ttcgaaacaa atgagaataa ggatactacc tatcaaaacc tttgggatgc 4440
ctttaaggcc gtctgcagag gcaagttcat cgccctcaac gcctataaaa gaaaacaaga 4500

gagatctaag atcgatactc tcacctctca gctgaaggag ttggagaaac aggaacagac 4560
ccactccaag gcgtcaagac ggcaggagat cacaaagatt cgcgccgagt tgaaagagat 4620
cgaaacccaa aagactcttc agaaaattaa cgagtctcgt agttggttct tcgagcggat 4680
taataagata gacagacctc tggcacgact gattaagaag aagcgcgaaa agaaccagat 4740
tgataccatc aagaacgaca agggcgacat cactactgac ccgaccgaga tccagaccac 4800
tattcgggag tattataagc atttgtatgc taacaagctt gagaacctgg aagagatgga 4860
cacttttctg gatacctata ctctgccacg gcttaatcaa gaggaagtcg agtccctcaa 4920

ccgcccaatt acaggaagcg agattgtggc cataattaac tccctgccga caaagaaatc 4980

tcctggtccg gacgggttta cagctgagtt ttatcaacgg tatatggaag agcttgtacc 5040

gtttctgctc aagctctttc agtctataga aaaggaaggc atcttgccca attccttcta 5100

cgaagcttct ataatactta ttcccaaacc aggacgcgat accacaaaga aggaaaactt 5160

ccggcccatt agtctcatga atatcgacgc taaaatattg aacaagattc tcgccaacag 5220

aatccaacaa catattaaga aattgataca tcacgaccag gtggggttta tacctggcat 5280

gcagggctgg tttaacatcc ggaagagtat taacgtcatt caacacatta atagagctaa 5340

ggataagaat catatgatca tctctataga cgcggaaaag gcattcgata agattcagca 5400

gccatttatg ctcaagactc tgaacaaact cggcatcgac ggaacatatt ttaagattat 5460

tcgcgcaatt tacgataagc cgactgctaa cattatcctt aacggccaaa agctcgaggc 5520

ctttccgctc aagactggaa cccgccaagg ctgtcccctc tccccgcttt tgtttaatat 5580

tgtactcgag gtgctggcta gggctattcg tcaagagaaa gagattaaag ggatacagct 5640

cgggaaggaa gaggtcaagc tttccttgtt cgccgatgat atgattgtgt acctggagaa 5700

tcctattgtg tctgctcaga accttcttaa acttatttct aactttagca aggtcagcgg 5760

ctataagatt aacgtccaga aatctcaggc ctttctgtac acaaataatc gacagaccga 5820

atcccagata atgggtgagc ttccgtttgt catagccagc aaaaggataa agtatctcgg 5880

aatccagctg acacgagacg ttaaagattt gtttaaggaa aattacaagc ctctcctgaa 5940

agagattaag gaagatacta ataagtggaa gaatatcccc tgttcatggg ttggcagaat 6000

caacatagtg aagatggcaa tacttcctaa agtgatatat cgctttaacg ccatcccaat 6060

taaactgcct atgaccttct ttacggagct cgagaaaaca acccttaaat ttatatggaa 6120

tcaaaagaga gcaagaatag cgaagtccat cttgagccag aagaataagg ccggtgggat 6180

tactttgcct gattttaagt tgtattataa agccacagta actaagacag cctggtattg 6240

gtatcagaat agagacatcg accagtggaa tcggaccgaa ccatcagaga taatgcccca 6300

catctataat taccttatat tcgataagcc agaaaagaat aaacagtggg gcaaagacag 6360

cctcttcaac aagtggtgtt gggagaattg gctggccata tgccggaaac tcaagctcga 6420

ccccttttctt acaccctaca ctaaaatcaa cagtaggtgg atcaaggact tgaatgtcaa 6480

gccaaagact ataaagacac tggaagagaa tcttgggatc acaatacaag atataggcgt 6540

cggcaaagat tttatgtcaa agacgcccaa ggccatggcc actaaggata agattgataa 6600

gtgggacctt attaagctca aaagcttctg tactgccaag gagaccacga tcagagttaa 6660

taggcagccc actacatggg aaaagatttt cgccacttat tcatcagata aggggttgat 6720

aagcagaata tataacgagc tgaagcagat ctacaagaag aaaacgaata atcccatcaa 6780

gaagtgggca aaagatatga acaggcattt tagcaaagag gatatctacg ccgcgaagaa 6840

gcatatgaag aagtgtagtt caagcttggc cattcgtgag atgcagatta agacgaccat 6900

gcgataccac cttaccccag tgaggatggc aattatcaag aaatctggca ataatagatg 6960

ttggcggggc tgtggcgaga ttggcaccct gctccattgc tggtgggatt gcaagctggt 7020

gcagccgctt tggaaatcag tctggcgctt tctgagggac ctcgagcttg agattccctt 7080

cgatcccgca attcccttgc tcggaatcta tcctaacgaa tacaagagct gttgttacaa 7140

ggatacgtgt acccggatgt tcatcgcggc cttgtttacg atagctaaga cgtggaatca 7200

gcctaagtgc cccacaatga tcgattggat caagaaaatg tggcatattt ataccatgga 7260

gtattacgca gcaattaaga atgacgaatt tatttccttc gttgggacct ggatgaagct 7320

ggagactatt attctgagca agctgtctca ggagcaaaag acaaagcata gaatcttctc 7380

tctcattggt ggtaacgact acaaagacga tgacgacaag taaagcgctt ctagaagttg 7440

tctcctcctg cactgactga ctgatacaat cgatttctgg atccgcaggc ctaatcaacc 7500

tctggattac aaaatttgtg aaagattgac tggtattctt aactatgttg ctccttttac 7560

gctatgtgga tacgctgctt taatgccttt gtatcatgct attgcttccc gtatggcttt 7620

cattttctcc tccttgtata aatcctggtt gctgtctctt tatgaggagt tgtggcccgt 7680

tgtcaggcaa cgtggcgtgg tgtgcactgt gtttgctgac gcaaccccca ctggttgggg 7740

cattgccacc acctgtcagc tcctttccgg gactttcgct ttccccctcc ctattgccac 7800

ggcggaactc atcgccgcct gccttgcccg ctgctggaca ggggctcggc tgttgggcac 7860

tgacaattcc gtggtgttgt cggggaagct gacgtccttt ccatggctgc tcgcctgtgt 7920

tgccacctgg attctgcgcg ggacgtcctt ctgctacgtc ccttcggccc tcaatccagc 7980

ggaccttcct tcccgctgag agacacaaaa aattccaaca cactattgca atgaaaataa 8040

atttccttta ttagccagaa gtcagatgct caaggggctt catgatgtcc ccataatttt 8100

tggcagaggg aaaaagatct cagtggtatt tgtgagccag ggcattggcc ttctgatagg 8160

cagcctgcac ctgaggagtg cggccgcttt acttgtacag ctcgtccatg ccgagagtga 8220

tcccggcggc ggtcacgaac tccagcagga ccatgtgatc gcgcttctcg ttggggtctt 8280

tgctcagggc ggactgggtg ctcaggtagt ggttgtcggg cagcagcacg gggccgtcgc 8340

cgatgggggt gttctgctgg tagtggtcgg cgagctgcac gctgccgtcc tcgatgttgt 8400

ggcggatctt gaagttcacc ttgatgccgt tcttctgctt gtcggccatg atatagacgt 8460

tgtggctgtt gtagttgtac tccagcttgt gccccaggat gttgccgtcc tccttgaagt 8520

cgatgccctt cagctcgatg cggttcacca gggtgtcgcc ctcgaacttc acctcggcgc 8580

gggtcttgta gttgccgtcg tccttgaaga agatggtgcg ctcctggacg tagccttcgg 8640

gcatggcgga cttgaagaag tcgtgctgct tcatgtggtc ggggtagcgg ctgaagcact 8700

gcacgccgta ggtcagggtg gtcacgaggg tgggccaggg cacgggcagc ttgccggtgg 8760
tgcagatgaa cttcagggtc agcttgccgt aggtggcatc gccctcgccc tcgccggaca 8820
cgctgaactt gtggccgttt acgtcgccgt ccagctcgac caggatgggc accaccccgg 8880
tgaacagctc ctcgcccttg ctcaccatgg tggcgggatc tgacggttca ctaaaccagc 8940
tctgcttata tagacctccc accgtacacg cctaccgccc atttgcgtca atggggcgga 9000
gttgttacga cattttggaa agtcccgttg attttggtgc caaaacaaac tcccattgac 9060
gtcaatgggg tggagacttg gaaatccccg tgagtcaaac cgctatccac gcccattgat 9120

gtactgccaa aaccgcatca ccatggtaat agcgatgact aatacgtaga tgtactgcca 9180
agtaggaaag tcccataagg tcatgtactg ggcataatgc caggcgggcc atttaccgtc 9240
attgacgtca ataggggggcg tacttggcat atgatacact tgatgtactg ccaagtgggc 9300
agtttaccgt aaatactcca cccattgacg tcaatggaaa gtccctattg gcgttactat 9360
gggaacatac gtcattattg acgtcaatgg gcgggggtcg ttgggcggtc agccaggcgg 9420
gccatttacc gtaagttatg taacgggcct gctgccggct ctgcggcctc ttccgcgtct 9480
tcgccttcgc cctcagacga gtcggatctc cctttgggcc gcctccccgc ctgtctagct 9540

tgactgactg agatacagcg taccttcagc tcacagacat gataagatac attgatgagt 9600
ttggacaaac cacaactaga atgcagtgaa aaaaatgctt tatttgtgaa atttgtgatg 9660
ctattgcttt atttgtaacc attataagct gcaataaaca agttaacaac aacaattgca 9720
ttcattttat gtttcaggtt caggggggagg tgtgggaggt tttttaaagc aagtaaaacc 9780
tctacaaatg tggtattggc ccatctctat cggtatcgta gcataacccc ttggggcctc 9840
taaacgggtc ttgaggggtt ttttgtgccc ctcgggccgg attgctatct accggcattg 9900
gcgcagaaaa aaatgcctga tgcgacgctg cgcgtcttat actcccacat atgccagatt 9960

cagcaacgga tacggcttcc ccaacttgcc cacttccata cgtgtcctcc ttaccagaaa 10020
tttatcctta aggtcgtcag ctatcctgca ggcgatctct cgatttcgat caagacattc 10080
ctttaatggt cttttctgga caccactagg ggtcagaagt agttcatcaa actttcttcc 10140
ctccctaatc tcattggtta ccttgggcta tcgaaactta attaagcgat ctgcatctca 10200
attagtcagc aaccatagtc ccgcccctaa ctccgcccat cccgccccta actccgccca 10260
gttccgccca ttctccgccc catcgctgac taatttttt tatttatgca gaggccgagg 10320
ccgcctcggc ctctgagcta ttccagaagt agtgaggagg ctttttgga ggcctaggct 10380

tttgcaaagg aggtagccaa catgattgaa caagatggat tgcacgcagg ttctcccgcc 10440

gcttgggtgg agaggctatt cggctatgac tgggcacaac agacaatcgg ctgctctgat 10500

gccgccgtgt tccggctgtc agcgcagggg cgcccggttc tttttgtcaa gaccgacctg 10560

tccggtgccc tgaatgaact ccaggacgag gcagcgcggc tatcgtggct ggccacgacg 10620

ggcgttcctt gcgcagctgt gctcgacgtt gtcactgaag cgggaaggga ctggctgcta 10680

ttgggcgaag tgccggggca ggatctcctg tcatctcacc ttgctcctgc cgagaaagta 10740

tccatcatgg ctgatgcaat gcggcggctg catacgcttg atccggctac ctgcccattc 10800

gaccaccaag cgaaacatcg catcgagcga gcacgtactc ggatggaagc cggtcttgtc 10860

gatcaggatg atctggacga agagcatcag gggctcgcgc cagccgaact gttcgccagg 10920

ctcaaggcgc ggatgcccga cggcgaggat ctcgtcgtga cccacggcga tgcctgcttg 10980

ccgaatatca tggtggaaaa tggccgcttt tctggattca tcgactgtgg ccggctgggt 11040

gtggcggacc gctatcagga catagcgttg gctacccgtg atattgctga agagcttggc 11100

ggcgaatggg ctgaccgctt cctcgtgctt tacggtatcg ccgctcccga ttcgcagcgc 11160

atcgccttct atcgccttct tgacgagttc ttctagtatg taagccctgt gccttctagt 11220

tgccagccat ctgttgtttg cccctccccc gtgccttcct tgaccctgga aggtgccact 11280

cccactgtcc tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat 11340

tctattctgg ggggtggggt ggggcaggac agcaaggggg aggattggga agacaatagc 11400

aggcatgctg gggatgcggt gggctctatg gttaattaac cagtcaagtc agctacttgg 11460

cgagatcgac ttgtctgggt ttcgactacg ctcagaattg cgtcagtcaa gttcgatctg 11520

gtccttgcta ttgcacccgt tctccgatta cgagtttcat ttaaatcatg tgagcaaaag 11580

gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 11640

gccccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 11700

gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 11760

ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc 11820

atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg 11880

tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt 11940

ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 12000

gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca 12060

ctagaagaac agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 12120

ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca 12180

agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 12240

ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa 12300

aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta 12360

tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag 12420

cgatctgtct atttcgttca tccatagttg catttaaatt tccgaactct ccaaggccct 12480

cgtcggaaaa tcttcaaacc tttcgtccga tccatcttgc aggctacctc tcgaacgaac 12540

tatcgcaagt ctct 12554

<210> 102

<211> 12569

<212> DNA

<213> Artificial Sequence

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

<400> 102

gacgatccat atcgttgctt ggcagtcagc cagtcggaat ccagcttggg acccaggaag 60

tccaatcgtc agatattgta ctcaagcctg gtcacggcag cgtaccgatc tgtttaaacc 120

tagatattga tagtctgatc ggtcaacgta taatcgagtc ctagcttttg caaacatcta 180

tcaagagaca ggatcagcag gaggctttcg catgagtatt caacatttcc gtgtcgccct 240

tattcccttt tttgcggcat tttgccttcc tgtttttgct cacccagaaa cgctggtgaa 300

agtaaaagat gctgaagatc agttgggtgc gcgagtgggt tacatcgaac tggatctcaa 360

cagcggtaag atccttgaga gttttcgccc cgaagaacgc tttccaatga tgagcacttt 420

taaagttctg ctatgtggcg cggtattatc ccgtattgac gccgggcaag agcaactcgg 480

tcgccgcata cactattctc agaatgactt ggttgagtat tcaccagtca cagaaaagca 540

tcttacggat ggcatgacag taagagaatt atgcagtgct gccataacca tgagtgataa 600

cactgcggcc aacttacttc tgacaacgat tggaggaccg aaggagctaa ccgcttttt 660

gcacaacatg ggggatcatg taactcgcct tgatcgttgg gaaccggagc tgaatgaagc 720

cataccaaac gacgagcgtg acaccacgat gcctgtagca atggcaacaa ccttgcgtaa 780

actattaact ggcgaactac ttactctagc ttcccggcaa cagttgatag actggatgga 840

ggcggataaa gttgcaggac cacttctgcg ctcggccctt ccggctggct ggtttattgc 900

tgataaatct ggagccggtg agcgtgggtc tcgcggtatc attgcagcac tggggccaga 960

tggtaagccc tcccgtatcg tagttatcta cacgacgggg agtcaggcaa ctatggatga 1020

acgaaataga cagatcgctg agataggtgc ctcactgatt aagcattggt aaccgattct 1080

aggtgcattg gcgcagaaaa aaatgcctga tgcgacgctg cgcgtcttat actcccacat 1140

atgccagatt cagcaacgga tacggcttcc ccaacttgcc cacttccata cgtgtcctcc 1200

ttaccagaaa tttatcctta agatcgttta aactcgactc tggctctatc gaatctccgt 1260

cgtttcgagc ttacgcgaac agccgtggcg ctcatttgct cgtcgggcat cgaatctcgt 1320

cagctatcgt cagcttacct ttttggcagc gatcgcggct cccgacatct tggaccatta 1380

gctccacagg tatcttcttc cctctagtgg tcataacagc agcttcagct acctctcaat 1440

tcaaaaaacc cctcaagacc cgtttagagg ccccaagggg ttatgctatc aatcgttgcg 1500

ttacacacac aaaaaaccaa cacacatcca tcttcgatgg atagcgattt tattatctaa 1560

ctgctgatcg agtgtagcca gatctagtaa tcaattacgg ggtcattagt tcatagccca 1620

tatatggagt tccgcgttac ataacttacg gtaaatggcc cgcctggctg accgcccaac 1680

gacccccgcc cattgacgtc aataatgacg tatgttccca tagtaacgcc aatagggact 1740

ttccattgac gtcaatgggt ggagtattta cggtaaactg cccacttggc agtacatcaa 1800

gtgtatcata tgccaagtac gccccctatt gacgtcaatg acggtaaatg gcccgcctgg 1860

cattatgccc agtacatgac cttatgggac tttcctactt ggcagtacat ctacgtatta 1920

gtcatcgcta ttaccatgct gatgcggttt tggcagtaca tcaatgggcg tggatagcgg 1980

tttgactcac ggggatttcc aagtctccac cccattgacg tcaatgggag tttgttttgg 2040

caccaaaatc aacgggactt tccaaaatgt cgtaacaact ccgccccatt gacgcaaatg 2100

ggcggtaggc gtgtacggtg ggaggtctat ataagcagag ctggtttagt gaaccgtcag 2160

atcagatctt tgtcgatcct accatccact cgacacaccc gccagcggcc gctaatacga 2220

ctcactatag ggagaagtac tgccaccatg ggcaagaagc aaaatcgcaa gacggggaat 2280

tccaagacac aatccgctag cccaccacct aaagagcgtt ctagctcccc tgctactgag 2340

cagtcctgga tggaaaacga cttcgatgaa ctccgggaag agggatttag gcgatccaac 2400

tattcagaac tccgcgaaga tatccagaca aaggggaagg aagtcgagaa tttcgagaag 2460

aacctcgagg agtgcatcac ccgtatcaca aacactgaga aatgtctcaa agaactcatg 2520

gaacttaaga caaaagccag ggagcttcga gaggagtgtc ggagtctgag atccaggtgt 2580

gaccagctcg aggagcgcgt gagcgcgatg gaagacgaga tgaacgagat gaaaagagag 2640

ggcaaattca gggagaagcg cattaagagg aacgaacaga gtctgcagga gatttgggat 2700

tacgtcaaga ggcctaacct gcggttgatc ggcgtccccg agagcgacgt agaaaacggg 2760

actaaactgg agaatacact tcaagacatc attcaagaaa attttccaaa cctggctcgg 2820

caagctaatg tgcaaatcca agagatccaa cgcacacccc agcggtatag ctctcggcgt 2880

gccaccccta ggcatattat cgtgcgcttt actaaggtgg agatgaaaga gaagatgctg 2940

cgagccgctc gggaaaaggg aagggtgact ttgaagggca aacctattcg gctgacggtt 3000

gaccttagcg ccgagacact ccaggcacgc cgggaatggg gccccatctt taatatcctg 3060

aaggagaaga acttccagcc acgaatctct taccctgcaa agttgagttt tatctccgag 3120

ggtgagatta agtatttcat cgataaacag atgctgcgag acttcgtgac aactcgccca 3180

gctctcaagg aactgctcaa agaggctctt aatatggagc gcaataatag atatcaaccc 3240

ttgcagaacc acgcaaagat gtgagacagc cgtcagacca tcaagactag gaagaaactg 3300

catcaactaa tgagcaaaat caccagctaa catcatagta tacatgaaaa ggccggcggc 3360

cacgaaaaag gccggccagg caaaaaagaa aaagaccggc tctaactcac atatcaccat 3420

ccttacactt aacattaacg gcctcaactc agctatcaag cgccatcggc tggccagctg 3480

gatcaaatca caggatccaa gcgtttgttg catccaagag acccacctga cctgtagaga 3540

tactcaccgc ctcaagatca agggatggcg aaagatttat caggcgaacg gtaagcagaa 3600

gaaagccgga gtcgcaattc tggtctcaga caagacggat ttcaagccca ccaaaattaa 3660

gcgtgataag gaaggtcact atattatggt gaaaggcagc atacagcagg aagaacttac 3720

catattgaac atctacgcgc caaacaccgg cgcacctcgc tttatcaaac aggtcctgtc 3780

cgatctgcag cgagatctgg attctcatac gttgattatg ggtgatttca atacaccatt 3840

gagcaccctg gatcgcagca ccaggcaaaa ggtaaataaa gacacgcaag agctcaatag 3900

cgcactgcat caggcagatc tcattgatat ttatcgcact cttcatccta agagtaccga 3960

gtacacattc ttcagcgccc cacatcatac atactcaaag atcgatcata tcgtcggctc 4020

aaaggctctg ctgtcaaagt gcaagcgcac agagataatt acaaattacc tgtcagatca 4080

tagcgcgatc aagctcgagc tgagaatcaa gaacctgacc cagagccgga gtaccacttg 4140

gaagcttaat aacctgctgc tcaacgatta ttgggtccac aatgagatga aggcagagat 4200

taaaatgttc ttcgaaacaa atgagaataa ggatactacc tatcaaaacc tttgggatgc 4260

ctttaaggcc gtctgcagag gcaagttcat cgccctcaac gcctataaaa gaaaacaaga 4320

gagatctaag atcgatactc tcacctctca gctgaaggag ttggagaaac aggaacagac 4380

ccactccaag gcgtcaagac ggcaggagat cacaaagatt cgcgccgagt tgaaagagat 4440

cgaaacccaa aagactcttc agaaaattaa cgagtctcgt agttggttct tcgagcggat 4500

taataagata gacagacctc tggcacgact gattaagaag aagcgcgaaa agaaccagat 4560

tgataccatc aagaacgaca agggcgacat cactactgac ccgaccgaga tccagaccac 4620

tattcgggag tattataagc atttgtatgc taacaagctt gagaacctgg aagagatgga 4680

cacttttctg gatacctata ctctgccacg gcttaatcaa gaggaagtcg agtccctcaa 4740

ccgcccaatt acaggaagcg agattgtggc cataattaac tccctgccga caaagaaatc 4800

tcctggtccg gacgggttta cagctgagtt ttatcaacgg tatatggaag agcttgtacc 4860
gtttctgctc aagctctttc agtctataga aaaggaaggc atcttgccca attccttcta 4920
cgaagcttct ataatactta ttcccaaacc aggacgcgat accacaaaga aggaaaactt 4980
ccggcccatt agtctcatga atatcgacgc taaaatattg aacaagattc tcgccaacag 5040
aatccaacaa catattaaga aattgataca tcacgaccag gtggggttta tacctggcat 5100
gcagggctgg tttaacatcc ggaagagtat taacgtcatt caacacatta atagagctaa 5160
ggataagaat catatgatca tctctataga cgcggaaaag gcattcgata agattcagca 5220

gccatttatg ctcaagactc tgaacaaact cggcatcgac ggaacatatt ttaagattat 5280
tcgcgcaatt tacgataagc cgactgctaa cattatcctt aacggccaaa agctcgaggc 5340
ctttccgctc aagactggaa cccgccaagg ctgtcccctc tccccgcttt tgtttaatat 5400
tgtactcgag gtgctggcta gggctattcg tcaagagaaa gagattaaag ggatacagct 5460
cgggaaggaa gaggtcaagc tttccttgtt cgccgatgat atgattgtgt acctggagaa 5520
tcctattgtg tctgctcaga accttcttaa acttatttct aactttagca aggtcagcgg 5580
ctataagatt aacgtccaga aatctcaggc ctttctgtac acaaataatc gacagaccga 5640

atcccagata atgggtgagc ttccgtttgt catagccagc aaaaggataa agtatctcgg 5700
aatccagctg acacgagacg ttaaagattt gtttaaggaa aattacaagc ctctcctgaa 5760
agagattaag gaagatacta ataagtggaa gaatatcccc tgttcatggg ttggcagaat 5820
caacatagtg aagatggcaa tacttcctaa agtgatatat cgctttaacg ccatcccaat 5880
taaactgcct atgaccttct ttacggagct cgagaaaaca acccttaaat ttatatggaa 5940
tcaaaagaga gcaagaatag cgaagtccat cttgagccag aagaataagg ccggtgggat 6000
tactttgcct gattttaagt tgtattataa agccacagta actaagacag cctggtattg 6060

gtatcagaat agagacatcg accagtggaa tcggaccgaa ccatcagaga taatgcccca 6120
catctataat taccttatat tcgataagcc agaaaagaat aaacagtggg gcaaagacag 6180
cctcttcaac aagtggtgtt gggagaattg gctggccata tgccggaaac tcaagctcga 6240
ccccttttctt acaccctaca ctaaaatcaa cagtaggtgg atcaaggact tgaatgtcaa 6300
gccaaagact ataaagacac tggaagagaa tcttgggatc acaatacaag atataggcgt 6360
cggcaaagat tttatgtcaa agacgcccaa ggccatggcc actaaggata agattgataa 6420
gtgggacctt attaagctca aaagcttctg tactgccaag gagaccacga tcagagttaa 6480

