



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0075865  
(43) 공개일자 2020년06월26일

(51) 국제특허분류(Int. Cl.)  
C12N 15/113 (2010.01) C12N 15/86 (2006.01)  
(52) CPC특허분류  
C12N 15/1138 (2013.01)  
C12N 15/86 (2013.01)  
(21) 출원번호 10-2020-7014683  
(22) 출원일자(국제) 2018년10월23일  
심사청구일자 없음  
(85) 번역문제출일자 2020년05월22일  
(86) 국제출원번호 PCT/US2018/057187  
(87) 국제공개번호 WO 2019/084068  
국제공개일자 2019년05월02일  
(30) 우선권주장  
62/575,795 2017년10월23일 미국(US)  
62/742,723 2018년10월08일 미국(US)

(71) 출원인  
프리베일 테라퓨틱스, 인크.  
미국 10016 뉴욕주 뉴욕 이스트 29티에이치 스트리트 430 스위트 940  
(72) 발명자  
아벨리오비치, 아사  
미국 10016 뉴욕주 뉴욕 이스트 35티에이치 스트리트 20 아파트먼트 15이  
헤크만, 로라  
미국 10016 뉴욕주 뉴욕 이스트 29티에이치 스트리트 430 스위트 940  
린, 허브  
미국 10016 뉴욕주 뉴욕 이스트 29티에이치 스트리트 430 스위트 940  
(74) 대리인  
양영준, 이상남

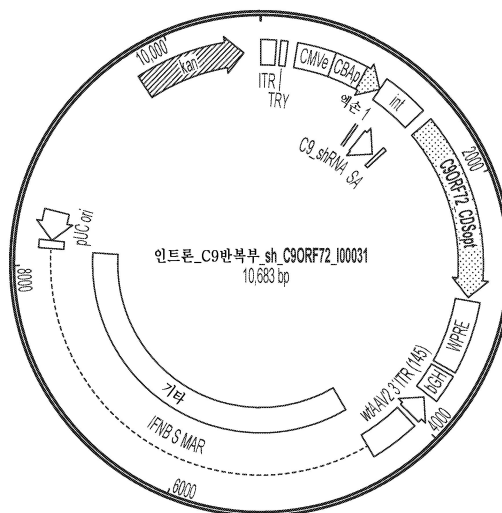
전체 청구항 수 : 총 90 항

(54) 발명의 명칭 신경변성 질환에 대한 유전자 요법

(57) 요약

본 개시내용은 일부 측면에서 신경변성 질환(예를 들어, 근위축성 측삭 경화증(ALS) 및/또는 전두측두엽 치매(FTD), 알츠하이머병, 고셔병, 파킨슨병, 루이 소체 치매 또는 리소솜 축적 질환)의 치료를 위한 조성물 및 방법에 관한 것이다. 일부 실시양태에서, 본 개시내용은 1개 이상의 억제 핵산(예를 들어, C9orf72, TMEM106B, ATNX2, RPS25 등을 표적화함), 야생형 C9orf72 단백질 또는 그의 부분, 또는 상기 중 임의의 조합을 코딩하는 트랜스젠을 포함하는 발현 구축물을 제공한다. 일부 실시양태에서, 본 개시내용은 치료를 필요로 하는 대상체에게 이러한 발현 구축물을 투여함으로써 ALS/FTD를 치료하는 방법을 제공한다.

대표도 - 도11



(52) CPC특허분류

*C12N 2330/51* (2013.01)

*C12N 2750/14143* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

AAV 역전된 말단 반복부 (ITR)가 플랭킹된 C9orf72의 발현 또는 활성을 억제하는 억제 핵산을 코딩하는 발현 구축물을 포함하며, 여기서 ITR 중 적어도 하나는 야생형 AAV2 ITR (서열식별번호: 32)에 비해 변형된 "D" 서열을 포함하는 것인 단리된 핵산.

#### 청구항 2

제1항에 있어서, 억제 핵산이 C9orf72의 디펩티드-반복 영역의 적어도 6개의 인접 뉴클레오타이드에 상보적인 단리된 핵산.

#### 청구항 3

제1항 또는 제2항에 있어서, 억제 핵산이 서열식별번호: 24에 제시된 핵산 서열 또는 그의 부분을 포함하는 억제 RNA인 단리된 핵산.

#### 청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 억제 핵산이 서열식별번호: 25에 제시된 서열 또는 그의 부분을 포함하는 것인 단리된 핵산.

#### 청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 변형된 "D" 영역이 발현 구축물 기준으로 ITR의 외부에 위치하는 "D" 서열인 단리된 핵산.

#### 청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, 변형된 "D" 서열을 포함하는 ITR이 3' ITR인 단리된 핵산.

#### 청구항 7

AAV 역전된 말단 반복부 (ITR)가 플랭킹된 TMEM106B의 발현 또는 활성을 억제하는 억제 핵산을 코딩하는 발현 구축물을 포함하며, 여기서 ITR 중 적어도 하나는 야생형 AAV2 ITR (서열식별번호: 32)에 비해 변형된 "D" 서열을 포함하는 것인 단리된 핵산.

#### 청구항 8

제7항에 있어서, 억제 핵산이 서열식별번호: 7에 제시된 서열의 적어도 6개의 인접 뉴클레오타이드에 상보적인 단리된 핵산.

#### 청구항 9

제7항 또는 제8항에 있어서, 억제 핵산이 서열식별번호: 33 또는 34에 제시된 핵산 서열을 포함하는 억제 RNA인 단리된 핵산.

#### 청구항 10

제7항 내지 제9항 중 어느 한 항에 있어서, 억제 핵산이 서열식별번호: 35 또는 36에 제시된 서열을 포함하는 것인 단리된 핵산.

#### 청구항 11

제7항 내지 제10항 중 어느 한 항에 있어서, 변형된 "D" 영역이 발현 구축물 기준으로 ITR의 외부에 위치하는

"D" 서열인 단리된 핵산.

## 청구항 12

제7항 내지 제11항 중 어느 한 항에 있어서, 변형된 "D" 서열을 포함하는 ITR이 3' ITR인 단리된 핵산.

## 청구항 13

AAV 역전된 말단 반복부 (ITR)가 플랭킹된 ATXN2의 발현 또는 활성을 억제하는 억제 핵산을 코딩하는 발현 구축물을 포함하며, 여기서 ITR 중 적어도 하나는 야생형 AAV2 ITR (서열식별번호: 32)에 비해 변형된 "D" 서열을 포함하는 것인 단리된 핵산.

## 청구항 14

제13항에 있어서, 억제 핵산이 서열식별번호: 9에 제시된 서열의 적어도 6개의 인접 뉴클레오티드에 상보적인 단리된 핵산.

## 청구항 15

제13항 또는 제14항에 있어서, 억제 핵산이 서열식별번호: 10-23 중 어느 하나에 제시된 핵산 서열을 포함하는 억제 RNA인 단리된 핵산.

## 청구항 16

제13항 내지 제15항 중 어느 한 항에 있어서, 변형된 "D" 영역이 발현 구축물 기준으로 ITR의 외부에 위치하는 "D" 서열인 단리된 핵산.

## 청구항 17

제13항 내지 제16항 중 어느 한 항에 있어서, 변형된 "D" 서열을 포함하는 ITR이 3' ITR인 단리된 핵산.

## 청구항 18

- (i) C9orf72의 발현 또는 활성을 억제하는 제1 억제 핵산, 및
  - (ii) TMEM106B 또는 ATXN2의 발현 또는 활성을 억제하는 제2 억제 핵산
- 을 코딩하는 발현 카세트를 포함하는 단리된 핵산.

## 청구항 19

제1항에 있어서, 제1 억제 핵산이 C9orf72 mRNA 전사체의 디캡티드-반복 영역에 결합하고, 임의로 여기서 디캡티드-반복 영역이 1개 이상의 GGGGCC 반복부 또는 1개 이상의 CCCCGG 반복부를 포함하는 것인 단리된 핵산.

## 청구항 20

제18항 또는 제19항에 있어서, 제1 억제 핵산 및 제2 억제 핵산이 각각 독립적으로 siRNA, shRNA, miRNA 및 dsRNA로부터 선택된 것인 단리된 핵산.

## 청구항 21

제18항 내지 제20항 중 어느 한 항에 있어서, 제1 억제 핵산 및/또는 제2 억제 핵산이 발현 구축물의 비번역 영역 내에 위치하고, 임의로 여기서 비번역 영역이 인트론, 5' 비번역 영역 (5'UTR) 또는 3' 비번역 영역 (3'UTR)인 단리된 핵산.

## 청구항 22

제18항 내지 제21항 중 어느 한 항에 있어서, 야생형 C9orf72 단백질 코딩 서열 또는  $\beta$ -글루코세레브로시다제 (GBA) 단백질 코딩 서열을 추가로 포함하고, 임의로 여기서 GBA 단백질이 GBA1 단백질인 단리된 핵산.

## 청구항 23

제18항 내지 제20항 중 어느 한 항에 있어서, 1개 이상의 프로모터를 추가로 포함하고, 임의로 여기서 각각의 1개 이상의 프로모터가 독립적으로 RNA pol III 프로모터 (예를 들어, U6, H1 등), 닭-베타 액틴 (CBA) 프로모터, CAG 프로모터, CD68 프로모터 또는 JeT 프로모터인 단리된 핵산.

#### 청구항 24

제18항 내지 제23항 중 어느 한 항에 있어서, 발현 구축물에 2개의 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR) 서열이 플랜킹되고, 임의로 여기서 ITR 서열 중 하나에 기능적 말단 분해 부위가 결합된 것인 단리된 핵산.

#### 청구항 25

제18항 내지 제24항 중 어느 한 항에 있어서, 서열식별번호: 1-28 또는 33-36 중 어느 하나에 제시된 서열 또는 그의 부분을 갖는 단리된 핵산.

#### 청구항 26

제1항 내지 제25항 중 어느 한 항의 단리된 핵산을 포함하는 벡터.

#### 청구항 27

제26항에 있어서, 플라스미드인 벡터.

#### 청구항 28

제26항에 있어서, 벡터가 바이러스 벡터이고, 임의로 여기서 바이러스 벡터가 재조합 아데노-연관 바이러스 벡터 (rAAV) 또는 마콜로바이러스 벡터인 벡터.

#### 청구항 29

제1항 내지 제25항 중 어느 한 항의 단리된 핵산 또는 제26항 내지 제28항 중 어느 한 항의 벡터 및 임의로 제약상 허용되는 담체를 포함하는 조성물.

#### 청구항 30

제1항 내지 제8항 중 어느 한 항의 단리된 핵산 또는 제9항 내지 제11항 중 어느 한 항의 벡터를 포함하는 숙주 세포.

#### 청구항 31

(i) 캡시드 단백질; 및

(ii) 제1항 내지 제25항 중 어느 한 항의 단리된 핵산 또는 제26항의 벡터

를 포함하는 재조합 아데노-연관 바이러스 (rAAV).

#### 청구항 32

제31항에 있어서, 캡시드 단백질이 혈액-뇌 장벽을 가로지를 수 있고, 임의로 여기서 캡시드 단백질이 AAV9 캡시드 단백질, AAVrh.10 캡시드 단백질 또는 AAV-PHP.B 변이체인 rAAV.

#### 청구항 33

제31항 또는 제32항에 있어서, 중추 신경계 (CNS)의 뉴런 세포 및/또는 비-뉴런 세포를 형질도입시키는 rAAV.

#### 청구항 34

신경변성 질환을 갖거나 갖는 것으로 의심되는 대상체에게 제1항 내지 제25항 중 어느 한 항의 단리된 핵산, 제26항 내지 제28항 중 어느 한 항의 벡터, 제29항의 조성물 또는 제31항 내지 제33항 중 어느 한 항의 rAAV를 투여하는 것을 포함하는, 신경변성 질환을 갖거나 갖는 것으로 의심되는 대상체를 치료하는 방법.

#### 청구항 35

제34항에 있어서, 신경변성 질환이 근위축성 측삭 경화증 (ALS) 및/또는 전두측두엽 치매 (FTD), 알츠하이머병, 고셔병, 파킨슨병, 루이 소체 치매 또는 리소솜 축적 질환인 방법.

#### 청구항 36

제34항 또는 제35항에 있어서, 신경변성 질환이 ALS 및/또는 FTD인 방법.

#### 청구항 37

제34항 내지 제36항 중 어느 한 항에 있어서, 투여가 대상체의 CNS에 대한 직접 주사를 포함하고, 임의로 여기서 직접 주사가 뇌내 주사, 실질내 주사, 경막내 주사, 대수조내 주사 또는 그의 임의의 조합인 방법.

#### 청구항 38

제37항에 있어서, 직접 주사가 대상체의 뇌척수액 (CSF)에 대한 직접 주사이고, 임의로 여기서 직접 주사가 수조내 주사, 뇌실내 주사 및/또는 요추내 주사인 방법.

#### 청구항 39

제34항 내지 제38항 중 어느 한 항에 있어서, 대상체가 약 30 내지 약 5000개의 GGGGCC 디펩티드 반복부 및/또는 약 30 내지 5000개의 CCCCGG 반복부를 갖는 것을 특징으로 하는 것인 방법.

#### 청구항 40

- (i) C9orf72의 발현 또는 활성을 억제하는 억제 핵산, 및
- (ii)  $\beta$ -글루코세레브로시다제 (GBA) 단백질, 임의로 GBA1 단백질인 GBA 단백질을 코딩하는 발현 카세트를 포함하는 단리된 핵산.

#### 청구항 41

제40항에 있어서, 억제 핵산이 C9orf72 mRNA 전사체의 디펩티드-반복 영역에 결합하고, 임의로 여기서 디펩티드-반복 영역이 1개 이상의 GGGGCC 반복부 또는 1개 이상의 CCCCGG 반복부를 포함하는 것인 단리된 핵산.

#### 청구항 42

제40항 또는 제41항에 있어서, 억제 핵산이 siRNA, shRNA, miRNA 또는 dsRNA인 단리된 핵산.

#### 청구항 43

제40항 내지 제42항 중 어느 한 항에 있어서, 억제 핵산이 발현 구축물의 비번역 영역에 위치하고, 임의로 여기서 비번역 영역이 인트론, 5' 비번역 영역(5'UTR) 또는 3' 비번역 영역(3'UTR)인 단리된 핵산.

#### 청구항 44

제40항 내지 제43항 중 어느 한 항에 있어서, 1개 이상의 프로모터를 추가로 포함하고, 임의로 여기서 각각의 1개 이상의 프로모터가 독립적으로 RNA pol III 프로모터 (예를 들어, U6, H1 등), 닭-베타 액틴 (CBA) 프로모터, CAG 프로모터, CD68 프로모터 또는 JeT 프로모터인 단리된 핵산.

#### 청구항 45

제40항 내지 제44항 중 어느 한 항에 있어서, 발현 구축물에 2개의 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR) 서열이 플랭킹되고, 임의로 여기서 ITR 서열 중 하나에 기능적 말단 분해 부위가 결합된 것인 단리된 핵산.

#### 청구항 46

제40항 내지 제45항 중 어느 한 항에 있어서, 서열식별번호: 2에 제시된 서열 또는 그의 부분을 갖는 단리된 핵

산.

#### 청구항 47

제40항 내지 제46항 중 어느 한 항의 단리된 핵산을 포함하는 벡터.

#### 청구항 48

제47항에 있어서, 플라스미드인 벡터.

#### 청구항 49

제47항에 있어서, 벡터가 바이러스 벡터이고, 임의로 여기서 바이러스 벡터가 재조합 아데노-연관 바이러스 벡터 (rAAV) 또는 바쿨로바이러스 벡터인 벡터.

#### 청구항 50

제40항 내지 제46항 중 어느 한 항의 단리된 핵산 또는 제47항 내지 제49항 중 어느 한 항의 벡터 및 임의로 제약상 허용되는 담체를 포함하는 조성물.

#### 청구항 51

제40항 내지 제46항 중 어느 한 항의 단리된 핵산 또는 제47항 내지 제49항 중 어느 한 항의 벡터를 포함하는 숙주 세포.

#### 청구항 52

(i) 캡시드 단백질; 및

(ii) 제40항 내지 제46항 중 어느 한 항의 단리된 핵산 또는 제47항의 벡터

를 포함하는 재조합 아데노-연관 바이러스 (rAAV).

#### 청구항 53

제52항에 있어서, 캡시드 단백질이 혈액-뇌 장벽을 가로지를 수 있고, 임의로 여기서 캡시드 단백질이 AAV9 캡시드 단백질, AAVrh.10 캡시드 단백질 또는 AAV-PHP.B 캡시드 단백질인 rAAV.

#### 청구항 54

제52항 또는 제53항에 있어서, 중추 신경계 (CNS)의 뉴런 세포 및/또는 비-뉴런 세포를 형질도입시키는 rAAV.

#### 청구항 55

신경변성 질환을 갖거나 갖는 것으로 의심되는 대상체에게 제40항 내지 제46항 중 어느 한 항의 단리된 핵산, 제47항 내지 제49항 중 어느 한 항의 벡터, 제50항의 조성물 또는 제52항 내지 제54항 중 어느 한 항의 rAAV를 투여하는 것을 포함하는, 신경변성 질환을 갖거나 갖는 것으로 의심되는 대상체를 치료하는 방법.

#### 청구항 56

제55항에 있어서, 신경변성 질환이 근위축성 측삭 경화증 (ALS) 및/또는 전두측두엽 치매 (FTD), 알츠하이머병, 고셔병, 파킨슨병, 루이 소체 치매 또는 리소솜 축적 질환인 방법.

#### 청구항 57

제55항 또는 제56항에 있어서, 신경변성 질환이 ALS 및/또는 FTD인 방법.

#### 청구항 58

제55항 내지 제57항 중 어느 한 항에 있어서, 투여가 대상체의 CNS에 대한 직접 주사를 포함하고, 임의로 여기서 직접 주사가 뇌내 주사, 실질내 주사, 경막내 주사 또는 그의 임의의 조합인 방법.

#### 청구항 59

제58항에 있어서, 직접 주사가 대상체의 뇌척수액 (CSF)에 대한 직접 주사이고, 임의로 여기서 직접 주사가 수조내 주사, 뇌실내 주사 및/또는 요추내 주사인 방법.

#### 청구항 60

제55항 내지 제59항 중 어느 한 항에 있어서, 대상체가 약 30 내지 약 5000개의 GGGGCC 디펩티드 반복부 및/또는 약 30 내지 5000개의 CCCCGG 반복부를 갖는 것을 특징으로 하는 것인 방법.

#### 청구항 61

(i) C9orf72의 발현 또는 활성을 억제하는 억제 핵산, 및

(ii) 야생형 C9orf72 단백질

을 코딩하는 발현 카세트를 포함하는 단리된 핵산.

#### 청구항 62

제61항에 있어서, 억제 핵산이 C9orf72 mRNA 전사체의 디펩티드-반복 영역에 결합하고, 임의로 여기서 디펩티드-반복 영역이 1개 이상의 GGGGCC 반복부 또는 1개 이상의 CCCCGG 반복부를 포함하는 것인 단리된 핵산.

#### 청구항 63

제61항 또는 제62항에 있어서, 억제 핵산이 siRNA, shRNA, miRNA 또는 dsRNA인 단리된 핵산.

#### 청구항 64

제61항 내지 제63항 중 어느 한 항에 있어서, 억제 핵산이 발현 구축물의 비번역 영역에 위치하고, 임의로 여기서 비번역 영역이 인트론, 5' 비번역 영역 (5'UTR) 또는 3' 비번역 영역 (3'UTR)인 단리된 핵산.

#### 청구항 65

제61항 내지 제64항 중 어느 한 항에 있어서, 1개 이상의 프로모터를 추가로 포함하고, 임의로 여기서 각각의 1개 이상의 프로모터가 독립적으로 RNA pol III 프로모터 (예를 들어, U6, H1 등), 닭-베타 액틴 (CBA) 프로모터, CAG 프로모터, CD68 프로모터 또는 JeT 프로모터인 단리된 핵산.

#### 청구항 66

제61항 내지 제65항 중 어느 한 항에 있어서, 발현 구축물에 2개의 아테노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR) 서열이 플랭킹되고, 임의로 여기서 ITR 서열 중 하나에 기능적 말단 분해 부위가 결합된 것인 단리된 핵산.

#### 청구항 67

제61항 내지 제66항 중 어느 한 항에 있어서, C9orf72 단백질이 서열식별번호: 3에 제시된 서열 또는 그의 부분에 의해 코딩되는 것인 단리된 핵산.

#### 청구항 68

제67항에 있어서, 야생형 C9orf72 단백질이 서열식별번호: 4에 제시된 서열을 포함하거나 이로 이루어진 것인 단리된 핵산.

#### 청구항 69

제67항 또는 제68항에 있어서, 서열식별번호: 5에 제시된 서열 또는 그의 부분을 갖는 단리된 핵산.

#### 청구항 70

제61항 내지 제69항 중 어느 한 항의 단리된 핵산을 포함하는 벡터.



#### 청구항 71

제70항에 있어서, 플라스미드인 벡터.

#### 청구항 72

제70항에 있어서, 벡터가 바이러스 벡터이고, 임의로 여기서 바이러스 벡터가 재조합 아데노-연관 바이러스 벡터 (rAAV)인 벡터.

#### 청구항 73

제61항 내지 제69항 중 어느 한 항의 단리된 핵산 또는 제70항 내지 제72항 중 어느 한 항의 벡터 및 임의로 제약상 허용되는 담체를 포함하는 조성물.

#### 청구항 74

제61항 내지 제69항 중 어느 한 항의 단리된 핵산 또는 제70항 내지 제72항 중 어느 한 항의 벡터를 포함하는 숙주 세포.

#### 청구항 75

(i) 캡시드 단백질; 및

(ii) 제61항 내지 제69항 중 어느 한 항의 단리된 핵산 또는 제70항의 벡터

를 포함하는 재조합 아데노-연관 바이러스 (rAAV).

#### 청구항 76

제75항에 있어서, 캡시드 단백질이 혈액-뇌 장벽을 가로지를 수 있고, 임의로 여기서 캡시드 단백질이 AAV9 캡시드 단백질, AAVrh.10 캡시드 단백질 또는 AAV-PHP.B 캡시드 단백질인 rAAV.

#### 청구항 77

제75항 또는 제76항에 있어서, 중추 신경계 (CNS)의 뉴런 세포 및/또는 비-뉴런 세포를 형질도입시키는 rAAV.

#### 청구항 78

신경변성 질환을 갖거나 갖는 것으로 의심되는 대상체에게 제61항 내지 제69항 중 어느 한 항의 단리된 핵산, 제70항 내지 제72항 중 어느 한 항의 벡터, 제73항의 조성물 또는 제75항 내지 제77항 중 어느 한 항의 rAAV를 투여하는 것을 포함하는, 신경변성 질환을 갖거나 갖는 것으로 의심되는 대상체를 치료하는 방법.

#### 청구항 79

제78항에 있어서, 신경변성 질환이 근위축성 측삭 경화증 (ALS) 및/또는 전두측두엽 치매 (FTD), 알츠하이머병, 고셔병, 파킨슨병, 루이 소체 치매 또는 리소솜 축적 질환인 방법.

#### 청구항 80

제78항 또는 제79항에 있어서, 신경변성 질환이 ALS 및/또는 FTD인 방법.

#### 청구항 81

제78항 내지 제80항 중 어느 한 항에 있어서, 투여가 대상체의 CNS에 대한 직접 주사를 포함하고, 임의로 여기서 직접 주사가 뇌내 주사, 실질내 주사, 경막내 주사 또는 그의 임의의 조합인 방법.

#### 청구항 82

제81항에 있어서, 직접 주사가 대상체의 뇌척수액 (CSF)에 대한 직접 주사이고, 임의로 여기서 직접 주사가 수조내 주사, 뇌실내 주사 및/또는 요추내 주사인 방법.

#### 청구항 83

제78항 내지 제82항 중 어느 한 항에 있어서, 대상체가 약 30 내지 약 5000개의 GGGGCC 디펩티드 반복부 및/또는 약 30 내지 5000개의 CCCC GG 반복부를 갖는 것을 특징으로 하는 것인 방법.

#### 청구항 84

AAV 역전된 말단 반복부 (ITR)가 플랭킹된 C9orf72의 발현 또는 활성을 억제하는 억제 핵산을 코딩하는 발현 구축물을 포함하며, 여기서 억제 핵산은 서열식별번호: 24, 25 및 37-49 중 어느 하나에 제시된 서열을 포함하는 것인 단리된 핵산.

#### 청구항 85

AAV 역전된 말단 반복부 (ITR)가 플랭킹된 ATXN2의 발현 또는 활성을 억제하는 억제 핵산을 코딩하는 발현 구축물을 포함하며, 여기서 억제 핵산은 서열식별번호: 10-23 중 어느 하나에 제시된 서열을 포함하는 것인 단리된 핵산.

#### 청구항 86

AAV 역전된 말단 반복부 (ITR)가 플랭킹된 TMEM106B의 발현 또는 활성을 억제하는 억제 핵산을 코딩하는 발현 구축물을 포함하며, 여기서 억제 핵산은 서열식별번호: 33-36 중 어느 하나에 제시된 서열을 포함하는 것인 단리된 핵산.

#### 청구항 87

서열식별번호: 51-58 중 어느 하나에 제시된 서열을 포함하는 재조합 AAV (rAAV) 벡터.

#### 청구항 88

C9orf72 단백질을 코딩하는 발현 구축물을 포함하며, 여기서 발현 구축물은 서열식별번호: 50에 제시된 서열을 포함하는 것인 단리된 핵산.

#### 청구항 89

제87항에 있어서, 발현 구축물에 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR)가 플랭킹된 것인 단리된 핵산.

#### 청구항 90

제88항에 있어서, ITR 중 적어도 하나가 야생형 AAV2 ITR (서열식별번호: 32)에 비해 변형된 "D" 서열을 포함하는 것인 단리된 핵산.

### 발명의 설명

### 기술 분야

[0001] 관련 출원

[0002] 본 출원은 "신경변성 질환에 대한 유전자 요법"이라는 발명의 명칭으로 2018년 10월 8일에 출원된 미국 가출원 일련 번호 62/742,723 및 "신경변성 질환에 대한 유전자 요법"이라는 발명의 명칭으로 2017년 10월 23일에 출원된 62/575,795를 35 U.S.C. § 119(e) 하에 우선권 주장하며, 이들 각각의 전체 내용은 본원에 참조로 포함된다.

### 배경 기술

[0003] 근위축성 측삭 경화증 (ALS) 및 전두측두엽 치매 (FTD)는 인간에서 C9orf72 유전자 내의 헥사뉴클레오타이드 반복 영역의 확장과 연관된 신경변성 질환이다. 일반적으로, C9orf72 반복 영역의 확장과 연관된 병리상태는 C9orf72 단백질의 감소된 발현 및 독성 RNA 뭉침(foci)의 축적으로 인한 기능 획득에 의해 야기된다. 현재, ALS/FTD에 대한 치료 옵션은 제한적이다.

### 발명의 내용

[0004] 본 개시내용의 측면은 신경변성 질환, 예컨대 근위축성 측삭 경화증 (ALS) 및/또는 전두측두엽 치매 (FTD), 알

츠하이머병, 고셔병, 파킨슨병, 루이 소체 치매 또는 리소솜 축적 질환의 치료에 유용한 조성물 및 방법에 관한 것이다. 일부 실시양태에서, 본원에 기재된 방법 및 조성물은 C9orf72 유전자의 디펩티드 반복 영역의 확장을 특징으로 하는 ALS/FTD를 갖는 대상체를 치료하는데 유용하다.

[0005] 일부 측면에서, 본 개시내용은 C9orf72 및/또는 아탁신 2 (ATXN2) 및/또는 리보솜 단백질 25 (RPS25)의 발현 또는 활성을 억제하는 억제 핵산을 코딩하는 발현 카세트를 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, ATXN을 표적화하는 억제 핵산은 서열식별번호(SEQ ID NO): 10-25 중 어느 하나에 제시된 서열을 포함하거나 이로 이루어진다. 일부 실시양태에서, C9orf72를 표적화하는 억제 핵산은 서열식별번호: 37-50 중 어느 하나에 제시된 서열을 포함하거나 이로 이루어진다.

[0006] 일부 측면에서, 본 개시내용은 코돈-최적화된 C9orf72 단백질 (또는 그의 부분)을 코딩하는 발현 카세트를 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 코돈-최적화된 C9orf72 단백질은 서열식별번호: 4에 제시된 아미노산 서열을 포함한다. 일부 실시양태에서, 코돈-최적화된 C9orf72 단백질은 서열식별번호: 51에 제시된 서열을 갖는 핵산에 의해 코딩된다.

[0007] 일부 측면에서, 본 개시내용은 C9orf72 및/또는 ATXN2 및/또는 RPS25, 및 야생형 C9orf72 단백질 (예를 들어, 병원성 디펩티드 반복부 확장이 결여된 C9orf72 단백질)의 발현 또는 활성을 억제하는 억제 핵산을 코딩하는 발현 카세트를 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 야생형 C9orf72 단백질은 서열식별번호: 3 또는 그의 부분에 의해 코딩된다. 일부 실시양태에서, 야생형 C9orf72 단백질은 서열식별번호: 4에 제시된 서열 또는 그의 부분을 포함하거나 이로 이루어진다.

[0008] 일부 측면에서, 본 개시내용은 C9orf72의 발현 또는 활성을 억제하는 제1 억제 핵산 및 아탁신 2 (ATXN2)의 발현 또는 활성을 억제하는 제2 억제 핵산을 코딩하는 발현 카세트를 포함하는 단리된 핵산을 제공한다. 일부 측면에서, 본 개시내용은 C9orf72의 발현 또는 활성을 억제하는 제1 억제 핵산 및 막횡단 단백질 106B (TMEM106B)의 발현 또는 활성을 억제하는 제2 억제 핵산을 코딩하는 발현 카세트를 포함하는 단리된 핵산을 제공한다. 일부 측면에서, 본 개시내용은 C9orf72의 발현 또는 활성을 억제하는 제1 억제 핵산 및 RPS25의 발현 또는 활성을 억제하는 제2 억제 핵산을 코딩하는 발현 카세트를 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 야생형 C9orf72 단백질을 코딩하는 핵산 서열 (예를 들어, 서열식별번호: 3에 제시된 바와 같음)을 추가로 포함한다.

[0009] 일부 측면에서, 본 개시내용은 C9orf72의 발현 또는 활성을 억제하는 억제 핵산 및  $\beta$ -글루코세레브로시다제 (GBA) 단백질을 코딩하는 발현 카세트를 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, GBA 단백질은 GBA1 단백질 (예를 들어, GBA1 유전자 또는 그의 부분에 의해 코딩되는 단백질)이다. 일부 측면에서, 본 개시내용은 ATXN2의 발현 또는 활성을 억제하는 억제 핵산 및  $\beta$ -글루코세레브로시다제 (GBA) 단백질을 코딩하는 발현 카세트를 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, GBA 단백질은 GBA1 단백질 (예를 들어, GBA1 유전자 또는 그의 부분에 의해 코딩되는 단백질)이다. 일부 측면에서, 본 개시내용은 TMEM106B의 발현 또는 활성을 억제하는 억제 핵산 및  $\beta$ -글루코세레브로시다제 (GBA) 단백질을 코딩하는 발현 카세트를 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, GBA 단백질은 GBA1 단백질 (예를 들어, GBA1 유전자 또는 그의 부분에 의해 코딩되는 단백질)이다.

[0010] 일부 실시양태에서, 억제 핵산 (예를 들어, 제1 억제 핵산, 제2 억제 핵산, 제3 억제 핵산 등)은 C9orf72의 디펩티드-반복 영역을 코딩하는 핵산 (예를 들어, 디펩티드-반복 영역을 포함하는 C9orf72 mRNA 전사체)에 결합한다. 일부 실시양태에서, 디펩티드-반복 영역은 1개 이상의 GGGGCC 반복부 또는 1개 이상의 CCCC GG 반복부 (예를 들어, C9orf72의 디펩티드 반복 영역)를 포함한다. 일부 실시양태에서, 디펩티드-반복 영역은 23개 이상 (예를 들어, 23 내지 10,000 사이의 임의의 정수, 예를 들어 24, 25, 30, 50, 100, 1000, 5000 또는 10,000 개)의 GGGGCC 반복부 (예를 들어, C9orf72의 센스 가닥 디펩티드 반복 영역), 또는 23개 이상 (예를 들어, 23 내지 10,000 사이의 임의의 정수, 예를 들어 24, 25, 30, 50, 100, 1000, 5000 또는 10,000개)의 CCCC GG 반복부 (예를 들어, C9orf72의 안티센스 가닥 디펩티드 반복 영역)를 포함한다.

[0011] 일부 실시양태에서, 억제 핵산은 디펩티드-반복 영역이 아닌 C9orf72의 영역을 코딩하는 핵산 (예를 들어, C9orf72 디펩티드-반복 영역의 외부에 있는 핵산의 부분)에 결합한다. 일부 실시양태에서, 억제 핵산은 1개 핵산 (예를 들어, 디펩티드-반복 영역에 인접한) 내지 디펩티드-반복 영역의 약 500개 핵산 내에 있는 단리된 핵산 서열에 결합한다. 일부 실시양태에서, 억제 핵산은 C9orf72 단백질을 코딩하는 유전자의 인트론 영역을 표적화한다.

- [0012] 일부 실시양태에서, 억제 핵산 (예를 들어, 제1 억제 핵산, 제2 억제 핵산, 제3 억제 핵산 등)은, 예를 들어 서열식별번호: 9에 제시된 바와 같은 ATXN2를 코딩하는 핵산 서열 (예를 들어, ATXN2 mRNA 전사체)에 결합한다. 일부 실시양태에서, ATXN2를 표적화하는 억제 핵산은 ATXN2를 코딩하는 핵산 서열의 비번역 영역 (예를 들어, 5'UTR, 3'UTR 등)에 결합한다.
- [0013] 일부 실시양태에서, 억제 핵산 (예를 들어, 제1 억제 핵산, 제2 억제 핵산, 제3 억제 핵산 등)은, 예를 들어 서열식별번호: 7에 제시된 바와 같은 TMEM106B를 코딩하는 핵산 서열 (예를 들어, TMEM106B mRNA 전사체)에 결합한다. 일부 실시양태에서, TMEM106B를 표적화하는 억제 핵산은 TMEM106B를 코딩하는 핵산 서열의 비번역 영역 (예를 들어, 5'UTR, 3'UTR 등)에 결합한다.
- [0014] 일부 실시양태에서, 억제 핵산 (예를 들어, 제1 억제 핵산, 제2 억제 핵산, 제3 억제 핵산 등)은, 예를 들어 서열식별번호: 60에 제시된 바와 같은 RPS25를 코딩하는 핵산 서열 (예를 들어, RPS25 mRNA 전사체)에 결합한다. 일부 실시양태에서, RPS25를 표적화하는 억제 핵산은 RPS25를 코딩하는 핵산 서열의 비번역 영역 (예를 들어, 5'UTR, 3'UTR 등)에 결합한다.
- [0015] 일부 실시양태에서, 억제 핵산 (예를 들어, 제1 억제 핵산 및/또는 제2 억제 핵산)은 siRNA, shRNA, miRNA 및 dsRNA이다. 일부 실시양태에서, miRNA는 miRNA 스캐폴드 서열, 예를 들어 miR-155 스캐폴드 서열이 플랭킹된 억제 핵산 서열을 포함하는 인공 miRNA (amiRNA)이다.
- [0016] 일부 실시양태에서, 억제 핵산 (예를 들어, 제1 억제 핵산 및/또는 제2 억제 핵산)은 발현 구축물의 비번역 영역에 위치한다. 일부 실시양태에서, 비번역 영역은 인트론, 5' 비번역 영역 (5'UTR) 또는 3' 비번역 영역 (3'UTR)이다.
- [0017] 일부 실시양태에서, 단리된 핵산은 1개 이상의 프로모터를 포함한다. 일부 실시양태에서, 프로모터는 RNA pol III 프로모터 (예를 들어, U6 또는 H1), RNA pol II 프로모터, 닭-베타 액틴 (CBA) 프로모터, CAG 프로모터, CD68 프로모터 또는 JeT 프로모터이다.
- [0018] 일부 실시양태에서, 발현 구축물에는 2개의 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR) 서열이 플랭킹되어 있다. 일부 실시양태에서, 발현 구축물에 플랭킹된 ITR 서열 중 하나는 기능적 말단 분해 부위가 결여되어 있다.
- [0019] 본 개시내용은 일부 측면에서 변형된 "D" 영역을 갖는 ITR (예를 들어, 야생형 AAV2 ITR (서열식별번호: 32)에 비해 변형된 D 서열)을 포함하는 rAAV 벡터에 관한 것이다. 일부 실시양태에서, 변형된 D 영역을 갖는 ITR은 rAAV 벡터의 5' ITR이다. 일부 실시양태에서, 변형된 "D" 영역은, 예를 들어 서열식별번호 29에 제시된 바와 같은 "S" 서열을 포함한다. 일부 실시양태에서, 변형된 "D" 영역을 갖는 ITR은 rAAV 벡터의 3' ITR이다. 일부 실시양태에서, 변형된 "D" 영역은 "D" 영역이 ITR의 3' 말단에 (예를 들어, 벡터의 트랜스진 삽입물 기준으로 ITR의 외부 또는 말단 단부에) 위치하는 3' ITR을 포함한다. 일부 실시양태에서, 변형된 "D" 영역은 서열식별번호: 29 또는 30에 제시된 바와 같은 서열을 포함한다.
- [0020] 일부 실시양태에서, 단리된 핵산 (예를 들어, rAAV 벡터)은 TRY 영역을 포함한다. 일부 실시양태에서, TRY 영역은 서열식별번호: 31에 제시된 서열을 포함한다.
- [0021] 일부 실시양태에서, 단리된 핵산은 서열식별번호: 1-62 중 어느 하나에 제시된 서열 또는 그의 부분을 포함한다 (또는 아미노산 서열을 코딩함).
- [0022] 일부 측면에서, 본 개시내용은 본 개시내용에 의해 기재된 바와 같은 단리된 핵산을 포함하는 벡터를 제공한다. 일부 실시양태에서, 벡터는 플라스미드 또는 바이러스 벡터이다. 일부 실시양태에서, 바이러스 벡터는 재조합 아데노-연관 바이러스 벡터 (rAAV) (예를 들어, 1개 이상의 억제 핵산을 코딩하는 단리된 핵산 서열 및/또는 1개 이상의 단백질, 예컨대 야생형 C9orf72 및/또는 GBA1을 코딩하는 단리된 핵산을 포함하는 트랜스진, AAV ITR 이 플랭킹됨) 또는 바쿨로바이러스 벡터이다. 일부 실시양태에서, rAAV 벡터는 단일 가닥 (예를 들어, 단일 가닥 DNA)이다.
- [0023] 일부 측면에서, 본 개시내용은 본 개시내용에 의해 기재된 바와 같은 단리된 핵산 또는 벡터를 포함하는 조성물을 제공한다. 일부 실시양태에서, 조성물은 제약상 허용되는 담체를 추가로 포함한다.
- [0024] 일부 측면에서, 본 개시내용은 본 개시내용에 의해 기재된 바와 같은 단리된 핵산 또는 벡터를 포함하는 숙주 세포를 제공한다. 일부 실시양태에서, 숙주 세포는 진핵 세포 (예를 들어, 포유동물 세포, 곤충 세포 등) 또는

원핵 세포 (예를 들어, 박테리아 세포)이다.

- [0025] 일부 측면에서, 본 개시내용은 본 개시내용에 의해 기재된 바와 같은 캡시드 단백질 및 단리된 핵산 또는 벡터를 포함하는 재조합 아데노-연관 바이러스 (rAAV)를 제공한다. 일부 실시양태에서, 캡시드 단백질은 혈액-뇌 장벽을 가로지를 수 있다. 일부 실시양태에서, 캡시드 단백질은 AAV9 캡시드 단백질, AAVrh.10 캡시드 단백질 또는 AAV-PHP.B 캡시드 단백질이다. 일부 실시양태에서, rAAV는 중추 신경계 (CNS)의 뉴런 세포 및/또는 비-뉴런 세포를 형질도입시킨다.
- [0026] 일부 측면에서, 본 개시내용은 신경변성 장애 (예를 들어, 근위축성 측삭 경화증 (ALS) 및/또는 전두측두엽 치매 (FTD), 알츠하이머병, 고셔병, 파킨슨병, 루이 소체 치매 또는 리소솜 축적 질환)를 갖거나 갖는 것으로 의심되는 대상체에게 본 개시내용에 의해 기재된 바와 같은 단리된 핵산, 벡터, 조성물 또는 rAAV를 투여하는 것을 포함하는, 상기 대상체를 치료하는 방법을 제공한다.
- [0027] 일부 실시양태에서, 투여는 대상체의 CNS에 대한 직접 주사를 포함한다. 일부 실시양태에서, CNS에 대한 직접 주사는 대상체의 뇌척수액 (CSF)에 대한 직접 주사, 예를 들어 수조내 주사, 뇌실내 주사, 요추내 주사 또는 그의 임의의 조합을 포함한다. 일부 실시양태에서, 직접 주사는 뇌내 주사, 실질내 주사, 경막내 주사, 대수조내 주사 또는 그의 임의의 조합이다. 일부 실시양태에서, 직접 주사는 대류 증강 전달 (CED)을 포함한다.
- [0028] 일부 실시양태에서, 대상체는 포유동물, 예를 들어 인간 대상체이다. 일부 실시양태에서, 대상체는 약 30 내지 약 5000개 (예를 들어, 30 내지 5000 사이의 임의의 정수개)의 GGGGCC 디펩티드 반복부 및/또는 약 30 내지 5000개 (예를 들어, 30 내지 5000 사이의 임의의 정수개)의 CCCCGG 반복부를 갖는 것을 특징으로 한다. 일부 실시양태에서, 대상체는 5000개 초과 GGGGCC 디펩티드 반복부 및/또는 CCCCGG 반복부를 갖는 것을 특징으로 한다.

### 도면의 간단한 설명

- [0029] 도 1은 C9orf72의 반복부 확장을 표적화하는 억제 핵산, 막횡단 단백질 106B (TMEM106B)를 표적화하는 억제 핵산 및 야생형 C9orf72 코딩 서열을 코딩하는 발현 구축물을 포함하는 rAAV 벡터를 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다. rAAV 벡터는 발현 구축물에 플랭킹된 AAV 역전된 말단 반복부를 추가로 포함한다.
- 도 2는 C9orf72의 반복부 확장을 표적화하는 억제 핵산 및  $\beta$ -글루코세레브로시다제 (GBA1) 코딩 서열을 코딩하는 발현 구축물을 포함하는 rAAV를 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다. rAAV 벡터는 발현 구축물에 플랭킹된 AAV 역전된 말단 반복부를 추가로 포함한다.
- 도 3은 C9orf72의 반복부 확장을 표적화하는 억제 핵산 및 야생형 C9orf72 코딩 서열을 코딩하는 발현 구축물을 포함하는 rAAV 벡터를 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다. rAAV 벡터는 발현 구축물에 플랭킹된 AAV 역전된 말단 반복부를 추가로 포함한다.
- 도 4는 pol III (H1) 프로모터에 작동가능하게 연결된 ATXN2 (예를 들어, ATXN2를 코딩하는 유전자)를 표적화하는 억제 핵산을 코딩하는 발현 구축물을 포함하는 rAAV 벡터를 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다. rAAV 벡터는 발현 구축물에 플랭킹된 AAV 역전된 말단 반복부를 추가로 포함한다. 3'UTR의 "D" 서열은 "외부" 위치에 위치한다.
- 도 5는 pol II (CBA) 프로모터에 작동가능하게 연결된 ATXN2 (예를 들어, ATXN2를 코딩하는 유전자)를 표적화하는 억제 핵산을 코딩하는 발현 구축물을 포함하는 rAAV 벡터를 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다. rAAV 벡터는 발현 구축물에 플랭킹된 AAV 역전된 말단 반복부를 추가로 포함한다. 3'UTR의 "D" 서열은 "외부" 위치에 위치한다.
- 도 6은 pol II (CBA) 프로모터에 작동가능하게 연결된 ATXN2 (예를 들어, ATXN2를 코딩하는 유전자)를 표적화하는 억제 핵산을 코딩하는 발현 구축물을 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다.
- 도 7은 pol II (CBA) 프로모터에 작동가능하게 연결된, 각각의 1개가 ATXN2 (예를 들어, ATXN2를 코딩하는 유전자)를 표적화하는 2개의 억제 핵산을 코딩하는 발현 구축물을 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다.
- 도 8은 pol II (CBA) 프로모터에 작동가능하게 연결된 ATXN2 (예를 들어, ATXN2를 코딩하는 유전자)를 표적화하는 억제 핵산을 코딩하는 발현 구축물을 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다.



도 9는 pol II (CBA) 프로모터에 작동가능하게 연결된 ATXN2 (예를 들어, ATXN2를 코딩하는 유전자)를 표적화하는 억제 핵산 및 야생형 C9orf72 단백질을 코딩하는 코돈-최적화된 핵산 서열을 코딩하는 발현 구축물을 포함하는 rAAV 벡터를 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다. rAAV 벡터는 발현 구축물에 플랜킹된 AAV 역전된 말단 반복부를 추가로 포함한다. 3'UTR의 "D" 서열은 "외부" 위치에 위치한다.

도 10은 pol II (CBA) 프로모터에 작동가능하게 연결된 C9orf72를 표적화하는 억제 핵산 및 야생형 C9orf72 단백질을 코딩하는 코돈-최적화된 핵산 서열을 코딩하는 발현 구축물을 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다.

도 11은 pol II (CBA) 프로모터에 작동가능하게 연결된 C9orf72를 표적화하는 억제 핵산 및 야생형 C9orf72 단백질을 코딩하는 코돈-최적화된 핵산 서열을 코딩하는 발현 구축물을 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다.

도 12는 pol II (CBA) 프로모터에 작동가능하게 연결된 C9orf72를 표적화하는 억제 핵산 및 야생형 C9orf72 단백질을 코딩하는 코돈-최적화된 핵산 서열을 코딩하는 발현 구축물을 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다.

도 13은 RNA pol III (예를 들어, H1) 프로모터에 작동가능하게 연결된 C9orf72를 표적화하는 억제 핵산을 코딩하는 발현 구축물을 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다.

도 14는 pol II (CBA) 프로모터에 작동가능하게 연결된 C9orf72를 표적화하는 억제 핵산을 코딩하는 발현 구축물을 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다.

도 15는 pol II (CBA) 프로모터에 작동가능하게 연결된 C9orf72를 표적화하는 2개의 억제 핵산 및 야생형 C9orf72 단백질을 코딩하는 코돈-최적화된 핵산 서열을 코딩하는 발현 구축물을 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다.

도 16은 pol II (CBA) 프로모터에 작동가능하게 연결된 C9orf72를 표적화하는 2개의 억제 핵산 및 야생형 C9orf72 단백질을 코딩하는 코돈-최적화된 핵산 서열을 코딩하는 발현 구축물을 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다.

도 17은 pol II (CBA) 프로모터에 작동가능하게 연결된 C9orf72를 표적화하는 2개의 억제 핵산 및 야생형 C9orf72 단백질을 코딩하는 코돈-최적화된 핵산 서열을 코딩하는 발현 구축물을 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다.

도 18은 ITR의 "외부"에 (예를 들어, 트랜스진 삽입물 또는 발현 구축물 기준으로 ITR의 말단에 근접하게) 위치하는 "D" 영역을 포함하는 rAAV 벡터 (상단) 및 벡터의 "내부"에 (예를 들어, 벡터의 트랜스진 삽입물에 근접하게) ITR을 갖는 야생형 rAAV 벡터를 도시하는 개략도이다.

도 19a-19b는 시험관내 C9orf72 발현 및 녹다운 검정에 대한 대표적인 데이터를 보여준다. 도 19a는 rAAV 벡터에 의한 내인성 C9orf72의 통계적으로 유의한 침묵을 나타내는 대표적인 데이터를 보여준다. 도 19b는 rAAV 벡터에 의한 형질감염 후 야생형 C9orf72 발현의 통계적으로 유의한 증가를 나타내는 대표적인 데이터를 보여준다.

도 20은 pol II (CBA) 프로모터에 작동가능하게 연결된 RPS25를 표적화하는 억제 핵산 및 야생형 C9orf72 단백질을 코딩하는 코돈-최적화된 핵산 서열을 코딩하는 발현 구축물을 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다.

도 21은 pol II (CBA) 프로모터에 작동가능하게 연결된 RPS25를 표적화하는 억제 핵산 및 야생형 C9orf72 단백질을 코딩하는 코돈-최적화된 핵산 서열을 코딩하는 발현 구축물을 포함하는 플라스미드의 한 실시양태를 도시하는 개략도이다.

### 발명을 실시하기 위한 구체적인 내용

본 개시내용의 측면은 신경변성 질환, 예컨대 ALS/FTD, 파킨슨병, 알츠하이머병, 리소솜 축적 질환 및 루이 소체 치매의 치료를 위한 조성물 및 방법에 관한 것이다. 본 개시내용은 부분적으로 대상체에서 ALS/FTD-연관 유전자 산물 (예를 들어, C9orf72, ATXN2, TMEM106B, 상기 유전자를 표적화하는 억제 핵산 등) 및 그의 조합을 코딩하는 발현 구축물에 기초한다. 유전자 산물은 단백질, 단백질의 단편 (예를 들어, 부분), ALS/FTD-연관 유전

[0030]

자를 억제하는 간섭 핵산 동일 수 있다. 일부 실시양태에서, 유전자 산물은 ALS/FTD-연관 유전자에 의해 코딩되는 단백질 또는 단백질 단편이다. 일부 실시양태에서, 유전자 산물은 ALS/FTD-연관 유전자를 억제하는 간섭 핵산 (예를 들어, shRNA, siRNA, miRNA, amiRNA 등)이다.

[0031] ALS/FTD-연관 유전자는 근위축성 측삭 경화증 (ALS), 전두측두엽 치매 (FTD), 또는 ALS 및 FTD (ALS/FTD)와 유전자적으로, 생화학적으로 또는 기능적으로 연관된 유전자 산물을 코딩하는 유전자를 지칭한다. 예를 들어, C9orf72 유전자 내 23개 초과 GGGGCC 핵산뉴클레오타이드 반복부를 갖는 개체는 반복 영역 확장을 갖지 않는 개체와 비교하여 ALS/FTD가 발생할 증가된 위험을 갖는 것으로 관찰되었다. 일부 실시양태에서, 본원에 기재된 발현 카세트는 야생형 또는 비-돌연변이체 형태의 ALS/FTD-연관 유전자 (또는 그의 코딩 서열)를 코딩한다. 일반적으로, "야생형" 또는 "비-돌연변이체" 형태의 유전자는 정상 또는 비-병원성 활성과 연관된 단백질 (예를 들어, 신경변성 질환의 발병 또는 진행을 일으키는 반복 영역 확장과 같은 돌연변이 또는 변화가 결여된 단백질)을 코딩하는 핵산을 지칭한다. 예를 들어, 일부 실시양태에서, 야생형 C9orf72 단백질은 서열식별번호: 4에 제시된 서열을 포함하거나 이로 이루어진다.

[0032] 단리된 핵산 및 벡터

[0033] 단리된 핵산은 DNA 또는 RNA일 수 있다. 일부 측면에서, 본 개시내용은 1개 이상의 ALS/FTD-연관 유전자, 예를 들어 C9orf72 (예를 들어, C9orf72의 디펩티드-반복 영역), ATXN2, TMEM106B, RPS25 등을 표적화하는 1개 이상의 억제 핵산을 코딩하는 단리된 핵산 (예를 들어, rAAV 벡터)을 제공한다. 억제 핵산은 유전자의 센스 가닥 (예를 들어, 유전자로부터 전사된 mRNA), 유전자의 안티센스 가닥 (예를 들어, 유전자로부터 전사된 mRNA), 또는 유전자의 센스 및 안티센스 가닥 둘 다 (예를 들어, 유전자로부터 전사된 mRNA)를 표적화할 수 있다.

[0034] 일반적으로, 본원에 기재된 바와 같은 단리된 핵산은 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 그 초과 억제 핵산 (예를 들어, dsRNA, siRNA, shRNA, miRNA, amiRNA 등)을 코딩할 수 있다. 일부 실시양태에서, 단리된 핵산은 10개 초과 억제 핵산을 코딩한다. 일부 실시양태에서, 각각의 1개 이상의 억제 핵산은 상이한 유전자 또는 유전자의 부분을 표적화한다 (예를 들어, 제1 miRNA는 유전자의 제1 표적 서열을 표적화하고, 제2 miRNA는 제1 표적 서열과 상이한 유전자의 제2 표적 서열을 표적화함). 일부 실시양태에서, 각각의 1개 이상의 억제 핵산은 동일한 유전자의 동일한 표적 서열을 표적화한다 (예를 들어, 단리된 핵산은 동일한 miRNA의 다중 카피를 코딩함).

[0035] 본 개시내용의 측면은 C9orf72 단백질 (예를 들어, C9orf72 mRNA 전사체의 디펩티드-반복 영역)을 표적화하는 1개 이상의 간섭 핵산 (예를 들어, dsRNA, siRNA, miRNA, amiRNA 등)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. 일부 실시양태에서, 디펩티드-반복 영역은 핵산뉴클레오타이드 반복 서열 GGGGCC의 5개 이상의 중합체 단위 (예를 들어, GGGGCC 반복 서열의 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 200, 500, 1000개 또는 그 초과 반복부를 포함하는 영역)에 의해 코딩된다.

[0036] 일반적으로, C9orf72 단백질은 소형 GTPase, 예컨대 Rab에 대한 교환 인자로서 수반되는 것으로 생각되는 뉴런 및 시냅스 전 말단의 세포질에서 발견되는 단백질을 지칭한다. 인간에서, C9orf72 유전자는 염색체 9 상에 위치한다. 일부 실시양태에서, C9orf72 유전자는 NCBI 참조 서열 NP\_060795.1로 나타내어지는 펩티드를 코딩한다. 일부 실시양태에서, C9orf72 유전자는 서열식별번호: 3에 제시된 서열을 포함하거나 또는 서열식별번호: 4에 제시된 아미노산 서열을 코딩한다.

[0037] C9orf72를 표적화하는 억제 핵산은 6 내지 50개 뉴클레오타이드 길이인 상보성 영역 (예를 들어, 표적 유전자, 예컨대 C9orf72, 또는 표적 유전자의 부분, 예를 들어 C9orf72의 디펩티드 반복 영역, 또는 디펩티드-반복 영역의 외부 영역에 혼성화하는 억제 핵산의 영역)을 포함할 수 있다. 일부 실시양태에서, 억제 핵산은 약 6 내지 30, 약 8 내지 20 또는 약 10 내지 19개 뉴클레오타이드 길이인 C9orf72와의 상보성 영역을 포함한다. 일부 실시양태에서, 억제 핵산은 C9orf72 서열의 적어도 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 또는 25개의 인접 뉴클레오타이드와 상보적이다. 일부 실시양태에서, 억제 핵산에 의해 표적화된 (예를 들어, 결합된) C9orf72 서열은 C9orf72의 디펩티드-반복 영역 (이를 기준으로 5' 또는 3'로) 1개 뉴클레오타이드 내지 500개 뉴클레오타이드 (예를 들어, 1 내지 500 사이의 임의의 정수개의 뉴클레오타이드) 떨어져 있다. 일부 실시양태에서, 억제 핵산은 C9orf72 단백질을 코딩하는 유전자의 인트론 영역 (예를 들어, 비-단백질 코딩 영역)을 표적화한다.

[0038] 본 개시내용의 측면은 TMEM106B 단백질 (예를 들어, TMEM106B 유전자의 유전자 산물)을 표적화하는 1개 이상의 간섭 핵산 (예를 들어, dsRNA, siRNA, miRNA, amiRNA 등)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관

한 것이다. TMEM106B 단백질은 수상돌기 형태발생 및 리소솜 트래픽킹의 조절에 수반되는 단백질인 막횡단 단백질 106B를 지칭한다. 인간에서, TMEM106B 유전자는 염색체 7 상에 위치한다. 일부 실시양태에서, TMEM106B 유전자는 NCBI 참조 서열 NP\_060844.2로 나타내어지는 펩티드를 코딩한다. 일부 실시양태에서, TMEM106B 유전자는 서열식별번호: 7에 제시된 서열을 포함하거나 또는 서열식별번호: 6에 제시된 아미노산 서열을 코딩한다.

[0039] TMEM106B를 표적화하는 억제 핵산은 6 내지 50개 뉴클레오타이드 길이인 상보성 영역 (예를 들어, 표적 유전자, 예컨대 TMEM106B에 혼성화하는 억제 핵산의 영역)을 포함할 수 있다. 일부 실시양태에서, 억제 핵산은 약 6 내지 30, 약 8 내지 20 또는 약 10 내지 19개 뉴클레오타이드 길이인 TMEM106B와의 상보성 영역을 포함한다. 일부 실시양태에서, 억제 핵산은 TMEM106B 서열의 적어도 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 또는 25개의 인접 뉴클레오타이드와 상보적이다.

[0040] 본 개시내용의 측면은 ATXN2 단백질 (예를 들어, SCA2 유전자로도 지칭되는 ATXN2 유전자의 유전자 산물)을 표적화하는 1개 이상의 간접 핵산 (예를 들어, dsRNA, siRNA, miRNA, amiRNA 등)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. ATXN2 단백질은 폴리(A)-결합 단백질과의 상호작용을 통해 mRNA 번역을 조절하는데 수반되는 단백질인 아탁신 2를 지칭한다. 인간에서, ATXN2 유전자는 염색체 12 상에 위치한다. 일부 실시양태에서, ATXN2 유전자는 NCBI 참조 서열 NP\_002964.3으로 나타내어지는 펩티드를 코딩한다. 일부 실시양태에서, ATXN2 유전자는 서열식별번호: 9에 제시된 서열을 포함하거나 또는 서열식별번호: 8에 제시된 아미노산 서열을 코딩한다.

[0041] ATXN2를 표적화하는 억제 핵산은 6 내지 50개 뉴클레오타이드 길이인 상보성 영역 (예를 들어, 표적 유전자, 예컨대 ATXN2에 혼성화하는 억제 핵산의 영역)을 포함할 수 있다. 일부 실시양태에서, 억제 핵산은 약 6 내지 30, 약 8 내지 20 또는 약 10 내지 19개 뉴클레오타이드 길이인 ATXN2와의 상보성 영역을 포함한다. 일부 실시양태에서, 억제 핵산은 ATXN2 서열의 적어도 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 또는 25개의 인접 뉴클레오타이드와 상보적이다.

[0042] 본 개시내용의 측면은 리보솜 단백질 s25 (RPS25) (예를 들어, RPS25의 유전자 산물)를 표적화하는 1개 이상의 간접 핵산 (예를 들어, dsRNA, siRNA, miRNA, amiRNA 등)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. RPS25 단백질은 단백질 합성에 수반되는 단백질 복합체인 s40 리보솜의 서브유닛인 리보솜 단백질 을 지칭한다. 인간에서, RPS25 유전자는 염색체 11 상에 위치한다. 일부 실시양태에서, RPS25 유전자는 NCBI 참조 서열 NP\_001019.1로 나타내어지는 펩티드를 코딩한다. 일부 실시양태에서, RPS25 유전자는 서열식별번호: 60에 제시된 서열을 포함한다.

[0043] RPS25를 표적화하는 억제 핵산은 6 내지 50개 뉴클레오타이드 길이인 상보성 영역 (예를 들어, 표적 유전자, 예컨대 RPS25에 혼성화하는 억제 핵산의 영역)을 포함할 수 있다. 일부 실시양태에서, 억제 핵산은 약 6 내지 30, 약 8 내지 20 또는 약 10 내지 19개 뉴클레오타이드 길이인 RPS25와의 상보성 영역을 포함한다. 일부 실시양태에서, 억제 핵산은 RPS25 서열의 적어도 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 또는 25개의 인접 뉴클레오타이드와 상보적이다.

[0044] 본 개시내용의 측면은 1개 이상의 억제 핵산 (예를 들어, C9orf72의 디펩티드 반복 영역을 표적화하는 억제 핵산, C9orf72의 비-디펩티드 반복 영역을 표적화하는 억제 핵산, 및/또는 TMEM106B를 표적화하는 억제 핵산, 및/또는 ATXN2를 표적화하는 억제 핵산, 및/또는 RPS25를 표적화하는 억제 핵산 등)을 코딩하는 제1 유전자 산물 및 단백질, 예컨대 야생형 C9orf72 단백질 또는 GBA 단백질을 코딩하는 제2 유전자 산물을 포함하는 발현 구축물에 관한 것이다.

[0045] 일부 실시양태에서, 단리된 핵산은 C9orf72의 발현 또는 활성을 억제하는 제1 억제 핵산 및 TMEM106B의 발현 또는 활성을 억제하는 제2 억제 핵산을 코딩하는 발현 카세트를 포함한다.

[0046] 일부 실시양태에서, 단리된 핵산은 C9orf72의 발현 또는 활성을 억제하는 제1 억제 핵산 및 ATXN2의 발현 또는 활성을 억제하는 제2 억제 핵산을 코딩하는 발현 카세트를 포함한다.

[0047] 일부 실시양태에서, 단리된 핵산은 C9orf72의 발현 또는 활성을 억제하는 제1 억제 핵산 및 RPS25의 발현 또는 활성을 억제하는 제2 억제 핵산을 코딩하는 발현 카세트를 포함한다.

[0048] 일부 실시양태에서, 단리된 핵산은 C9orf72의 발현 또는 활성을 억제하는 억제 핵산 및  $\beta$ -글루코세레브로시다제 (GBA) 단백질을 코딩하는 발현 카세트를 포함한다. 일부 실시양태에서, GBA 단백질은 GBA1 단백질 (예를 들어, GBA1 유전자 또는 그의 부분에 의해 코딩되는 단백질)이다.



- [0049] 일부 실시양태에서, 단리된 핵산은 C9orf72의 발현 또는 활성을 억제하는 억제 핵산 및 야생형 C9orf72 단백질 (예를 들어, 병원성 디펩티드 반복부 확장이 결여된 C9orf72 단백질)을 코딩하는 발현 카세트를 포함한다. 일부 실시양태에서, 야생형 C9orf72 단백질을 코딩하는 핵산 서열 또는 그의 부분은 코돈-최적화된 핵산 서열이다. 일부 실시양태에서, 야생형 C9orf72 단백질은 서열식별번호: 3에 제시된 핵산 서열 또는 그의 부분에 의해 코딩된다. 일부 실시양태에서, 야생형 C9orf72 단백질은 서열식별번호: 4에 제시된 서열 또는 그의 부분을 포함하거나 이로 이루어진다. 일부 실시양태에서, 코돈-최적화된 C9orf72를 코딩하는 단리된 핵산은 서열 식별번호: 51에 제시된 서열을 포함하거나 이로 이루어진다.
- [0050] 관련 기술분야의 통상의 기술자는 제1 유전자 산물 (예를 들어, C9orf72 단백질 또는 GBA 단백질을 코딩하는 핵산 서열) 및 제2 유전자 산물 (예를 들어, C9orf72, ATXN2, TMEM106B 등을 표적화하는 억제 RNA)의 발현 순서가 일반적으로 역전될 수 있다는 것 (예를 들어, 억제 RNA가 제1 유전자 산물이고, 단백질 코딩 서열이 제2 유전자 산물임)을 인식한다. 일부 실시양태에서, 유전자 산물은 유전자 (예를 들어, C9orf72, TMEM106B, ATXN2, GBA1 등)의 단편 (예를 들어, 부분)이다. 단백질 단편은 단백질 (예를 들어, C9orf72 단백질, GBA 단백질 등)의 약 50%, 약 60%, 약 70%, 약 80%, 약 90% 또는 약 99%를 차지할 수 있다. 일부 실시양태에서, 단백질 단편은 C9orf72 단백질 또는 GBA 단백질의 50% 내지 99.9% (예를 들어, 50% 내지 99.9% 사이의 임의의 값)를 차지한다. 일부 실시양태에서, 유전자 산물 (예를 들어, 억제 RNA)은 표적 유전자의 부분에 혼성화한다 (예를 들어, 표적 유전자, 예를 들어 C9orf72, ATXN2 또는 TMEM106B의 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21개 또는 그 초과인 인접 뉴클레오타이드에 상보적임).
- [0051] 일부 실시양태에서, 발현 구축물은 모노시스트론이다 (예를 들어, 발현 구축물은 제1 유전자 산물 및 제2 유전자 산물을 포함하는 단일 융합 단백질을 코딩함). 일부 실시양태에서, 발현 구축물은 폴리시스트론이다 (예를 들어, 발현 구축물은 2개의 별개의 유전자 산물, 예를 들어 2개의 상이한 단백질 또는 단백질 단편을 코딩함).
- [0052] 폴리시스트론 발현 벡터는 1개 이상 (예를 들어, 1, 2, 3, 4, 5개 또는 그 초과)의 프로모터를 포함할 수 있다. 임의의 적합한 프로모터, 예를 들어 구성적 프로모터, 유도성 프로모터, 내인성 프로모터, 조직-특이적 프로모터 (예를 들어, CNS-특이적 프로모터) 등이 사용될 수 있다. 일부 실시양태에서, 프로모터는 닭 베타-액틴 프로모터 (CBA 프로모터), CAG 프로모터 (예를 들어, 문헌 [Alexopoulou et al. (2008) BMC Cell Biol. 9:2; doi: 10.1186/1471-2121-9-2]에 기재된 바와 같음), CD68 프로모터, 또는 JeT 프로모터 (예를 들어, 문헌 [Tornøe et al. (2002) Gene 297(1-2):21-32, 또는 Karumuthil-Melethil et al. (2016) Human Gene Therapy 27(7):509-521]에 기재된 바와 같음)이다. 일부 실시양태에서, 프로모터는 RNA pol II 프로모터 또는 RNA pol III 프로모터 (예를 들어, U6, H1 등)이다. 일부 실시양태에서, 프로모터는 제1 유전자 산물, 제2 유전자 산물 또는 제1 유전자 산물 및 제2 유전자 산물을 코딩하는 핵산 서열에 작동가능하게 연결된다. 일부 실시양태에서, 발현 카세트는 전사 인자 결합 서열, 인트론 스플라이스 부위, 폴리(A) 부가 부위, 인핸서 서열, 리프레서 결합 부위 또는 상기의 임의의 조합을 포함하나 이에 제한되지 않는 1개 이상의 추가의 조절 서열을 포함한다.
- [0053] 일부 실시양태에서, 제1 유전자 산물을 코딩하는 핵산 서열 및 제2 유전자 산물을 코딩하는 핵산 서열은 내부 리보솜 진입 부위 (IRES)를 코딩하는 핵산 서열에 의해 분리된다. IRES 부위의 예는, 예를 들어 문헌 [Mokrejs et al. (2006) Nucleic Acids Res. 34(Database issue):D125-30]에 기재되어 있다. 일부 실시양태에서, 제1 유전자 산물을 코딩하는 핵산 서열 및 제2 유전자 산물을 코딩하는 핵산 서열은 자기-절단 펩티드를 코딩하는 핵산 서열에 의해 분리된다. 자기-절단 펩티드의 예는 T2A, P2A, E2A, F2A, BmCPV 2A 및 BmIFV 2A, 및 문헌 [Liu et al. (2017) Sci Rep. 7: 2193]에 기재된 것을 포함하나 이에 제한되지 않는다. 일부 실시양태에서, 자기-절단 펩티드는 T2A 펩티드이다.
- [0054] 병리학적으로, ALS 및 FTD와 같은 장애는 C9orf72 유전자로부터 유래된 반복부-연관 비-ATG (RAN) 번역 단백질로 주로 구성된 단백질 응집체의 축적과 연관된다. 따라서, 일부 실시양태에서, 본원에 기재된 단리된 핵산은 C9orf72 단백질 (예를 들어, 병원성 디펩티드-반복부 확장을 갖는 유전자에 의해 코딩되는 C9orf72 단백질)의 발현을 감소시키거나 방지하는 억제 핵산을 포함한다. 억제 핵산을 코딩하는 서열은 발현 구축물의 비번역 영역 (예를 들어, 인트론, 5'UTR, 3'UTR 등)에 위치할 수 있다.
- [0055] 일부 실시양태에서, 억제 핵산은 발현 구축물의 인트론, 예를 들어 제1 유전자 산물을 코딩하는 서열의 상류의 인트론에 위치한다. 억제 핵산은 이중 가닥 RNA (dsRNA), siRNA, 마이크로 RNA (miRNA), 인공 miRNA (amiRNA), 또는 RNA aptamer일 수 있다. 일반적으로, 억제 핵산은 표적 RNA (예를 들어, mRNA)의 약 6 내지 약 30개 (예를 들어, 6 내지 30 사이의 임의의 정수개)의 인접 뉴클레오타이드에 결합한다 (예를 들어, 이와 혼성화

함). 일부 실시양태에서, 억제 핵산 분자는 miRNA 또는 amiRNA, 예를 들어 C9orf72 (병원성 C9orf72 단백질을 코딩하는 유전자)를 표적화하는 miRNA이다. 일부 실시양태에서, miRNA는 그것이 혼성화하는 C9orf72 mRNA의 영역과의 어떠한 미스매치도 포함하지 않는다 (예를 들어, miRNA는 "완벽함"). 일부 실시양태에서, miRNA는 그것이 혼성화하는 C9orf72 mRNA의 영역과의 2 내지 20개 (예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 또는 20개)의 미스매치, 예컨대 "돌출부"를 포함한다. 일부 실시양태에서, miRNA는 그것이 혼성화하는 C9orf72 mRNA의 영역과 20개 초과 미스매치를 포함한다.

[0056] 일부 실시양태에서, 억제 핵산은 shRNA (예를 들어, C9orf72를 표적화하는 shRNA)이다. 일부 실시양태에서, 억제 핵산은 miRNA (예를 들어, C9orf72를 표적화하는 miRNA)이다. 일부 실시양태에서, 발현 구축물의 1개 이상의 억제 RNA의 발현은 1개 이상의 RNA pol III 프로모터, 예를 들어 H1 프로모터 또는 U6 프로모터에 의해 구동된다. 각각의 억제 RNA는 상이한 프로모터 또는 동일한 프로모터에 의해 구동될 수 있다.

[0057] 일부 실시양태에서, 억제 핵산은 인공 마이크로RNA (amiRNA)이다. 마이크로RNA (miRNA)는 전형적으로 식물 및 동물에서 발견되는 소형 비-코딩 RNA를 지칭하고, 유전자 발현의 전사 및 번역후 조절에서 기능한다. miRNA는 RNA 폴리머라제에 의해 전사되어 프리-miRNA로 지칭되는 헤어핀-루프 구조를 형성하고, 이는 후속적으로 효소 (예를 들어, 드로샤, 파샤, 스플라이세오솜 등)에 의해 프리-miRNA 헤어핀 구조로 프로세싱되고, 이는 이어서 다이스에 의해 프로세싱되어 miRNA/miRNA\* 듀플렉스 (여기서 \*는 miRNA 듀플렉스의 패신저 가닥을 나타냄)를 형성하고, 이들 중 한 가닥은 이어서 RNA-유도된 침묵 복합체 (RISC) 내로 혼입된다. 일부 실시양태에서, 본원에 기재된 바와 같은 억제 RNA는 C9orf72 (예를 들어, C9orf72의 디펩티드-반복 영역 또는 C9orf72의 비-디펩티드-반복 영역), ATXN2 또는 TMEM106B를 표적화하는 miRNA이다.

[0058] 일부 실시양태에서, C9orf72를 표적화하는 억제 핵산은 miRNA/miRNA\* 듀플렉스를 포함한다. 일부 실시양태에서, miRNA/miRNA\* 듀플렉스의 miRNA 가닥은 서열식별번호: 24 또는 25, 37 또는 38, 40 또는 41 중 어느 하나에 제시된 서열 또는 그의 부분을 포함하거나 이로 이루어진다. 일부 실시양태에서, miRNA/miRNA\* 듀플렉스의 miRNA\* 가닥은 서열식별번호: 24 또는 25, 37 또는 38, 40 또는 41에 제시된 서열 또는 그의 부분을 포함하거나 이로 이루어진다.