taggcagccc actacatggg aaaagatttt cgccacttat tcatcagata aggggttgat 6540

aagcagaata tataacgagc tgaagcagat ctacaagaag aaaacgaata atcccatcaa 6600

gaagtgggca aaagatatga acaggcattt tagcaaagag gatatctacg ccgcgaagaa 6660

gcatatgaag aagtgtagtt caagcttggc cattcgtgag atgcagatta agacgaccat 6720

gcgataccac cttaccccag tgaggatggc aattatcaag aaatctggca ataatagatg 6780

ttggcggggc tgtggcgaga ttggcaccct gctccattgc tggtgggatt gcaagctggt 6840

gcagccgctt tggaaatcag tctggcgctt tctgagggac ctcgagcttg agattccctt 6900

cgatcccgca attcccttgc tcggaatcta tcctaacgaa tacaagagct gttgttacaa 6960

ggatacgtgt acccggatgt tcatcgcggc cttgtttacg atagctaaga cgtggaatca 7020

gcctaagtgc cccacaatga tcgattggat caagaaaatg tggcatattt ataccatgga 7080

gtattacgca gcaattaaga atgacgaatt tatttccttc gttgggacct ggatgaagct 7140

ggagactatt attctgagca agctgtctca ggagcaaaag acaaagcata gaatcttctc 7200

tctcattggt ggtaacgact acaaagacga tgacgacaag taaagcgctt ctagaagttg 7260

tctcctcctg cactgactga ctgatacaat cgatttctgg atccgcaggc ctaatcaacc 7320

tctggattac aaaatttgtg aaagattgac tggtattctt aactatgttg ctccttttac 7380

gctatgtgga tacgctgctt taatgccttt gtatcatgct attgcttccc gtatggcttt 7440

cattttctcc tccttgtata aatcctggtt gctgtctctt tatgaggagt tgtggcccgt 7500

tgtcaggcaa cgtggcgtgg tgtgcactgt gtttgctgac gcaaccccca ctggttgggg 7560

cattgccacc acctgtcagc tcctttccgg gactttcgct ttccccctcc ctattgccac 7620

ggcggaactc atcgccgcct gccttgcccg ctgctggaca ggggctcggc tgttgggcac 7680

tgacaattcc gtggtgttgt cggggaagct gacgtccttt ccatggctgc tcgcctgtgt 7740

tgccacctgg attctgcgcg ggacgtcctt ctgctacgtc ccttcggccc tcaatccagc 7800

ggaccttcct tcccgctgag agacacaaaa aattccaaca cactattgca atgaaaataa 7860

atttccttta ttagccagaa gtcagatgct caaggggctt catgatgtcc ccataatttt 7920

tggcagaggg aaaaagatct cagtggtatt tgtgagccag ggcattggcc ttctgatagg 7980

cagcctgcac ctgaggagtg cggccgcttt acttgtacag ctcgtccatg ccgagagtga 8040

tcccggcggc ggtcacgaac tccagcagga ccatgtgatc gcgcttctcg ttggggtctt 8100

tgctcagggc ggactgggtg ctcaggtagt ggttgtcggg cagcagcacg gggccgtcgc 8160

cgatgggggt gttctgctgg tagtggtcgg cgagctgcac gctgccgtcc tcgatgttgt 8220

ggcggatctt gaagttcacc ttgatgccgt tcttctgctt gtcggccatg atatagacgt 8280

tgtggctgtt gtagttgtac tccagcttgt gccccaggat gttgccgtcc tccttgaagt 8340

cgatgccctt cagctcgatg cggttcacca gggtgtcgcc ctcgaacttc acctcggcgc 8400

gggtcttgta gttgccgtcg tccttgaaga agatggtgcg ctcctggacg tagccttcgg 8460

gcatggcgga cttgaagaag tcgtgctgct tcatgtggtc ggggtagcgg ctgaagcact 8520

gcacgccgta ggtcagggtg gtcacgaggg tgggccaggg cacgggcagc ttgccggtgg 8580

tgcagatgaa cttcagggtc agcttgccgt aggtggcatc gccctcgccc tcgccggaca 8640

cgctgaactt gtggccgttt acgtcgccgt ccagctcgac caggatgggc accaccccgg 8700

tgaacagctc ctcgcccttg ctcaccatgg tggcgggatc tgacggttca ctaaaccagc 8760

tctgcttata tagacctccc accgtacacg cctaccgccc atttgcgtca atggggcgga 8820

gttgttacga cattttggaa agtcccgttg attttggtgc caaaacaaac tcccattgac 8880

gtcaatgggg tggagacttg gaaatccccg tgagtcaaac cgctatccac gcccattgat 8940

gtactgccaa aaccgcatca ccatggtaat agcgatgact aatacgtaga tgtactgcca 9000

agtaggaaag tcccataagg tcatgtactg ggcataatgc caggcgggcc atttaccgtc 9060

attgacgtca ataggggggcg tacttggcat atgatacact tgatgtactg ccaagtgggc 9120

agtttaccgt aaatactcca cccattgacg tcaatggaaa gtccctattg gcgttactat 9180

gggaacatac gtcattattg acgtcaatgg gcgggggtcg ttgggcggtc agccaggcgg 9240

gccatttacc gtaagttatg taacgggcct gctgccggct ctgcggcctc ttccgcgtct 9300

tcgccttcgc cctcagacga gtcggatctc cctttgggcc gcctccccgc ctgtctagct 9360

tgactgactg agatacagcg taccttcagc tcacagacat gataagatac attgatgagt 9420

ttggacaaac cacaactaga atgcagtgaa aaaaatgctt tatttgtgaa atttgtgatg 9480

ctattgcttt atttgtaacc attataagct gcaataaaca agttaacaac aacaattgca 9540

ttcattttat gtttcaggtt cagggggagg tgtgggaggt tttttaaagc aagtaaaacc 9600

tctacaaatg tggtattggc ccatctctat cggtatcgta gcataacccc ttggggcctc 9660

taaacgggtc ttgaggggtt ttttgtgccc ctcgggccgg attgctatct accggcattg 9720

gcgcagaaaa aaatgcctga tgcgacgctg cgcgtcttat actcccacat atgccagatt 9780

cagcaacgga tacggcttcc ccaacttgcc cacttccata cgtgtcctcc ttaccagaaa 9840

tttatcctta aggtcgtcag ctatcctgca ggcgatctct cgatttcgat caagacattc 9900

ctttaatggt cttttctgga caccactagg ggtcagaagt agttcatcaa actttcttcc 9960

ctccctaatc tcattggtta ccttgggcta tcgaaactta attaagcgat ctgcatctca 10020

attagtcagc aaccatagtc ccgcccctaa ctccgcccat cccgccccta actccgccca 10080

gttccgccca ttctccgccc catcgctgac taatttttt tatttatgca gaggccgagg 10140

ccgcctcggc ctctgagcta ttccagaagt agtgaggagg ctttttttgga ggcctaggct 10200

tttgcaaagg aggtagccaa catgattgaa caagatggat tgcacgcagg ttctcccgcc 10260

gcttgggtgg agaggctatt cggctatgac tgggcacaac agacaatcgg ctgctctgat 10320
gccgccgtgt tccggctgtc agcgcagggg cgcccggttc tttttgtcaa gaccgacctg 10380
tccggtgccc tgaatgaact ccaggacgag gcagcgcggc tatcgtggct ggccacgacg 10440
ggcgttcctt gcgcagctgt gctcgacgtt gtcactgaag cgggaaggga ctggctgcta 10500
ttgggcgaag tgccggggca ggatctcctg tcatctcacc ttgctcctgc cgagaaagta 10560
tccatcatgg ctgatgcaat gcggcggctg catacgcttg atccggctac ctgcccattc 10620
gaccaccaag cgaaacatcg catcgagcga gcacgtactc ggatggaagc cggtcttgtc 10680

gatcaggatg atctggacga agagcatcag gggctcgcgc cagccgaact gttcgccagg 10740
ctcaaggcgc ggatgcccga cggcgaggat ctcgtcgtga cccacggcga tgcctgcttg 10800
ccgaatatca tggtggaaaa tggccgcttt tctggattca tcgactgtgg ccggctgggt 10860
gtggcggacc gctatcagga catagcgttg gctacccgtg atattgctga agagcttggc 10920
ggcgaatggg ctgaccgctt cctcgtgctt tacggtatcg ccgctcccga ttcgcagcgc 10980
atcgccttct atcgccttct tgacgagttc ttctagtatg taagccctgt gccttctagt 11040
tgccagccat ctgttgtttg cccctccccc gtgccttcct tgaccctgga aggtgccact 11100

cccactgtcc tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat 11160
tctattctgg ggggtggggt ggggcaggac agcaaggggg aggattggga agacaatagc 11220
aggcatgctg gggatgcggt gggctctatg gttaattaac cagtcaagtc agctacttgg 11280
cgagatcgac ttgtctgggt ttcgactacg ctcagaattg cgtcagtcaa gttcgatctg 11340
gtccttgcta ttgcacccgt tctccgatta cgagtttcat ttaaatcatg tgagcaaaag 11400
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 11460
gccccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 11520

gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 11580
ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc 11640
atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg 11700
tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt 11760
ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 11820
gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca 11880
ctagaagaac agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 11940

ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca 12000

agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 12060

ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa 12120

aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta 12180

tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag 12240

cgatctgtct atttcgttca tccatagttg catttaaatt tccgaactct ccaaggccct 12300

cgtcggaaaa tcttcaaacc tttcgtccga tccatcttgc aggctacctc tcgaacgaac 12360

tatcgcaagt ctcttggccg gccttgcgcc ttggctattg cttggcagcg cctatcgcca 12420

ggtattactc caatcccgaa tatccgagat cgggatcacc cgagagaagt tcaacctaca 12480

tcctcaatcc cgatctatcc gagatccgag gaatatcgaa atcggggcgc gcctggtgta 12540

ccgagaacga tcctctcagt gcgagtctc 12569

<210> 103

<211> 12581

<212> DNA

<213> Artificial Sequence

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

<400> 103

agatccgagg aatatcgaaa tcggggcgcg cctggtgtac cgagaacgat cctctcagtg 60

cgagtctcga cgatccatat cgttgcttgg cagtcagcca gtcggaatcc agcttgggac 120

ccaggaagtc caatcgtcag atattgtact caagcctggt cacggcagcg taccgatctg 180

tttaaaccta gatattgata gtctgatcgg tcaacgtata atcgagtcct agcttttgca 240

aacatctatc aagagacagg atcagcagga ggctttcgca tgagtattca acatttccgt 300

gtcgcccttа ttccctttt tgcggcattt tgccttcctg tttttgctca cccagaaacg 360

ctggtgaaag taaaagatgc tgaagatcag ttgggtgcgc gagtgggtta catcgaactg 420

gatctcaaca gcggtaagat ccttgagagt tttcgccccg aagaacgctt tccaatgatg 480

agcactttta aagttctgct atgtggcgcg gtattatccc gtattgacgc cgggcaagag 540

caactcggtc gccgcataca ctattctcag aatgacttgg ttgagtattc accagtcaca 600

gaaaagcatc ttacggatgg catgacagta agagaattat gcagtgctgc cataaccatg 660

agtgataaca ctgcggccaa cttacttctg acaacgattg gaggaccgaa ggagctaacc 720

gctttttgc acaacatggg ggatcatgta actcgccttg atcgttggga accggagctg 780

aatgaagcca taccaaacga cgagcgtgac accacgatgc ctgtagcaat ggcaacaacc 840

ttgcgtaaac tattaactgg cgaactactt actctagctt cccggcaaca gttgatagac 900

tggatggagg cggataaagt tgcaggacca cttctgcgct cggcccttcc ggctggctgg 960
tttattgctg ataaatctgg agccggtgag cgtgggtctc gcggtatcat tgcagcactg 1020
gggccagatg gtaagccctc ccgtatcgta gttatctaca cgacggggag tcaggcaact 1080
atggatgaac gaaatagaca gatcgctgag ataggtgcct cactgattaa gcattggtaa 1140
ccgattctag gtgcattggc gcagaaaaaa atgcctgatg cgacgctgcg cgtcttatac 1200
tcccacatat gccagattca gcaacggata cggcttcccc aacttgccca cttccatacg 1260
tgtcctcctt accagaaatt tatccttaag atcgtttaaa ctcgactctg gctctatcga 1320

atctccgtcg tttcgagctt acgcgaacag ccgtggcgct catttgctcg tcgggcatcg 1380
aatctcgtca gctatcgtca gcttaccttt ttggcagcga tcgcggctcc cgacatcttg 1440
gaccattagc tccacaggta tcttcttccc tctagtggtc ataacagcag cttcagctac 1500
ctctcaattc aaaaaacccc tcaagacccg tttagaggcc ccaaggggtt atgctatcaa 1560
tcgttgcgtt acacacacaa aaaaccaaca cacatccatc ttcgatggat agcgatttta 1620
ttatctaact gctgatcgag tgtagccaga tctagtaatc aattacgggg tcattagttc 1680
atagcccata tatggagttc cgcgttacat aacttacggt aaatggcccg cctggctgac 1740

cgcccaacga cccccgccca ttgacgtcaa taatgacgta tgttcccata gtaacgccaa 1800
tagggacttt ccattgacgt caatgggtgg agtatttacg gtaaactgcc cacttggcag 1860
tacatcaagt gtatcatatg ccaagtacgc cccctattga cgtcaatgac ggtaaatggc 1920
ccgcctggca ttatgcccag tacatgacct tatgggactt tcctacttgg cagtacatct 1980
acgtattagt catcgctatt accatgctga tgcggttttg gcagtacatc aatgggcgtg 2040
gatagcggtt tgactcacgg ggatttccaa gtctccaccc cattgacgtc aatgggagtt 2100
tgttttggca ccaaaatcaa cgggactttc caaaatgtcg taacaactcc gccccattga 2160

cgcaaatggg cggtaggcgt gtacggtggg aggtctatat aagcagagct ggtttagtga 2220
accgtcagat cagatctttg tcgatcctac catccactcg acacacccgc cagcggccgc 2280
taatacgact cactataggg agaagtactg ccaccatggg caagaagcaa aatcgcaaga 2340
cggggaattc caagacacaa tccgctagcc caccacctaa agagcgttct agctcccctg 2400
ctactgagca gtcctggatg gaaaacgact tcgatgaact ccgggaagag ggatttaggc 2460
gatccaacta ttcagaactc cgcgaagata tccagacaaa ggggaaggaa gtcgagaatt 2520
tcgagaagaa cctcgaggag tgcatcaccc gtatcacaaa cactgagaaa tgtctcaaag 2580