[0059] 일부 실시양태에서, TMEM106B를 표적화하는 억제 핵산은 miRNA/miRNA\* 듀플렉스를 포함한다. 일부 실시양태에서, miRNA/miRNA\* 듀플렉스의 miRNA 가닥은 서열식별번호: 1 또는 7에 제시된 서열 또는 그의 부분을 포함하거나 이로 이루어진다. 일부 실시양태에서, miRNA/miRNA\* 듀플렉스의 miRNA\* 가닥은 서열식별번호: 1 또는 7에 제시된 서열 또는 그의 부분을 포함하거나 이로 이루어진다.

[0060] 일부 실시양태에서, ATXN2를 표적화하는 억제 핵산은 miRNA/miRNA\* 듀플렉스를 포함한다. 일부 실시양태에서, miRNA/miRNA\* 듀플렉스의 miRNA 가닥은 서열식별번호: 10-23 중 어느 하나에 제시된 서열 또는 그의 부분을 포함하거나 이로 이루어진다. 일부 실시양태에서, miRNA/miRNA\* 듀플렉스의 miRNA\* 가닥은 서열식별번호: 10-23 중 어느 하나에 제시된 서열 또는 그의 부분을 포함하거나 이로 이루어진다.

[0061] 인공 마이크로RNA (amiRNA)는 천연 miRNA를 변형하여 프리-mRNA의 천연 표적화 영역을 관심 표적화 영역으로 대체함으로써 유도된다. 예를 들어, 자연 발생적으로 발현된 miRNA는 스캐폴드 또는 백본 (예를 들어, 프리-miRNA 스캐폴드)으로서 사용될 수 있고, 스템 서열은 관심 유전자를 표적화하는 miRNA의 것으로 대체된다. 인공 전구체 마이크로RNA (프리-amiRNA)는 보통 하나의 단일 안정한 소형 RNA가 우선적으로 생성되도록 프로세싱된다. 일부 실시양태에서, 본원에 기재된 scAAV 벡터 및 scAAV는 amiRNA를 코딩하는 핵산을 포함한다. 일부 실시양태에서, amiRNA의 프리-miRNA 스캐폴드는 프리-MIR-21, 프리-MIR-22, 프리-MIR-26a, 프리-MIR-30a, 프리-MIR-33, 프리-MIR-122, 프리-MIR-375, 프리-MIR-199, 프리-MIR-99, 프리-MIR-194, 프리-MIR-155 및 프리-MIR-451로 이루어진 군으로부터 선택된 프리-miRNA로부터 유도된다. 일부 실시양태에서, amiRNA는 C9orf72, ATXN2 또는 TMEM106B를 표적화하는 핵산 서열 및 예를 들어 문헌 [Fowler et al. Nucleic Acids Res. 2016 Mar 18; 44(5): e48]에 기재된 바와 같은 eSIBR amiRNA 스캐폴드를 포함한다.

[0062] 일부 측면에서, 본 개시내용은 신경변성 질환 (예를 들어, ALS/FTD)의 치료를 위한 억제 RNA의 조합을 포함하는 발현 구축물에 관한 것이다. 예를 들어 일부 실시양태에서, 본 개시내용에 의해 기재된 발현 구축물은 C9orf72를 표적화하는 억제 RNA 및 막형단 단백질 106B (TMEM106B)를 표적화하는 억제 RNA를 포함한다. 단리된 핵산이 억제 핵산의 서열을 코딩하는 순서는 달라질 수 있다. 예를 들어, 단리된 핵산은 5' 말단에서 3' 말단으로, C9orf72 및 TMEM106B 또는 TMEM106B 및 C9orf72를 표적화하는 shRNA를 코딩할 수 있다.

[0063] 본원에 기재된 바와 같은 단리된 핵산은 그 자체로 또는 벡터의 일부로서 존재할 수 있다. 일반적으로, 벡터는

플라스미드, 코스미드, 파지미드, 박테리아 인공 염색체 (BAC), 또는 바이러스 벡터 (예를 들어, 아데노바이러스 벡터, 아데노-연관 바이러스 (AAV) 벡터, 레트로바이러스 벡터, 바쿨로바이러스 벡터 등)일 수 있다. 일부 실시양태에서, 벡터는 플라스미드 (예를 들어, 본원에 기재된 바와 같은 단리된 핵산을 포함하는 플라스미드)이다. 일부 실시양태에서, 벡터는 재조합 AAV (rAAV) 벡터 (예를 들어, AAV ITR이 플랭킹된 트랜스진을 코딩하는 발현 구축물)이다. 일부 실시양태에서, rAAV 벡터는 단일 가닥 (예를 들어, 단일 가닥 DNA)이다. 일부 실시양태에서, 벡터는 바쿨로바이러스 벡터 (예를 들어, 아우토그라파 칼리포르니카(*Autographa californica*) 핵 다각체병 (AcNPV) 벡터)이다.

[0064] 전형적으로 rAAV 벡터는 2개의 AAV 역전된 말단 반복부 (ITR) 서열이 플랭킹된 트랜스진을 포함한다. 일부 실시양태에서, rAAV 벡터의 트랜스진은 본 개시내용에 의해 기재된 바와 같은 단리된 핵산을 포함한다. 일부 실시양태에서, rAAV 벡터의 2개의 ITR 서열 각각은 전장 ITR (예를 들어, 대략 145 bp 길이이고, 기능적 Rep 결합 부위 (RBS) 및 말단 분해 부위 (trs)를 함유함)이다. 일부 실시양태에서, rAAV 벡터의 ITR 중 하나는 말단절단된다 (예를 들어, 단축되거나 전장이 아님). 일부 실시양태에서, 말단절단된 ITR은 기능적 말단 분해 부위 (trs)가 결여되고, 자기-상보적 AAV 벡터 (scAAV 벡터)의 생산에 사용된다. 일부 실시양태에서, 말단절단된 ITR은, 예를 들어 문헌 [McCarty et al. (2003) Gene Ther. 10(26):2112-8]에 기재된 바와 같은  $\Delta$ ITR이다.

[0065] 본 개시내용의 측면은 야생형 AAV ITR에 비해, 예를 들어 야생형 AAV2 ITR (예를 들어, 서열식별번호: 32)에 비해 1개 이상의 변형 (예를 들어, 핵산 부가, 결실, 치환 등)을 갖는 ITR을 포함하는 단리된 핵산 (예를 들어, rAAV 벡터)에 관한 것이다. 야생형 AAV2 ITR의 구조가 도 18에 제시된다. 일반적으로, 야생형 ITR은, 2개의 교차 아암 (각각 B/B' 및 C/C'로 지칭되는 서열에 의해 형성됨), 보다 긴 스템 영역 (서열 A/A'에 의해 형성됨), 및 "D" 영역으로 지칭되는 단일-가닥 말단 영역으로 이루어진 회문식 이중 가닥 T 형상의 헤어핀 구조를 형성하도록 자기 어닐링하는 125개의 뉴클레오티드 영역을 포함한다 (도 18). 일반적으로, ITR의 "D" 영역은 A/A' 서열에 의해 형성된 스템 영역과 rAAV 벡터의 트랜스진을 함유하는 삽입물 사이에 위치한다 (예를 들어, ITR의 말단 기준으로 ITR의 "내부"에 또는 rAAV 벡터의 트랜스진 삽입물 또는 발현 구축물에 근접하게 위치함). 일부 실시양태에서, "D" 영역은 서열식별번호: 30에 제시된 서열을 포함한다. "D" 영역은, 예를 들어 문헌 [Ling et al. (2015) J Mol Genet Med 9(3)]에 개시된 바와 같이, 캡시드 단백질에 의한 rAAV 벡터의 캡시드화에서 중요한 역할을 하는 것으로 관찰되었다.

[0066] 본 개시내용은 부분적으로 ITR의 "외부"에 (예를 들어, 트랜스진 삽입물 또는 발현 구축물 기준으로 ITR의 말단에 근접하게) 위치하는 "D" 영역을 포함하는 rAAV 벡터가 비변형된 (예를 들어, 야생형) ITR로 ITR을 갖는 rAAV 벡터보다 AAV 캡시드 단백질에 의해 효율적으로 캡시드화된다는 놀라운 발견에 기초한다. 일부 실시양태에서, 변형된 "D" 서열 (예를 들어, "외부" 위치에 있는 "D" 서열)을 갖는 rAAV 벡터는 야생형 ITR 서열을 갖는 rAAV 벡터에 비해 감소된 독성을 갖는다.

[0067] 일부 실시양태에서, 변형된 "D" 서열은 야생형 "D" 서열 (예를 들어, 서열식별번호: 30)에 비해 적어도 1개의 뉴클레오티드 치환을 포함한다. 변형된 "D" 서열은 야생형 "D" 서열 (예를 들어, 서열식별번호: 30)에 비해 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 10개 초과 뉴클레오티드 치환을 가질 수 있다. 일부 실시양태에서, 변형된 "D" 서열은 야생형 "D" 서열 (예를 들어, 서열식별번호: 30)에 비해 적어도 10, 11, 12, 13, 14, 15, 16, 17, 18 또는 19개의 핵산 치환을 포함한다. 일부 실시양태에서, 변형된 "D" 서열은 야생형 "D" 서열 (예를 들어, 서열식별번호: 30)에 대해 약 10% 내지 약 99% (예를 들어, 10%, 15%, 20%, 25%, 30%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% 또는 99%) 동일하다. 일부 실시양태에서, 변형된 "D" 서열은 문헌 [Wang et al. (1995) J Mol Biol 250(5):573-80]에 기재된 바와 같은 "S" 서열로도 지칭되는 서열식별번호: 29에 제시된 서열을 포함한다.

[0068] 본 개시내용에 의해 기재된 바와 같은 단리된 핵산 또는 rAAV 벡터는, 문헌 [Francois, et al. 2005. The Cellular TATA Binding Protein Is Required for Rep-Dependent Replication of a Minimal Adeno-Associated Virus Type 2 p5 Element. J Virol]에 기재된 바와 같이, 예를 들어 서열식별번호: 31에 제시된 바와 같은 "TRY" 서열을 추가로 포함할 수 있다. 일부 실시양태에서, TRY 서열은 단리된 핵산 또는 rAAV 벡터의 ITR (예를 들어, 5' ITR)과 발현 구축물 (예를 들어, 트랜스진-코딩 삽입물) 사이에 위치한다.

[0069] 일부 측면에서, 본 개시내용은 본 개시내용에 의해 기재된 바와 같은 단리된 핵산 또는 rAAV 벡터를 포함하는 바쿨로바이러스 벡터에 관한 것이다. 일부 실시양태에서, 바쿨로바이러스 벡터는, 예를 들어 문헌 [Urabe et al. (2002) Hum Gene Ther 13(16):1935-43 및 Smith et al. (2009) Mol Ther 17(11):1888-1896]에 기재된 바와 같은 아우토그라파 칼리포르니카 핵 다각체병 (AcNPV) 벡터이다.

- [0070] 일부 측면에서, 본 개시내용은 본원에 기재된 바와 같은 단리된 핵산 또는 벡터를 포함하는 숙주 세포를 제공한다. 숙주 세포는 원핵 세포 또는 진핵 세포일 수 있다. 예를 들어, 숙주 세포는 포유동물 세포, 박테리아 세포, 효모 세포, 곤충 세포 등일 수 있다. 일부 실시양태에서, 숙주 세포는 포유동물 세포, 예를 들어 HEK293T 세포이다. 일부 실시양태에서, 숙주 세포는 박테리아 세포, 예를 들어 이. 콜라이(*E. coli*) 세포이다.
- [0071] rAAV
- [0072] 일부 측면에서, 본 개시내용은 본원에 기재된 바와 같은 핵산을 코딩하는 트랜스진을 포함하는 재조합 AAV (rAAV) (예를 들어, 본원에 기재된 바와 같은 rAAV 벡터)에 관한 것이다. 용어 "rAAV"는 일반적으로 1개 이상의 AAV 캡시드 단백질에 의해 캡시드화된 rAAV 벡터를 포함하는 바이러스 입자를 지칭한다. 본 개시내용에 의해 기재된 rAAV는 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9 및 AAV10으로부터 선택된 혈청형을 갖는 캡시드 단백질을 포함할 수 있다. 일부 실시양태에서, rAAV는 비-인간 숙주로부터의 캡시드 단백질, 예를 들어 레서스 AAV 캡시드 단백질, 예컨대 AAVrh.10, AAVrh.39 등을 포함한다. 일부 실시양태에서, 본 개시내용에 의해 기재된 rAAV는 야생형 캡시드 단백질의 변이체인 캡시드 단백질, 예컨대 그것이 유도된 야생형 AAV 캡시드 단백질에 비해 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 10개 초과 (예를 들어, 15, 20, 25, 50, 100개 등)의 아미노산 치환 (예를 들어, 돌연변이)을 포함하는 캡시드 단백질 변이체를 포함한다.
- [0073] 일부 실시양태에서, 본 개시내용에 의해 기재된 rAAV는 특히 CSF 공간 내로 또는 직접 뇌 실질 내로 도입될 때 CNS를 통해 용이하게 확산된다. 따라서, 일부 실시양태에서, 본 개시내용에 의해 기재된 rAAV는 혈액-뇌 장벽 (BBB)을 가로지를 수 있는 캡시드 단백질을 포함한다. 예를 들어, 일부 실시양태에서, rAAV는 AAV9 또는 AAVrh.10 혈청형을 갖는 캡시드 단백질을 포함한다. 일부 실시양태에서, rAAV는 문헌 [Deverman et al. (2016) *Nature Biotechnology* 34:204-209]에 기재된 바와 같이, 혈액-뇌 장벽을 가로지르는 AAV9 변이체, 예를 들어 AAV-PHP.B 혈청형을 포함한다. 일반적으로, rAAV의 생산은, 예를 들어 문헌 [Samulski et al. (1989) *J Virol.* 63(9):3822-8 및 Wright (2009) *Hum Gene Ther.* 20(7): 698-706]에 기재되어 있다.
- [0074] 일부 실시양태에서, 본 개시내용에 의해 기재된 바와 같은 rAAV (예를 들어, AAV 캡시드 단백질에 의해 캡시드화되어 rAAV 캡시드 입자를 형성하는 재조합 rAAV 게놈을 포함하는 것)는 바클로바이러스 벡터 발현 시스템 (BEVS)에서 생산된다. BEVS를 사용하는 rAAV의 생산은, 예를 들어 문헌 [Urabe et al. (2002) *Hum Gene Ther* 13(16):1935-43, Smith et al. (2009) *Mol Ther* 17(11):1888-1896], 미국 특허 번호 8,945,918, 미국 특허 번호 9,879,282 및 국제 PCT 공개 WO 2017/184879에 기재되어 있다. 그러나, rAAV는 임의의 적합한 방법을 사용하여 (예를 들어, 재조합 rep 및 cap 유전자를 사용하여) 생산될 수 있다.
- [0075] 제약 조성물
- [0076] 일부 측면에서, 본 개시내용은 본원에 기재된 바와 같은 단리된 핵산 또는 rAAV 및 제약상 허용되는 담체를 포함하는 제약 조성물을 제공한다. 본원에 사용된 용어 "제약상 허용되는"은 화합물의 생물학적 활성 또는 특성을 제거하지 않고 비교적 비독성인 물질, 예컨대 담체 또는 희석제를 지칭하며, 예를 들어 물질은 바람직하지 않은 생물학적 효과를 야기하지 않으면서 또는 그것이 함유된 조성물의 성분 중 임의의 성분과 유해한 방식으로 상호작용하지 않으면서 개체에게 투여될 수 있다.
- [0077] 본원에 사용된 용어 "제약상 허용되는 담체"는 본 발명 내에서 유용한 화합물을 환자 내에 또는 환자에게 운반 또는 수송하여 그의 의도된 기능을 수행할 수 있도록 하는데 수반되는 제약상 허용되는 물질, 조성물 또는 담체, 예컨대 액체 또는 고체 충전제, 안정화제, 분산제, 현탁화제, 희석제, 부형제, 증점제, 용매 또는 캡슐화 물질을 의미한다. 본 발명의 실시예에 사용되는 제약 조성물에 포함될 수 있는 추가의 성분은 관련 기술분야에 공지되어 있고, 예를 들어 문헌 [Remington's Pharmaceutical Sciences (Genaro, Ed., Mack Publishing Co., 1985, Easton, PA)]에 기재되어 있으며, 이는 본원에 참조로 포함된다.
- [0078] 본원에 제공된 조성물 (예를 들어, 제약 조성물)은 임의의 경로, 예컨대 경장 (예를 들어, 경구), 비경구, 정맥 내, 근육 내, 동맥 내, 수질 내, 경막 내, 대수조 내, 피하, 뇌실 내, 경피, 피내, 직장, 질 내, 복강 내, 국소 (분말, 연고, 크림 및/또는 점적제에 의해서와 같이), 점막, 비강, 협측, 설하에 의해; 기관 내 점적주입, 기관지 점적주입 및/또는 흡입에 의해; 및/또는 구강 스프레이, 비강 스프레이 및/또는 에어로졸로서 투여될 수 있다. 구체적으로 고려되는 경로는 경구 투여, 정맥 내 투여 (예를 들어, 전신 정맥 내 주사), 혈액 및/또는 림프 공급을 통한 국부 투여, 및/또는 이환 부위에 대한 직접 투여이다. 일반적으로, 가장 적절한 투여 경로는 작용제의 성질 (예를 들어, 위장관 환경에서의 그의 안정성) 및/또는 대상체의 상태 (예를 들어, 대상체가 경구 투여를 허용할 수 있는지의 여부)를 포함한 다양한 인자에 좌우될 것이다.



- [0079] 방법
- [0080] 본 개시내용은 부분적으로 대상체에서 신경변성 질환 (예를 들어, ALS/FTD 등)을 치료하기 위해 함께 작용하는 (예를 들어, 상승작용적으로) 1개 이상의 ALS-FTD-연관 유전자 산물 (또는 그의 조합)의 발현을 위한 조성물에 기초한다. 본원에 사용된 "치료하다" 또는 "치료하는"은 (a) 신경변성 질환 (예를 들어, ALS/FTD, 알츠하이머병, 고셔병, 파킨슨병, 루이 소체 치매, 리소솜 축적 질환 등)의 발병을 예방하거나 지연시키는 것; (b) 신경변성 질환의 중증도를 감소시키는 것; (c) 신경변성 질환에 특징적인 증상의 발생을 감소시키거나 예방하는 것; 및/또는 (d) 신경변성 질환에 특징적인 증상의 악화를 예방하는 것을 지칭한다. 예를 들어, ALS/FTD의 증상은, 예를 들어 운동 기능장애 (예를 들어, 마비, 떨림, 경직, 동작 둔화, 보행 곤란), 인지 기능장애 (예를 들어, 치매, 우울증, 불안증), 감정 및 행동 기능장애를 포함한다.
- [0081] 따라서, 일부 측면에서, 본 개시내용은 신경변성 질환을 갖거나 갖는 것으로 의심되는 대상체에게 본 개시내용에 의해 기재된 바와 같은 조성물 (예를 들어, 단리된 핵산 또는 벡터 또는 rAAV를 포함하는 조성물)을 투여하는 것을 포함하는, 상기 대상체를 치료하는 방법을 제공한다. 일부 실시양태에서, 신경변성 질환은 ALS/FTD, 알츠하이머병, 고셔병, 파킨슨병, 루이 소체 치매 또는 리소솜 축적 질환이다.
- [0082] 일부 실시양태에서, 조성물은 대상체의 CNS에 직접적으로, 예를 들어 대상체의 뇌 및/또는 척수 내로의 직접 주사에 의해 투여된다. CNS-직접 투여 양식의 예는 뇌내 주사, 뇌실내 주사, 수조내 주사, 실질내 주사, 경막내 주사 및 상기의 임의의 조합을 포함하나 이에 제한되지는 않는다. 일부 실시양태에서, 대상체의 CNS 내로의 직접 주사는 대상체의 중뇌, 선조체 및/또는 뇌 피질에서 트랜스진 발현 (예를 들어, 제1 유전자 산물, 제2 유전자 산물 및 적용가능한 경우 제3 유전자 산물의 발현)을 유발한다.
- [0083] 일부 실시양태에서, 본 개시내용에 의해 기재된 바와 같은 조성물은 대상체의 뇌척수액 (CSF)에 직접 투여된다. 일부 실시양태에서, CSF 내로의 직접 주사는 대상체의 척수 및/또는 CSF에서 트랜스진 발현 (예를 들어, 제1 유전자 산물, 제2 유전자 산물 및 적용가능한 경우 제3 유전자 산물의 발현)을 유발한다. 대상체의 CSF에 대한 직접 투여의 예는 수조내 주사, 뇌실내 주사, 요추내 주사 또는 그의 임의의 조합을 포함하나 이에 제한되지는 않는다.
- [0084] 일부 실시양태에서, 대상체의 CNS에 대한 직접 주사는 대류 증강 전달 (CED)을 포함한다. 대류 증강 전달은 뇌를 외과적으로 노출시키고, 뇌의 표적 영역 내로 직접 소-직경 카테터를 배치한 후, 이어서 대상체의 뇌로 직접 치료제 (예를 들어, 본원에 기재된 바와 같은 조성물 또는 rAAV)를 주입하는 것을 포함하는 치료 전략이다. CED는, 예를 들어 문헌 [Debinski et al. (2009) Expert Rev Neurother. 9(10):1519-27]에 기재되어 있다.
- [0085] 일부 실시양태에서, 조성물은 대상체에게 말초로, 예를 들어 말초 주사에 의해 투여된다. 말초 주사의 예는 피하 주사, 정맥내 주사, 동맥내 주사, 복강내 주사 또는 상기의 임의의 조합을 포함한다. 일부 실시양태에서, 말초 주사는 동맥내 주사, 예를 들어 대상체의 경동맥 내로의 주사이다.
- [0086] 일부 실시양태에서, 본 개시내용에 의해 기재된 바와 같은 조성물 (예를 들어, 단리된 핵산 또는 벡터 또는 rAAV를 포함하는 조성물)은 대상체의 CNS에 직접적으로 및 말초로 둘 다로 투여된다. 예를 들어, 일부 실시양태에서, 대상체에게 동맥내 주사 (예를 들어, 경동맥 내로의 주사) 및 실질내 주사 (예를 들어, CED에 의한 실질내 주사)에 의해 조성물을 투여한다. 일부 실시양태에서, CNS에 대한 직접 주사 및 말초 주사는 동시이다 (예를 들어, 동시에 일어난다). 일부 실시양태에서, 직접 주사는 말초 주사 전에 (예를 들어, 1분 내지 1주 또는 그 초과 전에) 이루어진다. 일부 실시양태에서, 직접 주사는 말초 주사 후에 (예를 들어, 1분 내지 1주 또는 그 초과 후에) 이루어진다.
- [0087] 대상체에게 투여되는 본 개시내용에 의해 기재된 바와 같은 조성물 (예를 들어, 단리된 핵산 또는 벡터 또는 rAAV를 포함하는 조성물)의 양은 투여 방법에 따라 달라질 것이다. 예를 들어, 일부 실시양태에서, 본원에 기재된 바와 같은 rAAV는 약  $10^9$  게놈 카피 (GC)/kg 내지 약  $10^{14}$  GC/kg (예를 들어, 약  $10^9$  GC/kg, 약  $10^{10}$  GC/kg, 약  $10^{11}$  GC/kg, 약  $10^{12}$  GC/kg, 약  $10^{13}$  GC/kg, 또는 약  $10^{14}$  GC/kg)의 역가로 대상체에게 투여된다. 일부 실시양태에서, 대상체에게 CSF 공간으로의 주사에 의해 또는 실질내 주사에 의해 높은 역가로 (예를 들어,  $>10^{12}$  게놈 카피 GC/kg (rAAV)) 투여된다.
- [0088] 본 개시내용에 의해 기재된 바와 같은 조성물 (예를 들어, 단리된 핵산 또는 벡터 또는 rAAV를 포함하는 조성물)은 대상체에게 1회 또는 다수회 (예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20회 또는 그 초과) 투여될 수 있다. 일부 실시양태에서, 조성물은 대상체에게, 예를 들어 주입 펌프를 통해 연속적으로 (예를 들어, 만성

적으로) 투여된다.

[0089] 실시예

[0090] 실시예 1: rAAV 벡터

[0091] 삼중-플라스미드 형질감염을 위해 세포, 예컨대 HEK293 세포를 사용하여 AAV 벡터를 생성한다. ITR 서열은 각각의 관심 트랜스진에 대한 프로모터/인핸서 요소, 3' 폴리A 신호, 및 번역 후 신호, 예컨대 WPRE 요소를 포함하는 발현 구축물에 플랜킹된다. 예를 들어, 단일 발현 카세트 또는 개별 발현 카세트를 사용한 발현에 의해 C9orf72 단백질 또는 GBA1 단백질 및 1개 이상의 억제 핵산 (예를 들어, C9orf72 및/또는 TMEM106B를 표적화하는 억제 핵산)과 같은 다중 유전자 산물이 동시에 발현될 수 있다. 발현되는 유전자의 상류에 효율적으로 스폰 라이싱되는 짧은 인트론 서열의 존재는 발현 수준을 개선시킬 수 있다. 이들 서열 내에 억제 RNA (예를 들어, miRNA, shRNA 등) 및 다른 조절 RNA가 잠재적으로 포함될 수 있다. 본 개시내용에 의해 기재된 발현 구축물의 예는 도 1-17 및 20-21 및 하기 표 1에 제시된다.

[0092] 표 1.

| 명칭                                                                | 프로모터 1 | shRNA            | CDS1    | 폴리A 1    | 프로모터 2 | CDS2 | 폴리A 2 | ITR 사이의 길이 |
|-------------------------------------------------------------------|--------|------------------|---------|----------|--------|------|-------|------------|
| 프리베일벡터_13_CMVe_CBAp_shRNAC9_mRNAiTMEM106B_C9ORF72_WPRE_bGH_4993nt | CBA    | C9반복부 & TMEM106B | C9ORF72 | WPRE_bGH |        |      |       | 4993       |
| 프리베일벡터_13H_H1_C9sh_CMVe_CBAp_c9ORF72_WPRE_bGH_3944nt              | H1     | C9반복부            | C9orf72 |          | CBA    |      |       | 3944       |
| 프리베일벡터_0_CMVe_CBAp_shRNAC9_GBA1_WPRE_bGH_3892nt                   | CBA    | C9orf72          | GBA1    | WPRE_bGH |        |      |       | 3892       |
| H1_ATXN2_sh_sen                                                   | H1     | ATXN2            |         |          |        |      |       | 899        |
| 인트론_ATXN2_sh_sen                                                  | CBA    | ATXN2            |         | WPRE_bGH |        |      |       | 2547       |
| 인트론_C9반복부_ATNX2_sh_C9ORF71_I00037                                 | CBA    | C9반복부_ATXN2      | C9Orf72 | WPRE_bGH |        |      |       | 4145       |
| 인트론_eSIBR_항_C9_인트론ShL3_인트론ShR3_CMVe_CBAp_C9ORF72_WPRE_bGH_I00135  | CBA    | C9orf72          | C9orf72 | WPRE_bGH |        |      |       | 4140       |
| 인트론_eSIBR_항_C9_인트론ShL1_인트론ShR1_CMVe_CBAp_C9ORF72_WPRE_bGH_I00133  | CBA    | C9orf72          | C9orf72 | WPRE_bGH |        |      |       | 4143       |
| 인트론_C9반복부shRNA_I00066                                             | CBA    | C9orf72          |         | WPRE_bGH |        |      |       | 2250       |
| 인트론_eSIBR_항_C9_인트론ShL2_인트론ShR2_CMVe_CBAp_C9ORF72_WPRE_bGH_I00134  | CBA    | C9orf72          | C9orf72 | WPRE_bGH |        |      |       | 4142       |
| 인트론_C9반복부_sh_C9ORF72_I00031                                       | CBA    | C9orf72          | C9orf72 | WPRE_bGH |        |      |       | 3997       |
| 인트론_C9유효_sh_C9ORF72_I00032                                        | CBA    | C9orf72          | C9orf72 | WPRE_bGH |        |      |       | 3994       |
| H1_C9반복부shRNA_I00069                                              | H1     | C9orf72          |         |          |        |      |       | 902        |
| 인트론_C9보존_sh_C9ORF72_I00030                                        | CBA    | C9orf72          | C9orf72 | WPRE_bGH |        |      |       | 3994       |

[0093]

[0094] 실시예 2: ALS/FTD 세포 내로의 바이러스 형질도입의 세포 기반 검정

[0095] 예를 들어 ALS/FTD 환자로부터의 섬유모세포, 단핵구 또는 hES 세포, 또는 환자-유래 유도 만능 줄기 세포

(iPSC)로서 C9orf72의 반복부 확장을 특징으로 하는 세포를 수득한다. 이들 세포는 RNA 뭉침 및 RAN 번역된 단백질 축적을 축적한다.

[0096] 이러한 세포 모델을 사용하여, 세포 병리학을 RAN 단백질과 같은 단백질 응집체의 축적에 관하여 항-RAN 단백질 항체로 정량화한 후, 이어서 형광 현미경검사를 사용하여 영상화한다. 웨스턴 블롯팅 및/또는 ELISA를 사용하여 RAN 단백질의 비정상적인 축적을 정량화한다.

[0097] 활성 및 기능을 확인하고 정량화하기 위해, AAV 벡터의 형질도입의 발현과 관련하여 치료 종점 (예를 들어, ALS/FTD-연관 병리상태의 감소)을 측정한다. 예를 들어 정량적 RT-PCT (qRT-PCR)을 사용하여 내인성 (예를 들어, 병원성, 반복부 확장-함유) C9orf72 mRNA 수준의 감소를 정량화할 수 있다.

[0098] 실시예 3: 시험관내 연구

[0099] 본 실시예는 본 개시내용에 의해 기재된 C9orf72 rAAV 벡터의 시험관내 시험을 기재한다. 포유동물 세포에서 C9orf72 녹다운 및 과다발현의 효과를 연구하였다. 시험된 구축물의 예가 표 2에 열거되어 있다.

| ID     | 프로모터    | 녹다운             | 프로모터 | 과다발현   |
|--------|---------|-----------------|------|--------|
| I00017 | H1      | C9_sh           | CMV  | opt-C9 |
| I00018 | CMV     | C9_sh, TMEM_mi  | CMV  | opt-C9 |
| I00030 | CMV_인트론 | C9_sh(보존)       | CMV  | opt-C9 |
| I00031 | CMV_인트론 | C9반복부_sh        | CMV  | opt-C9 |
| I00032 | CMV_인트론 | C9_sh(유효)       | CMV  | opt-C9 |
| I00037 | CMV_인트론 | C9r_sh, ATXN_sh | CMV  | opt-C9 |

[0100]

[0101] 정량적 PCR (qPCR) 및 ELISA에 의해 유전자 녹다운 및 과다발현을 검정하였다. 도 19a는 rAAV 벡터에 의한 C9orf72의 통계적으로 유의한 침묵을 나타내는 대표적인 데이터를 보여준다. 도 19b는 rAAV 벡터에 의한 형질 감염 후 야생형 C9orf72 발현의 통계적으로 유의한 증가를 나타내는 대표적인 데이터를 보여준다.

[0102] 서열

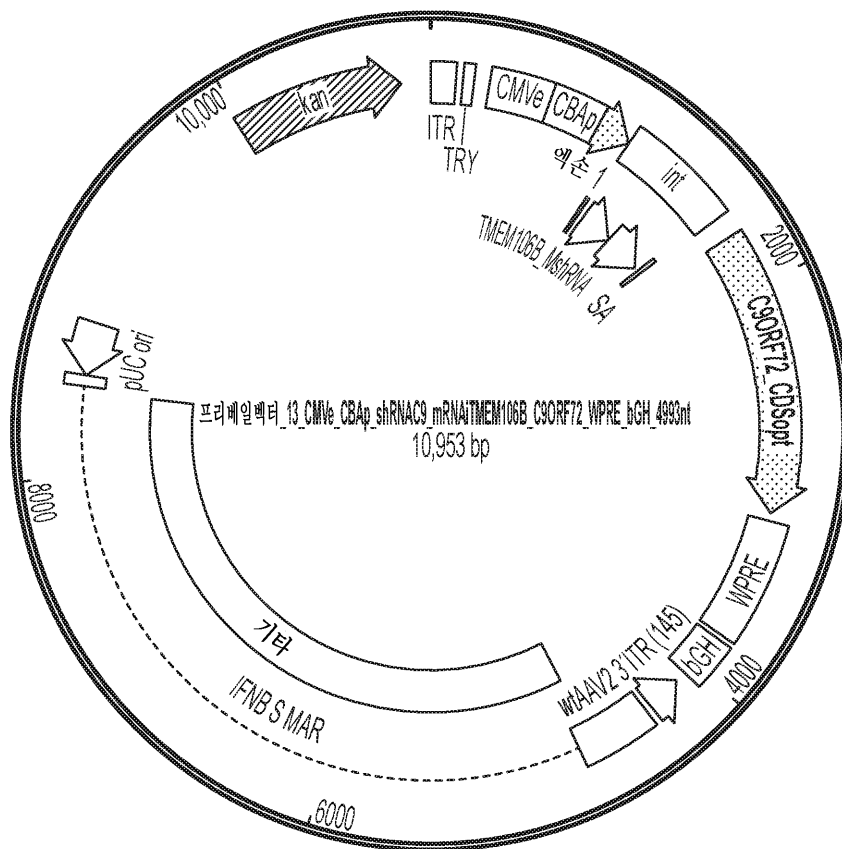
[0103] 일부 실시양태에서, 1개 이상의 유전자 산물 (예를 들어, 제1, 제2 및/또는 제3 유전자 산물)을 코딩하는 발현 카세트는 서열식별번호 1-62 중 어느 하나에 제시된 서열을 포함하거나 이로 이루어진다 (또는 이를 코딩함). 일부 실시양태에서, 유전자 산물은 서열식별번호: 1-62 중 어느 하나의 부분 (예를 들어, 단편)을 포함하거나 이로 이루어진다 (또는 이에 의해 코딩됨). 일부 실시양태에서, 하기 서열 내의 "T" 뉴클레오티드는, 예를 들어 RNA 분자와 관련하여 "U" 뉴클레오티드로 대체된다.

[0104] 관련 기술분야의 통상의 기술자는 상기 서열의 "부분"이 플라스미드 백본 (예를 들어, 복제 기점 서열, 선택 마커 서열 등)이 결합된 발현 카세트의 서열 (예를 들어, ITR을 코딩하는 서열, 간섭 RNA, 코딩 서열, 조절 서열 등)일 수 있음을 인지할 것이다.