aactcatgga acttaagaca aaagccaggg agcttcgaga ggagtgtcgg agtctgagat 2640

ccaggtgtga ccagctcgag gagcgcgtga gcgcgatgga agacgagatg aacgagatga 2700

aaagagaggg caaattcagg gagaagcgca ttaagaggaa cgaacagagt ctgcaggaga 2760

tttgggatta cgtcaagagg cctaacctgc ggttgatcgg cgtccccgag agcgacgtag 2820

aaaacgggac taaactggag aatacacttc aagacatcat tcaagaaaat tttccaaacc 2880

tggctcggca agctaatgtg caaatccaag agatccaacg cacaccccag cggtatagct 2940

ctcggcgtgc caccctagg catattatcg tgcgctttac taaggtggag atgaaagaga 3000

agatgctgcg agccgctcgg gaaaagggaa gggtgacttt gaagggcaaa cctattcggc 3060

tgacggttga ccttagcgcc gagacactcc aggcacgccg ggaatggggc cccatcttta 3120

atatcctgaa ggagaagaac ttccagccac gaatctctta ccctgcaaag ttgagtttta 3180

tctccgaggg tgagattaag tatttcatcg ataaacagat gctgcgagac ttcgtgacaa 3240

ctcgcccagc tctcaaggaa ctgctcaaag aggctcttaa tatggagcgc aataatagat 3300

atcaaccctt gcagaaccac gcaaagatgt gagacagccg tcagaccatc aagactagga 3360

agaaactgca tcaactaatg agcaaaatca ccagctaaca tcatagtata catgaaaagg 3420

ccggcggcca cgaaaaaggc cggccaggca aaaaagaaaa agggcggcgg cagcaccggc 3480

tctaactcac atatcaccat ccttacactt aacattaacg gcctcaactc agctatcaag 3540

cgccatcggc tggccagctg gatcaaatca caggatccaa gcgtttgttg catccaagag 3600

acccacctga cctgtagaga tactcaccgc ctcaagatca agggatggcg aaagatttat 3660

caggcgaacg gtaagcagaa gaaagccgga gtcgcaattc tggtctcaga caagacggat 3720

ttcaagccca ccaaaattaa gcgtgataag gaaggtcact atattatggt gaaaggcagc 3780

atacagcagg aagaacttac catattgaac atctacgcgc caaacaccgg cgcacctcgc 3840

tttatcaaac aggtcctgtc cgatctgcag cgagatctgg attctcatac gttgattatg 3900

ggtgatttca atacaccatt gagcaccctg gatcgcagca ccaggcaaaa ggtaaataaa 3960

gacacgcaag agctcaatag cgcactgcat caggcagatc tcattgatat ttatcgcact 4020

cttcatccta agagtaccga gtacacattc ttcagcgccc cacatcatac atactcaaag 4080

atcgatcata tcgtcggctc aaaggctctg ctgtcaaagt gcaagcgcac agagataatt 4140

acaaattacc tgtcagatca tagcgcgatc aagctcgagc tgagaatcaa gaacctgacc 4200

cagagccgga gtaccacttg gaagcttaat aacctgctgc tcaacgatta ttgggtccac 4260

aatgagatga aggcagagat taaaatgttc ttcgaaacaa atgagaataa ggatactacc 4320

tatcaaaacc tttgggatgc ctttaaggcc gtctgcagag gcaagttcat cgccctcaac 4380

gcctataaaa gaaaacaaga gagatctaag atcgatactc tcacctctca gctgaaggag 4440

ttggagaaac aggaacagac ccactccaag gcgtcaagac ggcaggagat cacaaagatt 4500

cgcgccgagt tgaaagagat cgaaacccaa aagactcttc agaaaattaa cgagtctcgt 4560

agttggttct tcgagcggat taataagata gacagacctc tggcacgact gattaagaag 4620

aagcgcgaaa agaaccagat tgataccatc aagaacgaca agggcgacat cactactgac 4680

ccgaccgaga tccagaccac tattcgggag tattataagc atttgtatgc taacaagctt 4740

gagaacctgg aagagatgga cacttttctg gatacctata ctctgccacg gcttaatcaa 4800

gaggaagtcg agtccctcaa ccgcccaatt acaggaagcg agattgtggc cataattaac 4860

tccctgccga caaagaaatc tcctggtccg gacgggttta cagctgagtt ttatcaacgg 4920

tatatggaag agcttgtacc gtttctgctc aagctctttc agtctataga aaaggaaggc 4980

atcttgccca attccttcta cgaagcttct ataatactta ttcccaaacc aggacgcgat 5040

accacaaaga aggaaaactt ccggcccatt agtctcatga atatcgacgc taaaatattg 5100

aacaagattc tcgccaacag aatccaacaa catattaaga aattgataca tcacgaccag 5160

gtggggttta tacctggcat gcagggctgg tttaacatcc ggaagagtat taacgtcatt 5220

caacacatta atagagctaa ggataagaat catatgatca tctctataga cgcggaaaag 5280

gcattcgata agattcagca gccatttatg ctcaagactc tgaacaaact cggcatcgac 5340

ggaacatatt ttaagattat tcgcgcaatt tacgataagc cgactgctaa cattatcctt 5400

aacggccaaa agctcgaggc ctttccgctc aagactggaa cccgccaagg ctgtcccctc 5460

tccccgcttt tgtttaatat tgtactcgag gtgctggcta gggctattcg tcaagagaaa 5520

gagattaaag ggatacagct cgggaaggaa gaggtcaagc tttccttgtt cgccgatgat 5580

atgattgtgt acctggagaa tcctattgtg tctgctcaga accttcttaa acttatttct 5640

aactttagca aggtcagcgg ctataagatt aacgtccaga aatctcaggc ctttctgtac 5700

acaaataatc gacagaccga atcccagata atgggtgagc ttccgtttgt catagccagc 5760

aaaaggataa agtatctcgg aatccagctg acacgagacg ttaaagattt gtttaaggaa 5820

aattacaagc ctctcctgaa agagattaag gaagatacta ataagtggaa gaatatcccc 5880

tgttcatggg ttggcagaat caacatagtg aagatggcaa tacttcctaa agtgatatat 5940

cgctttaacg ccatcccaat taaactgcct atgaccttct ttacggagct cgagaaaaca 6000

acccttaaat ttatatggaa tcaaaagaga gcaagaatag cgaagtccat cttgagccag 6060

aagaataagg ccggtgggat tactttgcct gattttaagt tgtattataa agccacagta 6120

actaagacag cctggtattg gtatcagaat agagacatcg accagtggaa tcggaccgaa 6180

ccatcagaga taatgcccca catctataat taccttatat tcgataagcc agaaaagaat 6240

aaacagtggg gcaaagacag cctcttcaac aagtggtgtt gggagaattg gctggccata 6300

tgccggaaac tcaagctcga ccccttttctt acaccctaca ctaaaatcaa cagtaggtgg 6360

atcaaggact tgaatgtcaa gccaaagact ataaagacac tggaagagaa tcttgggatc 6420

acaatacaag atataggcgt cggcaaagat tttatgtcaa agacgcccaa ggccatggcc 6480

actaaggata agattgataa gtgggacctt attaagctca aaagcttctg tactgccaag 6540

gagaccacga tcagagttaa taggcagccc actacatggg aaaagatttt cgccacttat 6600

tcatcagata aggggttgat aagcagaata tataacgagc tgaagcagat ctacaagaag 6660

aaaacgaata atcccatcaa gaagtgggca aaagatatga acaggcattt tagcaaagag 6720

gatatctacg ccgcgaagaa gcatatgaag aagtgtagtt caagcttggc cattcgtgag 6780

atgcagatta agacgaccat gcgataccac cttaccccag tgaggatggc aattatcaag 6840

aaatctggca ataatagatg ttggcggggc tgtggcgaga ttggcaccct gctccattgc 6900

tggtgggatt gcaagctggt gcagccgctt tggaaatcag tctggcgctt tctgagggac 6960

ctcgagcttg agattccctt cgatcccgca attcccttgc tcggaatcta tcctaacgaa 7020

tacaagagct gttgttacaa ggatacgtgt acccggatgt tcatcgcggc cttgtttacg 7080

atagctaaga cgtggaatca gcctaagtgc cccacaatga tcgattggat caagaaaatg 7140

tggcatattt ataccatgga gtattacgca gcaattaaga atgacgaatt tatttccttc 7200

gttgggacct ggatgaagct ggagactatt attctgagca agctgtctca ggagcaaaag 7260

acaaagcata gaatcttctc tctcattggt ggtaacgact acaaagacga tgacgacaag 7320

taaagcgctt ctagaagttg tctcctcctg cactgactga ctgatacaat cgatttctgg 7380

atccgcaggc ctaatcaacc tctggattac aaaatttgtg aaagattgac tggtattctt 7440

aactatgttg ctccttttac gctatgtgga tacgctgctt taatgccttt gtatcatgct 7500

attgcttccc gtatggcttt cattttctcc tccttgtata aatcctggtt gctgtctctt 7560

tatgaggagt tgtggcccgt tgtcaggcaa cgtggcgtgg tgtgcactgt gtttgctgac 7620

gcaacccccа ctggttgggg cattgccacc acctgtcagc tcctttccgg gactttcgct 7680

ttccccctcc ctattgccac ggcggaactc atcgccgcct gccttgcccg ctgctggaca 7740

ggggctcggc tgttgggcac tgacaattcc gtggtgttgt cggggaagct gacgtccttt 7800

ccatggctgc tcgcctgtgt tgccacctgg attctgcgcg ggacgtcctt ctgctacgtc 7860

ccttcggccc tcaatccagc ggaccttcct tcccgctgag agacacaaaa aattccaaca 7920

cactattgca atgaaaataa atttccttta ttagccagaa gtcagatgct caaggggctt 7980

catgatgtcc ccataatttt tggcagaggg aaaaagatct cagtggtatt tgtgagccag 8040

ggcattggcc ttctgatagg cagcctgcac ctgaggagtg cggccgcttt acttgtacag 8100

ctcgtccatg ccgagagtga tcccggcggc ggtcacgaac tccagcagga ccatgtgatc 8160

gcgcttctcg ttggggtctt tgctcagggc ggactgggtg ctcaggtagt ggttgtcggg 8220

cagcagcacg gggccgtcgc cgatgggggt gttctgctgg tagtggtcgg cgagctgcac 8280

gctgccgtcc tcgatgttgt ggcggatctt gaagttcacc ttgatgccgt tcttctgctt 8340

gtcggccatg atatagacgt tgtggctgtt gtagttgtac tccagcttgt gccccaggat 8400

gttgccgtcc tccttgaagt cgatgcccct cagctcgatg cggttcacca gggtgtcgcc 8460

ctcgaacttc acctcggcgc gggtcttgta gttgccgtcg tccttgaaga agatggtgcg 8520

ctcctggacg tagccttcgg gcatggcgga cttgaagaag tcgtgctgct tcatgtggtc 8580

ggggtagcgg ctgaagcact gcacgccgta ggtcagggtg gtcacgaggg tgggccaggg 8640

cacgggcagc ttgccggtgg tgcagatgaa cttcagggtc agcttgccgt aggtggcatc 8700

gccctcgccc tcgccggaca cgctgaactt gtggccgttt acgtcgccgt ccagctcgac 8760

caggatgggc accaccccgg tgaacagctc ctcgcccttg ctcaccatgg tggcgggatc 8820

tgacggttca ctaaaccagc tctgcttata tagacctccc accgtacacg cctaccgccc 8880

atttgcgtca atggggcgga gttgttacga cattttggaa agtcccgttg attttggtgc 8940

caaaacaaac tcccattgac gtcaatgggg tggagacttg gaaatccccg tgagtcaaac 9000

cgctatccac gcccattgat gtactgccaa aaccgcatca ccatggtaat agcgatgact 9060

aatacgtaga tgtactgcca agtaggaaag tcccataagg tcatgtactg ggcataatgc 9120

caggcgggcc atttaccgtc attgacgtca ataggggcgc tacttggcat atgatacact 9180

tgatgtactg ccaagtgggc agtttaccgt aaatactcca cccattgacg tcaatggaaa 9240

gtccctattg gcgttactat gggaacatac gtcattattg acgtcaatgg gcgggggtcg 9300

ttgggcggtc agccaggcgg gccatttacc gtaagttatg taacgggcct gctgccggct 9360

ctgcggcctc ttccgcgtct tcgccttcgc cctcagacga gtcggatctc cctttgggcc 9420

gcctccccgc ctgtctagct tgactgactg agatacagcg taccttcagc tcacagacat 9480

gataagatac attgatgagt ttggacaaac cacaactaga atgcagtgaa aaaaatgctt 9540

tatttgtgaa atttgtgatg ctattgcttt atttgtaacc attataagct gcaataaaca 9600

agttaacaac aacaattgca ttcattttat gtttcaggtt caggggggagg tgtgggaggt 9660

tttttaaagc aagtaaaacc tctacaaatg tggtattggc ccatctctat cggtatcgta 9720

gcataacccc ttggggcctc taaacgggtc ttgaggggtt ttttgtgccc ctcgggccgg 9780

attgctatct accggcattg gcgcagaaaa aaatgcctga tgcgacgctg cgcgtcttat 9840

actcccacat atgccagatt cagcaacgga tacggcttcc ccaacttgcc cacttccata 9900

cgtgtcctcc ttaccagaaa tttatcctta aggtcgtcag ctatcctgca ggcgatctct 9960

cgatttcgat caagacattc ctttaatggt cttttctgga caccactagg ggtcagaagt 10020

agttcatcaa actttcttcc ctccctaatc tcattggtta ccttgggcta tcgaaactta 10080

attaagcgat ctgcatctca attagtcagc aaccatagtc ccgcccctaa ctccgcccat 10140

cccgccccta actccgccca gttccgccca ttctccgccc catcgctgac taatttttt 10200

tatttatgca gaggccgagg ccgcctcggc ctctgagcta ttccagaagt agtgaggagg 10260

ctttttgga ggcctaggct tttgcaaagg aggtagccaa catgattgaa caagatggat 10320

tgcacgcagg ttctcccgcc gcttgggtgg agaggctatt cggctatgac tgggcacaac 10380

agacaatcgg ctgctctgat gccgccgtgt tccggctgtc agcgcagggg cgcccggttc 10440

tttttgtcaa gaccgacctg tccggtgccc tgaatgaact ccaggacgag gcagcgcggc 10500

tatcgtggct ggccacgacg ggcgttcctt gcgcagctgt gctcgacgtt gtcactgaag 10560

cgggaaggga ctggctgcta ttgggcgaag tgccggggca ggatctcctg tcatctcacc 10620

ttgctcctgc cgagaaagta tccatcatgg ctgatgcaat gcggcggctg catacgcttg 10680

atccggctac ctgcccattc gaccaccaag cgaaacatcg catcgagcga gcacgtactc 10740

ggatggaagc cggtcttgtc gatcaggatg atctggacga agagcatcag gggctcgcgc 10800

cagccgaact gttcgccagg ctcaaggcgc ggatgcccga cggcgaggat ctcgtcgtga 10860

cccacggcga tgcctgcttg ccgaatatca tggtggaaaa tggccgcttt tctggattca 10920

tcgactgtgg ccggctgggt gtggcggacc gctatcagga catagcgttg gctacccgtg 10980

atattgctga agagcttggc ggcgaatggg ctgaccgctt cctcgtgctt tacggtatcg 11040

ccgctcccga ttcgcagcgc atcgccttct atcgccttct tgacgagttc ttctagtatg 11100

taagccctgt gccttctagt tgccagccat ctgttgtttg cccctccccc gtgccttcct 11160

tgaccctgga aggtgccact cccactgtcc tttcctaata aaatgaggaa attgcatcgc 11220

attgtctgag taggtgtcat tctattctgg ggggtggggt ggggcaggac agcaaggggg 11280

aggattggga agacaatagc aggcatgctg gggatgcggt gggctctatg gttaattaac 11340

cagtcaagtc agctacttgg cgagatcgac ttgtctgggt ttcgactacg ctcagaattg 11400

cgtcagtcaa gttcgatctg gtccttgcta ttgcacccgt tctccgatta cgagtttcat 11460

ttaaatcatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct 11520

ggcgtttttc cataggctcc gccccccctga cgagcatcac aaaaatcgac gctcaagtca 11580

gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct 11640

cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc 11700

gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt 11760

tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc 11820

cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc 11880
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 11940
gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc 12000
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 12060
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 12120
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 12180
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 12240

ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat 12300
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg catttaaatt 12360
tccgaactct ccaaggccct cgtcggaaaa tcttcaaacc tttcgtccga tccatcttgc 12420
aggctacctc tcgaacgaac tatcgcaagt ctcttggccg gccttgcgcc ttggctattg 12480
cttggcagcg cctatcgcca ggtattactc caatcccgaa tatccgagat cgggatcacc 12540
cgagagaagt tcaacctaca tcctcaatcc cgatctatcc g 12581

<210> 104
<211> 12542
<212> DNA
<213> Artificial Sequence

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

<400> 104
agaacgatcc tctcagtgcg agtctcgacg atccatatcg ttgcttggca gtcagccagt 60
cggaatccag cttgggaccc aggaagtcca atcgtcagat attgtactca agcctggtca 120
cggcagcgta ccgatctgtt taaacctaga tattgatagt ctgatcggtc aacgtataat 180
cgagtcctag cttttgcaaa catctatcaa gagacaggat cagcaggagg ctttcgcatg 240
agtattcaac atttccgtgt cgcccttatt ccctttttg cggcattttg ccttcctgtt 300
tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt gggtgcgcga 360

gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa 420
gaacgctttc caatgatgag cacttttaaa gttctgctat gtggcgcggt attatcccgt 480
attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt 540
gagtattcac agtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc 600

agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac aacgattgga 660

ggaccgaagg agctaaccgc ttttttgcac aacatggggg atcatgtaac tcgccttgat 720

cgttgggaac cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct 780

gtagcaatgg caacaacctt gcgtaaacta ttaactggcg aactacttac tctagcttcc 840

cggcaacagt tgatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg 900

gcccttccgg ctggctggtt tattgctgat aaatctggag ccggtgagcg tgggtctcgc 960

ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg 1020

acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca 1080

ctgattaagc attggtaacc gattctaggt gcattggcgc agaaaaaaat gcctgatgcg 1140

acgctgcgcg tcttatactc ccacatatgc cagattcagc aacggatacg gcttccccaa 1200

cttgcccact tccatacgtg tcctccttac cagaaattta tccttaagat cgtttaaact 1260

cgactctggc tctatcgaat ctccgtcgtt tcgagcttac gcgaacagcc gtggcgctca 1320

tttgctcgtc gggcatcgaa tctcgtcagc tatcgtcagc ttacctttt ggcagcgatc 1380

gcggctcccg acatcttgga ccattagctc cacaggtatc ttcttccctc tagtggtcat 1440

aacagcagct tcagctacct ctcaattcaa aaaacccctc aagacccgtt tagaggcccc 1500

aaggggttat gctatcaatc gttgcgttac acacacaaaa aaccaacaca catccatctt 1560

cgatggatag cgattttatt atctaactgc tgatcgagtg tagccagatc tagtaatcaa 1620

ttacggggtc attagttcat agcccatata tggagttccg cgttacataa cttacggtaa 1680

atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata atgacgtatg 1740

ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggag tatttacggt 1800

aaactgccca cttggcagta catcaagtgt atcatatgcc aagtacgccc cctattgacg 1860

tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta tgggactttc 1920

ctacttggca gtacatctac gtattagtca tcgctattac catgctgatg cggttttggc 1980

agtacatcaa tgggcgtgga tagcggtttg actcacgggg atttccaagt ctccacccca 2040

ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg ggactttcca aaatgtcgta 2100

acaactccgc cccattgacg caaatgggcg gtaggcgtgt acggtgggag gtctatataa 2160

gcagagctgg tttagtgaac cgtcagatca gatctttgtc gatcctacca tccactcgac 2220

acacccgcca gcggccgcta atacgactca ctatagggag aagtactgcc accatgggca 2280

agaagcaaaa tcgcaagacg gggaattcca agacacaatc cgctagccca ccacctaaag 2340

agcgttctag ctcccctgct actgagcagt cctggatgga aaacgacttc gatgaactcc 2400

gggaagaggg atttaggcga tccaactatt cagaactccg cgaagatatc cagacaaagg 2460

ggaaggaagt cgagaatttc gagaagaacc tcgaggagtg catcacccgt atcacaaaca 2520
ctgagaaatg tctcaaagaa ctcatggaac ttaagacaaa agccagggag cttcgagagg 2580
agtgtcggag tctgagatcc aggtgtgacc agctcgagga gcgcgtgagc gcgatggaag 2640
acgagatgaa cgagatgaaa agagagggca aattcaggga gaagcgcatt aagaggaacg 2700
aacagagtct gcaggagatt tgggattacg tcaagaggcc taacctgcgg ttgatcggcg 2760
tccccgagag cgacgtagaa aacgggacta aactggagaa tacacttcaa gacatcattc 2820
aagaaaattt tccaaacctg gctcggcaag ctaatgtgca aatccaagag atccaacgca 2880

caccccagcg gtatagctct cggcgtgcca cccctaggca tattatcgtg cgctttacta 2940
aggtggagat gaaagagaag atgctgcgag ccgctcggga aaagggaagg gtgactttga 3000
agggcaaacc tattcggctg acggttgacc ttagcgccga gacactccag gcacgccggg 3060
aatggggccc catctttaat atcctgaagg agaagaactt ccagccacga atctcttacc 3120
ctgcaaagtt gagttttatc tccgagggtg agattaagta tttcatcgat aaacagatgc 3180
tgcgagactt cgtgacaact cgcccagctc tcaaggaact gctcaaagag gctcttaata 3240
tggagcgcaa taatagatat caacccttgc agaaccacgc aaagatgtga gacagccgtc 3300

agaccatcaa gactaggaag aaactgcatc aactaatgag caaaatcacc agctaacatc 3360
atagtataca tgaccggctc taactcacat atcaccatcc ttacacttaa cattaacggc 3420
ctcaactcag ctatcaagcg ccatcggctg gccagctgga tcaaatcaca ggatccaagc 3480
gtttgttgca tccaagagac ccacctgacc tgtagagata ctcaccgcct caagatcaag 3540
ggatggcgaa agatttatca ggcgaacggt aagcagaaga aagccggagt cgcaattctg 3600
gtctcagaca agacggattt caagcccacc aaaattaagc gtgataagga aggtcactat 3660
attatggtga aaggcagcat acagcaggaa gaacttacca tattgaacat ctacgcgcca 3720

aacaccggcg cacctcgctt tatcaaacag gtcctgtccg atctgcagcg agatctggat 3780
tctcatacgt tgattatggg tgatttcaat acaccattga gcaccctgga tcgcagcacc 3840
aggcaaaagg taaataaaga cacgcaagag ctcaatagcg cactgcatca ggcagatctc 3900
attgatattt atcgcactct tcatcctaag agtaccgagt acacattctt cagcgcccca 3960
catcatacat actcaaagat cgatcatatc gtcggctcaa aggctctgct gtcaaagtgc 4020
aagcgcacag agataattac aaattacctg tcagatcata gcgcgatcaa gctcgagctg 4080
agaatcaaga acctgaccca gagccggagt accacttgga agcttaataa cctgctgctc 4140

aacgattatt gggtccacaa tgagatgaag gcagagatta aaatgttctt cgaaacaaat 4200

gagaataagg atactaccta tcaaaacctt tgggatgcct ttaaggccgt ctgcagaggc 4260

aagttcatcg ccctcaacgc ctataaaaga aaacaagaga gatctaagat cgatactctc 4320

acctctcagc tgaaggagtt ggagaaacag gaacagaccc actccaaggc gtcaagacgg 4380

caggagatca caaagattcg cgccgagttg aaagagatcg aaacccaaaa gactcttcag 4440

aaaattaacg agtctcgtag ttggttcttc gagcggatta ataagataga cagacctctg 4500

gcacgactga ttaagaagaa gcgcgaaaag aaccagattg ataccatcaa gaacgacaag 4560

ggcgacatca ctactgaccc gaccgagatc cagaccacta ttcgggagta ttataagcat 4620

ttgtatgcta acaagcttga gaacctggaa gagatggaca cttttctgga tacctatact 4680

ctgccacggc ttaatcaaga ggaagtcgag tccctcaacc gcccaattac aggaagcgag 4740

attgtggcca taattaactc cctgccgaca aagaaatctc ctggtccgga cgggtttaca 4800

gctgagtttt atcaacggta tatggaagag cttgtaccgt ttctgctcaa gctctttcag 4860

tctatagaaa aggaaggcat cttgcccaat tccttctacg aagcttctat aatacttatt 4920

cccaaaccag gacgcgatac cacaaagaag gaaaacttcc ggcccattag tctcatgaat 4980

atcgacgcta aaatattgaa caagattctc gccaacagaa tccaacaaca tattaagaaa 5040

ttgatacatc acgaccaggt ggggtttata cctggcatgc agggctggtt taacatccgg 5100

aagagtatta acgtcattca acacattaat agagctaagg ataagaatca tatgatcatc 5160

tctatagacg cggaaaaggc attcgataag attcagcagc catttatgct caagactctg 5220

aacaaactcg gcatcgacgg aacatatttt aagattattc gcgcaattta cgataagccg 5280

actgctaaca ttatccttaa cggccaaaag ctcgaggcct ttccgctcaa gactggaacc 5340

cgccaaggct gtcccctctc cccgcttttg tttaatattg tactcgaggt gctggctagg 5400

gctattcgtc aagagaaaga gattaaaggg atacagctcg ggaaggaaga ggtcaagctt 5460

tccttgttcg ccgatgatat gattgtgtac ctggagaatc ctattgtgtc tgctcagaac 5520

cttcttaaac ttatttctaa ctttagcaag gtcagcggct ataagattaa cgtccagaaa 5580

tctcaggcct ttctgtacac aaataatcga cagaccgaat cccagataat gggtgagctt 5640

ccgtttgtca tagccagcaa aaggataaag tatctcggaa tccagctgac acgagacgtt 5700

aaagatttgt ttaaggaaaa ttacaagcct ctcctgaaag agattaagga agatactaat 5760

aagtggaaga atatcccctg ttcatgggtt ggcagaatca acatagtgaa gatggcaata 5820

cttcctaaag tgatatatcg ctttaacgcc atcccaatta aactgcctat gaccttcttt 5880

acggagctcg agaaaacaac ccttaaattt atatggaatc aaaagagagc aagaatagcg 5940

aagtccatct tgagccagaa gaataaggcc ggtgggatta ctttgcctga ttttaagttg 6000

tattataaag ccacagtaac taagacagcc tggtattggt atcagaatag agacatcgac 6060

cagtggaatc ggaccgaacc atcagagata atgccccaca tctataatta ccttatattc 6120

gataagccag aaaagaataa acagtggggc aaagacagcc tcttcaacaa gtggtgttgg 6180

gagaattggc tggccatatg ccggaaactc aagctcgacc cctttcttac accctacact 6240

aaaatcaaca gtaggtggat caaggacttg aatgtcaagc caaagactat aaagacactg 6300

gaagagaatc ttgggatcac aatacaagat ataggcgtcg gcaaagattt tatgtcaaag 6360

acgcccaagg ccatggccac taaggataag attgataagt gggaccttat taagctcaaa 6420

agcttctgta ctgccaagga gaccacgatc agagttaata ggcagcccac tacatgggaa 6480

aagattttcg ccacttattc atcagataag gggttgataa gcagaatata taacgagctg 6540

aagcagatct acaagaagaa aacgaataat cccatcaaga agtgggcaaa agatatgaac 6600

aggcatttta gcaaagagga tatctacgcc gcgaagaagc atatgaagaa gtgtagttca 6660

agcttggcca ttcgtgagat gcagattaag acgaccatgc gataccacct taccccagtg 6720

aggatggcaa ttatcaagaa atctggcaat aatagatgtt ggcggggctg tggcgagatt 6780

ggcaccctgc tccattgctg gtgggattgc aagctggtgc agccgctttg gaaatcagtc 6840

tggcgctttc tgagggacct cgagcttgag attcccttcg atcccgcaat tcccttgctc 6900

ggaatctatc ctaacgaata caagagctgt tgttacaagg atacgtgtac ccggatgttc 6960

atcgcggcct tgtttacgat agctaagacg tggaatcagc ctaagtgccc cacaatgatc 7020

gattggatca agaaaatgtg gcatatttat accatggagt attacgcagc aattaagaat 7080

gacgaattta tttccttcgt tgggacctgg atgaagctgg agactattat tctgagcaag 7140

ctgtctcagg agcaaaagac aaagcataga atcttctctc tcattggtgg taacgactac 7200

aaagacgatg acgacaagcc aaagaagaag cggaaggtct aaagcgcttc tagaagttgt 7260

ctcctcctgc actgactgac tgatacaatc gatttctgga tccgcaggcc taatcaacct 7320

ctggattaca aaatttgtga aagattgact ggtattctta actatgttgc tccttttacg 7380

ctatgtggat acgctgcttt aatgcctttg tatcatgcta ttgcttcccg tatggctttc 7440

attttctcct ccttgtataa atcctggttg ctgtctcttt atgaggagtt gtggcccgtt 7500

gtcaggcaac gtggcgtggt gtgcactgtg tttgctgacg caacccccac tggttggggc 7560

attgccacca cctgtcagct cctttccggg actttcgctt tccccctccc tattgccacg 7620

gcggaactca tcgccgcctg ccttgcccgc tgctggacag gggctcggct gttgggcact 7680

gacaattccg tggtgttgtc ggggaagctg acgtcctttc catggctgct cgcctgtgtt 7740

gccacctgga ttctgcgcgg gacgtccttc tgctacgtcc cttcggccct caatccagcg 7800

gaccttcctt cccgctgaga gacacaaaaa attccaacac actattgcaa tgaaaataaa 7860

tttcctttat tagccagaag tcagatgctc aaggggcttc atgatgtccc cataattttt 7920

ggcagaggga aaaagatctc agtggtattt gtgagccagg gcattggcct tctgataggc 7980

agcctgcacc tgaggagtgc ggccgcttta cttgtacagc tcgtccatgc cgagagtgat 8040

cccggcggcg gtcacgaact ccagcaggac catgtgatcg cgcttctcgt tggggtcttt 8100

gctcagggcg gactgggtgc tcaggtagtg gttgtcgggc agcagcacgg ggccgtcgcc 8160

gatgggggtg ttctgctggt agtggtcggc gagctgcacg ctgccgtcct cgatgttgtg 8220

gcggatcttg aagttcacct tgatgccgtt cttctgcttg tcggccatga tatagacgtt 8280

gtggctgttg tagttgtact ccagcttgtg ccccaggatg ttgccgtcct ccttgaagtc 8340

gatgcccttc agctcgatgc ggttcaccag ggtgtcgccc tcgaacttca cctcggcgcg 8400

ggtcttgtag ttgccgtcgt ccttgaagaa gatggtgcgc tcctggacgt agccttcggg 8460

catggcggac ttgaagaagt cgtgctgctt catgtggtcg gggtagcggc tgaagcactg 8520

cacgccgtag gtcagggtgg tcacgagggt gggccagggc acgggcagct tgccggtggt 8580

gcagatgaac ttcagggtca gcttgccgta ggtggcatcg ccctcgccct cgccggacac 8640

gctgaacttg tggccgttta cgtcgccgtc cagctcgacc aggatgggca ccaccccggt 8700

gaacagctcc tcgcccttgc tcaccatggt ggcgggatct gacggttcac taaaccagct 8760

ctgcttatat agacctccca ccgtacacgc ctaccgccca tttgcgtcaa tggggcggag 8820

ttgttacgac attttggaaa gtcccgttga ttttggtgcc aaaacaaact cccattgacg 8880

tcaatggggt ggagacttgg aaatccccgt gagtcaaacc gctatccacg cccattgatg 8940

tactgccaaa accgcatcac catggtaata gcgatgacta atacgtagat gtactgccaa 9000

gtaggaaagt cccataaggt catgtactgg gcataatgcc aggcgggcca tttaccgtca 9060

ttgacgtcaa taggggggcgt acttggcata tgatacactt gatgtactgc caagtgggca 9120

gtttaccgta aatactccac ccattgacgt caatggaaag tccctattgg cgttactatg 9180

ggaacatacg tcattattga cgtcaatggg cgggggtcgt tgggcggtca gccaggcggg 9240

ccatttaccg taagttatgt aacgggcctg ctgccggctc tgcggcctct tccgcgtctt 9300

cgccttcgcc ctcagacgag tcggatctcc ctttgggccg cctccccgcc tgtctagctt 9360

gactgactga gatacagcgt accttcagct cacagacatg ataagataca ttgatgagtt 9420

tggacaaacc acaactagaa tgcagtgaaa aaaatgcttt atttgtgaaa tttgtgatgc 9480

tattgcttta tttgtaacca ttataagctg caataaacaa gttaacaaca acaattgcat 9540

tcattttatg tttcaggttc agggggaggt gtgggaggtt ttttaaagca agtaaaacct 9600

ctacaaatgt ggtattggcc catctctatc ggtatcgtag cataacccct tggggcctct 9660

aaacgggtct tgaggggttt tttgtgcccc tcgggccgga ttgctatcta ccggcattgg 9720

cgcagaaaaa aatgcctgat gcgacgctgc gcgtcttata ctcccacata tgccagattc 9780

agcaacggat acggcttccc caacttgccc acttccatac gtgtcctcct taccagaaat 9840

ttatccttaa ggtcgtcagc tatcctgcag gcgatctctc gatttcgatc aagacattcc 9900

tttaatggtc ttttctggac accactaggg gtcagaagta gttcatcaaa ctttcttccc 9960

tccctaatct cattggttac cttgggctat cgaaacttaa ttaagcgatc tgcatctcaa 10020

ttagtcagca accatagtcc cgcccctaac tccgcccatc ccgcccctaa ctccgcccag 10080

ttccgcccat tctccgcccc atcgctgact aatttttttt atttatgcag aggccgaggc 10140

cgcctcggcc tctgagctat tccagaagta gtgaggaggc ttttttggag gcctaggctt 10200

ttgcaaagga ggtagccaac atgattgaac aagatggatt gcacgcaggt tctcccgccg 10260

cttgggtgga gaggctattc ggctatgact gggcacaaca gacaatcggc tgctctgatg 10320

ccgccgtgtt ccggctgtca gcgcaggggc gcccggttct ttttgtcaag accgacctgt 10380

ccggtgccct gaatgaactc caggacgagg cagcgcggct atcgtggctg gccacgacgg 10440

gcgttccttg cgcagctgtg ctcgacgttg tcactgaagc gggaagggac tggctgctat 10500

tgggcgaagt gccggggcag gatctcctgt catctcacct tgctcctgcc gagaaagtat 10560

ccatcatggc tgatgcaatg cggcggctgc atacgcttga tccggctacc tgcccattcg 10620

accaccaagc gaaacatcgc atcgagcgag cacgtactcg gatggaagcc ggtcttgtcg 10680

atcaggatga tctggacgaa gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc 10740

tcaaggcgcg gatgcccgac ggcgaggatc tcgtcgtgac ccacggcgat gcctgcttgc 10800

cgaatatcat ggtggaaaat ggccgctttt ctggattcat cgactgtggc cggctgggtg 10860

tggcggaccg ctatcaggac atagcgttgg ctacccgtga tattgctgaa gagcttggcg 10920

gcgaatgggc tgaccgcttc ctcgtgcttt acggtatcgc cgctcccgat tcgcagcgca 10980

tcgccttcta tcgccttctt gacgagttct tctagtatgt aagccctgtg ccttctagtt 11040

gccagccatc tgttgtttgc ccctcccccg tgccttcctt gaccctggaa ggtgccactc 11100

ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt 11160

ctattctggg gggtggggtg gggcaggaca gcaaggggga ggattgggaa gacaatagca 11220

ggcatgctgg ggatgcggtg ggctctatgg ttaattaacc agtcaagtca gctacttggc 11280

gagatcgact tgtctgggtt tcgactacgc tcagaattgc gtcagtcaag ttcgatctgg 11340

tccttgctat tgcacccgtt ctccgattac gagtttcatt taaatcatgt gagcaaaagg 11400

ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 11460

ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 11520

actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc ctgttccgac 11580

cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 11640

tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt 11700

gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 11760

caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 11820

agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 11880

tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 11940

tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa 12000

gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 12060

gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa 12120

aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat 12180

atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc 12240

gatctgtcta tttcgttcat ccatagttgc atttaaattt ccgaactctc caaggccctc 12300

gtcggaaaat cttcaaacct ttcgtccgat ccatcttgca ggctacctct cgaacgaact 12360

atcgcaagtc tcttggccgg ccttgcgcct tggctattgc ttggcagcgc ctatcgccag 12420

gtattactcc aatcccgaat atccgagatc gggatcaccc gagagaagtt caacctacat 12480

cctcaatccc gatctatccg agatccgagg aatatcgaaa tcggggcgcg cctggtgtac 12540

cg 12542

<210> 105

<211> 12554

<212> DNA

<213> Artificial Sequence

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

<400> 105

tcgacgatcc atatcgttgc ttggcagtca gccagtcgga atccagcttg ggacccagga 60

agtccaatcg tcagatattg tactcaagcc tggtcacggc agcgtaccga tctgtttaaa 120

cctagatatt gatagtctga tcggtcaacg tataatcgag tcctagcttt tgcaaacatc 180

tatcaagaga caggatcagc aggaggcttt cgcatgagta ttcaacattt ccgtgtcgcc 240

cttattccct tttttgcggc attttgcctt cctgtttttg ctcacccaga aacgctggtg 300

aaagtaaaag atgctgaaga tcagttgggt gcgcgagtgg gttacatcga actggatctc 360

aacagcggta agatccttga gagttttcgc cccgaagaac gctttccaat gatgagcact 420

tttaaagttc tgctatgtgg cgcggtatta tcccgtattg acgccgggca agagcaactc 480

ggtcgccgca tacactattc tcagaatgac ttggttgagt attcaccagt cacagaaaag 540

catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat 600

aacactgcgg ccaacttact tctgacaacg attggaggac cgaaggagct aaccgctttt 660

ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt gggaaccgga gctgaatgaa 720

gccataccaa acgacgagcg tgacaccacg atgcctgtag caatggcaac aaccttgcgt 780

aaactattaa ctggcgaact acttactcta gcttcccggc aacagttgat agactggatg 840

gaggcggata aagttgcagg accacttctg cgctcggccc ttccggctgg ctggtttatt 900

gctgataaat ctggagccgg tgagcgtggg tctcgcggta tcattgcagc actggggcca 960

gatggtaagc cctcccgtat cgtagttatc tacacgacgg ggagtcaggc aactatggat 1020

gaacgaaata gacagatcgc tgagataggt gcctcactga ttaagcattg gtaaccgatt 1080

ctaggtgcat tggcgcagaa aaaaatgcct gatgcgacgc tgcgcgtctt atactcccac 1140

atatgccaga ttcagcaacg gatacggctt ccccaacttg cccacttcca tacgtgtcct 1200

ccttaccaga aatttatcct taagatcgtt taaactcgac tctggctcta tcgaatctcc 1260

gtcgtttcga gcttacgcga acagccgtgg cgctcatttg ctcgtcgggc atcgaatctc 1320

gtcagctatc gtcagcttac ctttttggca gcgatcgcgg ctcccgacat cttggaccat 1380

tagctccaca ggtatcttct tccctctagt ggtcataaca gcagcttcag ctacctctca 1440

attcaaaaaa cccctcaaga cccgtttaga ggccccaagg ggttatgcta tcaatcgttg 1500

cgttacacac acaaaaaacc aacacacatc catcttcgat ggatagcgat tttattatct 1560

aactgctgat cgagtgtagc cagatctagt aatcaattac ggggtcatta gttcatagcc 1620

catatatgga gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgccca 1680

acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga 1740

ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 1800

aagtgtatca tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgcct 1860

ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat 1920

tagtcatcgc tattaccatg ctgatgcggt tttggcagta catcaatggg cgtggatagc 1980

ggtttgactc acggggattt ccaagtctcc accccattga cgtcaatggg agtttgtttt 2040

ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa ctccgcccca ttgacgcaaa 2100

tgggcggtag gcgtgtacgg tgggaggtct atataagcag agctggttta gtgaaccgtc 2160

agatcagatc tttgtcgatc ctaccatcca ctcgacacac ccgccagcgg ccgctaatac 2220

gactcactat agggagaagt actgccacca tgggcaagaa gcaaaatcgc aagacgggga 2280

attccaagac acaatccgct agcccaccac ctaaagagcg ttctagctcc cctgctactg 2340

agcagtcctg gatggaaaac gacttcgatg aactccggga agagggattt aggcgatcca 2400

actattcaga actccgcgaa gatatccaga caaaggggaa ggaagtcgag aatttcgaga 2460

agaacctcga ggagtgcatc acccgtatca caaacactga gaaatgtctc aaagaactca 2520

tggaacttaa gacaaaagcc agggagcttc gagaggagtg tcggagtctg agatccaggt 2580

gtgaccagct cgaggagcgc gtgagcgcga tggaagacga gatgaacgag atgaaaagag 2640

agggcaaatt cagggagaag cgcattaaga ggaacgaaca gagtctgcag gagatttggg 2700

attacgtcaa gaggcctaac ctgcggttga tcggcgtccc cgagagcgac gtagaaaacg 2760

ggactaaact ggagaataca cttcaagaca tcattcaaga aaattttcca aacctggctc 2820

ggcaagctaa tgtgcaaatc caagagatcc aacgcacacc ccagcggtat agctctcggc 2880

gtgccacccc taggcatatt atcgtgcgct ttactaaggt ggagatgaaa gagaagatgc 2940

tgcgagccgc tcgggaaaag ggaagggtga ctttgaaggg caaacctatt cggctgacgg 3000

ttgaccttag cgccgagaca ctccaggcac gccgggaatg gggccccatc tttaatatcc 3060

tgaaggagaa gaacttccag ccacgaatct cttaccctgc aaagttgagt tttatctccg 3120

agggtgagat taagtatttc atcgataaac agatgctgcg agacttcgtg acaactcgcc 3180

cagctctcaa ggaactgctc aaagaggctc ttaatatgga gcgcaataat agatatcaac 3240

ccttgcagaa ccacgcaaag atgtgagaca gccgtcagac catcaagact aggaagaaac 3300

tgcatcaact aatgagcaaa atcaccagct aacatcatag tatacatgac cggctctaac 3360

tcacatatca ccatccttac acttaacatt aacggcctca actcagctat caagcgccat 3420

cggctggcca gctggatcaa atcacaggat ccaagcgttt gttgcatcca agagacccac 3480

ctgacctgta gagatactca ccgcctcaag atcaagggat ggcgaaagat ttatcaggcg 3540

aacggtaagc agaagaaagc cggagtcgca attctggtct cagacaagac ggatttcaag 3600

cccaccaaaa ttaagcgtga taaggaaggt cactatatta tggtgaaagg cagcatacag 3660

caggaagaac ttaccatatt gaacatctac gcgccaaaca ccggcgcacc tcgctttatc 3720

aaacaggtcc tgtccgatct gcagcgagat ctggattctc atacgttgat tatgggtgat 3780

ttcaatacac cattgagcac cctggatcgc agcaccaggc aaaaggtaaa taaagacacg 3840

caagagctca atagcgcact gcatcaggca gatctcattg atatttatcg cactcttcat 3900

cctaagagta ccgagtacac attcttcagc gccccacatc atacatactc aaagatcgat 3960

catatcgtcg gctcaaaggc tctgctgtca aagtgcaagc gcacagagat aattacaaat 4020

tacctgtcag atcatagcgc gatcaagctc gagctgagaa tcaagaacct gacccagagc 4080
cggagtacca cttggaagct taataacctg ctgctcaacg attattgggt ccacaatgag 4140
atgaaggcag agattaaaat gttcttcgaa acaaatgaga ataaggatac tacctatcaa 4200
aacctttggg atgcctttaa ggccgtctgc agaggcaagt tcatcgccct caacgcctat 4260
aaaagaaaac aagagagatc taagatcgat actctcacct ctcagctgaa ggagttggag 4320
aaacaggaac agacccactc caaggcgtca agacggcagg agatcacaaa gattcgcgcc 4380
gagttgaaag agatcgaaac ccaaaagact cttcagaaaa ttaacgagtc tcgtagttgg 4440

ttcttcgagc ggattaataa gatagacaga cctctggcac gactgattaa gaagaagcgc 4500
gaaaagaacc agattgatac catcaagaac gacaagggcg acatcactac tgacccgacc 4560
gagatccaga ccactattcg ggagtattat aagcatttgt atgctaacaa gcttgagaac 4620
ctggaagaga tggacacttt tctggatacc tatactctgc cacggcttaa tcaagaggaa 4680
gtcgagtccc tcaaccgccc aattacagga agcgagattg tggccataat taactccctg 4740
ccgacaaaga aatctcctgg tccggacggg tttacagctg agttttatca acggtatatg 4800
gaagagcttg taccgtttct gctcaagctc tttcagtcta tagaaaagga aggcatcttg 4860

cccaattcct tctacgaagc ttctataata cttattccca aaccaggacg cgataccaca 4920
aagaaggaaa acttccggcc cattagtctc atgaatatcg acgctaaaat attgaacaag 4980
attctcgcca acagaatcca acaacatatt aagaaattga tacatcacga ccaggtgggg 5040
tttatacctg gcatgcaggg ctggtttaac atccggaaga gtattaacgt cattcaacac 5100
attaatagag ctaaggataa gaatcatatg atcatctcta tagacgcgga aaaggcattc 5160
gataagattc agcagccatt tatgctcaag actctgaaca aactcggcat cgacggaaca 5220
tattttaaga ttattcgcgc aatttacgat aagccgactg ctaacattat ccttaacggc 5280

caaaagctcg aggcctttcc gctcaagact ggaacccgcc aaggctgtcc cctctccccg 5340
cttttgttta atattgtact cgaggtgctg gctagggcta ttcgtcaaga gaaagagatt 5400
aaagggatac agctcgggaa ggaagaggtc aagctttcct tgttcgccga tgatatgatt 5460
gtgtacctgg agaatcctat tgtgtctgct cagaaccttc ttaaacttat ttctaacttt 5520
agcaaggtca gcggctataa gattaacgtc cagaaatctc aggcctttct gtacacaaat 5580
aatcgacaga ccgaatccca gataatgggt gagcttccgt ttgtcatagc cagcaaaagg 5640
ataaagtatc tcggaatcca gctgacacga gacgttaaag atttgtttaa ggaaaattac 5700

aagcctctcc tgaaagagat taaggaagat actaataagt ggaagaatat cccctgttca 5760

tgggttggca gaatcaacat agtgaagatg gcaatacttc ctaaagtgat atatcgcttt 5820

aacgccatcc caattaaact gcctatgacc ttctttacgg agctcgagaa aacaaccctt 5880

aaatttatat ggaatcaaaa gagagcaaga atagcgaagt ccatcttgag ccagaagaat 5940

aaggccggtg ggattacttt gcctgatttt aagttgtatt ataaagccac agtaactaag 6000

acagcctggt attggtatca gaatagagac atcgaccagt ggaatcggac cgaaccatca 6060

gagataatgc cccacatcta taattacctt atattcgata agccagaaaa gaataaacag 6120

tggggcaaag acagcctctt caacaagtgg tgttgggaga attggctggc catatgccgg 6180

aaactcaagc tcgacccctt tcttacaccc tacactaaaa tcaacagtag gtggatcaag 6240

gacttgaatg tcaagccaaa gactataaag acactggaag agaatcttgg gatcacaata 6300

caagatatag gcgtcggcaa agattttatg tcaaagacgc ccaaggccat ggccactaag 6360

gataagattg ataagtggga ccttattaag ctcaaaagct tctgtactgc caaggagacc 6420

acgatcagag ttaataggca gcccactaca tgggaaaaga ttttcgccac ttattcatca 6480

gataaggggt tgataagcag aatatataac gagctgaagc agatctacaa gaagaaaacg 6540

aataatccca tcaagaagtg ggcaaaagat atgaacaggc attttagcaa agaggatatc 6600

tacgccgcga agaagcatat gaagaagtgt agttcaagct tggccattcg tgagatgcag 6660

attaagacga ccatgcgata ccaccttacc ccagtgagga tggcaattat caagaaatct 6720

ggcaataata gatgttggcg gggctgtggc gagattggca ccctgctcca ttgctggtgg 6780

gattgcaagc tggtgcagcc gctttggaaa tcagtctggc gctttctgag ggacctcgag 6840

cttgagattc ccttcgatcc cgcaattccc ttgctcggaa tctatcctaa cgaatacaag 6900

agctgttgtt acaaggatac gtgtacccgg atgttcatcg cggccttgtt tacgatagct 6960

aagacgtgga atcagcctaa gtgccccaca atgatcgatt ggatcaagaa aatgtggcat 7020

atttatacca tggagtatta cgcagcaatt aagaatgacg aatttatttc cttcgttggg 7080

acctggatga agctggagac tattattctg agcaagctgt ctcaggagca aaagacaaag 7140

catagaatct tctctctcat tggtggtaac gactacaaag acgatgacga caagggcggc 7200

ggcagcccaa agaagaagcg gaaggtctaa agcgcttcta gaagttgtct cctcctgcac 7260

tgactgactg atacaatcga tttctggatc cgcaggccta atcaacctct ggattacaaa 7320

atttgtgaaa gattgactgg tattcttaac tatgttgctc cttttacgct atgtggatac 7380

gctgctttaa tgcctttgta tcatgctatt gcttcccgta tggctttcat tttctcctcc 7440

ttgtataaat cctggttgct gtctctttat gaggagttgt ggcccgttgt caggcaacgt 7500

ggcgtggtgt gcactgtgtt tgctgacgca acccccactg gttggggcat tgccaccacc 7560

tgtcagctcc tttccgggac tttcgctttc cccctcccta ttgccacggc ggaactcatc 7620

gccgcctgcc ttgcccgctg ctggacaggg gctcggctgt tgggcactga caattccgtg 7680

gtgttgtcgg ggaagctgac gtcctttcca tggctgctcg cctgtgttgc cacctggatt 7740

ctgcgcggga cgtccttctg ctacgtccct tcggccctca atccagcgga ccttccttcc 7800

cgctgagaga cacaaaaaat tccaacacac tattgcaatg aaaataaatt tcctttatta 7860

gccagaagtc agatgctcaa ggggcttcat gatgtcccca taatttttgg cagagggaaa 7920

aagatctcag tggtatttgt gagccagggc attggccttc tgataggcag cctgcacctg 7980

aggagtgcgg ccgctttact tgtacagctc gtccatgccg agagtgatcc cggcggcggt 8040

cacgaactcc agcaggacca tgtgatcgcg cttctcgttg gggtctttgc tcagggcgga 8100

ctgggtgctc aggtagtggt tgtcgggcag cagcacgggg ccgtcgccga tgggggtgtt 8160

ctgctggtag tggtcggcga gctgcacgct gccgtcctcg atgttgtggc ggatcttgaa 8220

gttcaccttg atgccgttct tctgcttgtc ggccatgata tagacgttgt ggctgttgta 8280

gttgtactcc agcttgtgcc ccaggatgtt gccgtcctcc ttgaagtcga tgcccttcag 8340

ctcgatgcgg ttcaccaggg tgtcgccctc gaacttcacc tcggcgcggg tcttgtagtt 8400

gccgtcgtcc ttgaagaaga tggtgcgctc ctggacgtag ccttcgggca tggcggactt 8460

gaagaagtcg tgctgcttca tgtggtcggg gtagcggctg aagcactgca cgccgtaggt 8520

cagggtggtc acgagggtgg gccagggcac gggcagcttg ccggtggtgc agatgaactt 8580

cagggtcagc ttgccgtagg tggcatcgcc ctcgccctcg ccggacacgc tgaacttgtg 8640

gccgtttacg tcgccgtcca gctcgaccag gatgggcacc accccggtga acagctcctc 8700

gcccttgctc accatggtgg cgggatctga cggttcacta aaccagctct gcttatatag 8760

acctcccacc gtacacgcct accgcccatt tgcgtcaatg gggcggagtt gttacgacat 8820

tttggaaagt cccgttgatt ttggtgccaa aacaaactcc cattgacgtc aatggggtgg 8880

agacttggaa atccccgtga gtcaaaccgc tatccacgcc cattgatgta ctgccaaaac 8940

cgcatcacca tggtaatagc gatgactaat acgtagatgt actgccaagt aggaaagtcc 9000

cataaggtca tgtactgggc ataatgccag gcgggccatt taccgtcatt gacgtcaata 9060

gggggcgtac ttggcatatg atacacttga tgtactgcca agtgggcagt ttaccgtaaa 9120

tactccaccc attgacgtca atggaaagtc cctattggcg ttactatggg aacatacgtc 9180

attattgacg tcaatgggcg ggggtcgttg ggcggtcagc caggcgggcc atttaccgta 9240

agttatgtaa cgggcctgct gccggctctg cggcctcttc cgcgtcttcg ccttcgccct 9300

cagacgagtc ggatctccct ttgggccgcc tccccgcctg tctagcttga ctgactgaga 9360

tacagcgtac cttcagctca cagacatgat aagatacatt gatgagtttg gacaaaccac 9420

aactagaatg cagtgaaaaa aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt 9480

tgtaaccatt ataagctgca ataaacaagt taacaacaac aattgcattc attttatgtt 9540
tcaggttcag ggggaggtgt gggaggtttt ttaaagcaag taaaacctct acaaatgtgg 9600
tattggccca tctctatcgg tatcgtagca taacccctg gggcctctaa acgggtcttg 9660
aggggttttt tgtgcccctc gggccggatt gctatctacc ggcattggcg cagaaaaaaa 9720
tgcctgatgc gacgctgcgc gtcttatact cccacatatg ccagattcag caacggatac 9780
ggcttcccca acttgcccac ttccatacgt gtcctcctta ccagaaattt atccttaagg 9840
tcgtcagcta tcctgcaggc gatctctcga tttcgatcaa gacattcctt taatggtctt 9900

ttctggacac cactaggggt cagaagtagt tcatcaaact ttcttccctc cctaatctca 9960
ttggttacct tgggctatcg aaacttaatt aagcgatctg catctcaatt agtcagcaac 10020
catagtcccg cccctaactc cgcccatccc gcccctaact ccgcccagtt ccgcccattc 10080
tccgccccat cgctgactaa ttttttttat ttatgcagag gccgaggccg cctcggcctc 10140
tgagctattc cagaagtagt gaggaggctt ttttggaggc ctaggctttt gcaaaggagg 10200
tagccaacat gattgaacaa gatggattgc acgcaggttc tcccgccgct tgggtggaga 10260
ggctattcgg ctatgactgg gcacaacaga caatcggctg ctctgatgcc gccgtgttcc 10320

ggctgtcagc gcaggggcgc ccggttcttt ttgtcaagac cgacctgtcc ggtgccctga 10380
atgaactcca ggacgaggca gcgcggctat cgtggctggc cacgacgggc gttccttgcg 10440
cagctgtgct cgacgttgtc actgaagcgg gaagggactg gctgctattg ggcgaagtgc 10500
cggggcagga tctcctgtca tctcaccttg ctcctgccga gaaagtatcc atcatggctg 10560
atgcaatgcg gcggctgcat acgcttgatc cggctacctg cccattcgac caccaagcga 10620
aacatcgcat cgagcgagca cgtactcgga tggaagccgg tcttgtcgat caggatgatc 10680
tggacgaaga gcatcagggg ctcgcgccag ccgaactgtt cgccaggctc aaggcgcgga 10740

tgcccgacgg cgaggatctc gtcgtgaccc acggcgatgc ctgcttgccg aatatcatgg 10800
tggaaaatgg ccgcttttct ggattcatcg actgtggccg gctgggtgtg gcggaccgct 10860
atcaggacat agcgttggct acccgtgata ttgctgaaga gcttggcggc gaatgggctg 10920
accgcttcct cgtgctttac ggtatcgccg ctcccgattc gcagcgcatc gccttctatc 10980
gccttcttga cgagttcttc tagtatgtaa gccctgtgcc ttctagttgc cagccatctg 11040
ttgtttgccc ctccccgtg ccttccttga ccctggaagg tgccactccc actgtccttt 11100
cctaataaaa tgaggaaatt gcatcgcatt gtctgagtag gtgtcattct attctggggg 11160

gtggggtggg gcaggacagc aaggggggagg attgggaaga caatagcagg catgctgggg 11220

atgcggtggg ctctatggtt aattaaccag tcaagtcagc tacttggcga gatcgacttg 11280

tctgggtttc gactacgctc agaattgcgt cagtcaagtt cgatctggtc cttgctattg 11340

cacccgttct ccgattacga gtttcattta aatcatgtga gcaaaaggcc agcaaaaggc 11400

caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga 11460

gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 11520

ccaggcgttt cccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac 11580

cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 11640

taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 11700

cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag 11760

acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt 11820

aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaagaacagt 11880

atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 11940

atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 12000

gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca 12060

gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 12120

ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 12180

ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 12240

tcgttcatcc atagttgcat ttaaatttcc gaactctcca aggccctcgt cggaaaatct 12300

tcaaaccttt cgtccgatcc atcttgcagg ctacctctcg aacgaactat cgcaagtctc 12360

ttggccggcc ttgcgccttg gctattgctt ggcagcgcct atcgccaggt attactccaa 12420

tcccgaatat ccgagatcgg gatcacccga gagaagttca acctacatcc tcaatcccga 12480

tctatccgag atccgaggaa tatcgaaatc ggggcgcgcc tggtgtaccg agaacgatcc 12540

tctcagtgcg agtc 12554

<210> 106

<211> 12569

<212> DNA

<213> Artificial Sequence

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

<400> 106

accgagaacg atcctctcag tgcgagtctc gacgatccat atcgttgctt ggcagtcagc 60

cagtcggaat ccagcttggg acccaggaag tccaatcgtc agatattgta ctcaagcctg 120

gtcacggcag cgtaccgatc tgtttaaacc tagatattga tagtctgatc ggtcaacgta 180
taatcgagtc ctagcttttg caaacatcta tcaagagaca ggatcagcag gaggctttcg 240
catgagtatt caacatttcc gtgtcgccct tattcccttt tttgcggcat tttgccttcc 300
tgtttttgct cacccagaaa cgctggtgaa agtaaaaagat gctgaagatc agttgggtgc 360
gcgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga gttttcgccc 420
cgaagaacgc tttccaatga tgagcacttt taaagttctg ctatgtggcg cggtattatc 480
ccgtattgac gccgggcaag agcaactcgg tcgccgcata cactattctc agaatgactt 540

ggttgagtat tcaccagtca cagaaaagca tcttacggat ggcatgacag taagagaatt 600
atgcagtgct gccataacca tgagtgataa cactgcggcc aacttacttc tgacaacgat 660
tggaggaccg aaggagctaa ccgcttttt gcacaacatg ggggatcatg taactcgcct 720
tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg acaccacgat 780
gcctgtagca atggcaacaa ccttgcgtaa actattaact ggcgaactac ttactctagc 840
ttcccggcaa cagttgatag actggatgga ggcggataaa gttgcaggac cacttctgcg 900
ctcggccctt ccggctggct ggtttattgc tgataaatct ggagccggtg agcgtgggtc 960

tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg tagttatcta 1020
cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg agataggtgc 1080
ctcactgatt aagcattggt aaccgattct aggtgcattg gcgcagaaaa aaatgcctga 1140
tgcgacgctg cgcgtcttat actcccacat atgccagatt cagcaacgga tacggcttcc 1200
ccaacttgcc cacttccata cgtgtcctcc ttaccagaaa tttatcctta agatcgttta 1260
aactcgactc tggctctatc gaatctccgt cgtttcgagc ttacgcgaac agccgtggcg 1320
ctcatttgct cgtcgggcat cgaatctcgt cagctatcgt cagcttacct ttttggcagc 1380

gatcgcggct cccgacatct tggaccatta gctccacagg tatcttcttc cctctagtgg 1440
tcataacagc agcttcagct acctctcaat tcaaaaaacc cctcaagacc cgtttagagg 1500
ccccaagggg ttatgctatc aatcgttgcg ttacacacac aaaaaaccaa cacacatcca 1560
tcttcgatgg atagcgattt tattatctaa ctgctgatcg agtgtagcca gatctagtaa 1620
tcaattacgg ggtcattagt tcatagccca tatatggagt tccgcgttac ataacttacg 1680
gtaaatggcc cgcctggctg accgcccaac gacccccgcc cattgacgtc aataatgacg 1740
tatgttccca tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta 1800