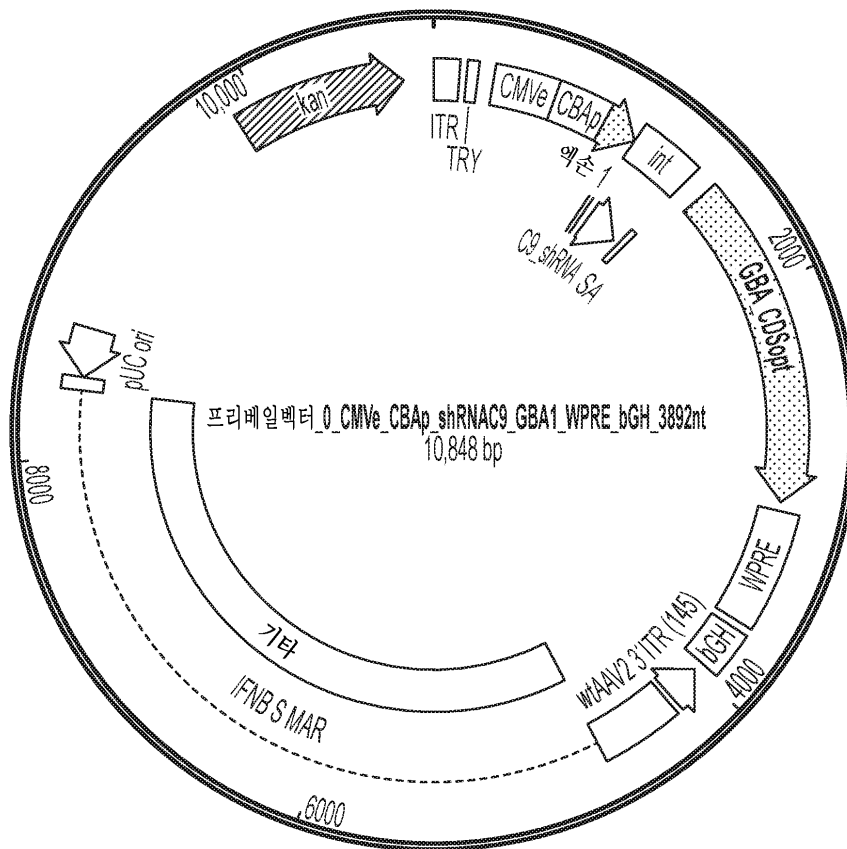


도면

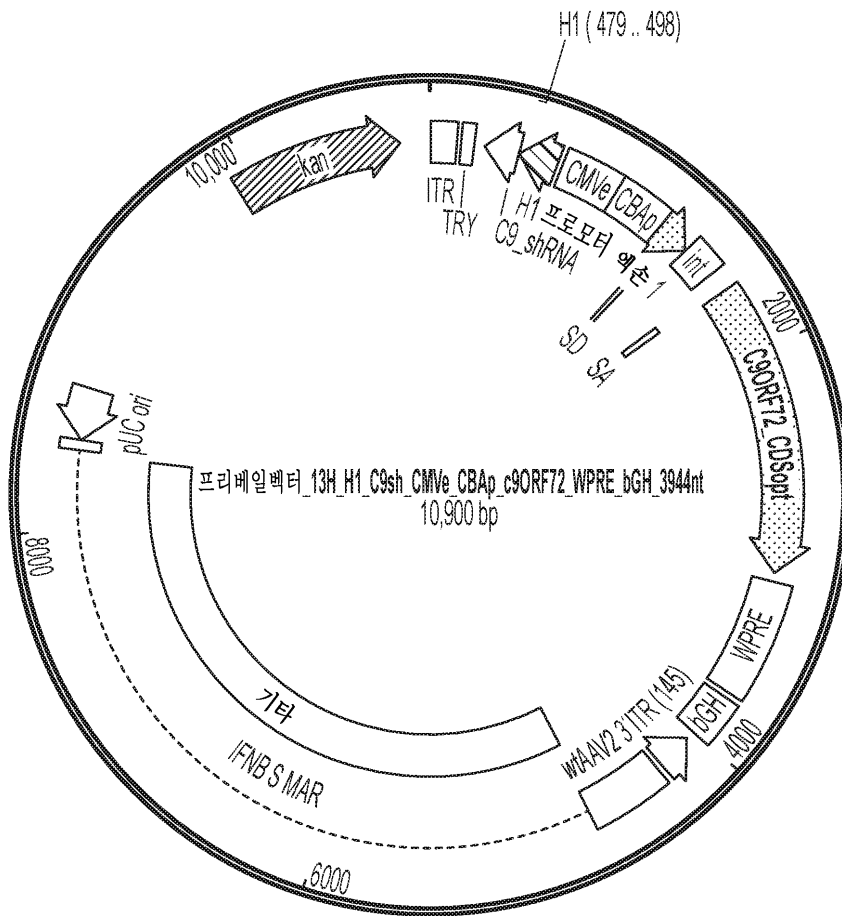
도면1



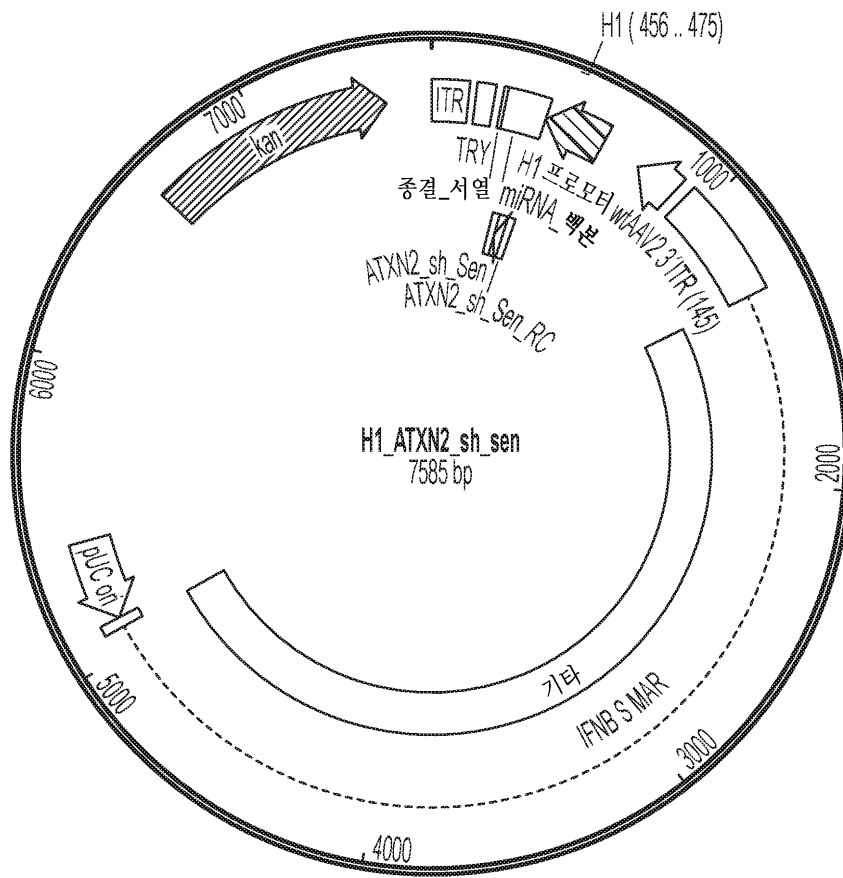
도면2



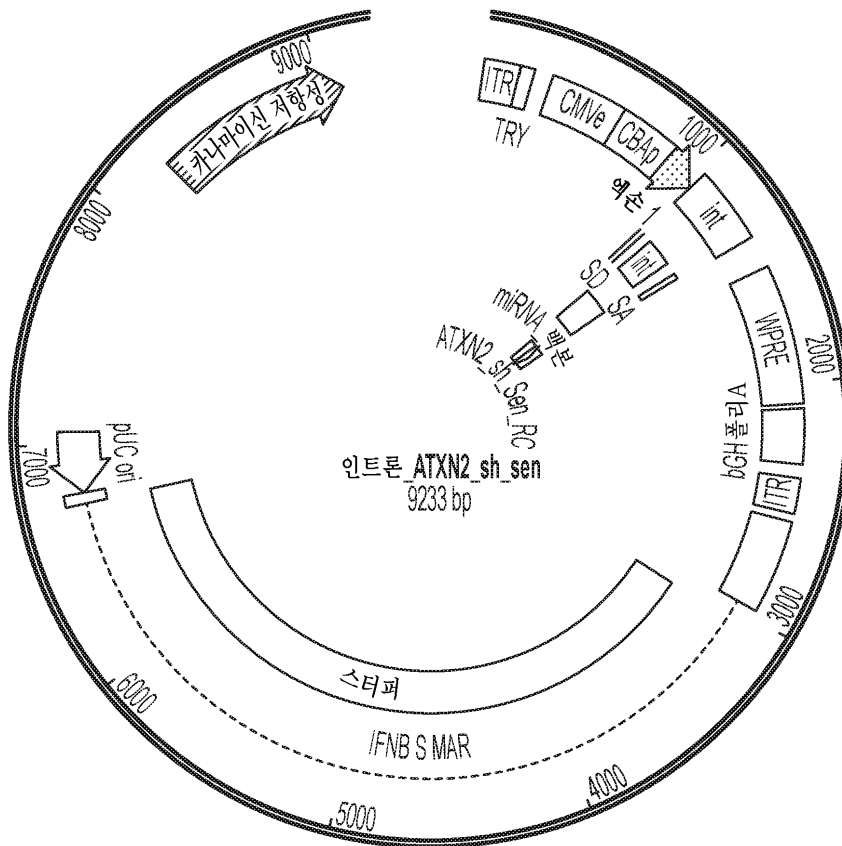
도면3



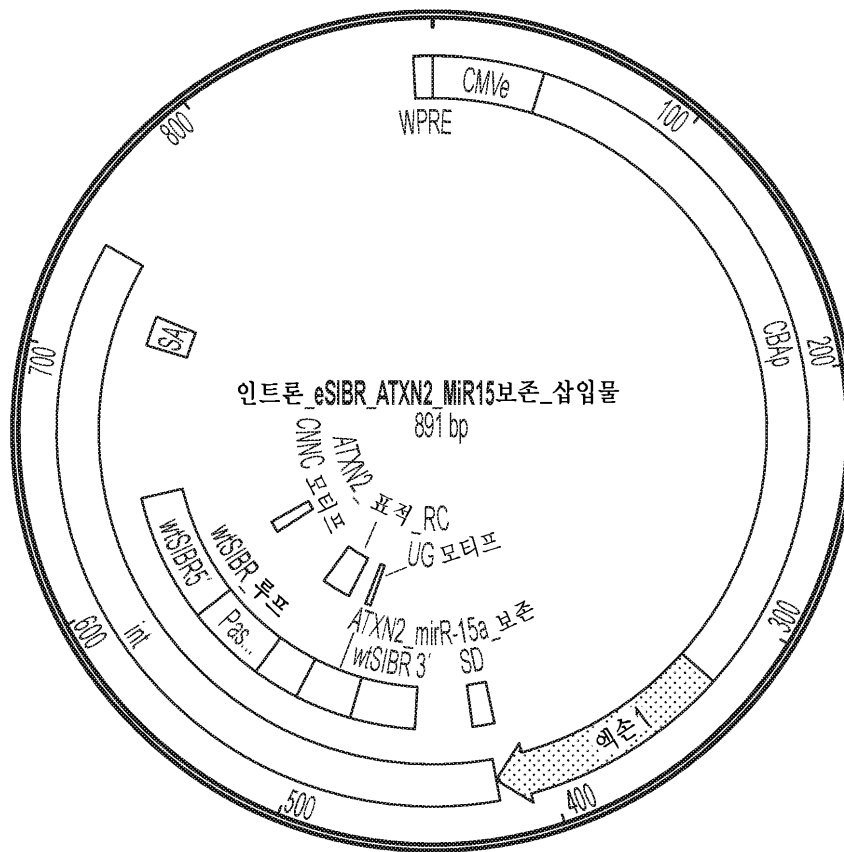
도면4



도면5

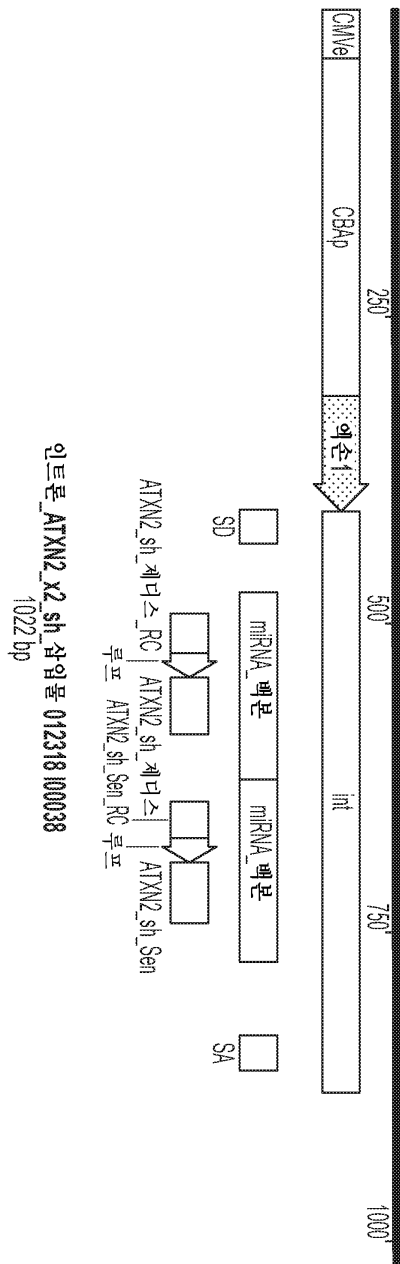


도면6

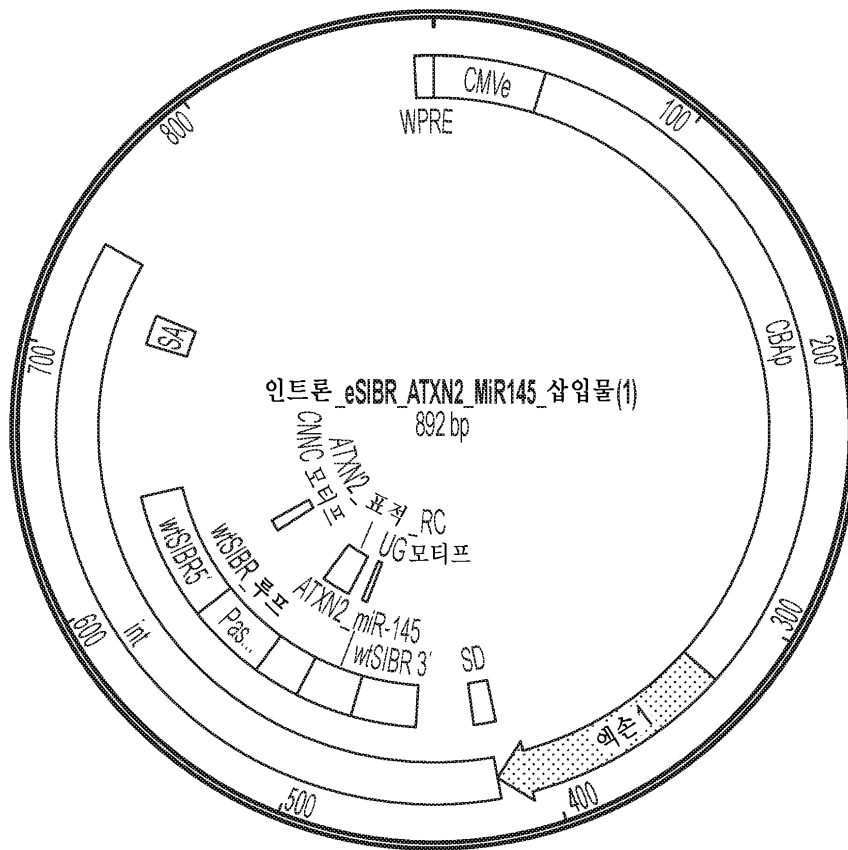




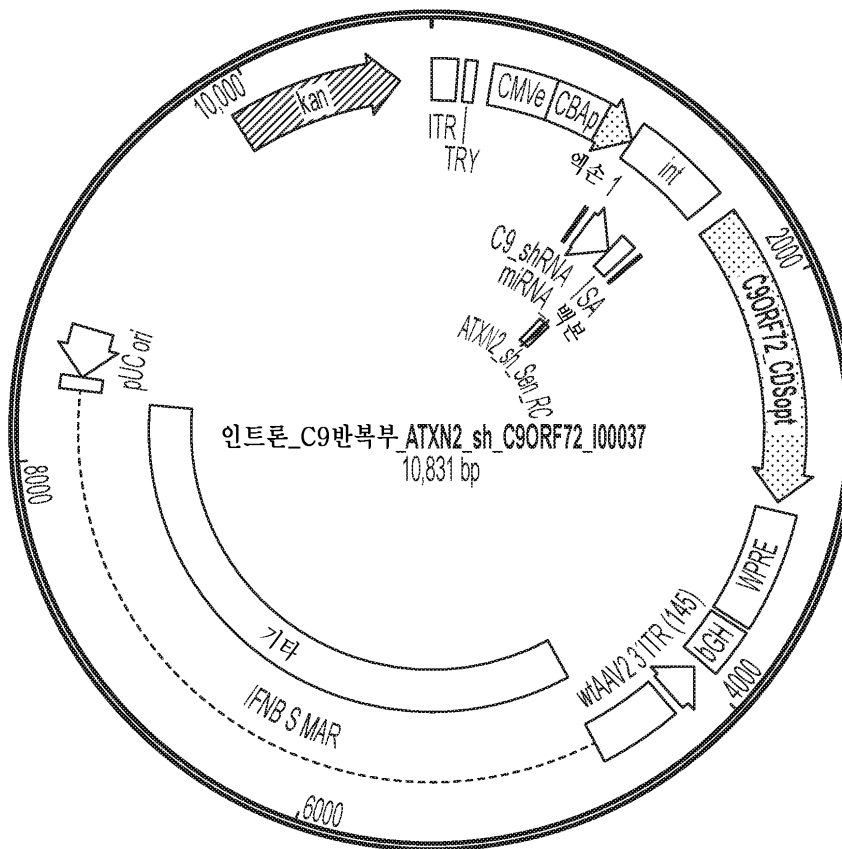
도면7



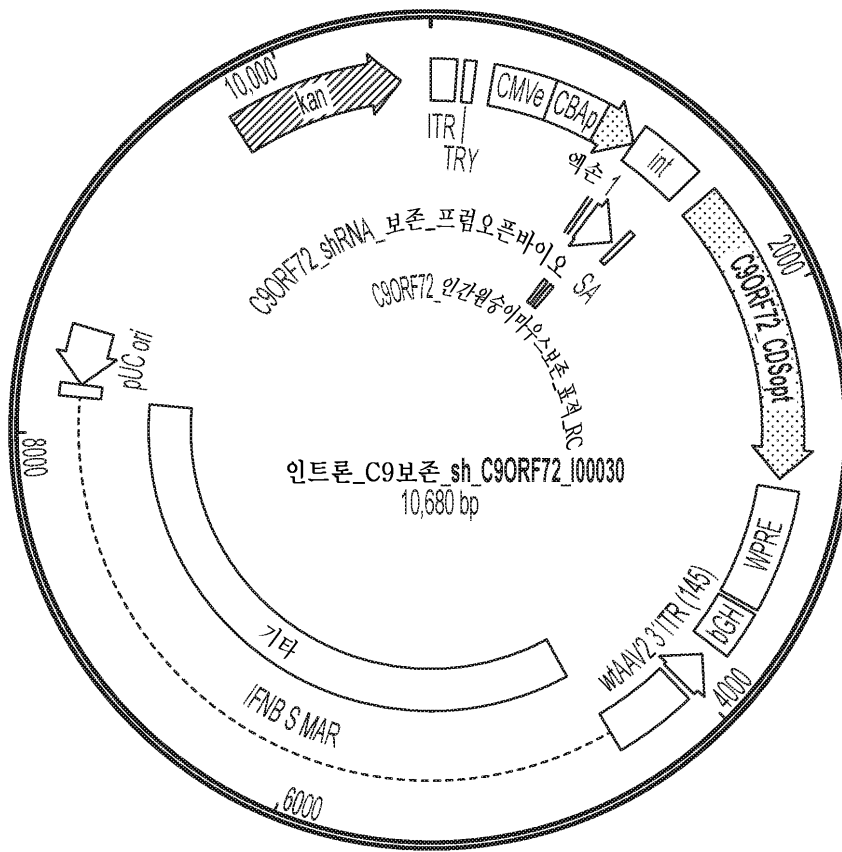
도면8



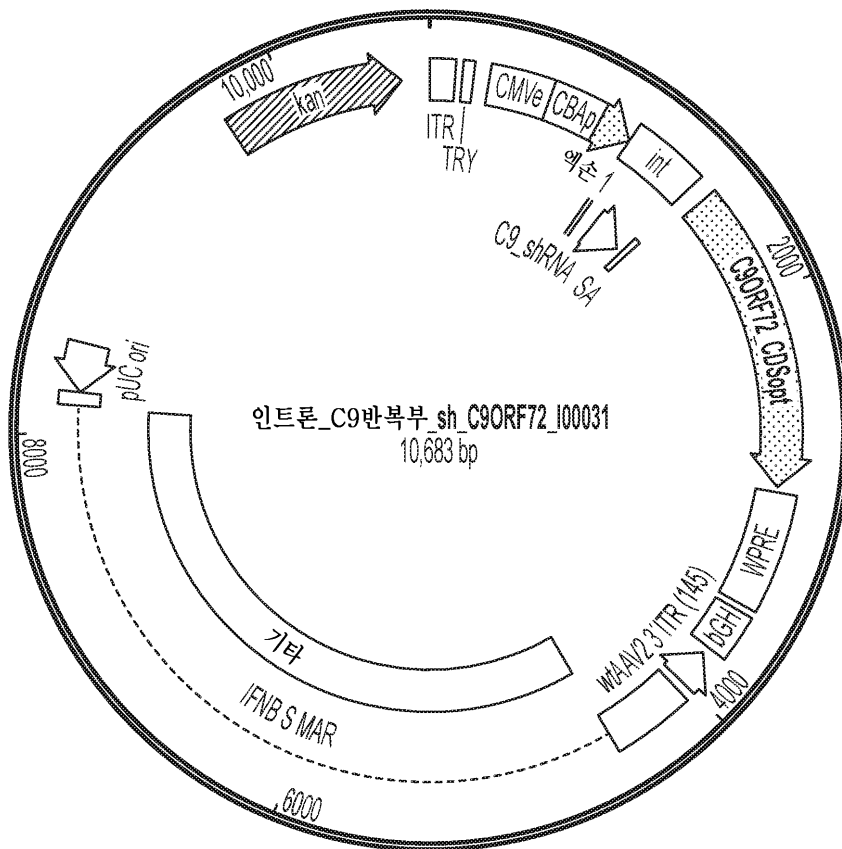
도면9



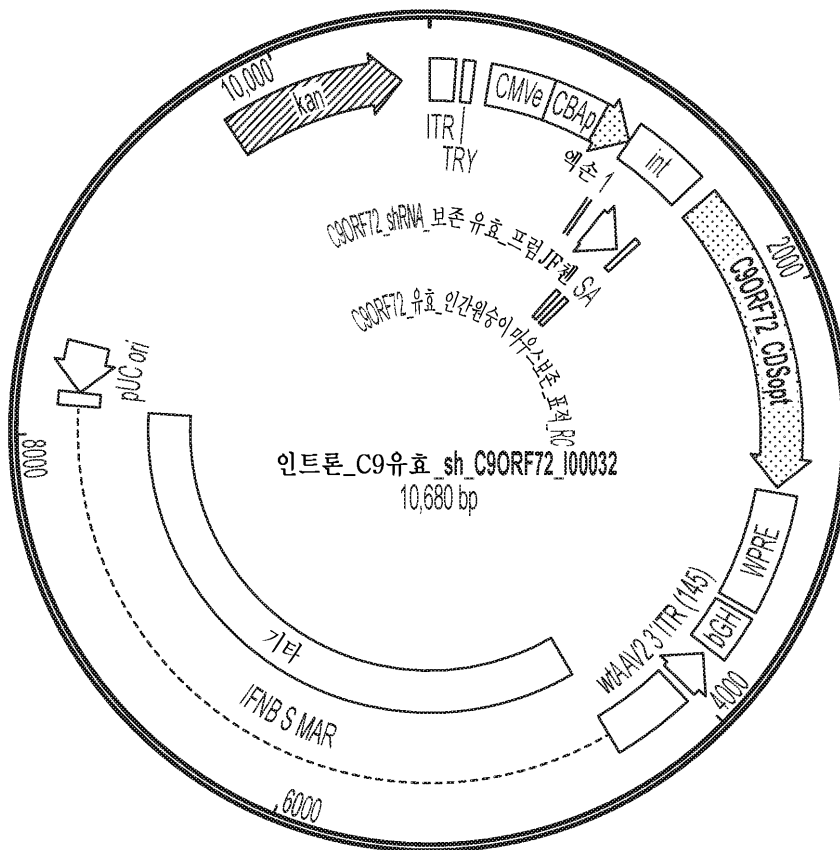
도면10



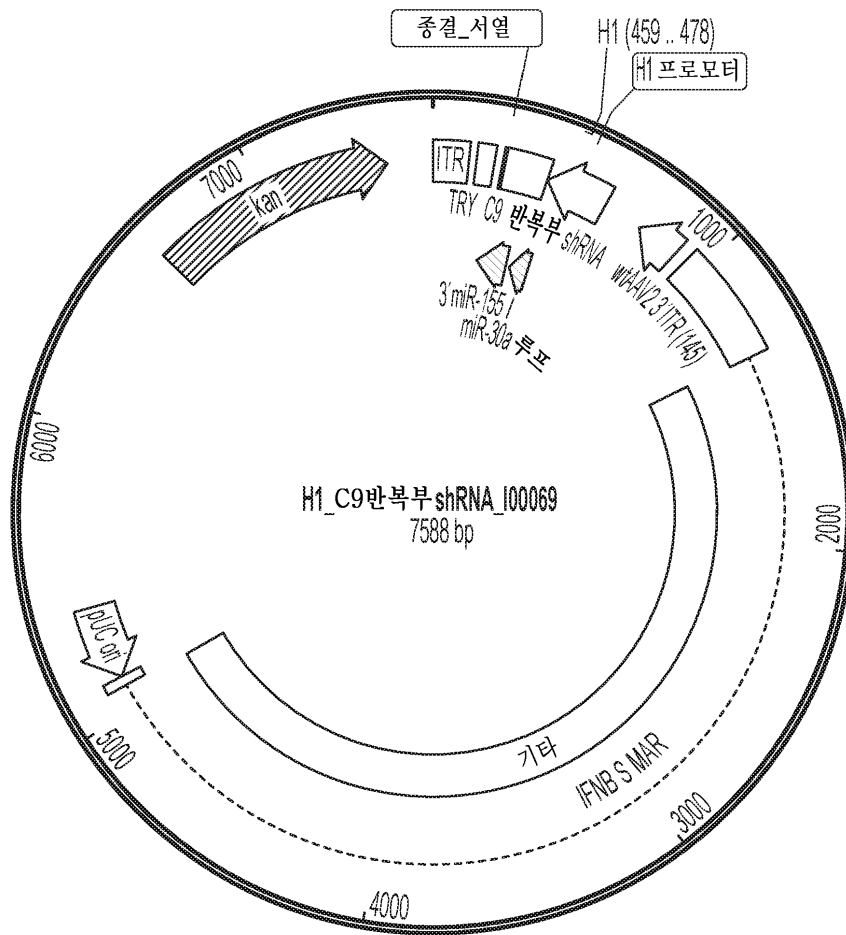
도면11



도면12

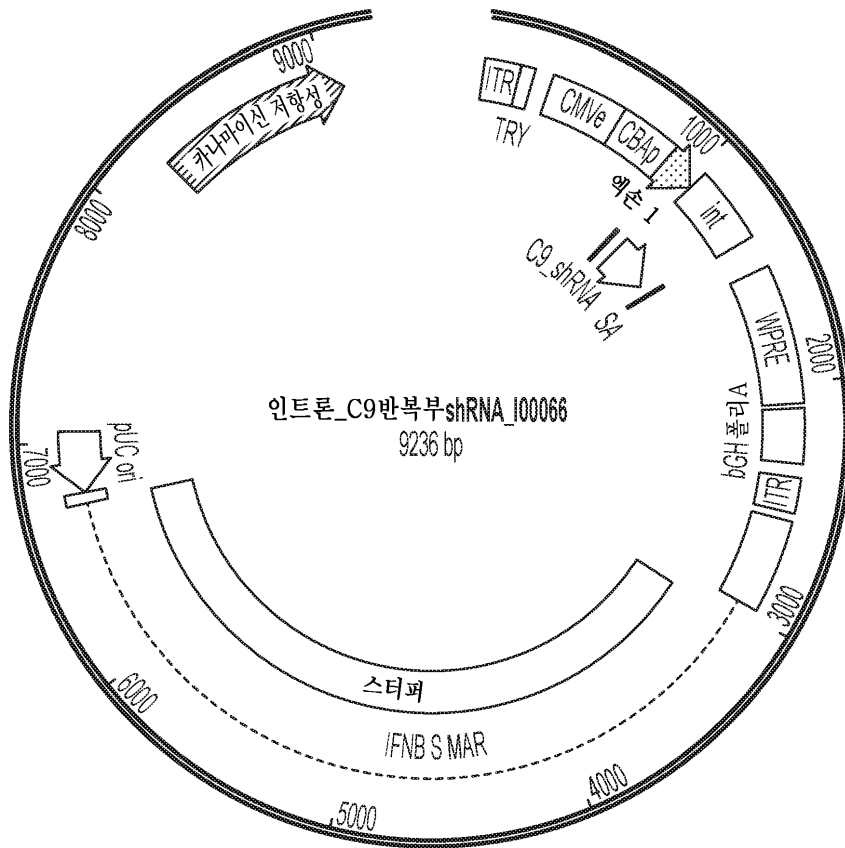


도면13

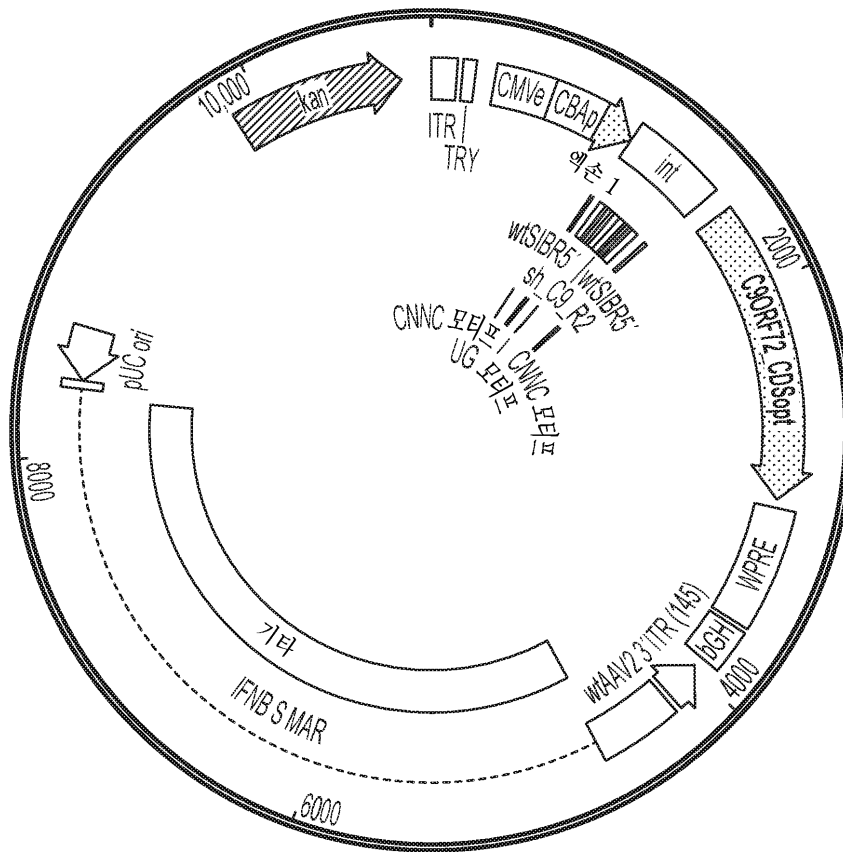




도면14

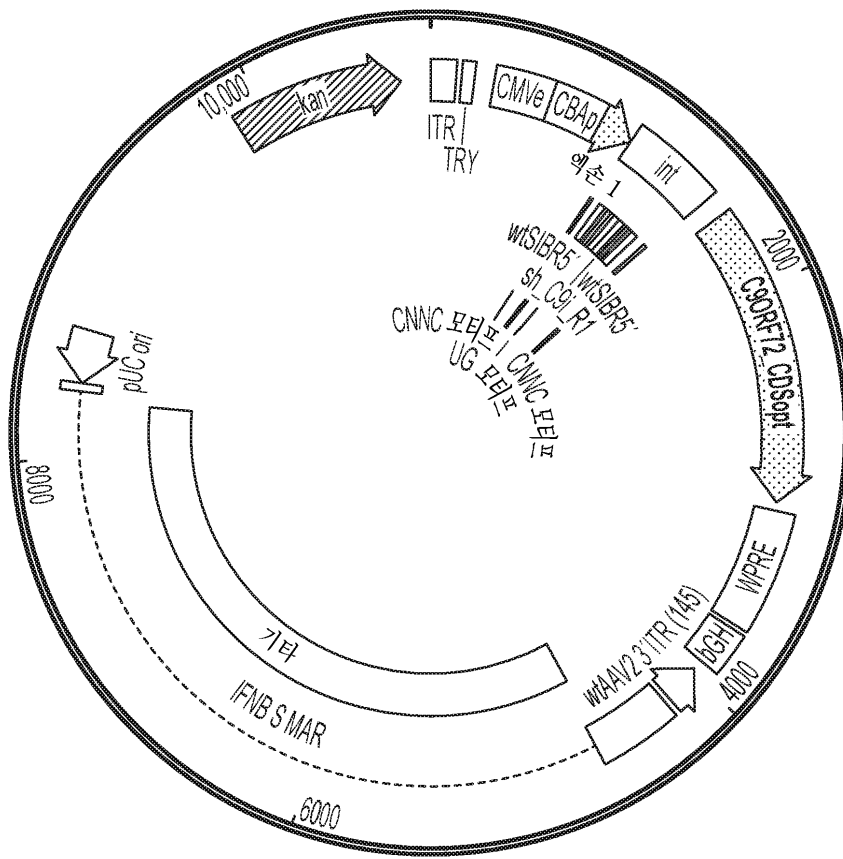


도면 15



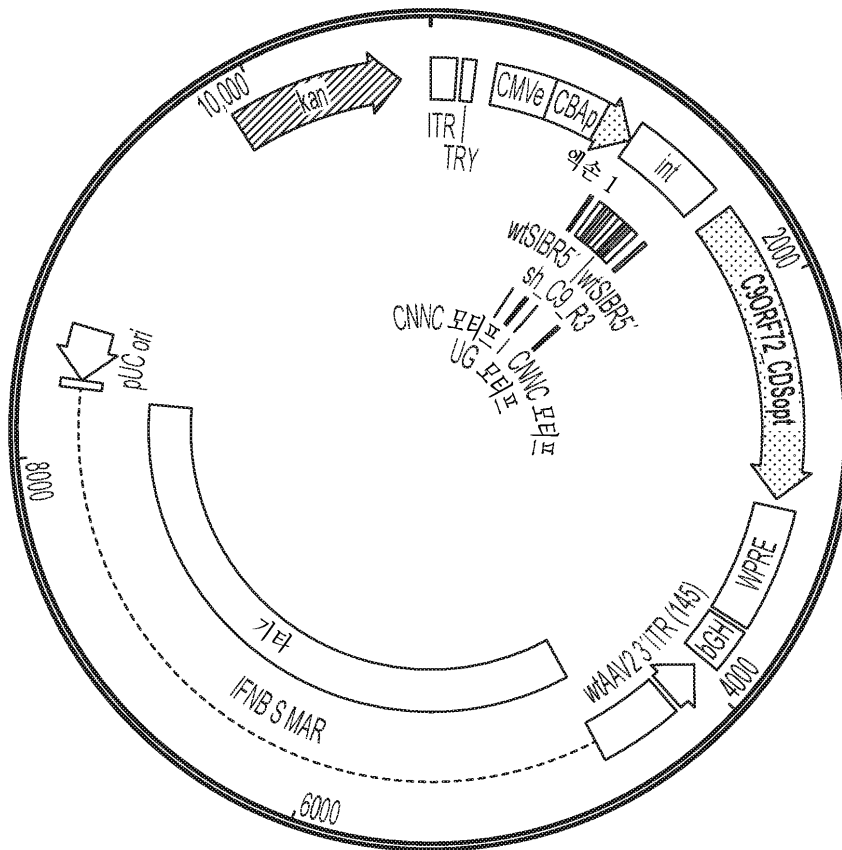
인트론\_eSIBR\_항\_C9\_인트론ShL2\_인트론SHR2\_CMVe\_CBAP\_C9ORF72\_WPRE\_bGH\_I00134  
10,828 bp

도면16



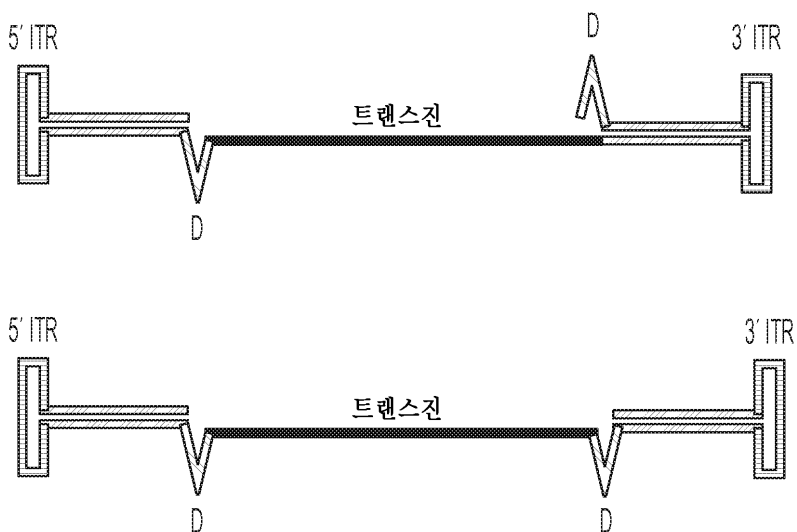
인트론\_eSIBR\_항\_C9\_인트론ShL1\_인트론ShR1\_CMVe\_CBAP\_C9ORF72\_WPRE\_bGH\_I00133  
10,829 bp

도면17

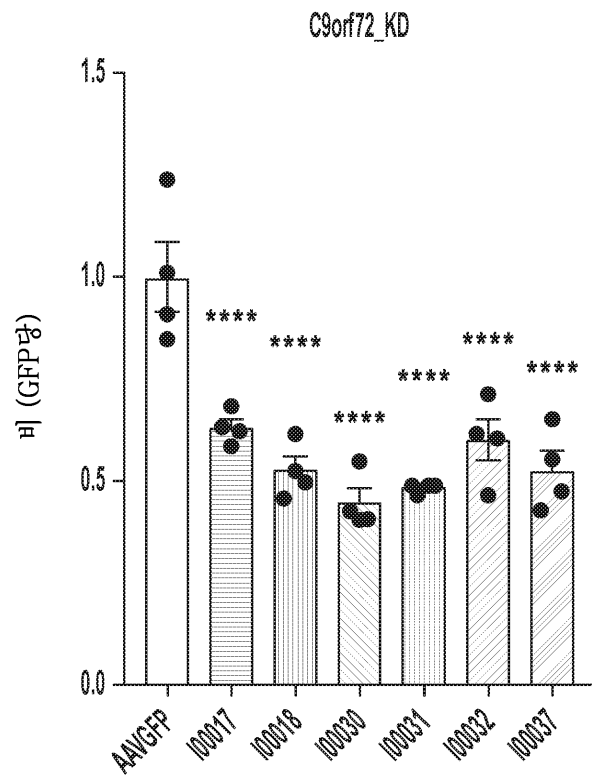


인트론\_eSIBR\_항\_C9\_인트론ShL3\_인트론ShR3\_CMVe\_CBAp\_C9ORF72\_WPRE\_bGH\_I00135  
10,826 bp

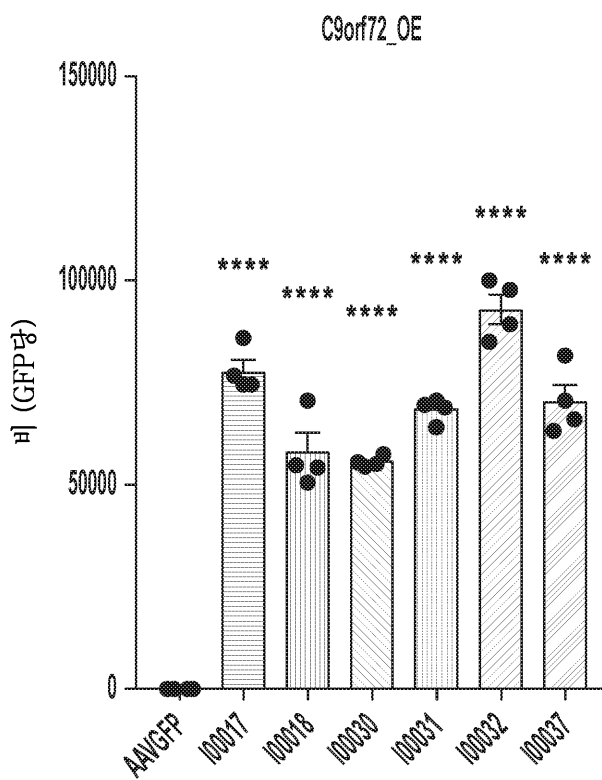
도면18



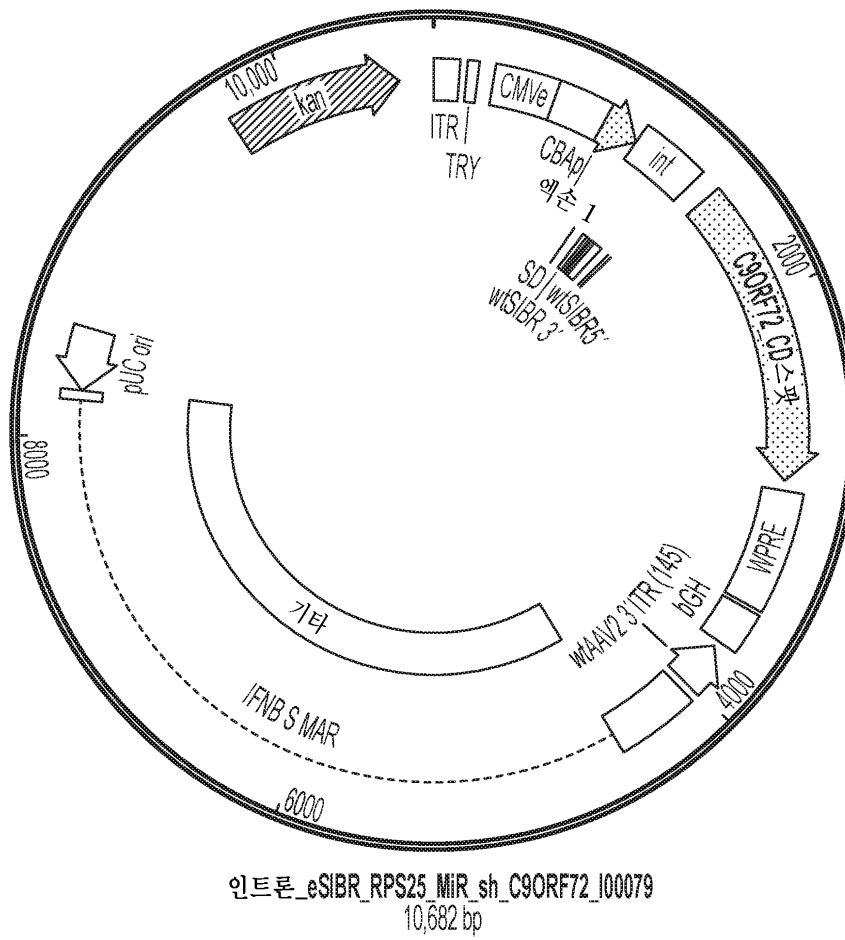
도면19a



도면19b



도면 20





<211> 10953

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 1

|                                                                     |      |
|---------------------------------------------------------------------|------|
| ttggccactc cctctctgcg cgctcgctcg ctcaactgagg cggggcgacc aaaggtcgcc  | 60   |
|                                                                     |      |
| cgacgcccgg gctttgcccc ggcggcctca gtgagcgagc gagcgcgag agagggagtg    | 120  |
| gccaaactcca tcaactagggg ttcttgctag ctctgggtat ttaagcccga gtgagcacgc | 180  |
| agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc   | 240  |
| tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac   | 300  |
| cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc   | 360  |
| cgcgttacat aacttacggt aaatggcccc cctggctgac cgcccaacga cccccccca    | 420  |
| ttgacgtcaa taatgacgta tgttccata gtaacgcaa tagggacttt ccattgacgt     | 480  |
|                                                                     |      |
| caatgggtgg actatttacg gtaaaactgcc cacttggcag tacatcaagt gtatcatatg  | 540  |
| ccaagtacgc ccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgccag     | 600  |
| tacatgacct tatgggactt tctacttgg cagtacatct acgtattagt catcgctatt    | 660  |
| accatggtcg aggtgagccc cacgttctgc ttactctcc ccatctcccc cccctccca     | 720  |
| cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg   | 780  |
| gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg   | 840  |
| agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccitt tatggcgagg   | 900  |
|                                                                     |      |
| cggcggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg   | 960  |
| ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cgcccgcccc ggctctgact   | 1020 |
| gaccgcgtta ctccacagg tgagcgggcg ggacggccct tctctccgg gctgtaatta     | 1080 |
| gcgcttggtt taatgacggc ttgttgagg cttgctgaag gctgtatgct gttgtcgggg    | 1140 |
| ccggtttcgg ggccttagtg aagccacaga tgtacggccc cgaaaccggc cccaggacac   | 1200 |
| aaggcctgtt actagcactc acatggaaca aatggccacc gtgggaggat gacaagtgat   | 1260 |
| atcacaaggt cccagggtcg gggtcagaaa ttctctcccg agggaaatgaa gccacaggag  | 1320 |
|                                                                     |      |
| ccaagagcag gaggaccaag gccctggcga aggccgtggc ctggttcaag taaaagatcc   | 1380 |
| tagtacagtg caggtcccaa tgtgtactag gatcttttac ttgaacgggg acgccggcat   | 1440 |
| ccgggctcag gacccccctc tctgccagag gcaccaacac cagagttcac aaatcagttc   | 1500 |
| cctgcccttt gcatgtagca aatttctgtg gctgcgtgaa agccttgagg ggctccggga   | 1560 |



|                                                                    |      |
|--------------------------------------------------------------------|------|
| gctagagcct ctgctaacca tgttcatgcc ttcttctttt tcctacagct cctgggcaac  | 1620 |
| gtgctggtta ttgtgctgtc tcatcathtt ggcaaagaat tcctcgaaga tccgaaggga  | 1680 |
| aagtcttcca cgactgtggg atccgttcga agatatcacc ggttgagcca ccatgagcac  | 1740 |
|                                                                    |      |
| cctgtgcccc cccccagcc ccgccgtggc caagaccgag atcgccctga gcggcaagag   | 1800 |
| ccccctgtg gccgccacct tcgcctactg ggacaacatc ctgggcccc gcgtgcgcca    | 1860 |
| catctgggcc cccaagaccg agcaggtgct gctgagcgac ggcgagatca ctttctggc   | 1920 |
| caaccacacc ctgaacggcg agatcctgcg caacgccgag agcggcgcca tcgacgtgaa  | 1980 |
| gttcttcgtg ctgagcgaga agggcgtgat catcgtgagc ctgatcttcg acggcaactg  | 2040 |
| gaacggcgac cgcagcacct acggcctgag catcatcctg cccagaccg agctgagctt   | 2100 |
| ctacctgccc ctgcaccgcg tgtgcgtgga ccgctgacc cacatcatcc gcaagggccg   | 2160 |
|                                                                    |      |
| catctggatg cacaaggagc gccaggagaa cgtgcagaag atcatcctgg agggcaccca  | 2220 |
| gcgcatggag gaccagggcc agagcatcat ccccatgctg accggcgagg tgatccccgt  | 2280 |
| gatggagctg ctgagcagca tgaagagcca cagcgtgccc gaggagatcg acatcgccga  | 2340 |
| caccgtgctg aacgacgacg acatcggcga cagctgccac gagggcttcc tgctgaacgc  | 2400 |
| catcagcagc cactgcaga cctgcggctg cagcgtggtg gtgggcagca gcgccgagaa   | 2460 |
| ggtgaacaag atcgtgcga cctgtgcct gttctgacc cccgccgagc gcaagtgcag     | 2520 |
| ccgctgtgc gaggcgaga gcagcttcaa gtacgagagc ggctgttcg tgcagggcct     | 2580 |
|                                                                    |      |
| gctgaaggac agcaccggca gttcgtgct gcccttcgc caggtagatg acgcccccta    | 2640 |
| ccccaccacc cacatcgacg tggacgtgaa caccgtgaag cagatgcccc cctgccacga  | 2700 |
| gcacatctac aaccagcgcc gctacatgcg cagcgagctg accgccttct ggcgcgccac  | 2760 |
| cagcgaggag gacatggccc aggacacat catctacacc gacgagagct tcacccccga   | 2820 |
| cctgaacatc ttccaggacg tgctgcaccg cgacaccctg gtgaaggcct tcttggaacca | 2880 |
| ggtgttccag ctgaagcccc gcctgagcct gcgcagcacc ttcttgcccc agttctgct   | 2940 |
| ggtgtgcac cgcaaggccc tgacctgat caagtacatc gaggacgaca cccagaaggg    | 3000 |
|                                                                    |      |
| caagaagccc ttcaagagcc tgcgcaacct gaagatcgac ctggacctga ccgccgaggg  | 3060 |
| cgacctgaac atcatcatgg ccctggccga gaagatcaag cccggcctgc acagcttcat  | 3120 |
| cttcggccgc cctttctaca ccagcgtgca ggagcgcgac gtgctgatga ctttctaatg  | 3180 |
| acaattgtta attaaagttta aacctcgag gccgcaagct tatcgataat caacctctgg  | 3240 |
| attacaaaat ttgtgaaaga ttgactggta ttcttaacta tgttgctcct ttacgctat   | 3300 |
| gtggatacgc tgccttaatg cctttgtatc atgctattgc ttcccgatg gctttcattt   | 3360 |

|                                                                   |      |
|-------------------------------------------------------------------|------|
| tctctcctt gtataaatcc tggttgctgt ctctttatga ggagtgtgg cccgttgtca   | 3420 |
|                                                                   |      |
| ggcaacgtgg cgtggtgtgc actgtgtttg ctgacgcaac cccactggt tggggcattg  | 3480 |
| ccaccacctg tcagctcctt tccgggactt tcgctttccc cctccctatt gccacggcgg | 3540 |
| aactcatcgc cgctgcctt gcccgtgct ggacaggggc tcggtgttg ggcactgaca    | 3600 |
| attccgtggt gttgtcggg aaatcatcgt cttttccttg gctgctcgcc tgtgttgcca  | 3660 |
| cctggattct gcgcgggacg tccttctgct acgtcccttc ggccctcaat ccagcggacc | 3720 |
| ttccttcccg cggcctgctg ccggctctgc ggctcttcc gcgtcttcgc cttcgccctc  | 3780 |
| agacgagtcg gatctccctt tgggccgcct ccccgcatcg ataccgtcga ctagagctcg | 3840 |
|                                                                   |      |
| ctgatcagcc tcgactgtgc cttctagtig ccagccatct gttgtttgcc cctccccgt  | 3900 |
| gccttccttg accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat | 3960 |
| tgcacgcat tgtctgagta ggtgtcattc tattctgggg ggtggggtgg ggcaggacag  | 4020 |
| caaggggag gattgggaag acaatagcag gcatgtggg gagagatcca cgataacaaa   | 4080 |
| cagctttttt ggggtgaaca tattgactga attccctgca ggttggccac tccctctctg | 4140 |
| cgcgctcgt cgctcactga ggccgcccgg gcaaagcccg ggcgctgggc gacctttggt  | 4200 |
| cggccgcct cagtgagcga gcgagcgcg agagaggag tggccaactc catcactagg    | 4260 |
|                                                                   |      |
| ggttcctcgc gccgctcgta cggctcagag gaattcctgc aggataactt gccaacctca | 4320 |
| ttctaaaatg tatatagaag cccaaaagac aataacaaaa atattcttgt agaacaaaat | 4380 |
| gggaaagaat gttccactaa atatcaagat ttagagcaaa gcatgagatg tgtggggata | 4440 |
| gacagtgagg ctgataaaat agagtagagc tcagaaacag acccattgat atatgtaagt | 4500 |
| gacctatgaa aaaaatatgg cattttacia tgggaaaatg atggtctttt tcttttttag | 4560 |
| aaaaacaggg aaatatattt atatgtaaaa aataaaaggg aaccatatg tcataccata  | 4620 |
| cacacaaaaa aattccagtg aattataagt ctaaatggag aaggcaaac tttaaatctt  | 4680 |
|                                                                   |      |
| ttagaaaata atatagaagc atgcagacca gcctggccaa catgatgaaa ccctctctac | 4740 |
| taataataaa atcagtagaa ctactcagga ctactttgag tgggaagtcc tttctatga  | 4800 |
| agacttcttt ggccaaaatt aggctctaaa tgcaaggaga tagtgcacat tgcctggctg | 4860 |
| cacttactga taaatgatgt tatcaccatc tttaaccaaa tgcacaggaa caagttatgg | 4920 |
| tactgatgtg ctggattgag aaggagctct acttcttga caggacacat ttgtatcaac  | 4980 |
| ttaaaaaagc agatttttgc cagcagaact attcattcag aggtaggaaa cttagaatag | 5040 |
| atgatgtcac tgattagcat ggcttcccca tctccacagc tgcttccac ccaggttgcc  | 5100 |

cacagttgag ttgtccagt gctcagggt gccactctc agtaagaagc cccacaccag 5160  
ccctctcca aatatgttg ctgttccttc cattaaagt accccactt agagcagcaa 5220  
gtggatttct gtttcttaca gttcaggaag gaggagtcag ctgtgagaac ctggagcctg 5280  
agatgcttct aagtcacct gctactgggg tcagggaagc cagactccag catcagcagt 5340  
caggagcact aagcccttgc caacatcctg ttctcagag aaactgcttc cattataatg 5400  
gttgtccttt ttaagctat caagccaaac aaccagtgc taccattatt ctcacacct 5460  
gaagccaagg gtctagcaa aagtcaagct gtcttgtaat ggttgatgtg cctccagctt 5520  
  
ctgtcttcag tcactccact cttagcctgc tctgaatcaa ctctgaccac agttccctgg 5580  
agccctgcc acctgctgcc cctgccacct tctccatctg cagtgtgtg cagccttctg 5640  
cactcttgca gagctaatag gtggagactt gaaggaagag gaggaaagt tctcataata 5700  
gccttctgc aagctcaaat gggaggtggg cactgtgccc aggagccttg gagcaaaggc 5760  
tgtgccaac ctctgactgc atccaggttt ggtcttgaca gagataagaa gccctggctt 5820  
ttggagccaa aatctaggtc agacttaggc aggattctca aagtttatca gcagaacatg 5880  
aggcagaaga cctttctgc tccagcttct tcaggctcaa ccttcacag aatagataga 5940  
  
aagagaggct gtgagggttc ttaaacaga agcaaatctg actcagagaa taaacaacct 6000  
cctagtaaac tacagcttag acagagcatc tgggtgtgag tgtgtcagt gtcctactca 6060  
actgtctggt atcagccctc atgaggactt ctcttcttc cctcatagac ctccatctct 6120  
gttttctta gcctgcagaa atctggatgg ctattcacag aatgcctgtg ctttcagagt 6180  
tgcatTTTT ctctggtatt ctggttcaag catTTgaagg taggaaagg tctccaagt 6240  
caagaaaggc agccctgagc ctcaactgcc tggctagtgt ggtcagtagg atgcaaaggc 6300  
tgttgaatgc cacaaggcca aactttaacc tgtgtaccac aagcctagca gcagaggcag 6360  
  
ctctgtcac tggaactctc tgttctttt ctctgagcc ttttcttctc ctgagtttct 6420  
tagctctct caaccttacc tctgcctac ccaggacaaa cccaagagcc actgttctg 6480  
tgatgtctc tccagcccta attagcatc atgacttcag cctgacctc catgtcaga 6540  
agcagtgcta atccacttca gatgagctgc tctatgcaac acaggcagag cctacaaacc 6600  
tttgaccag agccctccac atatcagtgt ttgttcatac tcacttcaac agcaaatgtg 6660  
actgtgaga ttaagatttt acacaagatg gtctgtaatt tcacagttag ttttatccca 6720  
ttaggtatga aagaattagc ataattcccc ttaaacatga atgaatctta gatTTTTaa 6780  
  
taaatagttt tggaagtaaa gacagagaca tcaggagcac aaggaatagc ctgagaggac 6840  
aaacagaaca agaaagagtc tggaataaca caggatgttc ttggcctctc caaagcaagt 6900  
gcaagcagat agtaccagca gcccaggct atcagagccc agtgaagaga agtaccatga 6960

|                                                                    |      |
|--------------------------------------------------------------------|------|
| aagccacagc tctaaccacc ctgttccaga gtgacagaca gtccccaaga caagccagcc  | 7020 |
| tgagccagag agagaactgc aagagaaagt ttctaattta ggttctgtta gattcagaca  | 7080 |
| agtgcaggtc atcctctctc cacagctact cacctctcca gcctaacaaa gcctgcagtc  | 7140 |
| cacactccaa ccctgggtgc tcacctccia gcctctccca acatcctgct ctctgaccat  | 7200 |
| cttctgcate tctcatctca ccatctccca ctgtctacag cctactcttg caactacat   | 7260 |
| ctcattttct gacatectgt ctacatcttc tgccatactc tgccatctac cataccacct  | 7320 |
| cttaccatct accacacat cttttatctc catccctctc agaagcctcc aagctgaatc   | 7380 |
| ctgctttatg tgttcatctc agccccctga tggaaaagctg accccagagg cagaactatt | 7440 |
| cccagagagc ttggccaaga aaaacaaaac taccagcctg gccaggctca ggagtagtaa  | 7500 |
| gctgcagtggt ctgttggtt ctagcttcaa cagctgcagg agttccactc tcaaagtctc  | 7560 |
| cacatttctc acatcctcct gattctggtc actaccatc ttcaaagaac agaatatctc   | 7620 |
| acatcagcat actgtgaagg actagtcatg ggtgcagctg ctgagagctg caaagtcatt  | 7680 |
| ctggatgggtg gagagcttac aaacatttca tgatgctccc cccgctctga tggctggagc | 7740 |
| ccaatcccta cacagactcc tgctgtatgt gttttcttt cactctgagc cacagccaga   | 7800 |
| gggcaggcat tcagtctcct cttcaggctg gggctggggc actgagaact cacccaacac  | 7860 |
| cttgctctca ctcttctgc aaaacaagaa agagctttgt gctgcagtag ccatgaagaa   | 7920 |
| tgaaaggaag gctttaacta aaaaatgtca gagattatit tcaaccctt actgtggatc   | 7980 |
| accagcaagg aggaacaca acacagagac attttttccc ctcaaattat caaaagaatc   | 8040 |
| actgcatttg ttaaagagag caactgaatc aggaagcaga gttttgaaca tatcagaagt  | 8100 |
| taggaatctg catcagagac aaatgcagtc atggttggtt gctgcatacc agccctaate  | 8160 |
| attagaagcc tcatggactt caaacatcat tcctctgac aagatgctct agcctaate    | 8220 |
| catgagataa aataaatctg ctttcagag ccaaagaaga gtcaccagc ttcttctcag    | 8280 |
| tgtgaacaag agtccagtc aggttagtca gtccagtca gtagaggaga ccagtctgca    | 8340 |
| tcctctaatt ttcaaaggca agaagatttg ttaccctgg acaccaggca caagtgaggt   | 8400 |
| cacagagctc ttagatatgc agtctcatg agtgaggaga ctaaagcgca tgccatcaag   | 8460 |
| acttcagtgt agagaaaacc tccaaaaaag cctctcact acttctggaa tagctcagag   | 8520 |
| gccgaggcgg cctcgccctc tgcataaata aaaaaatta gtcagccatg gggcgagaa    | 8580 |
| tggcggaac tggcgaggt tagggcgagg atggcgagg ttagggcgagg gactatggtt    | 8640 |
| gctgactaat tgagatgcat gctttgcata ctctgcctg ctggggagcc tggggacttt   | 8700 |
| ccacacctgg ttgctgacta attgagatgc atgctttgca tacttctgcc tgctggggag  | 8760 |
| cctggggact ttccacacc taactgacac acattccaca gctgcattaa tgaatcgcc    | 8820 |

|                                                                    |       |
|--------------------------------------------------------------------|-------|
| aacgcgcggg gagaggcggg ttgcgtattg ggcgctcttc cgcttcctcg ctcaactgact | 8880  |
|                                                                    |       |
| cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtataac  | 8940  |
| ggttatccac agaatacagg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa  | 9000  |
| aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg   | 9060  |
| acgagcatca caaaaatcga cgctcaagtc agagggtggcg aaacccgaca ggactataaa | 9120  |
| gataaccaggc gtttccccct ggaagctccc tcgtgcgctc tctgtttccg accctgccgc | 9180  |
| ttaccggata ccgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac   | 9240  |
| gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac  | 9300  |
|                                                                    |       |
| ccccgttca gcccgaccgc tgcgccttat ccggtaaacta tcgtcttgag tccaacccgg  | 9360  |
| taagacacga cttatcgcca ctggcagcag ccaactgtaa caggattagc agagcgaggt  | 9420  |
| atgtaggcgg tgcacagag tttctgaagt ggtggcctaa ctacggctac actagaagaa   | 9480  |
| cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct  | 9540  |
| cttgatccgg caaacaacc accgctggta gcggtggttt tttgtttgc aagcagcaga    | 9600  |
| ttacgcgcag aaaaaagga tctcaagaag atcctttgat cttttctac gggctctgacg   | 9660  |
| ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct  | 9720  |
|                                                                    |       |
| tcacctagat ctttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt  | 9780  |
| aaacttggtc tgacagttac caatgcttaa tcagtgggc acctatctca gcgatctgtc   | 9840  |
| tatttcgttc atccatagtt gcctgactcc tgcaaaccac gttgtgtctc aaaatctctg  | 9900  |
| atgttacatt gcacaagata aaaatatatc atcatgaaca ataaaactgt ctgcttacat  | 9960  |
| aaacagtaat acaaggggtg ttatgagcca tattcaacgg gaaacgtctt gctcgaggcc  | 10020 |
| gcgattaaat tccaacatgg atgctgattt atatgggtat aaatgggctc gcgataatgt  | 10080 |
| cgggcaatca ggtgcgacaa tctatcgatt gtatgggaag cccgatgcgc cagagttgtt  | 10140 |
|                                                                    |       |
| tctgaaacat ggcaaaggta gcgttgccaa tgatgttaca gatgagatgg tcagactaaa  | 10200 |
| ctggctgacg gaatttatgc ctcttcgac catcaagcat tttatccgta ctctgatga    | 10260 |
| tgcatggtta ctcaactctg cgatccccgg gaaaacagca ttccaggat tagaagaata   | 10320 |
| tcctgattca ggtgaaaata ttgttgatgc gctggcagtg ttcctgcgcc ggttgcatc   | 10380 |
| gattcctgtt tgtaattgtc cttttaacag cgatcgcgta tttcgtctcg ctgagcgca   | 10440 |
| atcacgaatg aataacggtt tggttgatgc gagtgatttt gatgacgagc gtaatggctg  | 10500 |
| gcctgttgaa caagtctgga aagaaatgca taagcttttg ccattctcac cggattcagt  | 10560 |

cgtcactcat ggtgatttct cacttgataa ccttattttt gacgagggga aattaatagg 10620  
 ttgtattgat gttggacgag tcggaatcgc agaccgatac caggatcttg ccatcctatg 10680  
 gaactgcctc ggtgagtttt ctctttcatt acagaaacgg ctttttcaa aatatggtat 10740  
 tgataatcct gatatgaata aattgcagtt tcatttgatg ctcgatgagt ttttctaagg 10800  
 gcggcctgcc accataccca cgccgaaaca agcgcctcatg agcccgaagt ggcgagcccc 10860  
 atcttcccca tcggtgatgt cggcgatata ggccgagca accgcacctg tggcgccggt 10920  
 gatgagggcg cgccaagtcg acgtccggca gtc 10953

<210> 2

<211> 10848

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 2

ttggccaactc cctctctgcg cgctcgctcg ctcaactgagg ccgggcgacc aaaggtcgcc 60  
 cgacgccccg gctttgcccc ggccgcctca gtgagcgagc gagcgcgag agagggagtg 120  
 gccaaactcca tcaactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180  
 agggctccca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240  
 tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300  
 cctagtattt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc 360

cgcgttacat aacttacggt aaatggcccc cctggctgac cgcccaacga cccccccca 420  
 ttgacgtcaa taatgacgta tgttccata gtaacccaa tagggacttt ccattgacgt 480  
 caatgggtgg actatttacg gtaaaactgcc cacttggcag tacatcaagt gtatcatatg 540  
 ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgccag 600  
 tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt 660  
 accatggteg aggtgagccc cagtttctgc ttcaactctcc ccattctccc cccctcccca 720  
 cccccaattt tgiatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg 780

gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcgggcggg ggcgaggcgg 840  
 agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg 900  
 cggcgggcggc ggccggcccta taaaagcga agcgcgcggc gggcgggagt cgctgcgacg 960  
 ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cggccgcccc ggctctgact 1020  
 gaccgcgtta ctccacagg tgagcgggcg ggacggccct tctcctccgg gctgtaatta 1080  
 gcgcttggtt taatgacggc ttgttgagg cttgctgaag gctgtatgct gttgtcgggg 1140

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ccggttttcgg ggccttagtg aagccacaga tgtacggccc cgaaaccggc cccaggacac | 1200 |
|                                                                    |      |
| aaggcctgtt actagcactc acatggaaca aatggccacc gtgggaggat gacaatttct  | 1260 |
| gtggctgctg gaaagccttg aggggctccg ggagctagag cctctgctaa ccatgttcat  | 1320 |
| gcctttcttct ttttcttaca gctcctgggc aacgtgctgg ttattgtgct gtctcatcat | 1380 |
| tttggcaaag aattcctcga agatccgaag ggaaagtctt ccacgactgt gggatccgtt  | 1440 |
| cgaagatata accggttgag ccacatgga attcagcagc cccagcagag aggaatgccc   | 1500 |
| caagcctctg agccgggtgt caatcatggc cggatctctg acaggactgc tgctgcttca  | 1560 |
| ggccgtgtct tgggcttctg gcgctagacc ttgcatcccc aagagcttcg gctacagcag  | 1620 |
|                                                                    |      |
| cgctcgtgtc gtgtgcaatg ccacctactg cgacagcttc gacctccta cctttcctgc   | 1680 |
| tctgggcacc ttcagcagat acgagagcac cagatccggc agacggatgg aactgagcat  | 1740 |
| gggacccatc caggccaatc acacaggcac tggcctgctg ctgacactgc agcctgagca  | 1800 |
| gaaattccag aaagtgaag gcttcggcgg agccatgaca gatgccgcg ctttgaatat    | 1860 |
| cctggctctg tctccaccag ctccagaact gctgctcaag agctacttca gcgaggaagg  | 1920 |
| catcggttac aacatcatca gagtgcccat ggccagctgc gacttcagca tcaggaccta  | 1980 |
| cacctacgcc gacacaccgc acgatttcca gctgcacaac ttcagcctgc ctgaagagga  | 2040 |
|                                                                    |      |
| caccaagctg aagatccctc tgatccacag agccctgcag ctggcacaaa gaccctgtc   | 2100 |
| actgtctggc tctccatgga catctccac ctggctgaaa acaaatggcg ccgtgaatgg   | 2160 |
| caagggcagc ctgaaaggcc aacctggcga catctaccac cagacctggg ccagatactt  | 2220 |
| cgtgaagtgc ctggacgct atgccgagca caagctgcag ttttgggcg tgacagccga    | 2280 |
| gaacgaacct tctgctggac tgctgagcgg ctaccccttt cagtgcctgg gctttacacc  | 2340 |
| cgagcaccag cgggacttta tcgcccgtga tctgggaccc acactggcca atagcaccca  | 2400 |
| ccataatgtg cggtctgta tgctggacga ccagagactg cttctgcccc actgggctaa   | 2460 |
|                                                                    |      |
| agtggctgtg acagatcctg aggccgcaa atacgtgcac ggaatcgccg tgcactggta   | 2520 |
| tctggacttt ctggccctg ccaaggccac actgggagag acacacagac tgttcccaa    | 2580 |
| caccatgctg ttcgccagcg aagcctgtgt gggcagcaag ttttgggaac agagcgtgcg  | 2640 |
| gctcggcagc tgggatatag gcatgcagta cagccacagc atcatcacca acctgctgta  | 2700 |
| ccacgtcgtc ggctggaccg actggaatct ggccctgaat cctgaaggcg gccctaactg  | 2760 |
| ggtccgaaac ttcgtggaca gcccacatcat cgtggacatc accaaggaca ctttctacaa | 2820 |
| gcagcccatg tttaccacc tgggacactt cagcaagttc atccccgagg gctctcagcg   | 2880 |

|                                                                   |      |
|-------------------------------------------------------------------|------|
| cgttggactg gtggcttccc agaagaacga tctggacgcc gtggctctga tgcacctga  | 2940 |
| tggatctgct gtggtggagg tcctgaaccg cagcagcaaa gatgtgcccc tgaccatcaa | 3000 |
| ggatccccgc gtgggattcc tggaaacaat cagccctggc tactccatcc acacctacct | 3060 |
| gtggcgtaga cagtgacaat tgtaattaa gtttaaacc tggaggccgc aagcttatcg   | 3120 |
| ataatcaacc tctggattac aaaatttgt aaagattgac tggattctt aactatgttg   | 3180 |
| ctctttttac gctatgtgga tacgtctgtt taatgccttt gtatcatgct attgcttccc | 3240 |
| gtatggcttt cattttctcc tcttgtata aatcctgggt gctgtctctt tatgaggagt  | 3300 |
|                                                                   |      |
| tgtggcccggt tgtcaggcaa cgtggcgtgg tgtgactgt gtttctgac gcaaccccca  | 3360 |
| ctggttgggg cattgccacc acctgtcagc tctttccgg gactttcgt tccccctcc    | 3420 |
| ctattgccac ggcggaactc atcgccgcct gccttgccc ctgctggaca ggggctcggc  | 3480 |
| tgttgggcac tgacaattcc gtggtgttgt cggggaaatc atcgtcttt ccttggctgc  | 3540 |
| tcgcctgtgt tgccacctgg attctgcgcg ggacgtcctt ctgctacgtc ctttcggccc | 3600 |
| tcaatccagc ggaccttct tcccgcggcc tgcgcccgc tctgcgcct cttccgcgtc    | 3660 |
| ttgccttcg cctcagacg agtcggatct cctttgggc cgcctcccgc catcgatacc    | 3720 |
|                                                                   |      |
| gtcgactaga gctcgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt | 3780 |
| ttgccctcc ccggtgcctt ccttgacct ggaagggtgc actccactg tcttttcta     | 3840 |
| ataaaatgag gaaattgcat cgcattgtct gtaggtgt cattctattc tggggggtgg   | 3900 |
| ggtggggcag gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggagag | 3960 |
| atccacgata acaaacagct tttttgggt gaacatattg actgaattcc ctgcaggttg  | 4020 |
| gccactccct ctctgcgcgc tcgtcgtc actgaggccg cccgggcaaa gcccgggcgt   | 4080 |
| cgggcgacct ttggtcgccc ggcctcagt agcgagcgag cgcgcagaga gggagtggcc  | 4140 |
|                                                                   |      |
| aactccatca ctagggttc ctgcggccgc tcgtacggtc tcgaggaatt cctgcaggat  | 4200 |
| aacttgcaa cctatttcta aaatgtatat agaagccaa aagacaataa caaaaatatt   | 4260 |
| ctttagaac aaaaaggga agaattgtcc actaaatc aagatttaga gcaaagcatg     | 4320 |
| agatgtgtgg ggatagacag tgaggctgat aaaatagagt agagctcaga aacagacca  | 4380 |
| ttgatatatg taagtacct atgaaaaaa tatggcattt tacaatggga aaatgatggt   | 4440 |
| ctttttttt tttagaaaa cagggaata tatttatg taaaaataa aagggaacct       | 4500 |
| atatgcata ccatacacac aaaaaattc cagtgaatta taagtctaaa tggagaaggc   | 4560 |
|                                                                   |      |
| aaaacttta atcttttaga aaataatata gaagcatgca gaccagctg gccaacatga   | 4620 |
| tgaaacctc tctactaata ataaaatcag tagaactact caggactact ttgagtggga  | 4680 |
| agtcctttt tatgaagact tctttggcca aaattaggct ctaaatgcaa ggagatagt   | 4740 |