cggtaaactg cccacttggc agtacatcaa gtgtatcata tgccaagtac gccccctatt 1860

gacgtcaatg acggtaaatg gcccgcctgg cattatgccc agtacatgac cttatgggac 1920

tttcctactt ggcagtacat ctacgtatta gtcatcgcta ttaccatgct gatgcggttt 1980

tggcagtaca tcaatgggcg tggatagcgg tttgactcac ggggatttcc aagtctccac 2040

cccattgacg tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt 2100

cgtaacaact ccgccccatt gacgcaaatg ggcggtaggc gtgtacggtg ggaggtctat 2160

ataagcagag ctggtttagt gaaccgtcag atcagatctt tgtcgatcct accatccact 2220

cgacacaccc gccagcggcc gctaatacga ctcactatag gggaagtac tgccaccatg 2280

ggcaagaagc aaaatcgcaa gacggggaat tccaagacac aatccgctag cccaccacct 2340

aaagagcgtt ctagctcccc tgctactgag cagtcctgga tggaaaacga cttcgatgaa 2400

ctccgggaag agggatttag gcgatccaac tattcagaac tccgcgaaga tatccagaca 2460

aaggggaagg aagtcgagaa tttcgagaag aacctcgagg agtgcatcac ccgtatcaca 2520

aacactgaga aatgtctcaa agaactcatg gaacttaaga caaaagccag ggagcttcga 2580

gaggagtgtc ggagtctgag atccaggtgt gaccagctcg aggagcgcgt gagcgcgatg 2640

gaagacgaga tgaacgagat gaaaagagag ggcaaattca gggagaagcg cattaagagg 2700

aacgaacaga gtctgcagga gatttgggat tacgtcaaga ggcctaacct gcggttgatc 2760

ggcgtccccg agagcgacgt agaaaacggg actaaactgg agaatacact tcaagacatc 2820

attcaagaaa attttccaaa cctggctcgg caagctaatg tgcaaatcca agagatccaa 2880

cgcacacccc agcggtatag ctctcggcgt gccaccccta ggcatattat cgtgcgcttt 2940

actaaggtgg agatgaaaga gaagatgctg cgagccgctc gggaaaaggg aagggtgact 3000

ttgaagggca aacctattcg gctgacggtt gaccttagcg ccgagacact ccaggcacgc 3060

cgggaatggg gccccatctt taatatcctg aaggagaaga acttccagcc acgaatctct 3120

taccctgcaa agttgagttt tatctccgag ggtgagatta agtatttcat cgataaacag 3180

atgctgcgag acttcgtgac aactcgccca gctctcaagg aactgctcaa agaggctctt 3240

aatatggagc gcaataatag atatcaaccc ttgcagaacc acgcaaagat gtgagacagc 3300

cgtcagacca tcaagactag gaagaaactg catcaactaa tgagcaaaat caccagctaa 3360

catcatagta tacatgaccg gctctaactc acatatcacc atccttacac ttaacattaa 3420

cggcctcaac tcagctatca agcgccatcg gctggccagc tggatcaaat cacaggatcc 3480

aagcgtttgt tgcatccaag agacccacct gacctgtaga gatactcacc gcctcaagat 3540

caagggatgg cgaaagattt atcaggcgaa cggtaagcag aagaaagccg gagtcgcaat 3600

tctggtctca gacaagacgg atttcaagcc caccaaaatt aagcgtgata aggaaggtca 3660

ctatattatg gtgaaaggca gcatacagca ggaagaactt accatattga acatctacgc 3720

gccaaacacc ggcgcacctc gctttatcaa acaggtcctg tccgatctgc agcgagatct 3780

ggattctcat acgttgatta tgggtgattt caatacacca ttgagcaccc tggatcgcag 3840

caccaggcaa aaggtaaata aagacacgca agagctcaat agcgcactgc atcaggcaga 3900

tctcattgat atttatcgca ctcttcatcc taagagtacc gagtacacat tcttcagcgc 3960

cccacatcat acatactcaa agatcgatca tatcgtcggc tcaaaggctc tgctgtcaaa 4020

gtgcaagcgc acagagataa ttacaaatta cctgtcagat catagcgcga tcaagctcga 4080

gctgagaatc aagaacctga cccagagccg gagtaccact tggaagctta ataacctgct 4140

gctcaacgat tattgggtcc acaatgagat gaaggcagag attaaaatgt tcttcgaaac 4200

aaatgagaat aaggatacta cctatcaaaa cctttgggat gcctttaagg ccgtctgcag 4260

aggcaagttc atcgccctca acgcctataa aagaaaacaa gagagatcta agatcgatac 4320

tctcacctct cagctgaagg agttggagaa acaggaacag acccactcca aggcgtcaag 4380

acggcaggag atcacaaaga ttcgcgccga gttgaaagag atcgaaaccc aaaagactct 4440

tcagaaaatt aacgagtctc gtagttggtt cttcgagcgg attaataaga tagacagacc 4500

tctggcacga ctgattaaga agaagcgcga aaagaaccag attgatacca tcaagaacga 4560

caagggcgac atcactactg acccgaccga gatccagacc actattcggg agtattataa 4620

gcatttgtat gctaacaagc ttgagaacct ggaagagatg gacacttttc tggataccta 4680

tactctgcca cggcttaatc aagaggaagt cgagtccctc aaccgcccaa ttacaggaag 4740

cgagattgtg gccataatta actccctgcc gacaaagaaa tctcctggtc cggacgggtt 4800

tacagctgag ttttatcaac ggtatatgga agagcttgta ccgtttctgc tcaagctctt 4860

tcagtctata gaaaaggaag gcatcttgcc caattccttc tacgaagctt ctataatact 4920

tattcccaaa ccaggacgcg ataccacaaa gaaggaaaac ttccggccca ttagtctcat 4980

gaatatcgac gctaaaatat tgaacaagat tctcgccaac agaatccaac aacatattaa 5040

gaaattgata catcacgacc aggtggggtt tatacctggc atgcagggct ggtttaacat 5100

ccggaagagt attaacgtca ttcaacacat taatagagct aaggataaga atcatatgat 5160

catctctata gacgcggaaa aggcattcga taagattcag cagccattta tgctcaagac 5220

tctgaacaaa ctcggcatcg acggaacata ttttaagatt attcgcgcaa tttacgataa 5280

gccgactgct aacattatcc ttaacggcca aaagctcgag gcctttccgc tcaagactgg 5340

aacccgccaa ggctgtcccc tctccccgct tttgtttaat attgtactcg aggtgctggc 5400

tagggctatt cgtcaagaga aagagattaa agggatacag ctcgggaagg aagaggtcaa 5460

gctttccttg ttcgccgatg atatgattgt gtacctggag aatcctattg tgtctgctca 5520

gaaccttctt aaacttattt ctaactttag caaggtcagc ggctataaga ttaacgtcca 5580

gaaatctcag gcctttctgt acacaaataa tcgacagacc gaatcccaga taatgggtga 5640
gcttccgttt gtcatagcca gcaaaaggat aaagtatctc ggaatccagc tgacacgaga 5700
cgttaaagat ttgtttaagg aaaattacaa gcctctcctg aaagagatta aggaagatac 5760
taataagtgg aagaatatcc cctgttcatg ggttggcaga atcaacatag tgaagatggc 5820
aatacttcct aaagtgatat atcgctttaa cgccatccca attaaactgc ctatgacctt 5880
ctttacggag ctcgagaaaa caacccttaa atttatatgg aatcaaaaga gagcaagaat 5940
agcgaagtcc atcttgagcc agaagaataa ggccggtggg attactttgc ctgattttaa 6000

gttgtattat aaagccacag taactaagac agcctggtat tggtatcaga atagagacat 6060
cgaccagtgg aatcggaccg aaccatcaga gataatgccc cacatctata attaccttat 6120
attcgataag ccagaaaaga ataaacagtg gggcaaagac agcctcttca acaagtggtg 6180
ttgggagaat tggctggcca tatgccggaa actcaagctc gacccctttc ttacacccta 6240
cactaaaatc aacagtaggt ggatcaagga cttgaatgtc aagccaaaga ctataaagac 6300
actggaagag aatcttggga tcacaataca agatataggc gtcggcaaag attttatgtc 6360
aaagacgccc aaggccatgg ccactaagga taagattgat aagtgggacc ttattaagct 6420

caaaagcttc tgtactgcca aggagaccac gatcagagtt aataggcagc ccactacatg 6480
ggaaaagatt ttcgccactt attcatcaga taaggggttg ataagcagaa tatataacga 6540
gctgaagcag atctacaaga agaaaacgaa taatcccatc aagaagtggg caaaagatat 6600
gaacaggcat tttagcaaag aggatatcta cgccgcgaag aagcatatga agaagtgtag 6660
ttcaagcttg gccattcgtg agatgcagat taagacgacc atgcgatacc accttacccc 6720
agtgaggatg gcaattatca agaaatctgg caataataga tgttggcggg gctgtggcga 6780
gattggcacc ctgctccatt gctggtggga ttgcaagctg gtgcagccgc tttggaaatc 6840

agtctggcgc tttctgaggg acctcgagct tgagattccc ttcgatcccg caattccctt 6900
gctcggaatc tatcctaacg aatacaagag ctgttgttac aaggatacgt gtacccggat 6960
gttcatcgcg gccttgttta cgatagctaa gacgtggaat cagcctaagt gccccacaat 7020
gatcgattgg atcaagaaaa tgtggcatat ttataccatg gagtattacg cagcaattaa 7080
gaatgacgaa tttatttcct tcgttgggac ctggatgaag ctggagacta ttattctgag 7140
caagctgtct caggagcaaa agacaaagca tagaatcttc tctctcattg gtggtaacga 7200
ctacaaagac gatgacgaca agaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa 7260

aaagaaaaag taaagcgctt ctagaagttg tctcctcctg cactgactga ctgatacaat 7320

cgatttctgg atccgcaggc ctaatcaacc tctggattac aaaatttgtg aaagattgac 7380

tggtattctt aactatgttg ctccttttac gctatgtgga tacgctgctt taatgccttt 7440

gtatcatgct attgcttccc gtatggcttt cattttctcc tccttgtata aatcctggtt 7500

gctgtctctt tatgaggagt tgtggcccgt tgtcaggcaa cgtggcgtgg tgtgcactgt 7560

gtttgctgac gcaaccccca ctggttgggg cattgccacc acctgtcagc tcctttccgg 7620

gactttcgct ttccccctcc ctattgccac ggcggaactc atcgccgcct gccttgcccg 7680

ctgctggaca ggggctcggc tgttgggcac tgacaattcc gtggtgttgt cggggaagct 7740

gacgtccttt ccatggctgc tcgcctgtgt tgccacctgg attctgcgcg ggacgtcctt 7800

ctgctacgtc ccttcggccc tcaatccagc ggaccttcct tcccgctgag agacacaaaa 7860

aattccaaca cactattgca atgaaaataa atttccttta ttagccagaa gtcagatgct 7920

caaggggctt catgatgtcc ccataatttt tggcagaggg aaaaagatct cagtggtatt 7980

tgtgagccag ggcattggcc ttctgatagg cagcctgcac ctgaggagtg cggccgcttt 8040

acttgtacag ctcgtccatg ccgagagtga tcccggcggc ggtcacgaac tccagcagga 8100

ccatgtgatc gcgcttctcg ttggggtctt tgctcagggc ggactgggtg ctcaggtagt 8160

ggttgtcggg cagcagcacg gggccgtcgc cgatgggggt gttctgctgg tagtggtcgg 8220

cgagctgcac gctgccgtcc tcgatgttgt ggcggatctt gaagttcacc ttgatgccgt 8280

tcttctgctt gtcggccatg atatagacgt tgtggctgtt gtagttgtac tccagcttgt 8340

gccccaggat gttgccgtcc tccttgaagt cgatgcccct cagctcgatg cggttcacca 8400

gggtgtcgcc ctcgaacttc acctcggcgc gggtcttgta gttgccgtcg tccttgaaga 8460

agatggtgcg ctcctggacg tagccttcgg gcatggcgga cttgaagaag tcgtgctgct 8520

tcatgtggtc ggggtagcgg ctgaagcact gcacgccgta ggtcagggtg gtcacgaggg 8580

tgggccaggg cacgggcagc ttgccggtgg tgcagatgaa cttcagggtc agcttgccgt 8640

aggtggcatc gccctcgccc tcgccggaca cgctgaactt gtggccgttt acgtcgccgt 8700

ccagctcgac caggatgggc accaccccgg tgaacagctc ctcgcccttg ctcaccatgg 8760

tggcgggatc tgacggttca ctaaaccagc tctgcttata tagacctccc accgtacacg 8820

cctaccgccc atttgcgtca atggggcgga gttgttacga cattttggaa agtcccgttg 8880

attttggtgc caaaacaaac tcccattgac gtcaatgggg tggagacttg gaaatccccg 8940

tgagtcaaac cgctatccac gcccattgat gtactgccaa aaccgcatca ccatggtaat 9000

agcgatgact aatacgtaga tgtactgcca agtaggaaag tcccataagg tcatgtactg 9060

ggcataatgc caggcgggcc atttaccgtc attgacgtca ataggggcg tacttggcat 9120

atgatacact tgatgtactg ccaagtgggc agtttaccgt aaatactcca cccattgacg 9180

tcaatggaaa gtccctattg gcgttactat gggaacatac gtcattattg acgtcaatgg 9240

gcgggggtcg ttgggcggtc agccaggcgg gccatttacc gtaagttatg taacgggcct 9300

gctgccggct ctgcggcctc ttccgcgtct tcgccttcgc cctcagacga gtcggatctc 9360

cctttgggcc gcctccccgc ctgtctagct tgactgactg agatacagcg taccttcagc 9420

tcacagacat gataagatac attgatgagt ttggacaaac cacaactaga atgcagtgaa 9480

aaaaatgctt tatttgtgaa atttgtgatg ctattgcttt atttgtaacc attataagct 9540

gcaataaaca agttaacaac aacaattgca ttcattttat gtttcaggtt cagggggagg 9600

tgtgggaggt tttttaaagc aagtaaaacc tctacaaatg tggtattggc ccatctctat 9660

cggtatcgta gcataacccc ttggggcctc taaacgggtc ttgaggggtt ttttgtgccc 9720

ctcgggccgg attgctatct accggcattg gcgcagaaaa aaatgcctga tgcgacgctg 9780

cgcgtcttat actcccacat atgccagatt cagcaacgga tacggcttcc ccaacttgcc 9840

cacttccata cgtgtcctcc ttaccagaaa tttatcctta aggtcgtcag ctatcctgca 9900

ggcgatctct cgatttcgat caagacattc ctttaatggt cttttctgga caccactagg 9960

ggtcagaagt agttcatcaa actttcttcc ctccctaatc tcattggtta ccttgggcta 10020

tcgaaactta attaagcgat ctgcatctca attagtcagc aaccatagtc ccgcccctaa 10080

ctccgcccat cccgccccta actccgccca gttccgccca ttctccgccc catcgctgac 10140

taatttttt tatttatgca gaggccgagg ccgcctcggc ctctgagcta ttccagaagt 10200

agtgaggagg ctttttgga ggcctaggct tttgcaaagg aggtagccaa catgattgaa 10260

caagatggat tgcacgcagg ttctcccgcc gcttgggtgg agaggctatt cggctatgac 10320

tgggcacaac agacaatcgg ctgctctgat gccgccgtgt tccggctgtc agcgcagggg 10380

cgcccggttc tttttgtcaa gaccgacctg tccggtgccc tgaatgaact ccaggacgag 10440

gcagcgcggc tatcgtggct ggccacgacg ggcgttcctt gcgcagctgt gctcgacgtt 10500

gtcactgaag cgggaaggga ctggctgcta ttgggcgaag tgccggggca ggatctcctg 10560

tcatctcacc ttgctcctgc cgagaaagta tccatcatgg ctgatgcaat gcggcggctg 10620

catacgcttg atccggctac ctgcccattc gaccaccaag cgaaacatcg catcgagcga 10680

gcacgtactc ggatggaagc cggtcttgtc gatcaggatg atctggacga agagcatcag 10740

gggctcgcgc cagccgaact gttcgccagg ctcaaggcgc ggatgcccga cggcgaggat 10800

ctcgtcgtga cccacggcga tgcctgcttg ccgaatatca tggtggaaaa tggccgcttt 10860

tctggattca tcgactgtgg ccggctgggt gtggcggacc gctatcagga catagcgttg 10920

gctacccgtg atattgctga agagcttggc ggcgaatggg ctgaccgctt cctcgtgctt 10980

tacggtatcg ccgctcccga ttcgcagcgc atcgccttct atcgccttct tgacgagttc 11040

ttctagtatg taagccctgt gccttctagt tgccagccat ctgttgtttg cccctccccc 11100
gtgccttcct tgaccctgga aggtgccact cccactgtcc tttcctaata aaatgaggaa 11160
attgcatcgc attgtctgag taggtgtcat tctattctgg ggggtggggt ggggcaggac 11220
agcaaggggg aggattggga agacaatagc aggcatgctg gggatgcggt gggctctatg 11280
gttaattaac cagtcaagtc agctacttgg cgagatcgac ttgtctgggt ttcgactacg 11340
ctcagaattg cgtcagtcaa gttcgatctg gtccttgcta ttgcacccgt tctccgatta 11400
cgagtttcat ttaaatcatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg 11460

ccgcgttgct ggcgtttttc cataggctcc gccccccctga cgagcatcac aaaaatcgac 11520
gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg 11580
gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct 11640
ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg 11700
tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct 11760
gcgccttatc cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac 11820
tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt 11880

tcttgaagtg gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc 11940
tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca 12000
ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat 12060
ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac 12120
gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt 12180
aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc 12240
aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg 12300

catttaaatt tccgaactct ccaaggccct cgtcggaaaa tcttcaaacc tttcgtccga 12360
tccatcttgc aggctacctc tcgaacgaac tatcgcaagt ctcttggccg gccttgcgcc 12420
ttggctattg cttggcagcg cctatcgcca ggtattactc caatcccgaa tatccgagat 12480
cgggatcacc cgagagaagt tcaacctaca tcctcaatcc cgatctatcc gagatccgag 12540
gaatatcgaa atcggggcgc gcctggtgt 12569

<210> 107

<211> 12581

<212> DNA

<213> Artificial Sequence

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

<400> 107

ccgagaacga tcctctcagt gcgagtctcg acgatccata tcgttgcttg gcagtcagcc 60

agtcggaatc cagcttggga cccaggaagt ccaatcgtca gatattgtac tcaagcctgg 120

tcacggcagc gtaccgatct gtttaaacct agatattgat agtctgatcg gtcaacgtat 180

aatcgagtcc tagcttttgc aaacatctat caagagacag gatcagcagg aggctttcgc 240

atgagtattc aacatttccg tgtcgcccct attccctttt ttgcggcatt ttgccttcct 300

gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca gttgggtgcg 360

cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag ttttcgcccc 420

gaagaacgct ttccaatgat gagcactttt aaagttctgc tatgtggcgc ggtattatcc 480

cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca gaatgacttg 540

gttgagtatt caccagtcac agaaaagcat cttacggatg gcatgacagt aagagaatta 600

tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct gacaacgatt 660

ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgcctt 720

gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga caccacgatg 780

cctgtagcaa tggcaacaac cttgcgtaaa ctattaactg gcgaactact tactctagct 840

tcccggcaac agttgataga ctggatggag gcggataaag ttgcaggacc acttctgcgc 900

tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga gcgtgggtct 960

cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac 1020

acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc 1080

tcactgatta agcattggta accgattcta ggtgcattgg cgcagaaaaa aatgcctgat 1140

gcgacgctgc gcgtcttata ctcccacata tgccagattc agcaacggat acggcttccc 1200

caacttgccc acttccatac gtgtcctcct taccagaaat ttatccttaa gatcgtttaa 1260

actcgactct ggctctatcg aatctccgtc gtttcgagct tacgcgaaca gccgtggcgc 1320

tcatttgctc gtcgggcatc gaatctcgtc agctatcgtc agcttacctt tttggcagcg 1380

atcgcggctc ccgacatctt ggaccattag ctccacaggt atcttcttcc ctctagtggt 1440

cataacagca gcttcagcta cctctcaatt caaaaaaccc ctcaagaccc gtttagaggc 1500

cccaaggggt tatgctatca atcgttgcgt tacacacaca aaaaaccaac acacatccat 1560

cttcgatgga tagcgatttt attatctaac tgctgatcga gtgtagccag atctagtaat 1620

caattacggg gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg 1680

taaatggccc gcctggctga ccgcccaacg acccccgccc attgacgtca ataatgacgt 1740
atgttcccat agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac 1800
ggtaaactgc ccacttggca gtacatcaag tgtatcatat gccaagtacg ccccctattg 1860
acgtcaatga cggtaaatgg cccgcctggc attatgccca gtacatgacc ttatgggact 1920
ttcctacttg gcagtacatc tacgtattag tcatcgctat taccatgctg atgcggtttt 1980
ggcagtacat caatgggcgt ggatagcggt ttgactcacg gggatttcca agtctccacc 2040
ccattgacgt caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc 2100

gtaacaactc cgccccattg acgcaaatgg gcggtaggcg tgtacggtgg gaggtctata 2160
taagcagagc tggtttagtg aaccgtcaga tcagatcttt gtcgatccta ccatccactc 2220
gacacacccg ccagcggccg ctaatacgac tcactatagg gagaagtact gccaccatgg 2280
gcaagaagca aaatcgcaag acggggaatt ccaagacaca atccgctagc ccaccaccta 2340
aagagcgttc tagctcccct gctactgagc agtcctggat ggaaaacgac ttcgatgaac 2400
tccgggaaga gggatttagg cgatccaact attcagaact ccgcgaagat atccagacaa 2460
aggggaagga agtcgagaat ttcgagaaga acctcgagga gtgcatcacc cgtatcacaa 2520

acactgagaa atgtctcaaa gaactcatgg aacttaagac aaaagccagg gagcttcgag 2580
aggagtgtcg gagtctgaga tccaggtgtg accagctcga ggagcgcgtg agcgcgatgg 2640
aagacgagat gaacgagatg aaaagagagg gcaaattcag ggagaagcgc attaagagga 2700
acgaacagag tctgcaggag atttgggatt acgtcaagag gcctaacctg cggttgatcg 2760
gcgtccccga gagcgacgta gaaaacggga ctaaactgga gaatacactt caagacatca 2820
ttcaagaaaa ttttccaaac ctggctcggc aagctaatgt gcaaatccaa gagatccaac 2880
gcacacccca gcggtatagc tctcggcgtg ccacccctag gcatattatc gtgcgcttta 2940

ctaaggtgga gatgaaagag aagatgctgc gagccgctcg ggaaaaggga agggtgactt 3000
tgaagggcaa acctattcgg ctgacggttg accttagcgc cgagacactc caggcacgcc 3060
gggaatgggg ccccatcttt aatatcctga aggagaagaa cttccagcca cgaatctctt 3120
accctgcaaa gttgagtttt atctccgagg gtgagattaa gtatttcatc gataaacaga 3180
tgctgcgaga cttcgtgaca actcgcccag ctctcaagga actgctcaaa gaggctctta 3240
atatggagcg caataataga tatcaaccct tgcagaacca cgcaaagatg tgagacagcc 3300
gtcagaccat caagactagg aagaaactgc atcaactaat gagcaaaatc accagctaac 3360

atcatagtat acatgaccgg ctctaactca catatcacca tccttacact taacattaac 3420

ggcctcaact cagctatcaa gcgccatcgg ctggccagct ggatcaaatc acaggatcca 3480

agcgtttgtt gcatccaaga gacccacctg acctgtagag atactcaccg cctcaagatc 3540

aagggatggc gaaagattta tcaggcgaac ggtaagcaga agaaagccgg agtcgcaatt 3600

ctggtctcag acaagacgga tttcaagccc accaaaatta agcgtgataa ggaaggtcac 3660

tatattatgg tgaaaggcag catacagcag gaagaactta ccatattgaa catctacgcg 3720

ccaaacaccg gcgcacctcg ctttatcaaa caggtcctgt ccgatctgca gcgagatctg 3780

gattctcata cgttgattat gggtgatttc aatacaccat tgagcaccct ggatcgcagc 3840

accaggcaaa aggtaaataa agacacgcaa gagctcaata gcgcactgca tcaggcagat 3900

ctcattgata tttatcgcac tcttcatcct aagagtaccg agtacacatt cttcagcgcc 3960

ccacatcata catactcaaa gatcgatcat atcgtcggct caaaggctct gctgtcaaag 4020

tgcaagcgca cagagataat tacaaattac ctgtcagatc atagcgcgat caagctcgag 4080

ctgagaatca agaacctgac ccagagccgg agtaccactt ggaagcttaa taacctgctg 4140

ctcaacgatt attgggtcca caatgagatg aaggcagaga ttaaaatgtt cttcgaaaca 4200

aatgagaata aggatactac ctatcaaaac ctttgggatg cctttaaggc cgtctgcaga 4260

ggcaagttca tcgccctcaa cgcctataaa agaaaacaag agagatctaa gatcgatact 4320

ctcacctctc agctgaagga gttggagaaa caggaacaga cccactccaa ggcgtcaaga 4380

cggcaggaga tcacaaagat tcgcgccgag ttgaaagaga tcgaaaccca aaagactctt 4440

cagaaaatta acgagtctcg tagttggttc ttcgagcgga ttaataagat agacagacct 4500

ctggcacgac tgattaagaa gaagcgcgaa aagaaccaga ttgataccat caagaacgac 4560

aagggcgaca tcactactga cccgaccgag atccagacca ctattcggga gtattataag 4620

catttgtatg ctaacaagct tgagaacctg gaagagatgg acacttttct ggatacctat 4680

actctgccac ggcttaatca agaggaagtc gagtccctca accgcccaat tacaggaagc 4740

gagattgtgg ccataattaa ctccctgccg acaaagaaat ctcctggtcc ggacgggttt 4800

acagctgagt tttatcaacg gtatatggaa gagcttgtac cgtttctgct caagctcttt 4860

cagtctatag aaaaggaagg catcttgccc aattccttct acgaagcttc tataatactt 4920

attcccaaac caggacgcga taccacaaag aaggaaaact tccggcccat tagtctcatg 4980

aatatcgacg ctaaaatatt gaacaagatt ctcgccaaca gaatccaaca acatattaag 5040

aaattgatac atcacgacca ggtggggttt atacctggca tgcagggctg gtttaacatc 5100

cggaagagta ttaacgtcat tcaacacatt aatagagcta aggataagaa tcatatgatc 5160

atctctatag acgcggaaaa ggcattcgat aagattcagc agccatttat gctcaagact 5220

ctgaacaaac tcggcatcga cggaacatat tttaagatta ttcgcgcaat ttacgataag 5280

ccgactgcta acattatcct taacggccaa aagctcgagg cctttccgct caagactgga 5340

acccgccaag gctgtcccct ctccccgctt ttgtttaata ttgtactcga ggtgctggct 5400

agggctattc gtcaagagaa agagattaaa gggatacagc tcgggaagga agaggtcaag 5460

ctttccttgt tcgccgatga tatgattgtg tacctggaga atcctattgt gtctgctcag 5520

aaccttctta aacttatttc taactttagc aaggtcagcg gctataagat taacgtccag 5580

aaatctcagg cctttctgta cacaaataat cgacagaccg aatcccagat aatgggtgag 5640

cttccgtttg tcatagccag caaaaggata aagtatctcg gaatccagct gacacgagac 5700

gttaaagatt tgtttaagga aaattacaag cctctcctga aagagattaa ggaagatact 5760

aataagtgga agaatatccc ctgttcatgg gttggcagaa tcaacatagt gaagatggca 5820

atacttccta aagtgatata tcgctttaac gccatcccaa ttaaactgcc tatgaccttc 5880

tttacggagc tcgagaaaac aacccttaaa tttatatgga atcaaaagag agcaagaata 5940

gcgaagtcca tcttgagcca gaagaataag gccggtggga ttactttgcc tgattttaag 6000

ttgtattata aagccacagt aactaagaca gcctggtatt ggtatcagaa tagagacatc 6060

gaccagtgga atcggaccga accatcagag ataatgcccc acatctataa ttaccttata 6120

ttcgataagc cagaaaagaa taaacagtgg ggcaaagaca gcctcttcaa caagtggtgt 6180

tgggagaatt ggctggccat atgccggaaa ctcaagctcg acccctttct tacaccctac 6240

actaaaatca acagtaggtg gatcaaggac ttgaatgtca agccaaagac tataaagaca 6300

ctggaagaga atcttgggat cacaatacaa gatataggcg tcggcaaaga ttttatgtca 6360

aagacgccca aggccatggc cactaaggat aagattgata agtgggacct tattaagctc 6420

aaaagcttct gtactgccaa ggagaccacg atcagagtta ataggcagcc cactacatgg 6480

gaaaagattt tcgccactta ttcatcagat aaggggttga taagcagaat atataacgag 6540

ctgaagcaga tctacaagaa gaaaacgaat aatcccatca agaagtgggc aaaagatatg 6600

aacaggcatt ttagcaaaga ggatatctac gccgcgaaga agcatatgaa gaagtgtagt 6660

tcaagcttgg ccattcgtga gatgcagatt aagacgacca tgcgatacca ccttacccca 6720

gtgaggatgg caattatcaa gaaatctggc aataatagat gttggcgggg ctgtggcgag 6780

attggcaccc tgctccattg ctggtgggat tgcaagctgg tgcagccgct ttggaaatca 6840

gtctggcgct ttctgaggga cctcgagctt gagattccct tcgatcccgc aattcccttg 6900

ctcggaatct atcctaacga atacaagagc tgttgttaca aggatacgtg tacccggatg 6960

ttcatcgcgg ccttgtttac gatagctaag acgtggaatc agcctaagtg ccccacaatg 7020

atcgattgga tcaagaaaat gtggcatatt tataccatgg agtattacgc agcaattaag 7080

aatgacgaat ttatttcctt cgttgggacc tggatgaagc tggagactat tattctgagc 7140

aagctgtctc aggagcaaaa gacaaagcat agaatcttct ctctcattgg tggtaacgac 7200
tacaaagacg atgacgacaa gggcggcggc agcaaaaggc cggcggccac gaaaaaggcc 7260
ggccaggcaa aaaagaaaaa gtaaagcgct tctagaagtt gtctcctcct gcactgactg 7320
actgatacaa tcgatttctg gatccgcagg cctaatcaac ctctggatta caaaatttgt 7380
gaaagattga ctggtattct taactatgtt gctcctttta cgctatgtgg atacgctgct 7440
ttaatgcctt tgtatcatgc tattgcttcc cgtatggctt tcattttctc ctccttgtat 7500
aaatcctggt tgctgtctct ttatgaggag ttgtggcccg ttgtcaggca acgtggcgtg 7560

gtgtgcactg tgtttgctga cgcaaccccc actggttggg gcattgccac cacctgtcag 7620
ctcctttccg ggactttcgc tttccccctc cctattgcca cggcggaact catcgccgcc 7680
tgccttgccc gctgctggac aggggctcgg ctgttgggca ctgacaattc cgtggtgttg 7740
tcggggaagc tgacgtcctt tccatggctg ctcgcctgtg ttgccacctg gattctgcgc 7800
gggacgtcct tctgctacgt cccttcggcc ctcaatccag cggaccttcc ttcccgctga 7860
gagacacaaa aaattccaac acactattgc aatgaaaata aatttccttt attagccaga 7920
agtcagatgc tcaaggggct tcatgatgtc cccataattt ttggcagagg gaaaaagatc 7980

tcagtggtat ttgtgagcca gggcattggc cttctgatag gcagcctgca cctgaggagt 8040
gcggccgctt tacttgtaca gctcgtccat gccgagagtg atcccggcgg cggtcacgaa 8100
ctccagcagg accatgtgat cgcgcttctc gttggggtct ttgctcaggg cggactgggt 8160
gctcaggtag tggttgtcgg gcagcagcac ggggccgtcg ccgatggggg tgttctgctg 8220
gtagtggtcg gcgagctgca cgctgccgtc ctcgatgttg tggcggatct tgaagttcac 8280
cttgatgccg ttcttctgct tgtcggccat gatatagacg ttgtggctgt tgtagttgta 8340
ctccagcttg tgccccagga tgttgccgtc ctccttgaag tcgatgccct tcagctcgat 8400

gcggttcacc agggtgtcgc cctcgaactt cacctcggcg cgggtcttgt agttgccgtc 8460
gtccttgaag aagatggtgc gctcctggac gtagccttcg ggcatggcgg acttgaagaa 8520
gtcgtgctgc ttcatgtggt cggggtagcg gctgaagcac tgcacgccgt aggtcagggt 8580
ggtcacgagg gtgggccagg gcacgggcag cttgccggtg gtgcagatga acttcagggt 8640
cagcttgccg taggtggcat cgccctcgcc ctcgccggac acgctgaact tgtggccgtt 8700
tacgtcgccg tccagctcga ccaggatggg caccaccccg gtgaacagct cctcgcccct 8760
gctcaccatg gtggcgggat ctgacggttc actaaaccag ctctgcttat atagacctcc 8820

caccgtacac gcctaccgcc catttgcgtc aatggggcgg agttgttacg acattttgga 8880

aagtcccgtt gattttggtg ccaaaacaaa ctcccattga cgtcaatggg gtggagactt 8940

ggaaatcccc gtgagtcaaa ccgctatcca cgcccattga tgtactgcca aaaccgcatc 9000

accatggtaa tagcgatgac taatacgtag atgtactgcc aagtaggaaa gtcccataag 9060

gtcatgtact gggcataatg ccaggcgggc catttaccgt cattgacgtc aataggggggc 9120

gtacttggca tatgatacac ttgatgtact gccaagtggg cagtttaccg taaatactcc 9180

acccattgac gtcaatggaa agtccctatt ggcgttacta tgggaacata cgtcattatt 9240

gacgtcaatg ggcgggggtc gttgggcggt cagccaggcg ggccatttac cgtaagttat 9300

gtaacgggcc tgctgccggc tctgcggcct cttccgcgtc ttcgccttcg ccctcagacg 9360

agtcggatct ccctttgggc cgcctccccg cctgtctagc ttgactgact gagatacagc 9420

gtaccttcag ctcacagaca tgataagata cattgatgag tttggacaaa ccacaactag 9480

aatgcagtga aaaaaatgct ttatttgtga aatttgtgat gctattgctt tatttgtaac 9540

cattataagc tgcaataaac aagttaacaa caacaattgc attcatttta tgtttcaggt 9600

tcaggggggag gtgtgggagg ttttttaaag caagtaaaac ctctacaaat gtggtattgg 9660

cccatctcta tcggtatcgt agcataaccc cttggggcct ctaaacgggt cttgaggggt 9720

tttttgtgcc cctcgggccg gattgctatc taccggcatt ggcgcagaaa aaaatgcctg 9780

atgcgacgct gcgcgtctta tactcccaca tatgccagat tcagcaacgg atacggcttc 9840

cccaacttgc ccacttccat acgtgtcctc cttaccagaa atttatcctt aaggtcgtca 9900

gctatcctgc aggcgatctc tcgatttcga tcaagacatt cctttaatgg tcttttctgg 9960

acaccactag gggtcagaag tagttcatca aactttcttc cctccctaat ctcattggtt 10020

accttgggct atcgaaactt aattaagcga tctgcatctc aattagtcag caaccatagt 10080

cccgccccta actccgccca tcccgcccct aactccgccc agttccgccc attctccgcc 10140

ccatcgctga ctaatttttt ttatttatgc agaggccgag gccgcctcgg cctctgagct 10200

attccagaag tagtgaggag gctttttttgg aggcctaggc ttttgcaaag gaggtagcca 10260

acatgattga acaagatgga ttgcacgcag gttctcccgc cgcttgggtg gagaggctat 10320

tcggctatga ctgggcacaa cagacaatcg gctgctctga tgccgccgtg ttccggctgt 10380

cagcgcaggg gcgcccggtt ctttttgtca agaccgacct gtccggtgcc ctgaatgaac 10440

tccaggacga ggcagcgcgg ctatcgtggc tggccacgac gggcgttcct tgcgcagctg 10500

tgctcgacgt tgtcactgaa gcgggaaggg actggctgct attgggcgaa gtgccggggc 10560

aggatctcct gtcatctcac cttgctcctg ccgagaaagt atccatcatg gctgatgcaa 10620

tgcggcggct gcatacgctt gatccggcta cctgcccatt cgaccaccaa gcgaaacatc 10680

gcatcgagcg agcacgtact cggatggaag ccggtcttgt cgatcaggat gatctggacg 10740

aagagcatca ggggctcgcg ccagccgaac tgttcgccag gctcaaggcg cggatgcccg 10800

acggcgagga tctcgtcgtg acccacggcg atgcctgctt gccgaatatc atggtggaaa 10860

atggccgctt ttctggattc atcgactgtg gccggctggg tgtggcggac cgctatcagg 10920

acatagcgtt ggctacccgt gatattgctg aagagcttgg cggcgaatgg gctgaccgct 10980

tcctcgtgct ttacggtatc gccgctcccg attcgcagcg catcgccttc tatcgccttc 11040

ttgacgagtt cttctagtat gtaagccctg tgccttctag ttgccagcca tctgttgttt 11100

gcccctcccc cgtgccttcc ttgaccctgg aaggtgccac tcccactgtc ctttcctaat 11160

aaaatgagga aattgcatcg cattgtctga gtaggtgtca ttctattctg gggggtgggg 11220

tggggcagga cagcaagggg gaggattggg aagacaatag caggcatgct ggggatgcgg 11280

tgggctctat ggttaattaa ccagtcaagt cagctacttg gcgagatcga cttgtctggg 11340

tttcgactac gctcagaatt gcgtcagtca agttcgatct ggtccttgct attgcacccg 11400

ttctccgatt acgagtttca tttaaatcat gtgagcaaaa ggccagcaaa aggccaggaa 11460

ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccccctg acgagcatca 11520

caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc 11580

gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata 11640

cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 11700

tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca 11760

gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga 11820

cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 11880

tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 11940

tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 12000

caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag 12060

aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa 12120

cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 12180

ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 12240

tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc 12300

atccatagtt gcatttaaat ttccgaactc tccaaggccc tcgtcggaaa atcttcaaac 12360

ctttcgtccg atccatcttg caggctacct ctcgaacgaa ctatcgcaag tctcttggcc 12420

ggccttgcgc cttggctatt gcttggcagc gcctatcgcc aggtattact ccaatcccga 12480

atatccgaga tcgggatcac ccgagagaag ttcaacctac atcctcaatc ccgatctatc 12540

cgagatccga ggaatatcga aatcggggcg cgcctggtgt a 12581

<210> 108

<211> 12602

<212> DNA

<213> Artificial Sequence

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

<400> 108

tggccggcct tgcgccttgg ctattgcttg gcagcgccta tcgccaggta ttactccaat 60

cccgaatatc cgagatcggg atcacccgag agaagttcaa cctacatcct caatcccgat 120

ctatccgaga tccgaggaat atcgaaatcg gggcgcgcct ggtgtaccga gaacgatcct 180

ctcagtgcga gtctcgacga tccatatcgt tgcttggcag tcagccagtc ggaatccagc 240

ttgggaccca ggaagtccaa tcgtcagata ttgtactcaa gcctggtcac ggcagcgtac 300

cgatctgttt aaacctagat attgatagtc tgatcggtca acgtataatc gagtcctagc 360

ttttgcaaac atctatcaag agacaggatc agcaggaggc tttcgcatga gtattcaaca 420

tttccgtgtc gcccttattc cctttttgc ggcattttgc cttcctgttt ttgctcaccc 480

agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcgcgag tgggttacat 540

cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag aacgctttcc 600

aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgta ttgacgccgg 660

gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg agtattcacc 720

agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca gtgctgccat 780

aaccatgagt gataacactg cggccaactt acttctgaca acgattggag gaccgaagga 840

gctaaccgct tttttgcaca acatgggggga tcatgtaact cgccttgatc gttgggaacc 900

ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg tagcaatggc 960

aacaaccttg cgtaaactat taactggcga actacttact ctagcttccc ggcaacagtt 1020

gatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg cccttccggc 1080

tggctggttt attgctgata aatctggagc cggtgagcgt gggtctcgcg gtatcattgc 1140

agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga cggggagtca 1200

ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac tgattaagca 1260

ttggtaaccg attctaggtg cattggcgca gaaaaaaatg cctgatgcga cgctgcgcgt 1320

cttatactcc cacatatgcc agattcagca acggatacgg cttccccaac ttgcccactt 1380

ccatacgtgt cctccttacc agaaatttat ccttaagatc gtttaaactc gactctggct 1440

ctatcgaatc tccgtcgttt cgagcttacg cgaacagccg tggcgctcat ttgctcgtcg 1500

ggcatcgaat ctcgtcagct atcgtcagct tacctttttg gcagcgatcg cggctcccga 1560

catcttggac cattagctcc acaggtatct tcttccctct agtggtcata acagcagctt 1620

cagctacctc tcaattcaaa aaacccctca agacccgttt agaggcccca aggggttatg 1680

ctatcaatcg ttgcgttaca cacacaaaaa accaacacac atccatcttc gatggatagc 1740

gattttatta tctaactgct gatcgagtgt agccagatct agtaatcaat tacggggtca 1800

ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa tggcccgcct 1860

ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt tcccatagta 1920

acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac 1980

ttggcagtac atcaagtgta tcatatgcca agtacgcccc ctattgacgt caatgacggt 2040

aaatggcccg cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag 2100

tacatctacg tattagtcat cgctattacc atgctgatgc ggttttggca gtacatcaat 2160

gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat 2220

gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc 2280

ccattgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag cagagctggt 2340

ttagtgaacc gtcagatcag atctttgtcg atcctaccat ccactcgaca cacccgccag 2400

cggccgctaa tacgactcac tatagggaga agtactgcca ccatgggcaa gaagcaaaat 2460

cgcaagacgg ggaattccaa gacacaatcc gctagcccac cacctaaaga gcgttctagc 2520

tcccctgcta ctgagcagtc ctggatggaa aacgacttcg atgaactccg ggaagaggga 2580

tttaggcgat ccaactattc agaactccgc gaagatatcc agacaaaggg gaaggaagtc 2640

gagaatttcg agaagaacct cgaggagtgc atcacccgta tcacaaacac tgagaaatgt 2700

ctcaaagaac tcatggaact taagacaaaa gccagggagc ttcgagagga gtgtcggagt 2760

ctgagatcca ggtgtgacca gctcgaggag cgcgtgagcg cgatggaaga cgagatgaac 2820

gagatgaaaa gagagggcaa attcagggag aagcgcatta agaggaacga acagagtctg 2880

caggagattt gggattacgt caagaggcct aacctgcggt tgatcggcgt ccccgagagc 2940

gacgtagaaa acgggactaa actggagaat acacttcaag acatcattca agaaaatttt 3000

ccaaacctgg ctcggcaagc taatgtgcaa atccaagaga tccaacgcac accccagcgg 3060

tatagctctc ggcgtgccac ccctaggcat attatcgtgc gctttactaa ggtggagatg 3120

aaagagaaga tgctgcgagc cgctcgggaa aagggaaggg tgactttgaa gggcaaacct 3180

attcggctga cggttgacct tagcgccgag acactccagg cacgccggga atggggcccc 3240

atctttaata tcctgaagga gaagaacttc cagccacgaa tctcttaccc tgcaaagttg 3300
agttttatct ccgagggtga gattaagtat ttcatcgata aacagatgct gcgagacttc 3360
gtgacaactc gcccagctct caaggaactg ctcaaagagg ctcttaatat ggagcgcaat 3420
aatagatatc aacccttgca gaaccacgca aagatgtgag acagccgtca gaccatcaag 3480
actaggaaga aactgcatca actaatgagc aaaatcacca gctaacatca tagtatacat 3540
gccaaagaag aagcggaagg tcggcggcgg cagcaccggc tctaactcac atatcaccat 3600
ccttacactt aacattaacg gcctcaactc agctatcaag cgccatcggc tggccagctg 3660

gatcaaatca caggatccaa gcgtttgttg catccaagag acccacctga cctgtagaga 3720
tactcaccgc ctcaagatca agggatggcg aaagatttat caggcgaacg gtaagcagaa 3780
gaaagccgga gtcgcaattc tggtctcaga caagacggat ttcaagccca ccaaaattaa 3840
gcgtgataag gaaggtcact atattatggt gaaaggcagc atacagcagg aagaacttac 3900
catattgaac atctacgcgc caaacaccgg cgcacctcgc tttatcaaac aggtcctgtc 3960
cgatctgcag cgagatctgg attctcatac gttgattatg ggtgatttca atacaccatt 4020
gagcaccctg gatcgcagca ccaggcaaaa ggtaaataaa gacacgcaag agctcaatag 4080

cgcactgcat caggcagatc tcattgatat ttatcgcact cttcatccta agagtaccga 4140
gtacacattc ttcagcgccc cacatcatac atactcaaag atcgatcata tcgtcggctc 4200
aaaggctctg ctgtcaaagt gcaagcgcac agagataatt acaaattacc tgtcagatca 4260
tagcgcgatc aagctcgagc tgagaatcaa gaacctgacc cagagccgga gtaccacttg 4320
gaagcttaat aacctgctgc tcaacgatta ttgggtccac aatgagatga aggcagagat 4380
taaaatgttc ttcgaaacaa atgagaataa ggatactacc tatcaaaacc tttgggatgc 4440
ctttaaggcc gtctgcagag gcaagttcat cgccctcaac gcctataaaa gaaaacaaga 4500

gagatctaag atcgatactc tcacctctca gctgaaggag ttggagaaac aggaacagac 4560
ccactccaag gcgtcaagac ggcaggagat cacaaagatt cgcgccgagt tgaaagagat 4620
cgaaacccaa aagactcttc agaaaattaa cgagtctcgt agttggttct tcgagcggat 4680
taataagata gacagacctc tggcacgact gattaagaag aagcgcgaaa agaaccagat 4740
tgataccatc aagaacgaca agggcgacat cactactgac ccgaccgaga tccagaccac 4800
tattcgggag tattataagc atttgtatgc taacaagctt gagaacctgg aagagatgga 4860
cacttttctg gatacctata ctctgccacg gcttaatcaa gaggaagtcg agtccctcaa 4920