|                                                                    |      |
|--------------------------------------------------------------------|------|
| catcatgcct ggctgcactt actgataaat gatgttatca ccatctttta ccaaagtcac  | 4800 |
| aggaacaagt tatggtactg atgtgctgga ttgagaagga gctctacttc cttgacagga  | 4860 |
| cacatttgta tcaacttaaa aaagcagatt ttggccagca gaactattca ttcagaggta  | 4920 |
| ggaaacttag aatagatgat gtcactgatt agcatggcct ccccatctcc acagctgctt  | 4980 |
| ccccccagg ttgcccacag ttgagtttgt ccagtgtca gggctgcca ctctcagtaa     | 5040 |
| gaagccccac accagccct ctccaatat gttggctgtt ccttcatta aagtacccc      | 5100 |
| actttagagc agcaagtgga tttctgttct ttacagtca ggaaggagga gtcagctgtg   | 5160 |
| agaacctgga gcctgagatg cttctaagtc ccactgtac tggggtcagg gaagccagac   | 5220 |
| tccagcatca gcagtcagga gcactaagcc cttgccaaca tctgtttct cagagaaact   | 5280 |
| gcttcatta taatggtgt cttttttta gctatcaagc caaacaacca gtgtctacca     | 5340 |
| ttattctcat cacctgaagc caagggttct agcaaaagtc aagctgtctt gtaatggttg  | 5400 |
| atgtgcctcc agcttctgtc ttcagtcact ccactcttag cctgctctga atcaactctg  | 5460 |
| accacagttc cctggagccc ctgccacctg ctgccctgc caccttctcc atctgcagtg   | 5520 |
| ctgtgcagcc ttctgactc ttgcagagct aataggtgga gacttgaagg aagaggagga   | 5580 |
| aagttttca taatagcctt gctgcaagct caaatgggag gtgggactg tgcccaggag    | 5640 |
| ccttgagca aaggctgtgc ccaacctctg actgcatcca ggtttgtct tgacagagat    | 5700 |
| aagaagccct ggcttttga gccaaaatct aggtcagact taggcaggat tctcaaagtt   | 5760 |
| tatcagcaga acatgaggca gaagaccctt tctgtccag cttcttcagg ctcaaccttc   | 5820 |
| atcagaatag atagaaagag aggctgtgag ggttcttaaa acagaagcaa atctgactca  | 5880 |
| gagaataaac aacctctag taaactacag cttagacaga gcatctggtg gtgagtgtgc   | 5940 |
| tcagtgtcct actcaactgt ctggtatcag cctcatgag gacttctctt ctttccctca   | 6000 |
| tagacctcca tctctgtttt ccttagcctg cagaaatctg gatggctatt cacagaatgc  | 6060 |
| ctgtgtcttc agagtgtcat ttttctctg gtattctggt tcaagcattt gaaggtagga   | 6120 |
| aaggttctcc aagtgaaga aagccagccc tgagcctcaa ctgcctggct agtgtgtca    | 6180 |
| gtaggatgca aaggctgttg aatgccacaa ggccaaactt taacctgtgt accacaagcc  | 6240 |
| tagcagcaga ggcagctctg ctactggaa ctctctgtct ttttctctt gaccttttc     | 6300 |
| ttttcttag tttctagct ctctcaacc ttacctctgc cctaccagg acaaaccacaa     | 6360 |
| gagccactgt ttctgtgatg tctctccag ccctaattag gcatcatgac ttcagcctga   | 6420 |
| ccttccatgc tcagaagcag tgctaatacca cttcagatga gctgctctat gcaacacagg | 6480 |
| cagagcctac aaacctttgc accagagccc tccacatctc agtgtttgtt catactcact  | 6540 |
| tcaacagcaa atgtgactgc tgagattaag attttacaca agatggtctg taatttcaca  | 6600 |

gttagtttta tcccattagg tatgaaagaa ttagcataat tccccttaaa catgaatgaa 6660

tcttagattt ttttaataaat agtttttgaa gtaaagacag agacatcagg agcacaagga 6720

atagcctgag aggacaaaaca gaacaagaaa gagtctggaa atacacagga tgttcttggc 6780

ctcctcaaag caagtgaag cagatagtag cagcagcccc aggctatcag agcccagtga 6840

agagaagtag catgaaagcc acagctctaa ccacctgtt ccagagttag agacagtccc 6900

caagacaagc cagcctgagc cagagagaga actgcaagag aaagtttcta atttaggttc 6960

tgttagattc agacaagtgc aggtcatcct ctctccacag ctactcacct ctccagccta 7020

acaaagcctg cagtcacac tccaacctg gtgtctcacc tctagcctc tccaacatc 7080

ctgtctctg accatcttct gcatctctca tctcaccatc tcccactgtc tacagcctac 7140

tcttgcaact accatctcat tttctgacat cctgtctaca tcttctgcca tactctgcca 7200

tctaccatac cacctcttac catctaccac accatctttt atctccatcc ctctcagaag 7260

cctccaagct gaatcctgct ttatgtgttc atctcagccc ctgcatggaa agctgacccc 7320

agaggcagaa ctattcccag agagcttggc caagaaaaac aaaactacca gcctggccag 7380

gtcaggagt agtaagctgc agtgtctgtt gtgttctagc ttcaacagct gcaggagtgc 7440

cactctcaaa tgctccacat ttctcacatc ctctgattc tggtcactac ccatcttcaa 7500

agaacagaat atctcacatc agcatactgt gaaggactag tcatgggtgc agctgtctag 7560

agctgcaaag tcattctgga tgggtggagag cttacaaaca tttcatgatg ctccccccgc 7620

tctgatggct ggagcccaat ccctacacag actcctgctg tatgtgtttt cttttcactc 7680

tgagccacag ccagagggca ggcatctagt ctctcttca ggctggggct ggggcactga 7740

gaactcacc aacaccttgc tctcactcct tctgcaaaac aagaaagagc tttgtgtctgc 7800

agtagccatg aagaatgaaa ggaaggcttt aactaaaaaa tgtcagagat tattttcaac 7860

cccttactgt ggatcaccag caaggaggaa acacaacaca gagacatttt tccccctcaa 7920

attatcaaaa gaatcactgc atttgtaaa gagagcaact gaatcaggaa gcagagtttt 7980

gaacatatca gaagttagga atctgcatca gagacaaatg cagtcatggt tgtttgctgc 8040

ataccagccc taatcattag aagcctcatg gacttcaaac atcatccct ctgacaagat 8100

gctctagcct aactccatga gataaaataa atctgccttt cagagccaaa gaagagtcca 8160

ccagcttctt ctgagtga acaagagctc cagttaggtt agtcagtcca gtgagttaga 8220

ggagaccagt ctgcatctc taattttcaa aggcaagaag atttgtttac cctggacacc 8280

aggcacaagt gaggtcacag agctcttaga tatgcagtcc tcatgagtga ggagactaaa 8340

|                                                                     |       |
|---------------------------------------------------------------------|-------|
| gcgcattgcca tcaagacttc agtgttagaga aaacctccaa aaaagcctcc tcactacttc | 8400  |
| tggaatagct cagaggccga ggcgcctcg gcctctgcat aaataaaaa aattagtcag     | 8460  |
| ccatggggcg gagaatgggc ggaactgggc ggagttaggg gcgggatggg cggagttagg   | 8520  |
| ggcgggacta tggttgctga ctaattgaga tgcattgctt gcatacttct gcctgctggg   | 8580  |
| gagcctgggg actttccaca cctggttgct gactaattga gatgcatgct ttgcatactt   | 8640  |
| ctgcctgctg gggagcctgg ggactttcca caccctaact gacacacatt ccacagctgc   | 8700  |
| attaatgaat cggccaacgc gcggggagag gcggtttgcg tatigggcgc tcttccgctt   | 8760  |
|                                                                     |       |
| cctcgtcac tgactcgtg cgctcggtcg ttggctgcg gcgagcggta tcagctcact      | 8820  |
| caaaggcggg aatagcgtta tccacagaat caggggataa cgcaggaaag aacatgtgag   | 8880  |
| caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata   | 8940  |
| ggctccggcc cctgacgag catcacaaaa atcgacgctc aagtacaggg tggcgaaacc    | 9000  |
| cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg   | 9060  |
| ttccgacct gccgcttacc ggatacctgt ccgctttct ccttcggga agcgtggcgc      | 9120  |
| tttctcatag ctacgctgt aggtatctca gtccggtgta ggtcgttcgc tccaagctgg    | 9180  |
|                                                                     |       |
| gctgtgtgca cgaaccccc gttcagcccc accgctgctg cttatccggt aactatcgtc    | 9240  |
| ttgagtccaa cccggaaga cagacttat cgccactggc agcagccact ggtaacagga     | 9300  |
| ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtggagg cctaactacg   | 9360  |
| gctacactag aagaacagta tttggtatct gcgctctgct gaagccagtt accttcggaa   | 9420  |
| aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggg ggtttttttg   | 9480  |
| tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt   | 9540  |
| ctacggggtc tgacgctcag tggaaacgaa actcacgtta agggattttg gtcattgagat  | 9600  |
|                                                                     |       |
| tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aatcaatct    | 9660  |
| aaagtatata tgagtaact tggcttgaca gttaccaatg cttaatcagt gaggcaccta    | 9720  |
| tctcagcgat ctgtctatct cgttcacca tagttgctg actcctgcaa accacgttgt     | 9780  |
| gtctcaaaat ctctgatgtt acattgcaca agataaaaa atatcatcat gaacaataaa    | 9840  |
| actgtctgct tacataaaca gtaatacaag ggggtgttatg agccatattc aacgggaaac  | 9900  |
| gtcttgctcg aggcccgcat taaattccaa catggatgct gatttatatg ggtataaatg   | 9960  |
| ggctcgcgat aatgtcgggc aatcaggtgc gacaatctat cgattgtatg ggaagccga    | 10020 |
|                                                                     |       |
| tgccccagag ttgtttctga aacatggcaa aggtagcgtt gccaatgatg ttacagatga   | 10080 |
| gatggtcaga ctaaactggc tgacggaatt tatgcctctt ccgaccatca agcattttat   | 10140 |
| ccgtactcct gatgatgcat ggttactcac cactgcgac cccgggaaaa cagcattcca    | 10200 |

ggtattagaa gaatatcctg attcaggtga aaatatgtt gatgcgctgg cagtgttcct 10260  
 gcgccggttg cattcgattc ctgtttgtaa ttgtcctttt aacagcgatc gcgtatttcg 10320  
 tctcgctcag gcgcaatcac gaatgaataa cggtttggtt gatgcgagtg attttgatga 10380  
 cgagcgtaat ggctggcctg ttgaacaagt ctggaaagaa atgcataagc ttttgccatt 10440  
  
 ctccaccgat tcagtcgtca ctcatggtga tttctcactt gataacctta tttttgacga 10500  
 ggggaaatta ataggttgta ttgatgttgg acgagtcgga atcgagacc gataccagga 10560  
 tcttgccatc ctatggaact gcctcgggtga gttttctcct tcattacaga aacggctttt 10620  
 tcaaaaaatat ggtattgata atcctgatat gaataaattg cagtttcatt tgatgctcga 10680  
 tgagtttttc taaggcggcg ctgccacat acccacgccg aaacaagcgc tcatgagccc 10740  
 gaagtggcga gcccgatctt ccccatcggt gatgtcggcg atataggcgc cagcaaccgc 10800  
 acctgtggcg ccggtgatga gggcgcgcca agtcgacgtc cggcagtc 10848

<210> 3

<211> 1446

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 3

atgagcacc tgtgcccc cccagcccc gccgtggcca agaccgagat cgcctgagc 60  
 ggcaagagcc cctgctggc cgccaccttc gcctactggg acaacatcct gggccccgc 120  
 gtgcccaca tctgggcccc caagaccgag caggtgctgc tgagcgacgg cgagatcacc 180  
 ttcctggcca accacacct gaacggcgag atcctgcgca acgccgagag cggcgccatc 240  
 gacgtgaagt tcttcgtgct gagcgagaag ggctgatca tcgtgagcct gatcttcgac 300  
 ggcaactgga acggcgaccg cagcacctac ggctgagca tcatcctgcc ccagaccgag 360  
  
 ctgagcttct acctgccct gcaccgctg tgcgtggacc gcctgacca catcatccgc 420  
 aagggccgca tctggatgca caaggagcgc caggagaacg tgcagaagat catcctggag 480  
 ggcaccgagc gcatggagga ccagggccag agcatcatcc ccatgctgac cggcgaggtg 540  
 atccccgtga tggagctgct gagcagcatg aagagccaca gcgtgcccgga ggagatcgac 600  
 atcgccgaca ccgtgctgaa cgacgacgac atcgccgaca gctgccacga gggcttcctg 660  
 ctgaacgcca tcagcagcca cctgcagacc tgcggctgca gcgtgggtgt gggcagcagc 720  
 gccgagaagg tgaacaagat cgtgcgcacc ctgtgcctgt tctgacccc cgccgagcgc 780  
  
 aagtgcagcc gcctgtgcga ggccgagagc agcttcaagt acgagagcgg cctgttcgtg 840

cagggcctgc tgaaggacag caccggcagc ttcgtgctgc ctttcgcca ggtgatgtac 900  
 gccccctacc ccaccacca catcgacgtg gacgtgaaca ccgtgaagca gatgcccccc 960  
 tgccacgagc acatctacaa ccagcgccgc tacatgcgca gcgagctgac cgccttctgg 1020  
 cgcgccacca gcgaggagga catggcccag gacacatca tctacaccga cgagagcttc 1080  
 acccccgacc tgaacatctt ccaggacgtg ctgcaccgag acaccctggt gaaggccttc 1140  
 ctggaccagg tgttccagct gaagcccggc ctgagcctgc gcagcacctt cctggcccag 1200

ttcctgctgg tgctgcaccg caaggccctg accctgatca agtacatcga ggacgacacc 1260  
 cagaagggca agaagccctt caagagcctg cgcaacctga agatcgacct ggacctgacc 1320  
 gccgagggcg acctgaacat catcatggcc ctggccgaga agatcaagcc cggcctgcac 1380  
 agcttcatct tcggccgccc cttctacacc agcgtgcagg agcgcgacgt gctgatgacc 1440  
 ttctaa 1446

<210> 4

<211> 481

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Polypeptide

<400> 4

Met Ser Thr Leu Cys Pro Pro Pro Ser Pro Ala Val Ala Lys Thr Glu

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

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

115                                      120                                      125  
 Arg Val Cys Val Asp Arg Leu Thr His Ile Ile Arg Lys Gly Arg Ile  
  
 130                                      135                                      140  
 Trp Met His Lys Glu Arg Gln Glu Asn Val Gln Lys Ile Ile Leu Glu  
 145                                      150                                      155                                      160  
 Gly Thr Glu Arg Met Glu Asp Gln Gly Gln Ser Ile Ile Pro Met Leu  
 165                                      170                                      175  
 Thr Gly Glu Val Ile Pro Val Met Glu Leu Leu Ser Ser Met Lys Ser  
 180                                      185                                      190  
 His Ser Val Pro Glu Glu Ile Asp Ile Ala Asp Thr Val Leu Asn Asp  
  
 195                                      200                                      205  
 Asp Asp Ile Gly Asp Ser Cys His Glu Gly Phe Leu Leu Asn Ala Ile  
 210                                      215                                      220  
 Ser Ser His Leu Gln Thr Cys Gly Cys Ser Val Val Val Gly Ser Ser  
 225                                      230                                      235                                      240  
 Ala Glu Lys Val Asn Lys Ile Val Arg Thr Leu Cys Leu Phe Leu Thr  
 245                                      250                                      255  
 Pro Ala Glu Arg Lys Cys Ser Arg Leu Cys Glu Ala Glu Ser Ser Phe  
  
 260                                      265                                      270  
 Lys Tyr Glu Ser Gly Leu Phe Val Gln Gly Leu Leu Lys Asp Ser Thr  
 275                                      280                                      285  
 Gly Ser Phe Val Leu Pro Phe Arg Gln Val Met Tyr Ala Pro Tyr Pro  
 290                                      295                                      300  
 Thr Thr His Ile Asp Val Asp Val Asn Thr Val Lys Gln Met Pro Pro  
 305                                      310                                      315                                      320  
 Cys His Glu His Ile Tyr Asn Gln Arg Arg Tyr Met Arg Ser Glu Leu  
  
 325                                      330                                      335  
 Thr Ala Phe Trp Arg Ala Thr Ser Glu Glu Asp Met Ala Gln Asp Thr  
 340                                      345                                      350  
 Ile Ile Tyr Thr Asp Glu Ser Phe Thr Pro Asp Leu Asn Ile Phe Gln  
 355                                      360                                      365

Asp Val Leu His Arg Asp Thr Leu Val Lys Ala Phe Leu Asp Gln Val

370 375 380

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

385 390 395 400

Phe Leu Leu Val Leu His Arg Lys Ala Leu Thr Leu Ile Lys Tyr Ile

405 410 415

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

420 425 430

Leu Lys Ile Asp Leu Asp Leu Thr Ala Glu Gly Asp Leu Asn Ile Ile

435 440 445

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

450 455 460

Gly Arg Pro Phe Tyr Thr Ser Val Gln Glu Arg Asp Val Leu Met Thr

465 470 475 480

Phe

<210> 5

<211> 10900

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 5

ttggccactc ccctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcgcctca gtgagcgagc gagcgcgag agagggagtg 120

gccaaactcca tcaactagggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca tttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccttaggtt ctagaaccgg tgacgtcttg tcatectccc acggtggcca ttgtttccat 300

gtgagtgcta gtaacaggcc ttgtgtcctg gggccggttt cggggccgta catctgtggc 360

ttcactaagg ccccgaaacc ggccccgaca acagcataca gccttcagca agcctccagt 420

ggtctcatag agaacttata agattcccaa atccaaagac atttcacgtt tatggtgatt 480

tcccagaaca catagcgaca tgcaaatatt gcagggcgcc actcccctgt ccctcacagc 540

|                                                                    |      |
|--------------------------------------------------------------------|------|
| catcttcctg ccagggcgca cgcgcgctgg gtgttccgc ctagtgacac tgggcccgcg   | 600  |
|                                                                    |      |
| attccttgga gcgggttgat gacgtcagcg ttcccatgg tgaagcttgg atctgaattc   | 660  |
| ggtaccctag ttattaatag taatcaatta cggggtcatt agttcatagc ccataatagg  | 720  |
| agttccgctg tacataactt acggtaaatg gccgcctgg ctgaccgcc aacgaccccc    | 780  |
| gcccattgac gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt  | 840  |
| gacgtcaatg ggtggactat ttacggtaaa ctgccactt ggagtagcat caagtgtatc   | 900  |
| atatgccaag tacgccccct attgacgtca atgacggtaa atggcccgcc tggcattatg  | 960  |
| cccagtagat gaccttatgg gactttctta ctggcgagta catctacgta ttagtcatcg  | 1020 |
|                                                                    |      |
| ctattacat ggicgaggtg agccccacgt tctgttcac tctcccatc tccccccct      | 1080 |
| ccccacccc aattttgtat ttatttatit ttaattatt ttgtgcagcg atggggcgcg    | 1140 |
| gggggggggg ggggcgcgcg ccagggcgggg cggggcgggg cgagggcgcg ggcggggcga | 1200 |
| ggcgagaggg tgcggcgcca gccaatcaga gcggcgcgct ccgaaagttt ctttttatgg  | 1260 |
| cgagggcgcg gcggcgcgcg ccctataaaa agcgaagcgc gcggcgggcg ggagtcgctg  | 1320 |
| cgacgtgcc ttcgccccgt gccccgctcc gccgcccct cgcgccgcc gccccggctc     | 1380 |
| tgactgaccg cgttactccc acaggtgagc gggcgggagc gcccttctcc tccgggctgt  | 1440 |
|                                                                    |      |
| aattagcgt tggtttaatg acggcttgtt ttctgtggct gcgtgaaagc cttagggggc   | 1500 |
| tccgggagct agagcctctg ctaaccatgt tcatgccttc ttctttttcc tacagtcct   | 1560 |
| gggcaacgtg ctggttatgt tgctgtctca tcattttggc aaagaattcc tcgaagatcc  | 1620 |
| gaagggaag tcttccacga ctgtgggac cgttcgaaga tatcacgggt tgagccacca    | 1680 |
| tgagcacct gtgcccccc cccagccccg ccgtggccaa gaccgagatc gccctgagcg    | 1740 |
| gcaagagccc cctgctggcc gccaccttcg cctactggga caacatcctg ggcccccgcg  | 1800 |
| tgcgccacat ctgggcccc aagaccgagc aggtgctgct gagcgacggc gagatcacct   | 1860 |
|                                                                    |      |
| tcctggccaa ccacaccctg aacggcgaga tcctgcgcaa cgccgagagc ggcgccatcg  | 1920 |
| acgtgaagtt cttcgtgctg agcgagaagg gcgtgatcat cgtgagcctg atcttcgacg  | 1980 |
| gcaactggaa cggcgaccgc agcacctacg gcctgagcat catcctgccc cagaccgagc  | 2040 |
| tgagcttcta cctgcccctg caccgcgtgt gcgtggaccg cctgaccac atcatccga    | 2100 |
| agggccgcat ctggatgcac aaggagcgcc aggagaacgt gcagaagatc atcctggagg  | 2160 |
| gcaccgagcg catggaggac cagggccaga gcatcatccc catgctgacc ggcgaggtga  | 2220 |
| tccccgtgat ggagctgctg agcagcatga agagccacag cgtgcccag gagatcgaca   | 2280 |



|                                                                    |      |
|--------------------------------------------------------------------|------|
| tcgccgacac cgtgctgaac gacgacgaca tcggcgacag ctgccacgag ggcttcctgc  | 2340 |
| tgaacccat cagcagccac ctgcagacct gcggctgcag cgtggtggtg ggcagcagcg   | 2400 |
| ccgagaaggt gaacaagatc gtgcgcaccc tgtgcctgtt cctgaccccc gccgagcgca  | 2460 |
| agtgcagccg ccgtgctgag gccgagagca gcttcaagta cgagagcggc ctgttcgtgc  | 2520 |
| agggcctgct gaaggacagc accggcagct tcgtgctgcc cttccgccag gtgatgtacg  | 2580 |
| ccccctacc caccaccac atcgacgtgg acgtgaacac cgtgaagcag atgccccct     | 2640 |
| gccacgagca catctaac cagcgccgct acatgcgcag cgagctgacc gccttctggc    | 2700 |
| gcgccaccag cgaggaggac atggcccagg acaccatcat ctacaccgac gagagcttca  | 2760 |
| cccccgacct gaacatcttc caggacgtgc tgcaccgca caccctggtg aaggccttc    | 2820 |
| tggaccaggt gtccagctg aagcccggcc tgaacctgcg cagcaccttc ctggcccagt   | 2880 |
| tcctgtggt gtgcaccgc aaggccctga cctgatcaa gtacatcgag gacgacacc      | 2940 |
| agaagggcaa gaagcccttc aagacctgc gcaacctgaa gatcgacctg gacctgaccg   | 3000 |
| ccgagggcga cctgaacatc atcatggccc tggccgagaa gatcaagccc ggcctgcaca  | 3060 |
| gcttcatctt cgccgcccc ttctacacca gcgtgcagga gcgcgacgtg ctgatgacct   | 3120 |
| tctaattgaca attgttaatt aagtttaaac cctcgaggcc gcaagcttat cgataatcaa | 3180 |
| cctctggatt acaaaatttg tgaaagattg actggtattc ttaactatgt tgctcctttt  | 3240 |
| acgtatgtg gatagctgc tttaatgcct ttgtatcatg ctattgcttc ccgtatggct    | 3300 |
| ttcattttct cctccttgta taaatcctgg ttgctgtctc ttatgagga gttgtggccc   | 3360 |
| gttgtcaggc aacgtggcgt ggtgtgcact gtgtttgctg acgcaacccc cactggttgg  | 3420 |
| ggcattgcca ccacctgtca gctcctttcc gggactttcg cttccccct ccctattgcc   | 3480 |
| acggcggaac tcatcgccgc ctgccttgcc cgctgctgga caggggctcg gctgttgggc  | 3540 |
| actgacaatt ccgtggtgtt gtccgggaaa tcacgtcctt ttcttggct gctgcctgt    | 3600 |
| gttgccacct ggattctgcg cgggacgtcc ttctgctacg tcccttcggc cctcaatcca  | 3660 |
| gcggaccttc cttcccgcg cctgtgcgc gctctgcggc ctcttcgcg tcttcgcctt     | 3720 |
| cgccctcaga cgagtcggat ctccctttgg gccgcctccc cgcatcgata ccgtcgacta  | 3780 |
| gagctcgctg atcagcctcg actgtgcctt ctagtgtcca gccatctgtt gtttgccct   | 3840 |
| ccccgtgcc ttcttgacc ctggaagtg ccactccac tgtcctttcc taataaaatg      | 3900 |
| aggaaattgc atcgattgt ctgagtaggt gtcattctat tctgggggtt ggggtggggc   | 3960 |
| aggacagcaa gggggaggat tgggaagaca atagcaggca tgctggggag agatccacga  | 4020 |
| taacaaacag cttttttggg gtgaacatat tgactgaatt cctgcaggt tggccactcc   | 4080 |
| ctctctgcgc gctcgctgc tcaatgaggc cgccccggca aagcccgggc gtcgggcgac   | 4140 |

|                                                                     |      |
|---------------------------------------------------------------------|------|
| ctttggtcgc ccggcctcag tgagcgagcg agcgcgagcaga gagggagtgg ccaactccat | 4200 |
| cactagggtt tcttgcggcc gctcgtacgg tctcgaggaa ttcctgcagg ataacttgcc   | 4260 |
| aacctcattc taaaatgtat atagaagccc aaaagacaat aacaaaaata ttctttaga    | 4320 |
| acaaaatggg aaagaatgtt ccactaaata tcaagattta gagcaaagca tgagatgtgt   | 4380 |
| gggatagac agtgaggctg ataaaaataga gtagagctca gaaacagacc cattgatata   | 4440 |
| tgtaagtgc ctatgaaaaa aatatggcat ttacaatgg gaaaatgatg gtctttttct     | 4500 |
| tttttagaaa aacagggaat tataattata tgtaaaaaat aaaagggaac ccataatgca   | 4560 |
| taccatacac acaaaaaaat tccagtgaat tataagtcta aatggagaag gcaaaacttt   | 4620 |
| aaatctttta gaaaataata tagaagcatg cagaccagcc tggccaacat gatgaaaccc   | 4680 |
| tctctactaa taataaaatc agtagaacta ctcaggacta ctttgagtgg gaagtccttt   | 4740 |
| tctatgaaga cttcttttggc caaaattagg ctctaaatgc aaggagatag tgcacatgc   | 4800 |
| ctggctgcac ttactgataa atgatgttat caccatcttt aaccaaagtc acaggaacaa   | 4860 |
| gttatggtag tgatgtgctg gattgagaag gagctctact tccttgacag gacacatttg   | 4920 |
| tatcaactta aaaaagcaga tttttgccag cagaactatt cattcagagg taggaaactt   | 4980 |
| agaatagatg atgtcactga ttagcatggc ttcccatctt ccacagctgc ttccaccca    | 5040 |
| ggttgcccac agttgagttt gtccagtgtc cagggtgccc cactctcagt aagaagcccc   | 5100 |
| acaccagccc ctctccaaat atgttggtctg ttcccttccat taaagtgacc ccactttaga | 5160 |
| gcagcaagtg gatttctgtt tcttacagtt caggaaggag gagtacagctg tgagaacctg  | 5220 |
| gagcctgaga tgcttctaag tcccactgtc actggggtca gggaagccag actccagcat   | 5280 |
| cagcagtcag gagcactaag cccttgccaa catcctgttt ctcagagaaa ctgcttccat   | 5340 |
| tataatggtt gtcttttttt aagctatcaa gccaaacaac cagtgtctac cattattctc   | 5400 |
| atcacctgaa gccaaagggtt ctagcaaaaag tcaagctgtc ttgtaatggt tgatgtgcct | 5460 |
| ccagcttctg tcttcagtc ctcactctt agcctgtctt gaatcaactc tgaccacagt     | 5520 |
| tccttgagc ccctgccacc tgctgcccct gccacattct ccactgcag tgctgtgcag     | 5580 |
| ccttctgcac tcttgcagag ctaatagggt gagacttgaa ggaagaggag gaaagtttct   | 5640 |
| cataatagcc ttgctgcaag ctcaaatggg aggtgggcac tgtgcccagg agccttgag    | 5700 |
| caaagctgt gcccaacctc tgactgcac caggtttggg cttgacagag ataagaagcc     | 5760 |
| ctggcttttg gagccaaaat ctaggtcaga cttaggcagg attctcaaag ttatcagca    | 5820 |
| gaacatgagg cagaagaccc tttctgtcc agcttcttca ggctcaacct tcatcagaat    | 5880 |
| agatagaaag agaggctgtg agggttctta aaacagaagc aaatctgact cagagaataa   | 5940 |
| acaacctcct agtaaacac agcttagaca gagcatctgg tggtagtgt gtcagtgtc      | 6000 |

|                                                                   |      |
|-------------------------------------------------------------------|------|
| ctactcaact gtctggatc agccctcatg aggacttctc ttttttcct catagacctc   | 6060 |
|                                                                   |      |
| catctctgtt ttccttagcc tgcagaaatc tggatggcta ttcacagaat gcctgtgctt | 6120 |
| tcagagtgc atttttctc tggattctg gttcaagcat ttgaaggtag gaaaggttct    | 6180 |
| ccaagtgcaa gaaagccagc cctgagcctc aactgcctgg ctagtgtggt cagtaggatg | 6240 |
| caaagctgt tgaatgccac aaggccaaac tttaacctgt gtaccacaag cctagcagca  | 6300 |
| gaggcagctc tgcctactgg aactctctgt cttctttctc ctgagccttt tcttttctg  | 6360 |
| agttttctag ctctctcaa ccttacctct gcctaccca ggacaaacc aagagccact    | 6420 |
| gtttctgtga tgcctctcc agccctaatt aggcatcatg acttcagcct gacctccat   | 6480 |
|                                                                   |      |
| gctcagaagc agtgctaate cacttcagat gagctgctct atgcaacaca ggcagagcct | 6540 |
| acaaaccttt gcaccagagc cctccacata tcagtgttg ttcatactca cttcaacagc  | 6600 |
| aatgtgact gctgagatta agattttaca caagatggtc tgtaatttca cagttagttt  | 6660 |
| tatccatta ggtatgaaag aattagcata attccccta aacatgaatg aatcttagat   | 6720 |
| tttttaataa atagtgttg aagtaaagac agagacatca ggagcacaag gaatagcctg  | 6780 |
| agaggacaaa cagaacaaga aagagtctgg aaatacacag gatgttcttg gcctctcaa  | 6840 |
| agcaagtgc agcagatagt accagcagcc ccaggctatc agagcccagt gaagagaagt  | 6900 |
|                                                                   |      |
| accatgaaag ccacagctct aaccacctg ttccagagtg acagacagtc cccaagacaa  | 6960 |
| gccagcctga gccagagaga gaactgcaag agaaagtctc taatttaggt tctgttagat | 7020 |
| tcagacaagt gcaggtcatc ctctctccac agctactcac ctctccagcc taacaaagcc | 7080 |
| tgcagtcac actccaacc tgggtgtctc cctctagcc tctccaaca tctgtctc       | 7140 |
| tgaccatctt ctgcatctct catctacca tctccactg tctacagcct actcttgcaa   | 7200 |
| ctaccatctc atttctgac atctgtctc catcttctgc catactctgc catctacat    | 7260 |
| accactctt accatctacc acaccatctt ttatctccat cctctcaga agcctccaag   | 7320 |
|                                                                   |      |
| ctgaatcctg ctttatgtgt tcatctcagc cctgcatgg aaagctgacc ccagaggcag  | 7380 |
| aactatccc agagagcttg gccaaaaaa aaaaaactac cagcctggcc aggctcagga   | 7440 |
| glagtaagct gcagtgtctg ttgtgttcta gcttcaacag ctgcaggagt tccactctca | 7500 |
| aatgctccac atttctcaca tctctctgat tctggctact acccatctc aaagaacaga  | 7560 |
| atatctcaca tcagcatact gtgaaggact agtcatgggt gcagctgctc agagctgcaa | 7620 |
| agtcattctg gatgggtggag agcttcaaaa catttcatga tgctcccc gctctgatgg  | 7680 |
| ctggagccca atccctacac agactcctgc tgtatgtgtt ttcctttcac tctgagccac | 7740 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| agccagaggg caggcattca gtctcctctt caggctgggg ctggggcact gagaactcac  | 7800 |
| ccaacacctt gtctctactc cttctgcaaa acaagaaaga gctttgtgct gcagtagcca  | 7860 |
| tgaagaatga aaggaaggct ttaactaaaa aatgtcagag attatittca accccttact  | 7920 |
| gtggatcacc agcaaggagg aaacacaaca cagagacatt ttttccctc aaattatcaa   | 7980 |
| aagaatcact gcattttgta aagagagcaa ctgaatcagg aagcagagtt ttgaacatat  | 8040 |
| cagaagttag gaatctgcat cagagacaaa tgcagtcatt gttgtttgct gcataccagc  | 8100 |
| cctaattcatt agaagcctca tggacttcaa acatcattcc ctctgacaag atgctctagc | 8160 |
| ctaactccat gagataaaat aaatctgcct ttcagagcca aagaagagtc caccagcttc  | 8220 |
| ttctcagtgt gaacaagagc tccagtcagg ttagtcagtc cagtcagta gaggagacca   | 8280 |
| gtctgcatcc tctaattttc aaaggcaaga agattttgtt accctggaca ccaggcacia  | 8340 |
| gtgaggtcac agagctctta gatatgcagt cctcatgagt gaggagacta aagcgcatgc  | 8400 |
| catcaagact tcagtgtaga gaaaacctcc aaaaaagcct cctcactact tctggaatag  | 8460 |
| ctcagaggcc gaggcggcct cggcctctgc ataaataaaa aaaattagtc agccatgggg  | 8520 |
| cggagaatgg gcggaactgg gcggagttag gggcgggatg ggcgagttta ggggcgggac  | 8580 |
| tatggttgct gactaatga gatgcatgct ttgcatactt ctgcctgctg gggagcctgg   | 8640 |
| ggactttcca cacctggttg ctgactaatt gagatgcatg ctttgcatc ttctgcctgc   | 8700 |
| tggggagcct ggggactttc cacaccctaa ctgacacaca ttccacagct gcattaatga  | 8760 |
| atcgccaac gcgcggggag aggcggtttg cgtattgggc gctcttcgc ttctcgtc      | 8820 |
| actgactcgc tgcgctcgtt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg  | 8880 |
| gtaatacgtt tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc  | 8940 |
| cagcaaaagg ccaggaaccg taaaaggcc gcgttgctgg cgtttttcca taggtccgc    | 9000 |
| ccccctgacg agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga  | 9060 |
| ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttcgacc   | 9120 |
| ctgccgtta ccggatacct gtccgccttt ctcccttcgg gaagcgtgac gctttctcat   | 9180 |
| agctcacgct gtaggtatct cagttcgggt taggtcgttc gctccaagct gggctgtgtg  | 9240 |
| cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtc   | 9300 |
| aaccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga   | 9360 |
| gcgaggtatg taggcggtgc tacagagttc ttgaagtggg ggcctaacta cggctacact  | 9420 |
| agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt  | 9480 |
| ggtagctctt gatccggcaa acaaacacc gctggttagcg gtggtttttt tgtttgcaag  | 9540 |
| cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg  | 9600 |

tctgacgctc agtgggaacga aaactcacgt taagggatit tggatcatgag attatcaaaa 9660  
aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata 9720  
tatgagtaaa cttgggtctga cagttaccaa tgcctaatca gtgaggcacc tatctcagcg 9780  
atctgtctat ttcgtttcatc catagttgcc tgactcctgc aaaccacgtt gtgtctcaaa 9840

atctctgatg ttacattgca caagataaaa atatatcatc atgaacaata aaactgtctg 9900  
cttacataaa cagtaataca aggggtgtta tgagccatat tcaacgggaa acgtcttgct 9960  
cgaggccgcg attaaattcc aacatggatg ctgatttata tgggtataaa tgggctcgcg 10020  
ataatgtcgg gcaatcaggt gcgacaatct atcgattgta tgggaagccc gatgcgccag 10080  
agttgtttct gaaacatggc aaaggtagcg ttgccaatga tgttacagat gagatggtca 10140  
gactaaactg gctgacggaa tttatgcctc ttccgaccat caagcatttt atccgtactc 10200  
ctgatgatgc atggttactc accactgcga tccccgggaa aacagcatc caggtattag 10260

aagaatatcc tgattcaggt gaaaatattg ttgatgcgtt ggcagtgttc ctgcgccggt 10320  
tgattcagat tctgtttgt aattgtcctt ttaacagcga tcgctattt cgtctcgctc 10380  
aggcgcaatc acgaatgaat aacggtttgg ttgatgcgag tgattttgat gacgagcgta 10440  
atggctggcc tgttgaacaa gtctggaaag aaatgcataa gcttttgcca ttctcaccgg 10500  
attcagtcgt cactcatggt gatttctcac ttgataacct tatttttgac gaggggaaat 10560  
taataggttg tattgatgtt ggacgagtcg gaatgcgaga ccgataccag gatcttgcca 10620  
tcctatggaa ctgcctcggt gagttttctc cttcattaca gaaacggctt tttcaaaaat 10680

atggatttga taatcctgat atgaataaat tgcagtttca tttgatgctc gatgagtttt 10740  
tctaaggcgc gcctgccacc ataccacgc cgaacaagc gctcatgagc ccgaagtggc 10800  
gagcccgatc ttccccatcg gtgatgtcgg cgatatagcg gccagcaacc gcacctgtgg 10860  
cgccggtgat gagggcgcgc caagtcgacg tccggcagtc 10900

<210> 6

<211> 274

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Polypeptide

<400> 6

Met Gly Lys Ser Leu Ser His Leu Pro Leu His Ser Ser Lys Glu Asp

1 5 10 15

Ala Tyr Asp Gly Val Thr Ser Glu Asn Met Arg Asn Gly Leu Val Asn

|                                                                 |     |     |
|-----------------------------------------------------------------|-----|-----|
| 20                                                              | 25  | 30  |
| Ser Glu Val His Asn Glu Asp Gly Arg Asn Gly Asp Val Ser Gln Phe |     |     |
| 35                                                              | 40  | 45  |
| Pro Tyr Val Glu Phe Thr Gly Arg Asp Ser Val Thr Cys Pro Thr Cys |     |     |
| 50                                                              | 55  | 60  |
| Gln Gly Thr Gly Arg Ile Pro Arg Gly Gln Glu Asn Gln Leu Val Ala |     |     |
| 65                                                              | 70  | 75  |
|                                                                 |     | 80  |
|                                                                 |     |     |
| Leu Ile Pro Tyr Ser Asp Gln Arg Leu Arg Pro Arg Arg Thr Lys Leu |     |     |
| 85                                                              | 90  | 95  |
| Tyr Val Met Ala Ser Val Phe Val Cys Leu Leu Leu Ser Gly Leu Ala |     |     |
| 100                                                             | 105 | 110 |
| Val Phe Phe Leu Phe Pro Arg Ser Ile Asp Val Lys Tyr Ile Gly Val |     |     |
| 115                                                             | 120 | 125 |
| Lys Ser Ala Tyr Val Ser Tyr Asp Val Gln Lys Arg Thr Ile Tyr Leu |     |     |
| 130                                                             | 135 | 140 |
|                                                                 |     |     |
| Asn Ile Thr Asn Thr Leu Asn Ile Thr Asn Asn Asn Tyr Tyr Ser Val |     |     |
| 145                                                             | 150 | 155 |
| Glu Val Glu Asn Ile Thr Ala Gln Val Gln Phe Ser Lys Thr Val Ile |     |     |
| 165                                                             | 170 | 175 |
| Gly Lys Ala Arg Leu Asn Asn Ile Thr Ile Ile Gly Pro Leu Asp Met |     |     |
| 180                                                             | 185 | 190 |
| Lys Gln Ile Asp Tyr Thr Val Pro Thr Val Ile Ala Glu Glu Met Ser |     |     |
| 195                                                             | 200 | 205 |
|                                                                 |     |     |
| Tyr Met Tyr Asp Phe Cys Thr Leu Ile Ser Ile Lys Val His Asn Ile |     |     |
| 210                                                             | 215 | 220 |
| Val Leu Met Met Gln Val Thr Val Thr Thr Tyr Phe Gly His Ser     |     |     |
| 225                                                             | 230 | 235 |
| Glu Gln Ile Ser Gln Glu Arg Tyr Gln Tyr Val Asp Cys Gly Arg Asn |     |     |
| 245                                                             | 250 | 255 |
| Thr Thr Tyr Gln Leu Gly Gln Ser Glu Tyr Leu Asn Val Leu Gln Pro |     |     |
| 260                                                             | 265 | 270 |

Gln Gln

<210> 7

<211> 6514

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 7

```

aggcgcggac gcaggttaca gcagcgcttg gcctctgctg atgccgtcgt tatectaccc      60
ctccccctgc ccagctctac ggcgcccgcg cgctccaggc cggctcgtcc acccccggc      120
tccccggact gtggactcca cgacctgtc ctcgcccttg tccgcgccga agcagcccgg      180
gactgcgcag cgcccccggt gccgatcttt tcctaattca gcagcgattt aaccaagagc      240
ctggaatatt ttaaggagta ataagagaca ttacaaact attctctctg aagcctgcta      300

cctggaggca tcctctagat aatcagaacc ttggttcca catcctctc ccttgtctta      360
actacaaaca tttctttctg ctgacttcaa ctctcagac atgggaaagt ctctttctca      420
tttgctttg cattcaagca aagaagatgc ttatgatgga gtcacatctg aaaacatgag      480
gaatggactg gttaatagtg aagtccataa tgaagatgga agaaatggag atgtctctca      540
gtttccatat gtggaattta caggaagaga tagtgtcacc tgccctactt gtcagggaac      600
aggaagaatt ctaggggggc aagaaaacca actggtggca ttgattccat atagtatca      660
gagattaagg ccaagaagaa caaagctgta tgtgatggct tctgtgtttg tctgtctact      720

cctttctgga ttggctgtgt ttttctttt ccctcgctct atcgacgtga aataattgg      780
tgtaaaatca gcctatgtca gttatgatgt tcagaagcgt acaatttatt taaatatcac      840
aaacacacta aatataacaa acaataacta ttactctgtc gaagttgaaa acatcactgc      900
ccaagttcaa ttttcaaaaa cagttattgg aaaggcacgc ttaacaaca taaccattat      960
tggtcactt gatatgaaac aaattgatta cacagtacct accgttatag cagaggaaat      1020
gagttatatg tatgatttct gtactctgat atccatcaaa gtgcataaca tagtactcat      1080
gatgcaagtt actgtgacaa caacatactt tggccactct gaacagatat cccaggagag      1140

glatcagtat gtcgactgtg gaagaaacac aacttatcag ttggggcagt ctgaatattt      1200
aatgtactt cagccacaac agtaaaaact ggaagagatg gatttaaaga agaaatatct      1260
attgatattt cctatactct caatgaagag gtatttccta ataggagacc ttaaattgaa      1320
caaacctaaa gtttactt ctaagagtac agttaaagt atgtggacct gcagtcttg      1380
taactctcca ctctgtgtta atgatataat tgtactagga tcttttactt gaatctaaat      1440

```

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ttactggttg atttccttct ccagcctatc ccctacaggg aaaagctgat acttcccta   | 1500 |
| tagtacaata aataattatt taaaagtcag agctccagtc actactgaaa acataat     | 1560 |
|                                                                    |      |
| ggtgataaac ataatttgag aaacttaatt tctgaatgtt tttatagaaa attactgaaa  | 1620 |
| gtctattact catggaagac ttttaagaa taaccttttt tccgttttta taaattccca   | 1680 |
| ttgttatatg gtagtatttc agctacacaa tattttagct tttagctaga cttttatagc  | 1740 |
| ttttcatttg ttgaaatggt aatcatctgc atgtttttgt cacttatttc aggttagtga  | 1800 |
| ttgcctaaca cttataagcc aaaataatct ttgcaaaatt ccatacctaa aattttgaaa  | 1860 |
| gcccctaagtg ttttcacaca tctttctgta ttagttatag ttttgtagaa tctttgtgtg | 1920 |
| atcttcaaac attatcattt aatgtacaat actgtaaata aactgtgcat ggcttttata  | 1980 |
|                                                                    |      |
| cagcttttagt aaatgtcaaa taaagtggta cagactcatt acaacaagtt tctcataaaa | 2040 |
| atacaataaa taggaaaatg aaattcagaa acccatagac tgggaatagg ttccagttac  | 2100 |
| agcttggatc tggcataaaa taaatttgaa ataaaatatt ttgatgctcc atttttttat  | 2160 |
| gttgcttttc atactaaaga atgggtgtaga catgttttgc aactgttagg taccagtta  | 2220 |
| tcaattttat caatgtttta gaggaggaaa ttattttttt ggtagaaatt gttcaagaaa  | 2280 |
| tccttaattg aatgtcatta aatgatggtg gccaaaataa aacctattta gaaatttaat  | 2340 |
| cactttgcac atcacttggg atatgatgcc tctagtagtt acttttttat agttttctac  | 2400 |
|                                                                    |      |
| ttttggtttt atttaaaatt gttttcaa atagattatt gacttattca actttgctgt    | 2460 |
| tttatatttt cagtatcatt ttccattttt tttttttttt gtcttttcac ttaccaagtt  | 2520 |
| ctaggacat ttaaaatag tactaagtgt aggagtgggt atgatacca aaaatgtagc     | 2580 |
| tgggttgaga ttaatttcgt tctgttttct catgacagaa atcaggtttc ctttcccca   | 2640 |
| cccctaagtg cctaacttag gtctgaaaca gcctgtttat tagtctgact ctctcaacca  | 2700 |
| taaaacataa gctttattta attctgcctt taaacacact caggtttccc ctttaatttc  | 2760 |
| atattatttt ctgcaagttt tcttgagtat ctcaattcg ttgaatgtgg tttttggttt   | 2820 |
|                                                                    |      |
| tttttggttt taacactagt ctcccttaaa ttcatgtcta actcaagcca tccttactat  | 2880 |
| taaacccaaa tcagtccttt aagttcatta tggcctttct agtatttaaa aaaaaaaaaa  | 2940 |
| tcattttcat ttttctctg ctacgtttcc tgactactac tgcatacttc tctgatacag   | 3000 |
| gttctgtttg tattttttat atcattctca ttttctcatt tgacatgac tatgtctata   | 3060 |
| tatgatatag gtcccccttt gtctcaaaat ttttaattat gtgacttcaa aaatcacctg  | 3120 |
| tatctgtagt agggcttcca aatctgcttc tccatatgtg accagtcacc tgtctgcttt  | 3180 |



|                                                                     |      |
|---------------------------------------------------------------------|------|
| cacatttagc tagtgaacta cacatttact aaaatgtgta aattttacac atttagtgac   | 3240 |
|                                                                     |      |
| tgtgtaaaat aaaaaaaaaag ttatttttatc atatcctttc tattatgttc ccatcctgtc | 3300 |
| ctcatgtccc atttacttta ttatcaccat tcattttcttc aaaattatct tttagatacg  | 3360 |
| ctcatacaaa aatcaatcct tgttttcttg ctgtgtctct ttaaccttgg aaaattacat   | 3420 |
| cgtgtaaatt aaacagattt ttctgatgat ctgtgcttct tatatactat tagagtgcac   | 3480 |
| gatagtatct cctgaaaagg atggaaagta gaagcatttg cttttagtca ctttaatttg   | 3540 |
| aatctttttt ctcatctttt tgaattaatt ttttttatta tatctacttt tagtggagtt   | 3600 |
| tgagtcagaa aaaaacaaga atttgaaaca agtaaaaaga tagaagagaa ataaagatgg   | 3660 |
|                                                                     |      |
| tatgtgacta ctttcagaga gagttaagta actgtcagaa taagcctgga acaaaacagg   | 3720 |
| ctgtaaatta ataaaactac aaacacacat tcagggtgaag cagaagtata gccataaaac  | 3780 |
| atctagaag agtgaatgag gcttttagct tttcttaggt caatgtccag tgtgcttttt    | 3840 |
| tccatgggaa taggataggt attaatacgc ttttctaaac tgctctcaga cttatccag    | 3900 |
| aggacatggt aaagatatgt tacagaaatt tttctgatac ttcctggaat aactttaagt   | 3960 |
| tacaccctag tagactggtc attctaataa aatccagtac tataacaaac ctctgtatgt   | 4020 |
| tgatagcaca ttggcccttt ttagagttct ttcctatggt tttcttacgt gatttcccac   | 4080 |
|                                                                     |      |
| agttccatga gtccaacaaa ggagagtgat aggtctcctta tcttttagaa gaggaaggaa  | 4140 |
| aggcatgaag aagttgaggg actggctgaa gatcacgtac ttactaagta gtacaactgg   | 4200 |
| agcaagatca agtatctctg tctcccatat ctgtgttcta tcatttaaaa tatatatgg    | 4260 |
| aaatccctgc tgactcagat tggatgatt aaaaatgaga ggaaagttca aatagttagt    | 4320 |
| agtgacaaac taatactgct ggactaagat ttggttagca ttgttttcta aaatatatta   | 4380 |
| aatggagaat gaacacttat aaaaatgctt ggaacataat ctttagctta atttctgtt    | 4440 |
| aaaatttagt accccttcat cattccaata aagataagac tgatccattg tctaaggaaa   | 4500 |
|                                                                     |      |
| ttatttataa ataatagaga ttaatttatt tgagatttga aataagaata gtagaaaat    | 4560 |
| attagatacc acataaatg tttgaaatta ctgaataacc atcttaagta tggaacattt    | 4620 |
| aatggctat attttatttg tgtacagttt tttctgtcct tgtaggcca gtgaagcaat     | 4680 |
| tattttctct aagaaaatga caataaaata taacacactt cagattgtct gatttacagt   | 4740 |
| ttggaaagga caccgcaatg ttcaaatagg taggagacca tcaaaaacac aattaaagta   | 4800 |
| acataatagg agacttgaaa cttcagccta ataaatcctt catggttctt agccttatta   | 4860 |
| ttgtgatata attctagata ttttcttgga gggcatgtgc ccaactctcc cgcaccccat   | 4920 |

ttgtttgtc ttttaaagtt cttagaataa acagttcttt atataataat tatattttat 4980  
 ttaagaaaat agttttag gtacttttta aaagatgtaa atttttaaat ttacaaatac 5040  
 atatgggtct ttgataagca ataggaattg aattacaagt tactagggtt ataagcaaaa 5100  
 ggttgcttac cataatgtca ttaggtcacg atttttagct cacatctgga agcagcaact 5160  
 acttggtcag agtacatata agagtaatta gttttattct ctctttttta taaaatcggg 5220  
 tttcagatga gatgtttatc ttagactatt ttagggaaaa attttacatg ttigagatgg 5280  
 tggagtaaaa agactgttaa acatttcttt taaaaaatta tttttacatt acaacaatat 5340  
  
 atttatgatg tggtcagatc aaaaatttaa cttctgtgtc ccagatctac tttcaaagt 5400  
 agattttcac ttgtcagctt aaatttctga ctagaactaa catttgtgta tttttgtgct 5460  
 tagtcggaat acaaatttca cagtggattt ttgaagtttg tccttaaatt ggataaaatc 5520  
 aagtgattaa agttactaaa gagataaaaa tggtaatctt cattttttaa agtaatttgg 5580  
 ttgtgtttat agttatttgt acaagtattt atcacagact ctaaattgaa aaatgtagta 5640  
 tgatctatat ttgaccctaa aaatgttga ttaatttaac aaatatggca gatttttcat 5700  
 aactaagtct taagtcttct aaaaggaagc tgttaccctt ctgtttttta ttacattaat 5760  
  
 tgaaatgtgt tttagagat acaatttcag catattttat atattaaaa acaaaaaagg 5820  
 attagtattg agccagtggc caaaaggtaa tattactacc atgtagactg ttatagtcca 5880  
 aattgtccca cttcaccag aatttttagaa actagaagtc tgggaggtag tatatcagct 5940  
 gtagttgggt aattccaagt gctgatagta ctattcatct tttttattat tgtgtcagat 6000  
 gaaacaaatg ccaagttgca aaatatgcag atttttatta tataatggtt ttaggcataa 6060  
 attattaaca agccatgcct tatgtgtttc atcttatatt tttctttaga actaaactat 6120  
 aacagatttt ggaaatgat ttgacgtgct tgctcacttg attgacttgg tcagatattt 6180  
  
 gaatgatggt attacctaga ttctaatect tgattctagt tatataataa ataatataga 6240  
 atatgaaaat atgtttgggc atttactgtt tatattatgt agtagcctcc atcatgacac 6300  
 acttactaca tttatgaatt gagcagttct gtaattgtaa ttattattgc tgttcattga 6360  
 acaaaacatg cttataatag caaacaataa gaaatgcccc caaaatgcta tttttttaat 6420  
 tcagttataa ctgttactct tgtagttgtg tatgacgcaa taaaatttgt aaaaaattt 6480  
 cagcatgaaa aataaaattt gtatcactta tgta 6514  
  
 <210> 8  
 <211> 1313  
 <212> PRT  
 <213> Artificial Sequence

<220><223> Synthetic Polypeptide

<400> 8

Met Arg Ser Ala Ala Ala Ala Pro Arg Ser Pro Ala Val Ala Thr Glu

1 5 10 15

Ser Arg Arg Phe Ala Ala Ala Arg Trp Pro Gly Trp Arg Ser Leu Gln

20 25 30

Arg Pro Ala Arg Arg Ser Gly Arg Gly Gly Gly Gly Ala Ala Pro Gly

35 40 45

Pro Tyr Pro Ser Ala Ala Pro Pro Pro Pro Gly Pro Gly Pro Pro Pro

50 55 60

Ser Arg Gln Ser Ser Pro Pro Ser Ala Ser Asp Cys Phe Gly Ser Asn

65 70 75 80

Gly Asn Gly Gly Gly Ala Phe Arg Pro Gly Ser Arg Arg Leu Leu Gly

85 90 95

Leu Gly Gly Pro Pro Arg Pro Phe Val Val Leu Leu Leu Pro Leu Ala

100 105 110

Ser Pro Gly Ala Pro Pro Ala Ala Pro Thr Arg Ala Ser Pro Leu Gly

115 120 125

Ala Arg Ala Ser Pro Pro Arg Ser Gly Val Ser Leu Ala Arg Pro Ala

130 135 140

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

145 150 155 160

Met Ser Leu Lys Pro Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln

165 170 175

Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Pro Pro Pro Ala

180 185 190

Ala Ala Asn Val Arg Lys Pro Gly Gly Ser Gly Leu Leu Ala Ser Pro

195 200 205

Ala Ala Ala Pro Ser Pro Ser Ser Ser Ser Val Ser Ser Ser Ser Ala

210 215 220

Thr Ala Pro Ser Ser Val Val Ala Ala Thr Ser Gly Gly Gly Arg Pro

225 230 235 240

Gly Leu Gly Arg Gly Arg Asn Ser Asn Lys Gly Leu Pro Gln Ser Thr  
245 250 255

Ile Ser Phe Asp Gly Ile Tyr Ala Asn Met Arg Met Val His Ile Leu  
260 265 270

Thr Ser Val Val Gly Ser Lys Cys Glu Val Gln Val Lys Asn Gly Gly  
275 280 285

Ile Tyr Glu Gly Val Phe Lys Thr Tyr Ser Pro Lys Cys Asp Leu Val  
290 295 300

Leu Asp Ala Ala His Glu Lys Ser Thr Glu Ser Ser Ser Gly Pro Lys  
305 310 315 320

Arg Glu Glu Ile Met Glu Ser Ile Leu Phe Lys Cys Ser Asp Phe Val  
325 330 335

Val Val Gln Phe Lys Asp Met Asp Ser Ser Tyr Ala Lys Arg Asp Ala  
340 345 350

Phe Thr Asp Ser Ala Ile Ser Ala Lys Val Asn Gly Glu His Lys Glu  
355 360 365

Lys Asp Leu Glu Pro Trp Asp Ala Gly Glu Leu Thr Ala Asn Glu Glu  
370 375 380

Leu Glu Ala Leu Glu Asn Asp Val Ser Asn Gly Trp Asp Pro Asn Asp  
385 390 395 400

Met Phe Arg Tyr Asn Glu Glu Asn Tyr Gly Val Val Ser Thr Tyr Asp  
405 410 415

Ser Ser Leu Ser Ser Tyr Thr Val Pro Leu Glu Arg Asp Asn Ser Glu  
420 425 430

Glu Phe Leu Lys Arg Glu Ala Arg Ala Asn Gln Leu Ala Glu Glu Ile  
435 440 445

Glu Ser Ser Ala Gln Tyr Lys Ala Arg Val Ala Leu Glu Asn Asp Asp  
450 455 460

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

Arg Glu Gly His Ser Ile Asn Thr Arg Glu Asn Lys Tyr Ile Pro Pro

485                      490                      495  
 Gly Gln Arg Asn Arg Glu Val Ile Ser Trp Gly Ser Gly Arg Gln Asn  
 500                      505                      510  
  
 Ser Pro Arg Met Gly Gln Pro Gly Ser Gly Ser Met Pro Ser Arg Ser  
 515                      520                      525  
 Thr Ser His Thr Ser Asp Phe Asn Pro Asn Ser Gly Ser Asp Gln Arg  
 530                      535                      540  
 Val Val Asn Gly Gly Val Pro Trp Pro Ser Pro Cys Pro Ser Pro Ser  
 545                      550                      555                      560  
 Ser Arg Pro Pro Ser Arg Tyr Gln Ser Gly Pro Asn Ser Leu Pro Pro  
 565                      570                      575  
  
 Arg Ala Ala Thr Pro Thr Arg Pro Pro Ser Arg Pro Pro Ser Arg Pro  
 580                      585                      590  
 Ser Arg Pro Pro Ser His Pro Ser Ala His Gly Ser Pro Ala Pro Val  
 595                      600                      605  
 Ser Thr Met Pro Lys Arg Met Ser Ser Glu Gly Pro Pro Arg Met Ser  
 610                      615                      620  
 Pro Lys Ala Gln Arg His Pro Arg Asn His Arg Val Ser Ala Gly Arg  
 625                      630                      635                      640  
  
 Gly Ser Ile Ser Ser Gly Leu Glu Phe Val Ser His Asn Pro Pro Ser  
 645                      650                      655  
 Glu Ala Ala Thr Pro Pro Val Ala Arg Thr Ser Pro Ser Gly Gly Thr  
 660                      665                      670  
 Trp Ser Ser Val Val Ser Gly Val Pro Arg Leu Ser Pro Lys Thr His  
 675                      680                      685  
 Arg Pro Arg Ser Pro Arg Gln Asn Ser Ile Gly Asn Thr Pro Ser Gly  
 690                      695                      700  
  
 Pro Val Leu Ala Ser Pro Gln Ala Gly Ile Ile Pro Thr Glu Ala Val  
 705                      710                      715                      720  
 Ala Met Pro Ile Pro Ala Ala Ser Pro Thr Pro Ala Ser Pro Ala Ser  
 725                      730                      735

Asn Arg Ala Val Thr Pro Ser Ser Glu Ala Lys Asp Ser Arg Leu Gln  
740 745 750

Asp Gln Arg Gln Asn Ser Pro Ala Gly Asn Lys Glu Asn Ile Lys Pro  
755 760 765

Asn Glu Thr Ser Pro Ser Phe Ser Lys Ala Glu Asn Lys Gly Ile Ser  
770 775 780

Pro Val Val Ser Glu His Arg Lys Gln Ile Asp Asp Leu Lys Lys Phe  
785 790 795 800

Lys Asn Asp Phe Arg Leu Gln Pro Ser Ser Thr Ser Glu Ser Met Asp  
805 810 815

Gln Leu Leu Asn Lys Asn Arg Glu Gly Glu Lys Ser Arg Asp Leu Ile  
820 825 830

Lys Asp Lys Ile Glu Pro Ser Ala Lys Asp Ser Phe Ile Glu Asn Ser  
835 840 845

Ser Ser Asn Cys Thr Ser Gly Ser Ser Lys Pro Asn Ser Pro Ser Ile  
850 855 860

Ser Pro Ser Ile Leu Ser Asn Thr Glu His Lys Arg Gly Pro Glu Val  
865 870 875 880

Thr Ser Gln Gly Val Gln Thr Ser Ser Pro Ala Cys Lys Gln Glu Lys  
885 890 895

Asp Asp Lys Glu Glu Lys Lys Asp Ala Ala Glu Gln Val Arg Lys Ser  
900 905 910

Thr Leu Asn Pro Asn Ala Lys Glu Phe Asn Pro Arg Ser Phe Ser Gln  
915 920 925

Pro Lys Pro Ser Thr Thr Pro Thr Ser Pro Arg Pro Gln Ala Gln Pro  
930 935 940

Ser Pro Ser Met Val Gly His Gln Gln Pro Thr Pro Val Tyr Thr Gln  
945 950 955 960

Pro Val Cys Phe Ala Pro Asn Met Met Tyr Pro Val Pro Val Ser Pro  
965 970 975

Gly Val Gln Pro Leu Tyr Pro Ile Pro Met Thr Pro Met Pro Val Asn

|      |     |     |     |     |     |      |     |     |     |     |      |     |     |     |     |
|------|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|-----|
| 980  |     |     |     |     |     | 985  |     |     |     |     | 990  |     |     |     |     |
| Gln  | Ala | Lys | Thr | Tyr | Arg | Ala  | Val | Pro | Asn | Met | Pro  | Gln | Gln | Arg | Gln |
| 995  |     |     |     |     |     | 1000 |     |     |     |     | 1005 |     |     |     |     |
| Asp  | Gln | His | His | Gln | Ser | Ala  | Met | Met | His | Pro | Ala  | Ser | Ala | Ala |     |
| 1010 |     |     |     |     |     | 1015 |     |     |     |     | 1020 |     |     |     |     |
|      |     |     |     |     |     |      |     |     |     |     |      |     |     |     |     |
| Gly  | Pro | Pro | Ile | Ala | Ala | Thr  | Pro | Pro | Ala | Tyr | Ser  | Thr | Gln | Tyr |     |
| 1025 |     |     |     |     |     | 1030 |     |     |     |     | 1035 |     |     |     |     |
| Val  | Ala | Tyr | Ser | Pro | Gln | Gln  | Phe | Pro | Asn | Gln | Pro  | Leu | Val | Gln |     |
| 1040 |     |     |     |     |     | 1045 |     |     |     |     | 1050 |     |     |     |     |
| His  | Val | Pro | His | Tyr | Gln | Ser  | Gln | His | Pro | His | Val  | Tyr | Ser | Pro |     |
| 1055 |     |     |     |     |     | 1060 |     |     |     |     | 1065 |     |     |     |     |
| Val  | Ile | Gln | Gly | Asn | Ala | Arg  | Met | Met | Ala | Pro | Pro  | Thr | His | Ala |     |
| 1070 |     |     |     |     |     | 1075 |     |     |     |     | 1080 |     |     |     |     |
|      |     |     |     |     |     |      |     |     |     |     |      |     |     |     |     |
| Gln  | Pro | Gly | Leu | Val | Ser | Ser  | Ser | Ala | Thr | Gln | Tyr  | Gly | Ala | His |     |
| 1085 |     |     |     |     |     | 1090 |     |     |     |     | 1095 |     |     |     |     |
| Glu  | Gln | Thr | His | Ala | Met | Tyr  | Ala | Cys | Pro | Lys | Leu  | Pro | Tyr | Asn |     |
| 1100 |     |     |     |     |     | 1105 |     |     |     |     | 1110 |     |     |     |     |
| Lys  | Glu | Thr | Ser | Pro | Ser | Phe  | Tyr | Phe | Ala | Ile | Ser  | Thr | Gly | Ser |     |
| 1115 |     |     |     |     |     | 1120 |     |     |     |     | 1125 |     |     |     |     |
| Leu  | Ala | Gln | Gln | Tyr | Ala | His  | Pro | Asn | Ala | Thr | Leu  | His | Pro | His |     |
| 1130 |     |     |     |     |     | 1135 |     |     |     |     | 1140 |     |     |     |     |
|      |     |     |     |     |     |      |     |     |     |     |      |     |     |     |     |
| Thr  | Pro | His | Pro | Gln | Pro | Ser  | Ala | Thr | Pro | Thr | Gly  | Gln | Gln | Gln |     |
| 1145 |     |     |     |     |     | 1150 |     |     |     |     | 1155 |     |     |     |     |
| Ser  | Gln | His | Gly | Gly | Ser | His  | Pro | Ala | Pro | Ser | Pro  | Val | Gln | His |     |
| 1160 |     |     |     |     |     | 1165 |     |     |     |     | 1170 |     |     |     |     |
| His  | Gln | His | Gln | Ala | Ala | Gln  | Ala | Leu | His | Leu | Ala  | Ser | Pro | Gln |     |
| 1175 |     |     |     |     |     | 1180 |     |     |     |     | 1185 |     |     |     |     |
| Gln  | Gln | Ser | Ala | Ile | Tyr | His  | Ala | Gly | Leu | Ala | Pro  | Thr | Pro | Pro |     |
| 1190 |     |     |     |     |     | 1195 |     |     |     |     | 1200 |     |     |     |     |
|      |     |     |     |     |     |      |     |     |     |     |      |     |     |     |     |
| Ser  | Met | Thr | Pro | Ala | Ser | Asn  | Thr | Gln | Ser | Pro | Gln  | Asn | Ser | Phe |     |
| 1205 |     |     |     |     |     | 1210 |     |     |     |     | 1215 |     |     |     |     |

Pro Ala Ala Gln Gln Thr Val Phe Thr Ile His Pro Ser His Val  
 1220 1225 1230  
 Gln Pro Ala Tyr Thr Asn Pro Pro His Met Ala His Val Pro Gln  
 1235 1240 1245  
 Ala His Val Gln Ser Gly Met Val Pro Ser His Pro Thr Ala His  
 1250 1255 1260

Ala Pro Met Met Leu Met Thr Thr Gln Pro Pro Gly Gly Pro Gln  
 1265 1270 1275  
 Ala Ala Leu Ala Gln Ser Ala Leu Gln Pro Ile Pro Val Ser Thr  
 1280 1285 1290  
 Thr Ala His Phe Pro Tyr Met Thr His Pro Ser Val Gln Ala His  
 1295 1300 1305

His Gln Gln Gln Leu  
 1310

<210> 9

<211> 4712

<212> DNA

<213> Artificial Sequence

<220><

223> Synthetic Polynucleotide

<400> 9

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| acccccgaga aagcaacca gcgcgccgcc cgctcctcac gtgtccctcc cggccccggg   | 60  |
| gccacctcac gttctgcttc cgtctgacct ctccgacttc cggtaaagag tccctatccg  | 120 |
| cacctccgct cccaccgggc gcctcggcgc gcccgccctc cgatgcgctc agcggccgca  | 180 |
| gctcctcgga gtcccgcggt ggccaccgag tctcgccgct tcgccgcagc caggtggccc  | 240 |
| gggtggcgct cgctccagcg gccggcgcgcg cgagcggggc ggggcgggcg tggcgcggcc | 300 |
| ccgggaccgt atccctccgc cgcccctccc ccgccgggcc ccggcccccc tccctcccgg  | 360 |
| cagagctcgc ctccctccgc ctgagactgt ttggtagca acggcaacgg cggcggcgcg   | 420 |
| tttcggcccc gctcccgcg gctccttgggt ctccgggggc ctccccgcc ctctgctgct   | 480 |
| ctccttctcc cctcggccag cccgggcgcc cctccggccg cgccaaccg cgcctccccg   | 540 |
| ctcggcgccc gcgcgtcccc gccgcgttcc ggcgtctcct tggcgcgccc ggcctccggc  | 600 |
| tgtccccgcc cggcgtgcga gccggtgtat gggccccca ccatgtcgt gaagccccag    | 660 |
| cagcagcagc agcagcagca gcagcagcag cagcagcaac agcagcagca gcagcagcag  | 720 |



|                                                                   |      |
|-------------------------------------------------------------------|------|
| cagcagccgc cgcccgccgc tgccaatgtc cgcaagcccg gcggcagcgg ccttctagcg | 780  |
| tcgcccgcgc ccgcgccttc gccgtcctcg tcctcggtct cctcgctctc ggccacggct | 840  |
| ccctcctcgg tggtcgcggc gacctccggc ggccggaggc ccggcctggg cagaggtcga | 900  |
| aacagtaaca aaggactgcc tcagtctacg atttcttttg atggaatcta tgcaaatatg | 960  |
| aggatggttc atatacttac atcagttgtt ggctccaaat gtgaagtaca agtgaanaat | 1020 |
| ggaggtatat atgaaggagt ttttaaaact tacagtccga agtgtgattt ggtacttgat | 1080 |
| gccgcacatg agaaaaglac agaatccagt tcggggccga aacgtgaaga aataatggag | 1140 |
| agtattttgt tcaaatgttc agactttgtt gtggtacagt ttaaagatat ggactccagt | 1200 |
| tatgcaaaaa gagatgcttt tactgactct gctatcagtg ctaaagtga tggcgaacac  | 1260 |
| aaagagaagg acctggagcc ctgggatgca ggtgaactca cagccaatga ggaacttgag | 1320 |
| gctttggaaa atgacgtatc taatggatgg gatcccaatg atatgtttcg atataatgaa | 1380 |
| gaaaattatg gtgtagtgtc tacgtatgat agcagtttat cttcgtatac agtgcctta  | 1440 |
| gaaagagata actcagaaga atttttaaaa cgggaagcaa gggcaaacca gtagcagaa  | 1500 |
| gaaattgagt caagtgccta gtacaaagct cgagtggccc tggaaaatga tgaaggagt  | 1560 |
| gaggaagaaa aatacacagc agttcagaga aattccagtg aacgtgaggg gcacagcata | 1620 |
| aacactaggg aaaataaata tattcctcct ggacaaagaa atagagaagt catatcctgg | 1680 |
| ggaagtggga gacagaattc accgcgtatg ggccagcctg gatcgggctc catgccatca | 1740 |
| agatccactt ctcacattc agatttcaac ccgaattctg gttcagacca aagagtagtt  | 1800 |
| aatggagggt ttccttgccc atcgcttgc ccatctcctt cctctcgccc accttctgc   | 1860 |
| taccagttag gtcccaactc tcttcacct cgggcagcca cccctacag gccgcctcc    | 1920 |
| aggccccct cgcggccatc cagacccccg tctacccct ctgctcatgg ttctccagt    | 1980 |
| cctgtctcta ctatgcetaa acgcatgtct tcagaagggc ctccaaggat gtcccaaaag | 2040 |
| gcccagcgac atcctcgaaa tcacagagtt tctgctggga ggggttccat atccagtggc | 2100 |
| ctagaatttg tatcccacaa cccaccagt gaagcagcta ctctccagt agcaaggacc   | 2160 |
| agtcctcgg ggggaacgtg gtcacagtg gtcagtgggg ttccaagatt atcccctaaa   | 2220 |
| actcatagac ccaggtctcc cagacagaac agtattggaa atacccccag tgggccagtt | 2280 |
| cttgcttctc cccaagctgg tattattcca actgaagctg ttgccatgcc tattccagt  | 2340 |
| gcatctccta cgcctgctag tcctgcatcg aacagagctg ttacccttc tagtgaggct  | 2400 |
| aaagattcca ggcttcaaga tcagaggcag aactctcctg cagggaataa agaaaatatt | 2460 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| aaaccaatg aaacatcacc tagcttctca aaagctgaaa acaaaggtat atcaccagtt   | 2520 |
|                                                                    |      |
| gtttctgaac atagaaaaca gattgatgat ttaaagaaat ttaagaatga ttttaggtta  | 2580 |
| cagccaagtt ctacttctga atctatggat caactactaa acaaaaatag agaggagaa   | 2640 |
| aatcaagag atttgatcaa agacaaaatt gaaccaagtg ctaaggattc tttcattgaa   | 2700 |
| aatagcagca gcaactgtac cagtggcagc agcaagccga atagccccag catttccct   | 2760 |
| tcaatactta gtaacacgga gcacaagagg ggacctgagg tcacttccca aggggttcag  | 2820 |
| acttccagcc cagcatgtaa acaagagaaa gacgataagg aagagaagaa agacgcagct  | 2880 |
| gagcaagtta ggaaatcaac attgaatccc aatgcaaagg agttcaaccc acgttccttc  | 2940 |
|                                                                    |      |
| tctcagccaa agcctttctac taccccaact tcacctgggc ctcaagcaca acctagccca | 3000 |
| tctatggtgg gtcatcaaca gccaaactca gtttatactc agcctgtttg ttttgacca   | 3060 |
| aatatgatgt atccagtccc agtgagccca ggcggtgcaac ctttataccc aatacctatg | 3120 |
| acgcccattgc cagtgaatca agccaagaca tatagagcag taccaaatat gcccacacag | 3180 |
| cggcaagacc agcatcatca gagtgccatg atgcaccag cgtcagcagc gggcccaccg   | 3240 |
| attgcagcca ccccaccagc ttactccagc caatatgttg cctacagtcc tcagcagttc  | 3300 |
| ccaaatcagc cccttgttca gcatgtgcc cattatcagt ctacagcatcc tcatgtctat  | 3360 |
|                                                                    |      |
| agtctgtaa tacagggtaa tgctagaatg atggcaccac caacacacgc ccagcctggt   | 3420 |
| ttagtatctt cttcagcaac tcagtacggg gctcatgagc agacgcatgc gatgtatgca  | 3480 |
| tgtcccaaat taccatacaa caaggagaca agcccttctt tctactttgc catttccagc  | 3540 |
| ggctcccttg ctacagcagta tgcgcacct aacgtaccc tgcaccaca tactccacac    | 3600 |
| cctcagcctt cagctacccc cactggacag cagcaaagcc aacatggtgg aagtcatcct  | 3660 |
| gcaccagtc ctgttcagca ccatcagcac caggccgccc aggccttcca tctggccagt   | 3720 |
| ccacagcagc agtcagccat ttaccacgcg gggtttgcgc caactccacc ctccatgaca  | 3780 |
|                                                                    |      |
| cctgcctcca acacgcagtc gccacagaat agtttcccag cagcacaaca gactgtcttt  | 3840 |
| acgatccatc cttctcagct tcagccggcg tataccaacc caccacacat ggcccacgta  | 3900 |
| cctcaggctc atgtacagtc aggaatggtt ctttctcatc caactgccca tgcgccaatg  | 3960 |
| atgctaata gacacagcc acccggcggt cccagggccg ccttcgtca aagtgcacta     | 4020 |
| cagccattc cagtctcagc aacagcgcat tccccata tgacgcaccc ttcagtacaa     | 4080 |
| gcccaccacc aacagcagtt gtaaggctgc cctggaggaa ccgaaaggcc aaattccctc  | 4140 |
| ctcccttcta ctgcttctac caactggaag cacagaaaac tagaatttca tttattttgt  | 4200 |

ttttaaaata tatatgttga tttcttgtaa catccaatag gaatgctaac agttcacttg 4260  
cagtggaaga tacttggacc gagtagaggc atttaggaac ttgggggcta ttccataatt 4320  
ccatatgctg tttcagagtc ccgcaggtac cccagctctg cttgccgaaa ctggaagtta 4380  
tttatTTTT aataaccctt gaaagtcatg aacacatcag ctagcaaaaag aagtaacaag 4440  
agtgattctt gctgctatta ctgctaaaaa aaaaaaaaaa aaaaaatcaa gacttggAAC 4500  
gcccttttac taaacttgac aaagtttcag taaattctta ccgtcaaact gacggattat 4560  
tatttataaa tcaagtttga tgaggTgatc actgtctaca gtggttcaac ttttaagtta 4620

agggaaaaac ttttactttg tagataatat aaaataaaaa cttaaaaaaa atttaaaaaa 4680  
taaaaaaaagt tttaaaaact gaaaaaaaaa aa 4712