ccgcccaatt acaggaagcg agattgtggc cataattaac tccctgccga caaagaaatc 4980

tcctggtccg gacgggttta cagctgagtt ttatcaacgg tatatggaag agcttgtacc 5040

gtttctgctc aagctctttc agtctataga aaaggaaggc atcttgccca attccttcta 5100

cgaagcttct ataatactta ttcccaaacc aggacgcgat accacaaaga aggaaaactt 5160

ccggcccatt agtctcatga atatcgacgc taaaatattg aacaagattc tcgccaacag 5220

aatccaacaa catattaaga aattgataca tcacgaccag gtggggttta tacctggcat 5280

gcagggctgg tttaacatcc ggaagagtat taacgtcatt caacacatta atagagctaa 5340

ggataagaat catatgatca tctctataga cgcggaaaag gcattcgata agattcagca 5400

gccatttatg ctcaagactc tgaacaaact cggcatcgac ggaacatatt ttaagattat 5460

tcgcgcaatt tacgataagc cgactgctaa cattatcctt aacggccaaa agctcgaggc 5520

ctttccgctc aagactggaa cccgccaagg ctgtcccctc tccccgcttt tgtttaatat 5580

tgtactcgag gtgctggcta gggctattcg tcaagagaaa gagattaaag ggatacagct 5640

cgggaaggaa gaggtcaagc tttccttgtt cgccgatgat atgattgtgt acctggagaa 5700

tcctattgtg tctgctcaga accttcttaa acttatttct aactttagca aggtcagcgg 5760

ctataagatt aacgtccaga aatctcaggc ctttctgtac acaaataatc gacagaccga 5820

atcccagata atgggtgagc ttccgtttgt catagccagc aaaaggataa agtatctcgg 5880

aatccagctg acacgagacg ttaaagattt gtttaaggaa aattacaagc ctctcctgaa 5940

agagattaag gaagatacta ataagtggaa gaatatcccc tgttcatggg ttggcagaat 6000

caacatagtg aagatggcaa tacttcctaa agtgatatat cgctttaacg ccatcccaat 6060

taaactgcct atgaccttct ttacggagct cgagaaaaca acccttaaat ttatatggaa 6120

tcaaaagaga gcaagaatag cgaagtccat cttgagccag aagaataagg ccggtgggat 6180

tactttgcct gattttaagt tgtattataa agccacagta actaagacag cctggtattg 6240

gtatcagaat agagacatcg accagtggaa tcggaccgaa ccatcagaga taatgcccca 6300

catctataat taccttatat tcgataagcc agaaaagaat aaacagtggg gcaaagacag 6360

cctcttcaac aagtggtgtt gggagaattg gctggccata tgccggaaac tcaagctcga 6420

ccccttttctt acaccctaca ctaaaatcaa cagtaggtgg atcaaggact tgaatgtcaa 6480

gccaaagact ataaagacac tggaagagaa tcttgggatc acaatacaag atataggcgt 6540

cggcaaagat tttatgtcaa agacgcccaa ggccatggcc actaaggata agattgataa 6600

gtgggacctt attaagctca aaagcttctg tactgccaag gagaccacga tcagagttaa 6660

taggcagccc actacatggg aaaagatttt cgccacttat tcatcagata aggggttgat 6720

aagcagaata tataacgagc tgaagcagat ctacaagaag aaaacgaata atcccatcaa 6780

gaagtgggca aaagatatga acaggcattt tagcaaagag gatatctacg ccgcgaagaa 6840

gcatatgaag aagtgtagtt caagcttggc cattcgtgag atgcagatta agacgaccat 6900

gcgataccac cttaccccag tgaggatggc aattatcaag aaatctggca ataatagatg 6960

ttggcggggc tgtggcgaga ttggcaccct gctccattgc tggtgggatt gcaagctggt 7020

gcagccgctt tggaaatcag tctggcgctt tctgagggac ctcgagcttg agattccctt 7080

cgatcccgca attcccttgc tcggaatcta tcctaacgaa tacaagagct gttgttacaa 7140

ggatacgtgt acccggatgt tcatcgcggc cttgtttacg atagctaaga cgtggaatca 7200

gcctaagtgc cccacaatga tcgattggat caagaaaatg tggcatattt ataccatgga 7260

gtattacgca gcaattaaga atgacgaatt tatttccttc gttgggacct ggatgaagct 7320

ggagactatt attctgagca agctgtctca ggagcaaaag acaaagcata gaatcttctc 7380

tctcattggt ggtaacgact acaaagacga tgacgacaag aaaaggccgg cggccacgaa 7440

aaaggccggc caggcaaaaa agaaaaagta aagcgcttct agaagttgtc tcctcctgca 7500

ctgactgact gatacaatcg atttctggat ccgcaggcct aatcaacctc tggattacaa 7560

aatttgtgaa agattgactg gtattcttaa ctatgttgct ccttttacgc tatgtggata 7620

cgctgcttta atgcctttgt atcatgctat tgcttcccgt atggctttca ttttctcctc 7680

cttgtataaa tcctggttgc tgtctcttta tgaggagttg tggcccgttg tcaggcaacg 7740

tggcgtggtg tgcactgtgt ttgctgacgc aacccccact ggttggggca ttgccaccac 7800

ctgtcagctc ctttccggga ctttcgcttt ccccctccct attgccacgg cggaactcat 7860

cgccgcctgc cttgcccgct gctggacagg ggctcggctg ttgggcactg acaattccgt 7920

ggtgttgtcg gggaagctga cgtcctttcc atggctgctc gcctgtgttg ccacctggat 7980

tctgcgcggg acgtccttct gctacgtccc ttcggccctc aatccagcgg accttccttc 8040

ccgctgagag acacaaaaaa ttccaacaca ctattgcaat gaaaataaat ttcctttatt 8100

agccagaagt cagatgctca aggggcttca tgatgtcccc ataatttttg gcagagggaa 8160

aaagatctca gtggtatttg tgagccaggg cattggcctt ctgataggca gcctgcacct 8220

gaggagtgcg gccgctttac ttgtacagct cgtccatgcc gagagtgatc ccggcggcgg 8280

tcacgaactc cagcaggacc atgtgatcgc gcttctcgtt ggggtctttg ctcagggcgg 8340

actgggtgct caggtagtgg ttgtcgggca gcagcacggg gccgtcgccg atgggggtgt 8400

tctgctggta gtggtcggcg agctgcacgc tgccgtcctc gatgttgtgg cggatcttga 8460

agttcacctt gatgccgttc ttctgcttgt cggccatgat atagacgttg tggctgttgt 8520

agttgtactc cagcttgtgc cccaggatgt tgccgtcctc cttgaagtcg atgcccttca 8580

gctcgatgcg gttcaccagg gtgtcgccct cgaacttcac ctcggcgcgg gtcttgtagt 8640

tgccgtcgtc cttgaagaag atggtgcgct cctggacgta gccttcgggc atggcggact 8700

tgaagaagtc gtgctgcttc atgtggtcgg ggtagcggct gaagcactgc acgccgtagg 8760
tcagggtggt cacgagggtg ggccagggca cgggcagctt gccggtggtg cagatgaact 8820
tcagggtcag cttgccgtag gtggcatcgc cctcgccctc gccggacacg ctgaacttgt 8880
ggccgtttac gtcgccgtcc agctcgacca ggatgggcac caccccggtg aacagctcct 8940
cgcccttgct caccatggtg gcgggatctg acggttcact aaaccagctc tgcttatata 9000
gacctcccac cgtacacgcc taccgcccat ttgcgtcaat ggggcggagt tgttacgaca 9060
ttttggaaag tcccgttgat tttggtgcca aaacaaactc ccattgacgt caatggggtg 9120

gagacttgga aatccccgtg agtcaaaccg ctatccacgc ccattgatgt actgccaaaa 9180
ccgcatcacc atggtaatag cgatgactaa tacgtagatg tactgccaag taggaaagtc 9240
ccataaggtc atgtactggg cataatgcca ggcgggccat ttaccgtcat tgacgtcaat 9300
aggggggcgta cttggcatat gatacacttg atgtactgcc aagtgggcag tttaccgtaa 9360
atactccacc cattgacgtc aatggaaagt ccctattggc gttactatgg gaacatacgt 9420
cattattgac gtcaatgggc gggggtcgtt gggcggtcag ccaggcgggc catttaccgt 9480
aagttatgta acgggcctgc tgccggctct gcggcctctt ccgcgtcttc gccttcgccc 9540

tcagacgagt cggatctccc tttgggccgc ctccccgcct gtctagcttg actgactgag 9600
atacagcgta ccttcagctc acagacatga taagatacat tgatgagttt ggacaaacca 9660
caactagaat gcagtgaaaa aaatgcttta tttgtgaaat ttgtgatgct attgctttat 9720
ttgtaaccat tataagctgc aataaacaag ttaacaacaa caattgcatt cattttatgt 9780
ttcaggttca gggggaggtg tgggaggttt tttaaagcaa gtaaaacctc tacaaatgtg 9840
gtattggccc atctctatcg gtatcgtagc ataacccctt ggggcctcta aacgggtctt 9900
gaggggtttt ttgtgcccct cgggccggat tgctatctac cggcattggc gcagaaaaaa 9960

atgcctgatg cgacgctgcg cgtcttatac tcccacatat gccagattca gcaacggata 10020
cggcttcccc aacttgccca cttccatacg tgtcctcctt accagaaatt tatccttaag 10080
gtcgtcagct atcctgcagg cgatctctcg atttcgatca agacattcct ttaatggtct 10140
tttctggaca ccactagggg tcagaagtag ttcatcaaac tttcttccct ccctaatctc 10200
attggttacc ttgggctatc gaaacttaat taagcgatct gcatctcaat tagtcagcaa 10260
ccatagtccc gcccctaact ccgcccatcc cgcccctaac tccgcccagt tccgcccatt 10320
ctccgcccca tcgctgacta attttttta tttatgcaga ggccgaggcc gcctcggcct 10380

ctgagctatt ccagaagtag tgaggaggct tttttggagg cctaggcttt tgcaaaggag 10440

gtagccaaca tgattgaaca agatggattg cacgcaggtt ctcccgccgc ttgggtggag 10500

aggctattcg gctatgactg ggcacaacag acaatcggct gctctgatgc cgccgtgttc 10560

cggctgtcag cgcaggggcg cccggttctt tttgtcaaga ccgacctgtc cggtgccctg 10620

aatgaactcc aggacgaggc agcgcggcta tcgtggctgg ccacgacggg cgttccttgc 10680

gcagctgtgc tcgacgttgt cactgaagcg ggaagggact ggctgctatt gggcgaagtg 10740

ccggggcagg atctcctgtc atctcacctt gctcctgccg agaaagtatc catcatggct 10800

gatgcaatgc ggcggctgca tacgcttgat ccggctacct gcccattcga ccaccaagcg 10860

aaacatcgca tcgagcgagc acgtactcgg atggaagccg gtcttgtcga tcaggatgat 10920

ctggacgaag agcatcaggg gctcgcgcca gccgaactgt tcgccaggct caaggcgcgg 10980

atgcccgacg gcgaggatct cgtcgtgacc cacggcgatg cctgcttgcc gaatatcatg 11040

gtggaaaatg gccgcttttc tggattcatc gactgtggcc ggctgggtgt ggcggaccgc 11100

tatcaggaca tagcgttggc tacccgtgat attgctgaag agcttggcgg cgaatgggct 11160

gaccgcttcc tcgtgcttta cggtatcgcc gctcccgatt cgcagcgcat cgccttctat 11220

cgccttcttg acgagttctt ctagtatgta agccctgtgc cttctagttg ccagccatct 11280

gttgtttgcc cctcccccgt gccttccttg accctggaag gtgccactcc cactgtcctt 11340

tcctaataaa atgaggaaat tgcatcgcat tgtctgagta ggtgtcattc tattctgggg 11400

ggtggggtgg ggcaggacag caagggggag gattgggaag acaatagcag gcatgctggg 11460

gatgcggtgg gctctatggt taattaacca gtcaagtcag ctacttggcg agatcgactt 11520

gtctgggttt cgactacgct cagaattgcg tcagtcaagt tcgatctggt ccttgctatt 11580

gcacccgttc tccgattacg agtttcattt aaatcatgtg agcaaaaggc cagcaaaagg 11640

ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc ccccctgacg 11700

agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 11760

accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta 11820

ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct 11880

gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc 11940

ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa 12000

gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 12060

taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag 12120

tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt 12180

gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 12240

cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 12300

agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca 12360

cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa 12420

cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat 12480

ttcgttcatc catagttgca tttaaatttc cgaactctcc aaggccctcg tcggaaaatc 12540

ttcaaacctt tcgtccgatc catcttgcag gctacctctc gaacgaacta tcgcaagtct 12600

ct 12602

<210> 109

<211> 12541

<212> DNA

<213> Artificial Sequence

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

<400> 109

tggccggcct tgcgccttgg ctattgcttg gcagcgccta tcgccaggta ttactccaat 60

cccgaatatc cgagatcggg atcacccgag agaagttcaa cctacatcct caatcccgat 120

ctatccgaga tccgaggaat atcgaaatcg gggcgcgcct ggtgtaccga gaacgatcct 180

ctcagtgcga gtctcgacga tccatatcgt tgcttggcag tcagccagtc ggaatccagc 240

ttgggaccca ggaagtccaa tcgtcagata ttgtactcaa gcctggtcac ggcagcgtac 300

cgatctgttt aaacctagat attgatagtc tgatcggtca acgtataatc gagtcctagc 360

ttttgcaaac atctatcaag agacaggatc agcaggaggc tttcgcatga gtattcaaca 420

tttccgtgtc gcccttattc cctttttgc ggcattttgc cttcctgttt ttgctcaccc 480

agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcgcgag tgggttacat 540

cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag aacgctttcc 600

aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgta ttgacgccgg 660

gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg agtattcacc 720

agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca gtgctgccat 780

aaccatgagt gataacactg cggccaactt acttctgaca acgattggag gaccgaagga 840

gctaaccgct tttttgcaca acatgggggga tcatgtaact cgccttgatc gttgggaacc 900

ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg tagcaatggc 960

aacaaccttg cgtaaactat taactggcga actacttact ctagcttccc ggcaacagtt 1020

gatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg cccttccggc 1080

tggctggttt attgctgata aatctggagc cggtgagcgt gggtctcgcg gtatcattgc 1140

agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga cggggagtca 1200

ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac tgattaagca 1260

ttggtaaccg attctaggtg cattggcgca gaaaaaaatg cctgatgcga cgctgcgcgt 1320

cttatactcc cacatatgcc agattcagca acggatacgg cttccccaac ttgcccactt 1380

ccatacgtgt cctccttacc agaaatttat ccttaagatc gtttaaactc gactctggct 1440

ctatcgaatc tccgtcgttt cgagcttacg cgaacagccg tggcgctcat ttgctcgtcg 1500

ggcatcgaat ctcgtcagct atcgtcagct tacctttttg gcagcgatcg cggctcccga 1560

catcttggac cattagctcc acaggtatct tcttccctct agtggtcata acagcagctt 1620

cagctacctc tcaattcaaa aaacccctca agacccgttt agaggcccca aggggttatg 1680

ctatcaatcg ttgcgttaca cacacaaaaa accaacacac atccatcttc gatggatagc 1740

gattttatta tctaactgct gatcgagtgt agccagatct agtaatcaat tacggggtca 1800

ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa tggcccgcct 1860

ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt tcccatagta 1920

acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac 1980

ttggcagtac atcaagtgta tcatatgcca agtacgcccc ctattgacgt caatgacggt 2040

aaatggcccg cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag 2100

tacatctacg tattagtcat cgctattacc atgctgatgc ggttttggca gtacatcaat 2160

gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat 2220

gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc 2280

ccattgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag cagagctggt 2340

ttagtgaacc gtcagatcag atctttgtcg atcctaccat ccactcgaca cacccgccag 2400

cggccgctaa tacgactcac tatagggaga agtactgcca ccatgggcaa gaagcaaaat 2460

cgcaagacgg ggaattccaa gacacaatcc gctagcccac cacctaaaga gcgttctagc 2520

tcccctgcta ctgagcagtc ctggatggaa aacgacttcg atgaactccg ggaagaggga 2580

tttaggcgat ccaactattc agaactccgc gaagatatcc agacaaaggg gaaggaagtc 2640

gagaatttcg agaagaacct cgaggagtgc atcacccgta tcacaaacac tgagaaatgt 2700

ctcaaagaac tcatggaact taagacaaaa gccagggagc ttcgagagga gtgtcggagt 2760

ctgagatcca ggtgtgacca gctcgaggag cgcgtgagcg cgatggaaga cgagatgaac 2820

gagatgaaaa gagagggcaa attcagggag aagcgcatta agaggaacga acagagtctg 2880

caggagattt gggattacgt caagaggcct aacctgcggt tgatcggcgt ccccgagagc 2940

gacgtagaaa acgggactaa actggagaat acacttcaag acatcattca agaaaatttt 3000

ccaaacctgg ctcggcaagc taatgtgcaa atccaagaga tccaacgcac accccagcgg 3060

tatagctctc ggcgtgccac ccctaggcat attatcgtgc gctttactaa ggtggagatg 3120

aaagagaaga tgctgcgagc cgctcgggaa aagggaaggg tgactttgaa gggcaaacct 3180

attcggctga cggttgacct tagcgccgag acactccagg cacgccggga atggggcccc 3240

atctttaata tcctgaagga gaagaacttc cagccacgaa tctcttaccc tgcaaagttg 3300

agttttatct ccgagggtga gattaagtat ttcatcgata aacagatgct gcgagacttc 3360

gtgacaactc gcccagctct caaggaactg ctcaaagagg ctcttaatat ggagcgcaat 3420

aatagatatc aacccttgca gaaccacgca aagatgggct ccggcgaggg caggggaagc 3480

cttctaacat gcggggacgt ggaggaaaat cccggcccag gtagcggccc aaagaagaag 3540

cggaaggtcg gcggcggcag caccggctct aactcacata tcaccatcct tacacttaac 3600

attaacggcc tcaactcagc tatcaagcgc catcggctgg ccagctggat caaatcacag 3660

gatccaagcg tttgttgcat ccaagagacc cacctgacct gtagagatac tcaccgcctc 3720

aagatcaagg gatggcgaaa gatttatcag gcgaacggta agcagaagaa agccggagtc 3780

gcaattctgg tctcagacaa gacggatttc aagcccacca aaattaagcg tgataaggaa 3840

ggtcactata ttatggtgaa aggcagcata cagcaggaag aacttaccat attgaacatc 3900

tacgcgccaa acaccggcgc acctcgcttt atcaaacagg tcctgtccga tctgcagcga 3960

gatctggatt ctcatacgtt gattatgggt gatttcaata caccattgag caccctggat 4020

cgcagcacca ggcaaaaggt aaataaagac acgcaagagc tcaatagcgc actgcatcag 4080

gcagatctca ttgatattta tcgcactctt catcctaaga gtaccgagta cacattcttc 4140

agcgccccac atcatacata ctcaaagatc gatcatatcg tcggctcaaa ggctctgctg 4200

tcaaagtgca agcgcacaga gataattaca aattacctgt cagatcatag cgcgatcaag 4260

ctcgagctga gaatcaagaa cctgacccag agccggagta ccacttggaa gcttaataac 4320

ctgctgctca acgattattg ggtccacaat gagatgaagg cagagattaa aatgttcttc 4380

gaaacaaatg agaataagga tactacctat caaaaccttt gggatgcctt taaggccgtc 4440

tgcagaggca agttcatcgc cctcaacgcc tataaaagaa aacaagagag atctaagatc 4500

gatactctca cctctcagct gaaggagttg gagaaacagg aacagaccca ctccaaggcg 4560

tcaagacggc aggagatcac aaagattcgc gccgagttga aagagatcga aacccaaaag 4620

actcttcaga aaattaacga gtctcgtagt tggttcttcg agcggattaa taagatagac 4680

agacctctgg cacgactgat taagaagaag cgcgaaaaga accagattga taccatcaag 4740

aacgacaagg gcgacatcac tactgacccg accgagatcc agaccactat tcgggagtat 4800
tataagcatt tgtatgctaa caagcttgag aacctggaag agatggacac ttttctggat 4860
acctatactc tgccacggct taatcaagag gaagtcgagt ccctcaaccg cccaattaca 4920
ggaagcgaga ttgtggccat aattaactcc ctgccgacaa agaaatctcc tggtccggac 4980
gggtttacag ctgagtttta tcaacggtat atggaagagc ttgtaccgtt tctgctcaag 5040
ctctttcagt ctatagaaaa ggaaggcatc ttgcccaatt ccttctacga agcttctata 5100
atacttattc ccaaaccagg acgcgatacc acaaagaagg aaaacttccg gcccattagt 5160

ctcatgaata tcgacgctaa aatattgaac aagattctcg ccaacagaat ccaacaacat 5220
attaagaaat tgatacatca cgaccaggtg gggtttatac ctggcatgca gggctggttt 5280
aacatccgga agagtattaa cgtcattcaa cacattaata gagctaagga taagaatcat 5340
atgatcatct ctatagacgc ggaaaaggca ttcgataaga ttcagcagcc atttatgctc 5400
aagactctga acaaactcgg catcgacgga acatatttta agattattcg cgcaatttac 5460
gataagccga ctgctaacat tatccttaac ggccaaaagc tcgaggcctt tccgctcaag 5520
actggaaccc gccaaggctg tcccctctcc ccgcttttgt ttaatattgt actcgaggtg 5580

ctggctaggg ctattcgtca agagaaagag attaaaggga tacagctcgg gaaggaagag 5640
gtcaagcttt ccttgttcgc cgatgatatg attgtgtacc tggagaatcc tattgtgtct 5700
gctcagaacc ttcttaaact tatttctaac tttagcaagg tcagcggcta taagattaac 5760
gtccagaaat ctcaggcctt tctgtacaca aataatcgac agaccgaatc ccagataatg 5820
ggtgagcttc cgtttgtcat agccagcaaa aggataaagt atctcggaat ccagctgaca 5880
cgagacgtta aagatttgtt taaggaaaat tacaagcctc tcctgaaaga gattaaggaa 5940
gatactaata agtggaagaa tatcccctgt tcatgggttg gcagaatcaa catagtgaag 6000

atggcaatac ttcctaaagt gatatatcgc tttaacgcca tcccaattaa actgcctatg 6060
accttcttta cggagctcga gaaaacaacc cttaaattta tatggaatca aaagagagca 6120
agaatagcga agtccatctt gagccagaag aataaggccg gtgggattac tttgcctgat 6180
tttaagttgt attataaagc cacagtaact aagacagcct ggtattggta tcagaataga 6240
gacatcgacc agtggaatcg gaccgaacca tcagagataa tgccccacat ctataattac 6300
cttatattcg ataagccaga aaagaataaa cagtggggca aagacagcct cttcaacaag 6360
tggtgttggg agaattggct ggccatatgc cggaaactca agctcgaccc ctttcttaca 6420

ccctacacta aaatcaacag taggtggatc aaggacttga atgtcaagcc aaagactata 6480

aagacactgg aagagaatct tgggatcaca atacaagata taggcgtcgg caaagatttt 6540

atgtcaaaga cgcccaaggc catggccact aaggataaga ttgataagtg ggaccttatt 6600

aagctcaaaa gcttctgtac tgccaaggag accacgatca gagttaatag gcagcccact 6660

acatgggaaa agattttcgc cacttattca tcagataagg ggttgataag cagaatatat 6720

aacgagctga agcagatcta caagaagaaa acgaataatc ccatcaagaa gtgggcaaaa 6780

gatatgaaca ggcattttag caaagaggat atctacgccg cgaagaagca tatgaagaag 6840

tgtagttcaa gcttggccat tcgtgagatg cagattaaga cgaccatgcg ataccacctt 6900

accccagtga ggatggcaat tatcaagaaa tctggcaata atagatgttg gcggggctgt 6960

ggcgagattg gcaccctgct ccattgctgg tgggattgca agctggtgca gccgctttgg 7020

aaatcagtct ggcgctttct gagggacctc gagcttgaga ttcccttcga tcccgcaatt 7080

cccttgctcg gaatctatcc taacgaatac aagagctgtt gttacaagga tacgtgtacc 7140

cggatgttca tcgcggcctt gtttacgata gctaagacgt ggaatcagcc taagtgcccc 7200

acaatgatcg attggatcaa gaaaatgtgg catatttata ccatggagta ttacgcagca 7260

attaagaatg acgaatttat ttccttcgtt gggacctgga tgaagctgga gactattatt 7320

ctgagcaagc tgtctcagga gcaaaagaca aagcatagaa tcttctctct cattggtggt 7380

aacgactaca aagacgatga cgacaagtaa agcgcttcta gaagttgtct cctcctgcac 7440

tgactgactg atacaatcga tttctggatc cgcaggccta atcaacctct ggattacaaa 7500

atttgtgaaa gattgactgg tattcttaac tatgttgctc cttttacgct atgtggatac 7560

gctgctttaa tgcctttgta tcatgctatt gcttcccgta tggctttcat tttctcctcc 7620

ttgtataaat cctggttgct gtctctttat gaggagttgt ggcccgttgt caggcaacgt 7680

ggcgtggtgt gcactgtgtt tgctgacgca acccccactg gttggggcat tgccaccacc 7740

tgtcagctcc tttccgggac tttcgctttc cccctcccta ttgccacggc ggaactcatc 7800

gccgcctgcc ttgcccgctg ctggacaggg gctcggctgt tgggcactga caattccgtg 7860

gtgttgtcgg ggaagctgac gtcctttcca tggctgctcg cctgtgttgc cacctggatt 7920

ctgcgcggga cgtccttctg ctacgtccct tcggccctca atccagcgga ccttccttcc 7980

cgctgagaga cacaaaaaat tccaacacac tattgcaatg aaaataaatt tcctttatta 8040

gccagaagtc agatgctcaa ggggcttcat gatgtcccca taatttttgg cagagggaaa 8100

aagatctcag tggtatttgt gagccagggc attggccttc tgataggcag cctgcacctg 8160

aggagtgcgg ccgctttact tgtacagctc gtccatgccg agagtgatcc cggcggcggt 8220

cacgaactcc agcaggacca tgtgatcgcg cttctcgttg gggtctttgc tcagggcgga 8280

ctgggtgctc aggtagtggt tgtcgggcag cagcacgggg ccgtcgccga tgggggtgtt 8340

ctgctggtag tggtcggcga gctgcacgct gccgtcctcg atgttgtggc ggatcttgaa 8400

gttcaccttg atgccgttct tctgcttgtc ggccatgata tagacgttgt ggctgttgta 8460

gttgtactcc agcttgtgcc ccaggatgtt gccgtcctcc ttgaagtcga tgcccttcag 8520

ctcgatgcgg ttcaccaggg tgtcgccctc gaacttcacc tcggcgcggg tcttgtagtt 8580

gccgtcgtcc ttgaagaaga tggtgcgctc ctggacgtag ccttcgggca tggcggactt 8640

gaagaagtcg tgctgcttca tgtggtcggg gtagcggctg aagcactgca cgccgtaggt 8700

cagggtggtc acgagggtgg gccagggcac gggcagcttg ccggtggtgc agatgaactt 8760

cagggtcagc ttgccgtagg tggcatcgcc ctcgccctcg ccggacacgc tgaacttgtg 8820

gccgtttacg tcgccgtcca gctcgaccag gatgggcacc accccggtga acagctcctc 8880

gcccttgctc accatggtgg cgggatctga cggttcacta aaccagctct gcttatatag 8940

acctcccacc gtacacgcct accgcccatt tgcgtcaatg gggcggagtt gttacgacat 9000

tttggaaagt cccgttgatt ttggtgccaa aacaaactcc cattgacgtc aatggggtgg 9060

agacttggaa atccccgtga gtcaaaccgc tatccacgcc cattgatgta ctgccaaaac 9120

cgcatcacca tggtaatagc gatgactaat acgtagatgt actgccaagt aggaaagtcc 9180

cataaggtca tgtactgggc ataatgccag gcgggccatt taccgtcatt gacgtcaata 9240

gggggcgtac ttggcatatg atacacttga tgtactgcca agtgggcagt ttaccgtaaa 9300

tactccaccc attgacgtca atggaaagtc cctattggcg ttactatggg aacatacgtc 9360

attattgacg tcaatgggcg ggggtcgttg ggcggtcagc caggcgggcc atttaccgta 9420

agttatgtaa cgggcctgct gccggctctg cggcctcttc cgcgtcttcg ccttcgccct 9480

cagacgagtc ggatctccct ttgggccgcc tccccgcctg tctagcttga ctgactgaga 9540

tacagcgtac cttcagctca cagacatgat aagatacatt gatgagtttg gacaaaccac 9600

aactagaatg cagtgaaaaa aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt 9660

tgtaaccatt ataagctgca ataaacaagt taacaacaac aattgcattc attttatgtt 9720

tcaggttcag ggggaggtgt gggaggtttt ttaaagcaag taaaacctct acaaatgtgg 9780

tattggccca tctctatcgg tatcgtagca taacccccttg gggcctctaa acgggtcttg 9840

aggggttttt tgtgcccctc gggccggatt gctatctacc ggcattggcg cagaaaaaaa 9900

tgcctgatgc gacgctgcgc gtcttatact cccacatatg ccagattcag caacggatac 9960

ggcttcccca acttgcccac ttccatacgt gtcctcctta ccagaaattt atccttaagg 10020

tcgtcagcta tcctgcaggc gatctctcga tttcgatcaa gacattcctt taatggtctt 10080

ttctggacac cactaggggt cagaagtagt tcatcaaact ttcttccctc cctaatctca 10140

ttggttacct tgggctatcg aaacttaatt aagcgatctg catctcaatt agtcagcaac 10200

catagtcccg cccctaactc cgcccatccc gcccctaact ccgcccagtt ccgcccattc 10260
tccgccccat cgctgactaa ttttttttat ttatgcagag gccgaggccg cctcggcctc 10320
tgagctattc cagaagtagt gaggaggctt ttttggaggc ctaggctttt gcaaaggagg 10380
tagccaacat gattgaacaa gatggattgc acgcaggttc tcccgccgct tgggtggaga 10440
ggctattcgg ctatgactgg gcacaacaga caatcggctg ctctgatgcc gccgtgttcc 10500
ggctgtcagc gcaggggcgc ccggttcttt ttgtcaagac cgacctgtcc ggtgccctga 10560
atgaactcca ggacgaggca gcgcggctat cgtggctggc cacgacgggc gttccttgcg 10620

cagctgtgct cgacgttgtc actgaagcgg gaagggactg gctgctattg ggcgaagtgc 10680
cggggcagga tctcctgtca tctcaccttg ctcctgccga gaaagtatcc atcatggctg 10740
atgcaatgcg gcggctgcat acgcttgatc cggctacctg cccattcgac caccaagcga 10800
aacatcgcat cgagcgagca cgtactcgga tggaagccgg tcttgtcgat caggatgatc 10860
tggacgaaga gcatcagggg ctcgcgccag ccgaactgtt cgccaggctc aaggcgcgga 10920
tgcccgacgg cgaggatctc gtcgtgaccc acggcgatgc ctgcttgccg aatatcatgg 10980
tggaaaatgg ccgcttttct ggattcatcg actgtggccg gctgggtgtg gcggaccgct 11040

atcaggacat agcgttggct acccgtgata ttgctgaaga gcttggcggc gaatgggctg 11100
accgcttcct cgtgctttac ggtatcgccg ctcccgattc gcagcgcatc gccttctatc 11160
gccttcttga cgagttcttc tagtatgtaa gccctgtgcc ttctagttgc cagccatctg 11220
ttgtttgccc ctccccccgtg ccttccttga ccctggaagg tgccactccc actgtccttt 11280
cctaataaaa tgaggaaatt gcatcgcatt gtctgagtag gtgtcattct attctggggg 11340
gtggggtggg gcaggacagc aaggggggagg attgggaaga caatagcagg catgctgggg 11400
atgcggtggg ctctatggtt aattaaccag tcaagtcagc tacttggcga gatcgacttg 11460

tctgggtttc gactacgctc agaattgcgt cagtcaagtt cgatctggtc cttgctattg 11520
cacccgttct ccgattacga gtttcattta aatcatgtga gcaaaaggcc agcaaaaggc 11580
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga 11640
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 11700
ccaggcgttt cccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac 11760
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 11820
taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 11880

cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag 11940

acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt 12000

aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaagaacagt 12060

atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 12120

atccggcaaa caaaccaccg ctggtagcgg tggtttttt gtttgcaagc agcagattac 12180

gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca 12240

gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 12300

ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 12360

ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 12420

tcgttcatcc atagttgcat ttaaatttcc gaactctcca aggccctcgt cggaaaatct 12480

tcaaaccttt cgtccgatcc atcttgcagg ctacctctcg aacgaactat cgcaagtctc 12540

t 12541

<210> 110

<211> 12556

<212> DNA

<213> Artificial Sequence

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

<400> 110

accgagaacg atcctctcag tgcgagtctc gacgatccat atcgttgctt ggcagtcagc 60

cagtcggaat ccagcttggg acccaggaag tccaatcgtc agatattgta ctcaagcctg 120

gtcacggcag cgtaccgatc tgtttaaacc tagatattga tagtctgatc ggtcaacgta 180

taatcgagtc ctagcttttg caaacatcta tcaagagaca ggatcagcag gaggctttcg 240

catgagtatt caacatttcc gtgtcgccct tattcccttt tttgcggcat tttgccttcc 300

tgtttttgct cacccagaaa cgctggtgaa agtaaaagat gctgaagatc agttgggtgc 360

gcgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga gttttcgccc 420

cgaagaacgc tttccaatga tgagcacttt taaagttctg ctatgtggcg cggtattatc 480

ccgtattgac gccgggcaag agcaactcgg tcgccgcata cactattctc agaatgactt 540

ggttgagtat tcaccagtca cagaaaagca tcttacggat ggcatgacag taagagaatt 600

atgcagtgct gccataacca tgagtgataa cactgcggcc aacttacttc tgacaacgat 660

tggaggaccg aaggagctaa ccgcttttt gcacaacatg ggggatcatg taactcgcct 720

tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg acaccacgat 780

gcctgtagca atggcaacaa ccttgcgtaa actattaact ggcgaactac ttactctagc 840

ttcccggcaa cagttgatag actggatgga ggcggataaa gttgcaggac cacttctgcg 900

ctcggccctt ccggctggct ggtttattgc tgataaatct ggagccggtg agcgtgggtc 960

tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg tagttatcta 1020

cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg agataggtgc 1080

ctcactgatt aagcattggt aaccgattct aggtgcattg gcgcagaaaa aaatgcctga 1140

tgcgacgctg cgcgtcttat actcccacat atgccagatt cagcaacgga tacggcttcc 1200

ccaacttgcc cacttccata cgtgtcctcc ttaccagaaa tttatcctta agatcgttta 1260

aactcgactc tggctctatc gaatctccgt cgtttcgagc ttacgcgaac agccgtggcg 1320

ctcatttgct cgtcgggcat cgaatctcgt cagctatcgt cagcttacct ttttggcagc 1380

gatcgcggct cccgacatct tggaccatta gctccacagg tatcttcttc cctctagtgg 1440

tcataacagc agcttcagct acctctcaat tcaaaaaacc cctcaagacc cgtttagagg 1500

ccccaagggg ttatgctatc aatcgttgcg ttacacacac aaaaaaccaa cacacatcca 1560

tcttcgatgg atagcgattt tattatctaa ctgctgatcg agtgtagcca gatctagtaa 1620

tcaattacgg ggtcattagt tcatagccca tatatggagt tccgcgttac ataacttacg 1680

gtaaatggcc cgcctggctg accgcccaac gacccccgcc cattgacgtc aataatgacg 1740

tatgttccca tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta 1800

cggtaaactg cccacttggc agtacatcaa gtgtatcata tgccaagtac gccccctatt 1860

gacgtcaatg acggtaaatg gcccgcctgg cattatgccc agtacatgac cttatgggac 1920

tttcctactt ggcagtacat ctacgtatta gtcatcgcta ttaccatgct gatgcggttt 1980

tggcagtaca tcaatgggcg tggatagcgg tttgactcac ggggatttcc aagtctccac 2040

cccattgacg tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt 2100

cgtaacaact ccgccccatt gacgcaaatg ggcggtaggc gtgtacggtg ggaggtctat 2160

ataagcagag ctggtttagt gaaccgtcag atcagatctt tgtcgatcct accatccact 2220

cgacacaccc gccagcggcc gctaatacga ctcactatag ggagaagtac tgccaccatg 2280

ggcaagaagc aaaatcgcaa gacggggaat tccaagacac aatccgctag cccaccacct 2340

aaagagcgtt ctagctcccc tgctactgag cagtcctgga tggaaaacga cttcgatgaa 2400

ctccgggaag agggatttag gcgatccaac tattcagaac tccgcgaaga tatccagaca 2460

aaggggaagg aagtcgagaa tttcgagaag aacctcgagg agtgcatcac ccgtatcaca 2520

aacactgaga aatgtctcaa agaactcatg gaacttaaga caaaagccag ggagcttcga 2580

gaggagtgtc ggagtctgag atccaggtgt gaccagctcg aggagcgcgt gagcgcgatg 2640

gaagacgaga tgaacgagat gaaaagagag ggcaaattca gggagaagcg cattaagagg 2700

aacgaacaga gtctgcagga gatttgggat tacgtcaaga ggcctaacct gcggttgatc 2760

ggcgtccccg agagcgacgt agaaaacggg actaaactgg agaatacact tcaagacatc 2820

attcaagaaa attttccaaa cctggctcgg caagctaatg tgcaaatcca agagatccaa 2880

cgcacacccc agcggtatag ctctcggcgt gccaccccta ggcatattat cgtgcgcttt 2940

actaaggtgg agatgaaaga gaagatgctg cgagccgctc gggaaaaggg aagggtgact 3000

ttgaagggca aacctattcg gctgacggtt gaccttagcg ccgagacact ccaggcacgc 3060

cgggaatggg gccccatctt taatatcctg aaggagaaga acttccagcc acgaatctct 3120

taccctgcaa agttgagttt tatctccgag ggtgagatta agtatttcat cgataaacag 3180

atgctgcgag acttcgtgac aactcgccca gctctcaagg aactgctcaa agaggctctt 3240

aatatggagc gcaataatag atatcaaccc ttgcagaacc acgcaaagat gggctccggc 3300

gagggcaggg gaagccttct aacatgcggg gacgtggagg aaaatcccgg cccaggtagc 3360

ggcaccggct ctaactcaca tatcaccatc cttacactta acattaacgg cctcaactca 3420

gctatcaagc gccatcggct ggccagctgg atcaaatcac aggatccaag cgtttgttgc 3480

atccaagaga cccacctgac ctgtagagat actcaccgcc tcaagatcaa gggatggcga 3540

aagatttatc aggcgaacgg taagcagaag aaagccggag tcgcaattct ggtctcagac 3600

aagacggatt tcaagcccac caaaattaag cgtgataagg aaggtcacta tattatggtg 3660

aaaggcagca tacagcagga agaacttacc atattgaaca tctacgcgcc aaacaccggc 3720

gcacctcgct ttatcaaaca ggtcctgtcc gatctgcagc gagatctgga ttctcatacg 3780

ttgattatgg gtgatttcaa tacaccattg agcaccctgg atcgcagcac caggcaaaag 3840

gtaaataaag acacgcaaga gctcaatagc gcactgcatc aggcagatct cattgatatt 3900

tatcgcactc ttcatcctaa gagtaccgag tacacattct tcagcgcccc acatcataca 3960

tactcaaaga tcgatcatat cgtcggctca aaggctctgc tgtcaaagtg caagcgcaca 4020

gagataatta caaattacct gtcagatcat agcgcgatca agctcgagct gagaatcaag 4080

aacctgaccc agagccggag taccacttgg aagcttaata acctgctgct caacgattat 4140

tgggtccaca atgagatgaa ggcagagatt aaaatgttct tcgaaacaaa tgagaataag 4200

gatactacct atcaaaacct ttgggatgcc tttaaggccg tctgcagagg caagttcatc 4260

gccctcaacg cctataaaag aaaacaagag agatctaaga tcgatactct cacctctcag 4320

ctgaaggagt tggagaaaca ggaacagacc cactccaagg cgtcaagacg gcaggagatc 4380

acaaagattc gcgccgagtt gaaagagatc gaaacccaaa agactcttca gaaaattaac 4440

gagtctcgta gttggttctt cgagcggatt aataagatag acagacctct ggcacgactg 4500

attaagaaga agcgcgaaaa gaaccagatt gataccatca agaacgacaa gggcgacatc 4560

actactgacc cgaccgagat ccagaccact attcgggagt attataagca tttgtatgct 4620

aacaagcttg agaacctgga agagatggac acttttctgg atacctatac tctgccacgg 4680

cttaatcaag aggaagtcga gtccctcaac cgcccaatta caggaagcga gattgtggcc 4740

ataattaact ccctgccgac aaagaaatct cctggtccgg acgggtttac agctgagttt 4800

tatcaacggt atatggaaga gcttgtaccg tttctgctca agctctttca gtctatagaa 4860

aaggaaggca tcttgcccaa ttccttctac gaagcttcta taatacttat tcccaaacca 4920

ggacgcgata ccacaaagaa ggaaaacttc cggcccatta gtctcatgaa tatcgacgct 4980

aaaatattga acaagattct cgccaacaga atccaacaac atattaagaa attgatacat 5040

cacgaccagg tggggtttat acctggcatg cagggctggt ttaacatccg gaagagtatt 5100

aacgtcattc aacacattaa tagagctaag gataagaatc atatgatcat ctctatagac 5160

gcggaaaagg cattcgataa gattcagcag ccatttatgc tcaagactct gaacaaactc 5220

ggcatcgacg gaacatattt taagattatt cgcgcaattt acgataagcc gactgctaac 5280

attatcctta acggccaaaa gctcgaggcc tttccgctca agactggaac ccgccaaggc 5340

tgtcccctct ccccgctttt gtttaatatt gtactcgagg tgctggctag ggctattcgt 5400

caagagaaag agattaaagg gatacagctc gggaaggaag aggtcaagct ttccttgttc 5460

gccgatgata tgattgtgta cctggagaat cctattgtgt ctgctcagaa ccttcttaaa 5520

cttatttcta actttagcaa ggtcagcggc tataagatta acgtccagaa atctcaggcc 5580

tttctgtaca caaataatcg acagaccgaa tcccagataa tgggtgagct tccgtttgtc 5640

atagccagca aaaggataaa gtatctcgga atccagctga cacgagacgt taaagatttg 5700

tttaaggaaa attacaagcc tctcctgaaa gagattaagg aagatactaa taagtggaag 5760

aatatcccct gttcatgggt tggcagaatc aacatagtga agatggcaat acttcctaaa 5820

gtgatatatc gctttaacgc catcccaatt aaactgccta tgaccttctt tacggagctc 5880

gagaaaacaa cccttaaatt tatatggaat caaaagagag caagaatagc gaagtccatc 5940

ttgagccaga agaataaggc cggtgggatt actttgcctg attttaagtt gtattataaa 6000

gccacagtaa ctaagacagc ctggtattgg tatcagaata gagacatcga ccagtggaat 6060

cggaccgaac catcagagat aatgccccac atctataatt accttatatt cgataagcca 6120

gaaaagaata aacagtgggg caaagacagc ctcttcaaca agtggtgttg ggagaattgg 6180

ctggccatat gccggaaact caagctcgac ccctttctta caccctacac taaaatcaac 6240

agtaggtgga tcaaggactt gaatgtcaag ccaaagacta taaagacact ggaagagaat 6300

cttgggatca caatacaaga tataggcgtc ggcaaagatt ttatgtcaaa gacgcccaag 6360
gccatggcca ctaaggataa gattgataag tgggacctta ttaagctcaa aagcttctgt 6420
actgccaagg agaccacgat cagagttaat aggcagccca ctacatggga aaagattttc 6480
gccacttatt catcagataa ggggttgata agcagaatat ataacgagct gaagcagatc 6540
tacaagaaga aaacgaataa tcccatcaag aagtgggcaa aagatatgaa caggcatttt 6600
agcaaagagg atatctacgc cgcgaagaag catatgaaga agtgtagttc aagcttggcc 6660
attcgtgaga tgcagattaa gacgaccatg cgataccacc ttaccccagt gaggatggca 6720

attatcaaga aatctggcaa taatagatgt tggcggggct gtggcgagat tggcaccctg 6780
ctccattgct ggtgggattg caagctggtg cagccgcttt ggaaatcagt ctggcgcttt 6840
ctgagggacc tcgagcttga gattcccttc gatcccgcaa ttcccttgct cggaatctat 6900
cctaacgaat acaagagctg ttgttacaag gatacgtgta cccggatgtt catcgcggcc 6960
ttgtttacga tagctaagac gtggaatcag cctaagtgcc ccacaatgat cgattggatc 7020
aagaaaatgt ggcatattta taccatggag tattacgcag caattaagaa tgacgaattt 7080
atttccttcg ttgggacctg gatgaagctg gagactatta ttctgagcaa gctgtctcag 7140

gagcaaaaga caaagcatag aatcttctct ctcattggtg gtaacgacta caaagacgat 7200
gacgacaaga aaaggccggc ggccacgaaa aaggccggcc aggcaaaaaa gaaaaagtaa 7260
agcgcttcta gaagttgtct cctcctgcac tgactgactg atacaatcga tttctggatc 7320
cgcaggccta atcaacctct ggattacaaa atttgtgaaa gattgactgg tattcttaac 7380
tatgttgctc cttttacgct atgtggatac gctgctttaa tgcctttgta tcatgctatt 7440
gcttcccgta tggctttcat tttctcctcc ttgtataaat cctggttgct gtctctttat 7500
gaggagttgt ggcccgttgt caggcaacgt ggcgtggtgt gcactgtgtt tgctgacgca 7560

acccccactg gttggggcat tgccaccacc tgtcagctcc tttccgggac tttcgctttc 7620
cccctcccta ttgccacggc ggaactcatc gccgcctgcc ttgcccgctg ctggacaggg 7680
gctcggctgt tgggcactga caattccgtg gtgttgtcgg ggaagctgac gtcctttcca 7740
tggctgctcg cctgtgttgc cacctggatt ctgcgcggga cgtccttctg ctacgtccct 7800
tcggccctca atccagcgga ccttccttcc cgctgagaga cacaaaaaat tccaacacac 7860
tattgcaatg aaaataaatt tcctttatta gccagaagtc agatgctcaa ggggcttcat 7920
gatgtcccca taatttttgg cagagggaaa aagatctcag tggtatttgt gagccagggc 7980

attggccttc tgataggcag cctgcacctg aggagtgcgg ccgctttact tgtacagctc 8040

gtccatgccg agagtgatcc cggcggcggt cacgaactcc agcaggacca tgtgatcgcg 8100

cttctcgttg gggtctttgc tcagggcgga ctgggtgctc aggtagtggt tgtcgggcag 8160

cagcacgggg ccgtcgccga tgggggtgtt ctgctggtag tggtcggcga gctgcacgct 8220

gccgtcctcg atgttgtggc ggatcttgaa gttcaccttg atgccgttct tctgcttgtc 8280

ggccatgata tagacgttgt ggctgttgta gttgtactcc agcttgtgcc ccaggatgtt 8340

gccgtcctcc ttgaagtcga tgcccttcag ctcgatgcgg ttcaccaggg tgtcgccctc 8400

gaacttcacc tcggcgcggg tcttgtagtt gccgtcgtcc ttgaagaaga tggtgcgctc 8460

ctggacgtag ccttcgggca tggcggactt gaagaagtcg tgctgcttca tgtggtcggg 8520

gtagcggctg aagcactgca cgccgtaggt cagggtggtc acgagggtgg gccagggcac 8580

gggcagcttg ccggtggtgc agatgaactt cagggtcagc ttgccgtagg tggcatcgcc 8640

ctcgccctcg ccggacacgc tgaacttgtg gccgtttacg tcgccgtcca gctcgaccag 8700

gatgggcacc accccggtga acagctcctc gcccttgctc accatggtgg cgggatctga 8760

cggttcacta aaccagctct gcttatatag acctcccacc gtacacgcct accgcccatt 8820

tgcgtcaatg gggcggagtt gttacgacat tttggaaagt cccgttgatt ttggtgccaa 8880

aacaaactcc cattgacgtc aatggggtgg agacttggaa atccccgtga gtcaaaccgc 8940

tatccacgcc cattgatgta ctgccaaaac cgcatcacca tggtaatagc gatgactaat 9000

acgtagatgt actgccaagt aggaaagtcc cataaggtca tgtactgggc ataatgccag 9060

gcgggccatt taccgtcatt gacgtcaata gggggcgtac ttggcatatg atacacttga 9120

tgtactgcca agtgggcagt ttaccgtaaa tactccaccc attgacgtca atggaaagtc 9180

cctattggcg ttactatggg aacatacgtc attattgacg tcaatgggcg ggggtcgttg 9240

ggcggtcagc caggcgggcc atttaccgta agttatgtaa cgggcctgct gccggctctg 9300

cggcctcttc cgcgtcttcg ccttcgccct cagacgagtc ggatctccct ttgggccgcc 9360

tccccgcctg tctagcttga ctgactgaga tacagcgtac cttcagctca cagacatgat 9420

aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa aatgctttat 9480

ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca ataaacaagt 9540

taacaacaac aattgcattc attttatgtt tcaggttcag ggggaggtgt gggaggtttt 9600

ttaaagcaag taaaacctct acaaatgtgg tattggccca tctctatcgg tatcgtagca 9660

taacccccttg gggcctctaa acgggtcttg aggggttttt tgtgcccctc gggccggatt 9720

gctatctacc ggcattggcg cagaaaaaaa tgcctgatgc gacgctgcgc gtcttatact 9780

cccacatatg ccagattcag caacggatac ggcttcccca acttgcccac ttccatacgt 9840

gtcctcctta ccagaaattt atccttaagg tcgtcagcta tcctgcaggc gatctctcga 9900

tttcgatcaa gacattcctt taatggtctt ttctggacac cactaggggt cagaagtagt 9960

tcatcaaact ttcttccctc cctaatctca ttggttacct tgggctatcg aaacttaatt 10020

aagcgatctg catctcaatt agtcagcaac catagtcccg cccctaactc cgcccatccc 10080

gcccctaact ccgcccagtt ccgcccattc tccgccccat cgctgactaa tttttttttat 10140

ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt gaggaggctt 10200

ttttggaggc ctaggctttt gcaaaggagg tagccaacat gattgaacaa gatggattgc 10260

acgcaggttc tcccgccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 10320

caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 10380

ttgtcaagac cgacctgtcc ggtgccctga atgaactcca ggacgaggca gcgcggctat 10440

cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 10500

gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 10560

ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 10620

cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 10680

tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctcgcgccag 10740

ccgaactgtt cgccaggctc aaggcgcgga tgcccgacgg cgaggatctc gtcgtgaccc 10800

acggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 10860

actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 10920

ttgctgaaga gcttggcggc gaatgggctg accgcttcct cgtgctttac ggtatcgccg 10980

ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagttcttc tagtatgtaa 11040

gccctgtgcc ttctagttgc cagccatctg ttgtttgccc ctcccccgtg ccttccttga 11100

ccctggaagg tgccactccc actgtccttt cctaataaaa tgaggaaatt gcatcgcatt 11160

gtctgagtag gtgtcattct attctggggg gtggggtggg gcaggacagc aaggggggagg 11220

attgggaaga caatagcagg catgctgggg atgcggtggg ctctatggtt aattaaccag 11280

tcaagtcagc tacttggcga gatcgacttg tctgggtttc gactacgctc agaattgcgt 11340

cagtcaagtt cgatctggtc cttgctattg cacccgttct ccgattacga gtttcattta 11400

aatcatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc 11460

gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag 11520

gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccccctggaa gctccctcgt 11580

gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg 11640

aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt aggtcgttcg 11700

ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg 11760

taactatcgt cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac 11820
tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg 11880
gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc tgaagccagt 11940
taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg 12000
tggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc 12060
tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt 12120
ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt 12180

taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag 12240
tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcat ttaaatttcc 12300
gaactctcca aggccctcgt cggaaaatct tcaaaccttt cgtccgatcc atcttgcagg 12360
ctacctctcg aacgaactat cgcaagtctc ttggccggcc ttgcgccttg gctattgctt 12420
ggcagcgcct atcgccaggt attactccaa tcccgaatat ccgagatcgg gatcacccga 12480
gagaagttca acctacatcc tcaatcccga tctatccgag atccgaggaa tatcgaaatc 12540
ggggcgcgcc tggtgt 12556

<210> 111
<211> 48
<212> RNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 111
cuaggaaucu ggaaguaccg aggaaacucg guacuuccug uguccuag 48
<210> 112
<211> 37
<212> RNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 112
auauggaaga uccuggggaa cugggaucuu ccuaagu 37
<210> 113

<211> 76

<212> DNA

<213> Artificial Sequence

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

<220><223> Description of Combined DNA/RNA Molecule: Synthetic
oligonucleotide

<220><221> modified_base

<222> (16)..(17)

<223> Dihydrouridine

<220><221> modified_base

<222> (37)..(37)

<223> a, c, t, g or u

<220><221> modified_base

<222> (39)..(39)

<223> Pseudouridine

<220><221> modified_base

<222> (55)..(55)

<223> Pseudouridine

<400> 113

gcgcauuuag cucagnnggg agagcgccag acugaananc uggagcuccu gugtncgauc 60

cacagaauuc gcacca 76