<210> 10

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 10

ttggcctttc ggttctctcc 19

<210> 11

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 11

ggaggaaccg aaaggccaa 19

<210> 12

<211> 148

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 12

tggaggcttg ctgaaggctg tatgtctgtt tcttggcctt tcggttctc cagtgaagcc 60

acagatggga ggaaccgaaa ggccaaagga cacaaggcct gttactagca ctcacatgga 120

acaaatggcc accgtgggag gatgacaa 148

<210> 13

<211> 148  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic Polynucleotide  
 <400> 13  
 ttgtcaccct cccacgggtg ccatttggtc catgtgagtg ctagtaacag gccttggtgc 60  
  
 ctttggcctt tcggttcctc ccactgtgtg cttcactgga ggaaccgaaa ggccaagaca 120  
 acagcataca gccttcagca agcctcca 148  
 <210> 14  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic Polynucleotide  
 <400> 14  
 atagcagcaa gaatcactct tg 22  
 <210> 15  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic Polynucleotide  
 <400> 15  
 caagagtgat tcttgctgct at 22  
  
 <210> 16  
 <211> 151  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic Polynucleotide  
 <400> 16  
 ctggaggctt gccttgggct gtatgctgat agcagcaaga atcactcttg ttttggcctc 60  
 tgactgaaca agaggatctt gctgctatca ggacacaagg ccctttatca gcactcacat 120  
 ggaacaaatg gccaccgtgg gaggatgaca a 151  
 <210> 17  
 <211> 151

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 17

ttgtcaccct cccacgggtg ccatttggtc catgtgagtg ctgataaagg gccttggtgc 60

ctgatagcag caagatcctc ttgttcagtc agaggccaaa acaagagtga ttcttgctgc 120

tatcagcata cagcccaaag caagcctcca g 151

<210> 18

<211> 300

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 18

tggaggttg ctgaaggctg tatgtgttg tctactgtc tataatgtctt ggcagtgaag 60

ccacagatgg ccaagacata tagagcagta aggacacaag gcctgttact agcactcaca 120

tggaacaaat ggccaccgtg ggagatgac aatggaggct tgctgaaggc tgtatgtgt 180

tgtcttggcc ttctgggtcc tccagtgaag ccacagatgg gaggaaccga aaggccaaag 240

gacacaaggc ctgttactag cactcacatg gaacaaatgg ccaccgtggg aggatgacaa 300

<210> 19

<211> 300

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 19

ttgtcaccct cccacgggtg ccatttggtc catgtgagtg ctagtaacag gccttggtgc 60

ctttggcctt tcggttcctc ccatctgtgg cttcactgga ggaaccgaaa ggccaagaca 120

acagcatata gccttcagca agcctccatt gtcactctcc cacgtggcc atttgttcca 180

tgtgagtgt agtaacaggc cttgtgtcct tactgtctta tatgtcttgg ccatctgtgg 240

cttactgtcc aagacatata gagcagtaga caacagcata cagccttcag caagcctcca 300

<210> 20

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 20

ttccagttgg tagaagcagt aga 23

<210> 21

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 21

tctactgctt ctaccaactg gaa 23

<210> 22

<211> 152

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 22

ctggaggctt gctttgggct gtatgctgtt ccagttggta gaagcagtag attttgcct 60

ctgactgac tactgttcac caactggaac aggacacaag gccctttatc agcactcaca 120

tggaacaaat ggccaccgtg ggaggatgac aa 152

<210> 23

<211> 152

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 23

ttgtcatcct cccacggtag ccatttggtc catgtgagtg ctgataaagg gccttgtgtc 60

ctgttccagt tggatgaacag tagatcagtc agaggccaaa atctactgct tctaccaact 120

ggaacagcat acagcccaaa gcaagcctcc ag 152

<210> 24

<211> 151

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 24

tggaggcttg ctgaaggctg tatgctgttg tcggggccgg ttctggggcc ttagtgaagc 60  
cacagatgta cggccccgaa accggccccca ggacacaagg cctgttacta gcactcacat 120  
ggaacaaatg gccaccgtgg gaggatgaca a 151

<210> 25

<211> 151

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 25

ttgtcatcct cccacggtgg ccatttgttc catgtgagtg ctagtaacag gccttgtgtc 60  
  
ctggggccgg ttctggggcc gtacatctgt ggcttcacta agggccccgaa accggccccg 120  
acaacagcat acagccttca gcaagcctcc a 151

<210> 26

<211> 4145

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 26

ttggccactc cctctctcgc cgctcgctcg ctactgagg cggggcgacc aaaggtcgcc 60  
cgacggccgg gctttgcccg ggcgccctca gtgagcgagc gagcgcgag agagggagtg 120  
gccaaactca tactagggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc 180  
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240  
  
tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300  
cctagttatt aatagtaatc aattacgggg tcattagtgc atagcccata tatggagtgc 360  
cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga cccccgcca 420  
ttgacgtcaa taatgacgta tggtcccata gtaacgcaa tagggacttt ccattgacgt 480  
caatgggtgg actatttacg gtaaaactgcc cacttggcag tacatcaagt gtatcatatg 540  
ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgccag 600  
tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgtatt 660  
  
accatggtcg aggtgagccc cagttctgc ttactctcc ccattcccc cccctccca 720

cccccaattt tgiatttatt ttttttttaa ttattttgtg cagcgatggg ggcggggggg 780  
 gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg 840  
 agaggtgcgg cgccagccaa tcagagcggc gcgctccgaa agtttcttt tatggcgagg 900  
 cggcggcggc ggcgggccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg 960  
 ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cggccgcccc ggctctgact 1020  
 gaccgcgtta ctccacagg tgagcgggcg ggacggccct tctcctcagc gctgtaatta 1080  
  
 gcgcttggtt taatgacggc ttgttggagg ctgctgaag gctglatgct gttgtcgggg 1140  
 ccggtttcgg ggccttagtg aagccacaga tgtacggccc cgaaaccggc cccaggacac 1200  
 aaggcctgtt actagcactc acatggaaca aatggccacc gtgggaggat gacaatggag 1260  
 gcttgctgaa ggctgtatgc tgttgtcttg gcctttcggg tcctccagtg aagccacaga 1320  
 tgggaggaac cgaaaggcca aaggacacaa ggcctgttac tagcactcac atggaacaaa 1380  
 tggccaccgt gggaggatga caatttctgt ggctgcgtga aagccttgag gggtccggg 1440  
 agctagagcc tctgctaacc atgttcatgc cttcttcttt ttcctacagc tcctgggcaa 1500  
  
 cgtgctggtt attgtgctgt ctcatcattt tggcaaagaa ttctcgaag atccgaaggg 1560  
 aaagtcttc acgacttgg gatccgttcg aagatatcac cggttgagcc accatgagca 1620  
 ccctgtgccc cccccccagc cccgccgtgg ccaagaccga gatcgccctg agcggcaaga 1680  
 gccccctgct ggccgccacc ttgcctact gggacaacat cctgggcccc cgcgtgcgcc 1740  
 acatctgggc ccccaagacc gagcaggtgc tgctgagcga cggcgagatc accttcttg 1800  
 ccaaccacac cctgaacggc gagatcctgc gcaacgccga gagcggcgcc atcgacgtga 1860  
 agttcttcgt gctgagcgag aaggcggtga tcacgtgag cctgatcttc gacggcaact 1920  
  
 ggaacggcga ccgcagcacc tacggcctga gcatcatcct gcccagacc gagctgagct 1980  
 tctacctgcc cctgcaccgc gtgtcgtgg accgcctgac ccacatcatc cgcaagggcc 2040  
 gcatctggat gcacaaggag cgccaggaga acgtgcagaa gatcatcctg gagggcaccg 2100  
 agcgcattga ggaccagggc cagagcatca tccccatgct gaccggcgag gtgatccccg 2160  
 tgatggagct gctgagcagc atgaagagcc acagcgtgcc cgaggagatc gacatcgccg 2220  
 acaccgtgct gaacgacgac gacatcggcg acagctgcca cgagggttc ctgctgaacg 2280  
 ccatcagcag ccacctgcag acctcgggct gcagcgtggt ggtgggcagc agcggcgaga 2340  
  
 aggtgaacaa gatcgtgcgc accctgtgcc tgttctgac ccccgccgag cgcaagtgca 2400  
 gccgcctgtg cgaggccgag agcagcttca agtacgagag cggcctgttc gtgcagggcc 2460  
 tgctgaagga cagcaccggc agcttcgtgc tgcccttcgg ccaggtgatg tacgccccct 2520  
 accccaccac ccacatcgac gtggacgtga acaccgtgaa gcagatgccc cctgccacg 2580



agcacatcta caaccagcgc cgctacatgc gcagcgagct gaccgccttc tggcgcgcca 2640  
 ccagcgagga ggacatggcc caggacacca tcatctacac cgacgagagc ttcacccccg 2700  
 acctgaacat cttccaggac gtgctgcacc gcgacaccct ggtgaaggcc ttcctggacc 2760  
  
 aggtgttcca gctgaagccc ggcctgagcc tgcgcagcac cttcctggcc cagtctctgc 2820  
 tgggtgtgca ccgaaggcc ctgaccctga tcaagtacat cgaggacgac acccagaagg 2880  
 gcaagaagcc cttcaagagc ctgcgcaacc tgaagatcga cctggacctg accgccgagg 2940  
 gcgacctgaa catcatcatg gccctggccg agaagatcaa gcccggcctg cacagcttca 3000  
 tcttcggccg ccccttctac accagcgtgc aggagcgcca cgtgctgatg accttctaac 3060  
 aattgttaat taagtttaaa ccctcgaggc cgcaagctta tcgataatca acctctggat 3120  
 tacaaaattt gtgaaagatt gactggtatt cttaactatg ttgctccttt tacgctatgt 3180  
  
 ggatacgtg ctttaatgcc tttgtatcat gctattgctt cccgtatggc tttcattttc 3240  
 tctccttgt ataaatcctg gttgtgtgtc ctttatgagg agttgtggcc cgttgtcagg 3300  
 caacgtggcg tgggtgtcac tgtgtttgt gacgcaaccc ccactggttg gggcattgcc 3360  
 accacctgtc agctccttcc cgggactttc gctttccccc tccctattgc cacggcgga 3420  
 ctcatcgccg cctgccttgc ccgtgctgga acaggggctc ggctgttggg cactgacaat 3480  
 tccgtggtgt tgcggggaa atcatcgctc tttccttggc tgcctgctg tgttgccacc 3540  
 tggattctgc gggggacgtc cttctgtac gtcccttcgg cctcaatcc agcggacctt 3600  
  
 cttcccgcg gccgtgtgcc ggctctgcgg cctcttcgc gtcttcgcct tcgccctcag 3660  
 acgagtcgga tctcccttgg ggccgcctcc ccgcatcgat accgtcgact agagctcgct 3720  
 gatcagcctc gactgtgcct tctagtggc agccatctgt tgtttgcccc tccccgtgc 3780  
 cttccttgac cctggaaggt gccactcca ctgtccttcc ctaataaaat gaggaattg 3840  
 catcgattg tctgagtagg tgtcattcta tttctggggg tggggtgggg caggacagca 3900  
 agggggagga ttgggaagac aatagcaggc atgctgggga gagatccacg ataacaaca 3960  
 gcttttttgg ggtgaacata ttgactgaat tcctgcagg ttggccactc cctctctgcg 4020  
  
 cgctcgtcg ctactgagg ccgcccgggc aaagcccggg cgtcgggcga ctttggctcg 4080  
 cccggcctca gtgagcgagc gagcgcgcag agagggagtg gccaactcca tcactagggg 4140  
 ttct 4145  
  
 <210> 27  
 <211> 899  
 <212> DNA  
 <213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 27

|                                                                     |     |
|---------------------------------------------------------------------|-----|
| ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc    | 60  |
| cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg    | 120 |
| gccaaactcca tctactagggg ttcttgctag ctctgggtat ttaagcccga gtgagcacgc | 180 |
|                                                                     |     |
| agggtctcca ttttgaagcg ggaggttacg cgttcgctga ctactagtgg gtaccagagc   | 240 |
| aaaaaaattg tcatcctccc acggtggcca ttgtttccat gtgagtgcta gtaacaggcc   | 300 |
| ttgtgtcctt tggcctttcg gttcctccca tctgtggctt cactggagga accgaaaggc   | 360 |
| caagacaaca gcatacagcc ttacagcaagc ctccagtggc ctacatacaga acttataaga | 420 |
| ttcccaaattc caaagacatt tcacgtttat ggtgatttcc cagaacacat agcgacatgc  | 480 |
| aaatatgtca gggcgccact cccctgtccc tcacagccat cttcctgccca gggcgcacgc  | 540 |
| gcgctgggtg ttcccgccca gtgacactgg gcccgcgatt ccttggagcg ggttgatgac   | 600 |
|                                                                     |     |
| gtcagcgttt cccatgggtga agcttggatc tgatccctag gttctagaac cggtgaccaa  | 660 |
| ttgttaatta agtttaaacc ctgagggccg caagcagatc cagcataaca aacagctttt   | 720 |
| ttgggggtgaa catattgact gaattccctg caggttggcc actccctctc tgcgcgctcg  | 780 |
| ctcgctcact gaggcggccc gggcaaagcc cgggcgtcgg gcgacctttg gtcgcccggc   | 840 |
| ctcagtgagc gagcgagcgc gcagagaggg agtggccaac tccatcacta ggggttcct    | 899 |

<210> 28

<211> 2547

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 28

|                                                                     |     |
|---------------------------------------------------------------------|-----|
| ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgcccgggc aaagcccggg    | 60  |
|                                                                     |     |
| cgtcggggga cctttggctg cccggcctca gtgagcgagc gagcgcgag agagggagtg    | 120 |
| gccaaactcca tctactagggg ttcttgctag ctctgggtat ttaagcccga gtgagcacgc | 180 |
| agggtctcca ttttgaagcg ggaggttacg cgttcgctga ctactagtgg gtaccagagc   | 240 |
| tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac   | 300 |
| ctagttatta atagtaatca attacggggt cattagtcca tagcccatat atggagtacc   | 360 |
| gcgttacata acttacggtta aatggccgcg ctggctgacc gcccaacgac ccccgcccat  | 420 |
| tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc   | 480 |

aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc 540  
caagtacgcc ccctattgac gtcaatgacg gtaaattggcc cgcctggcat tatgcccagt 600  
acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta 660  
ccatggtcga ggtgagcccc acgttctgct tcactctccc catctcccc ccctccccac 720  
ccccaatttt gtatttattt attttttaat tattttgtgc agcgatgggg gcgggggggg 780  
ggggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga 840  
gaggtgcggc ggcagccaat cagagcggcg cgtccgaaa gtttcctttt atggcgaggc 900  
  
ggcggcgggc gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgacgc 960  
tgctttcgcc ccgtgccccg ctccgccgc gcctcgcgc gcccgccccg gctctgactg 1020  
accgcgttac tcccacaggt gagcggggcg gacggccctt ctctcagcg ctgtaattag 1080  
cgcttggttt aatgacggct tgttggaggc ttgctgaagg ctgtatgctg ttgtcttggc 1140  
ctttcggttc ctccagtga gccacagatg ggaggaaccg aaaggccaaa ggacacaagg 1200  
cctgttacta gcactcacat ggaacaaatg gccaccgtgg gaggatgaca atttctgtgg 1260  
ctgcgtgaaa gccttgaggg gctccgggag ctagagcctc tgctaaccat gttcatgcct 1320  
  
tcttcttttt cctacagctc ctgggcaacg tgctggttat tgtgctgtct catcattttg 1380  
gcaaagaatt cctcgaagat ccgaaggga agtcttccac gactgtggga tccgttcgaa 1440  
gatatcaccg gttgagccac ccaattgtta attaagtta aaccctcgag gccgcaagct 1500  
tatcgataat caacctctgg attacaaaat ttgtgaaaga ttgactggta ttcttaacta 1560  
tgttgctcct ttacgctat gtggatacgc tgctttaatg cctttgtatc atgctattgc 1620  
ttcccgatg gccttcattt tctctcctt gtataaatcc tgggtgctgt ctctttatga 1680  
ggagttgtgg ccggttgta ggcaacgtgg cgtggtgtgc actgtgttg ctgacgcaac 1740  
  
ccccactggt tggggcattg ccaccacctg tcagctcctt tccgggactt tcgctttccc 1800  
cctccctatt gccacggcgg aactcatgc cgcctgcctt gcccgctgct ggacaggggc 1860  
tcggctgttg ggcactgaca attccgtggt gtgtcgggg aatcatcgt cctttccttg 1920  
gtgctcgcc tgtgttgcca cctggattct gcgcgggacg tccttctgct acgtcccttc 1980  
ggccctcaat ccagcggacc ttcttcccg cggcctgctg ccggtctgc gccctcttc 2040  
gcgtcttcgc cttgccttc agacgagtcg gatctccctt tgggccgcct ccccgcatcg 2100  
ataccgtcga ctagagctcg ctgacagcc tcgactgtgc cttctagttg ccagccatct 2160  
  
gtgttttgc cctccccgt gccttcttg acctggaag gtgccactcc cactgtcctt 2220  
tcctaataaa atgaggaat tgcatcgcat tgtctagta ggtgtcattc tattctgggg 2280  
ggtgggttg ggcaggacag caagggggag gattgggaag acaatagcag gcatgctggg 2340

gagagatcca cgataacaaa cagctttttt ggggtgaaca tattgactga attccctgca 2400  
 ggttggccac tccctctctg cgcgctcgct cgctcactga ggccgcccgg gcaaagcccg 2460  
 ggcgtcgggc gaccttttgt cgcgggcct cagtgagcga gcgagcgcgc agagagggag 2520  
 tggccaactc catcactagg ggttcct 2547

<210> 29

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 29

tattagatct gatggccgc 19

<210> 30

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 30

ctccatcact aggggttcct 20

<210> 31

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 31

agctctgggt atttaagccc gagtgagcac gcagggtctc cattttgaag cgggaggtta 60

<210> 32

<211> 145

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 32

aggaaccctt agtgatggag ttggccactc cctctctgcg cgctcgctcg ctactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc 120

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| gagcgcgcag agagggagtg gccaa                                        | 145 |
| <210> 33                                                           |     |
| <211> 24                                                           |     |
| <212> DNA                                                          |     |
| <213> Artificial Sequence                                          |     |
| <220><223> Synthetic Polynucleotide                                |     |
| <400> 33                                                           |     |
| gtgtactagg atcttttact tgaa                                         | 24  |
| <210> 34                                                           |     |
| <211> 24                                                           |     |
| <212> DNA                                                          |     |
| <213> Artificial Sequence                                          |     |
| <220><223> Synthetic Polynucleotide                                |     |
| <400> 34                                                           |     |
| ttcaagtaaa agatcctagt acac                                         | 24  |
| <210> 35                                                           |     |
| <211> 267                                                          |     |
| <212> DNA                                                          |     |
| <213> Artificial Sequence                                          |     |
| <220><223> Synthetic Polynucleotide                                |     |
| <400> 35                                                           |     |
| gtgatatcac aaggtcccag ggctggggtc agaaattctc tcccaggga atgaagccac   | 60  |
| aggagccaag agcaggagga ccaaggccct ggcgaaggcc gtggcctcgt tcaagtaaaa  | 120 |
| gacccagta cagtgcaggt cccaatgtgt actaggatct tttacttgaa cggggacgcc   | 180 |
| ggcatccggg ctcaggaccc ccctctctgc cagaggcacc aacaccagag ttcacaaatc  | 240 |
| agtctcctgc cctttgcatg tagcaaa                                      | 267 |
| <210> 36                                                           |     |
| <211> 267                                                          |     |
| <212> DNA                                                          |     |
| <213> Artificial Sequence                                          |     |
| <220><223> Synthetic Polynucleotide                                |     |
| <400> 36                                                           |     |
| tttgctacat gcaaagggca ggagactgat ttgtgaactc tgggtgttggc gcctctggca | 60  |

gagagggggg tcttgagccc ggatgccggc gtccccgttc aagtaaaaga tcctagtaca 120  
cattgggacc tgcactgtac taggatcttt tacttgaacg aggccacggc cttcgccagg 180  
gccttggtec tctgtctctt ggctcctgtg gcttcattcc ctcgggagag aatttctgac 240

cccagccctg ggacctgtg atatcac 267

<210> 37  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Synthetic Polynucleotide  
<400> 37  
atgcatccat attcttcct 19  
<210> 38  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Synthetic Polynucleotide  
<400> 38  
aggaagaata tggatgcat 19  
<210> 39  
<211> 148  
<212> DNA  
<213> Artificial Sequence

<220><223> Synthetic Polynucleotide  
<400> 39  
tggaggcttg ctgaagcgtg tatgctgttg tcatgcatcc atattcttcc tagtgaagcc 60  
acagatgagg aagaatatgg atgcatagga cacaaggcct gttactagca ctacatgga 120  
acaaatggcc accgtgggag gatgacaa 148  
<210> 40  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Synthetic Polynucleotide  
<400> 40

ttatttactt tcctgcac 19

<210> 41

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 41

gtgcagagaa agtaaataa 19

<210> 42

<211> 148

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 42

tggaggcttg ctgaaggctg tatgctgttg tcttatttac tttctctgca cagtgaagcc 60

acagatggtg cagagaaagt aaataaagga cacaaggcct gttactagca ctcacatgga 120

acaaatggcc accgtgggag gatgacaa 148

<210> 43

<211> 151

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 43

tggaggcttg ctgaaggctg tatgctgttg tcggggccgg tttcggggcc ttagtgaagc 60

cacagatgta cgccccgaa accggcccca ggacacaagg cctgttacta gcactcacat 120

ggaacaaatg gccaccgtgg gaggatgaca a 151

<210> 44

<211> 151

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 44

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| tggaggcttg ctgaaggctg tatgctgttg tcggggccgg tttcggggcc ttagtgaagc  | 60  |
| cacagatgta cgccccgaa accggcccca ggacacaagg cctgttacta gcactcacat   | 120 |
| ggaacaaatg gccaccgtgg gaggatgaca a                                 | 151 |
| <210> 45                                                           |     |
| <211> 296                                                          |     |
| <212> DNA                                                          |     |
| <213> Artificial Sequence                                          |     |
| <220><223> Synthetic Polynucleotide                                |     |
| <400> 45                                                           |     |
| ttgtcatcct cccacggtgg ccatttggtc catgtgagtg ctgataaagg gccttgtgtc  | 60  |
| ctgtctgtgtt aatctttatc aggttcagtc agaggccaaa aacctgataa agattaacca | 120 |
| gacagcatac agcccaaagc aagcctccag ttgtcatcct cccacggtgg ccatttggtc  | 180 |
| catgtgagtg ctgataaagg gccttgtgtc ctgcggttgc ggtgcctgcg cctcagtcag  | 240 |
| aggccaaaag ggcaggcac cgcaaccgca gcatacagcc caaagcaagc ctccag       | 296 |
| <210> 46                                                           |     |
| <211> 296                                                          |     |
| <212> DNA                                                          |     |
| <213> Artificial Sequence                                          |     |
| <220><223> Synthetic Polynucleotide                                |     |
| <400> 46                                                           |     |
| ctggaggctt gctttgggct gtatctgctg gttgcggtgc ctgcgccttt tggcctctga  | 60  |
| ctgaggcgca ggcaccgcaa ccgcaggaca caaggccctt tatcagcact cacatggaac  | 120 |
| aatggccac cgtgggagga tgacaactgg aggccttgctt tgggctgtat gctgtctggt  | 180 |
| taatctttat caggtttttg gcctctgact gaacctgata aagattaacc agacaggaca  | 240 |
| caaggccctt tatcagcact cacatggaac aatggccac cgtgggagga tgacaa       | 296 |
| <210> 47                                                           |     |
| <211> 297                                                          |     |
| <212> DNA                                                          |     |
| <213> Artificial Sequence                                          |     |
| <220><223> Synthetic Polynucleotide                                |     |
| <400> 47                                                           |     |
| ttgtcatcct cccacggtgg ccatttggtc catgtgagtg ctgataaagg gccttgtgtc  | 60  |



ctgggttggt tccctccttg ttttcagtca gaggccaaaa taaacaagga gggaaacaac 120  
ccagcatata gcccaaagca agcctccagt tgcatacttc ccacggtggc catttggtcc 180  
atgtgagtgc tgataaaggc ctttgtgtcc tggctgcggt tgcggtgcct gctcagtcag 240  
aggccaaaag caggcacgcg aaccgcagcc agcatacagc ccaaagcaag cctccag 297  
<210> 48

<211> 297

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 48

ctggaggctt gctttgggct gtatgctggc tgcggttgcg gtgcctgctt ttggcctctg 60  
actgagcagg caccgcaacc gcagccagga cacaaggccc tttatcagca ctcacatgga 120  
acaaatggcc accgtgggag gatgacaact ggaggcttgc tttgggctgt atgctgggtt 180  
gtttccctcc ttgtttatit tggcctctga ctgaaaacaa ggagggaac aaccaggac 240  
acaaggccct ttatcagcac tcacatggaa caaatggcca ccgtgggagg atgacaa 297

<210> 49

<211> 294

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 49

ttgtcactct cccacggtgg ccatttggtc catgtgagtg ctgataaagg gccttggtgc 60  
ctgtttcttc tggtaatct ttatcagtca gaggccaaaa taaagattaa ccagaagaaa 120  
cagcatagag cccaaagcaa gcctccagtt gtcactctcc cacggtggcc atttggtcca 180  
tgtgagtgtc gataaagggc ctttgtgtct gcggtgcctg cgcccgcggc tcagtcagag 240  
gccaaaagcc gcgggcgcag gcaccgcagc atacagccca aagcaagcct ccag 294

<210> 50

<211> 294

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 50

ctggaggctt gctttgggct gtatgctgcg gtgcctgcgc ccgcggcttt tggcctctga 60  
ctgagccgcg ggcgaggca cgcaggaca caaggccctt tatcagcact cacatggaac 120  
aatggccac cgtgggagga tgacaactgg aggccttgctt tgggctgtat gctgtttctt 180  
ctggttaatc tttatatttg cctctgactg ataaagatta accagaagaa acaggacaca 240  
aggcccttta tcagcactca catggaacaa atggccaccg tgggaggatg acaa 294

<210> 51

<211> 1446

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 51

atgagcacc tgtgcccc cccagcccc gccgtggcca agaccgagat cggcctgagc 60  
ggcaagagcc cctgctggc cgccacctt gcctactggg acaacatcct gggccccgc 120  
gtgcgccaca tctgggcccc caagaccgag caggtgctgc tgagcgacgg cgagatcacc 180  
ttcctggcca accacacct gaacggcgag atcctgcgca acgccgagag cggcgccatc 240  
gacgtgaagt tcttcgtgt gagcgagaag ggctgatca tcgtgagcct gatcttcgac 300  
ggcaactgga acggcgaccg cagcacctac ggctgagca tcctcctgcc ccagaccgag 360

ctgagcttct acctgccct gcaccgcgtg tgcgtggacc gcctgacca catcatccgc 420  
aagggccgca tctggatgca caaggagcgc caggagaacg tgcagaagat catcctggag 480  
ggcaccgagc gcatggagga ccagggccag agcatcatcc ccatgctgac cggcgagggtg 540  
atccccgtga tggagctgt gagcagcatg aagagccaca gcgtgcccga ggagatcgac 600  
atgcccgaca ccgtgctgaa cgacgacgac atcgccgaca gctgccacga gggttctctg 660  
ctgaacgcca tcagcagcca cctgcagacc tgcggctgca gcgtgggtgt gggcagcagc 720  
gccgagaagg tgaacaagat cgtgcgcacc ctgtgcctgt tcctgacccc cgccgagcgc 780

aagtgcagcc gcctgtgcga ggccgagagc agcttcaagt acgagagcgg cctgttcgtg 840  
cagggcctgc tgaaggacag caccggcagc ttcgtgctgc ccttccgcca ggtgatgtac 900  
gccccctacc ccaccacca catcgacgtg gacgtgaaca cgtgaagca gatgcccc 960  
tgccacgagc acatctacaa ccagcgccgc tacatgcgca gcgagctgac cgccttctgg 1020  
cgcgccacca gcgaggagga catggcccag gacacatca tctacaccga cgagagcttc 1080  
acccccgacc tgaacatctt ccaggacgtg ctgcaccgag acacctggt gaaggccttc 1140  
ctggaccagg tgttccagct gaagcccggc ctgagcctgc gcagcacctt cctggcccag 1200

ttcctgctgg tgctgcaccg caaggccctg accctgatca agtacatcga ggacgacacc 1260  
cagaagggca agaagccctt caagagcctg cgcaacctga agatcgacct ggacctgacc 1320  
gccgagggcg acctgaacat catcatggcc ctggccgaga agatcaagcc cggcctgcac 1380  
agcttcatct tcggccgccc ctctacacc agcgtgcagg agcgcgacgt gctgatgacc 1440  
ttctaa 1446

<210> 52  
<211> 3994  
<212> DNA  
<213> Artificial Sequence  
<220><223> Synthetic Polynucleotide  
<400> 52

ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg 120  
gccaaactcca tctactagggg ttctgctag ctctgggtat ttaagcccga gtgagcacgc 180  
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240  
tccctagggt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300  
cctagtattt aatagtaate aattacgggg tcattagtgc atagcccata tatggagtgc 360  
cgcgttacat aacttacggt aaatggcccc cctggctgac cgccaacga cccccccca 420  
ttgacgtcaa taatgacgta tgttcccata gtaacgcaa tagggacttt ccattgacgt 480

caatgggtgg actatttacg gtaaaactgcc cacttggcag tacatcaagt gtatcatatg 540  
ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgcttgca ttatgccag 600  
tacatgacct tatgggactt tctacttgg cagtacatct acgtattagt catcgctatt 660  
accatggtcg aggtgagccc cacgttctgc ttactctcc ccatctcccc cccctccca 720  
cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg 780  
gggggggggc gcgcgccagg cggggcgggg cgggcgagg ggcggggcgg ggcgaggcgg 840  
agaggcgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg 900

cggcggcggc ggcggcccta taaaaagcga agcgcgcggc ggcggggagt cgctgcgacg 960  
ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cgccgcgcc ggctctgact 1020  
gaccgcgtta ctccacagg tgagcgggcg ggacggccct tctctcagc gctgtaatta 1080  
gcgcttggtt taatgacggc ttgttgagg cttgctgaag gctgtatgct gttgtcatgc 1140  
atccatattc ttctagtga agccacagat gaggaagaat atggatgcat aggacacaag 1200  
gcctgttact agcactcaca tggaacaaat ggccaccgtg ggaggatgac aatttctgtg 1260

|                                                                    |      |
|--------------------------------------------------------------------|------|
| gctgctgtaa agccttgagg ggctccggga gctagagcct ctgctaacca tgttcattgcc | 1320 |
|                                                                    |      |
| ttcttctttt tcttacagct cctgggcaac gtgctggtta ttgtgctgtc tcatcatttt  | 1380 |
| ggcaaagaat tctcgaaga tccgaaggga aagtcttcca cgactgtggg atccgttcga   | 1440 |
| agatacacc gggtgagcca ccatgagcac cctgtgcccc ccccccagcc ccgccgtggc   | 1500 |
| caagaccgag atcgccctga gcggcaagag cccctgctg gccgccacct tcgcctactg   | 1560 |
| ggacaacatc ctgggcccc gcgtgcgcca catctgggcc cccaagaccg agcaggtgct   | 1620 |
| gctgagcgac ggcgagatca ctttcttggc caaccacacc ctgaacggcg agatcctgcg  | 1680 |
| caacgccgag agcgcgcca tcgactgaa gttcttctgt ctgagcgaga agggcgtgat    | 1740 |
|                                                                    |      |
| catcgtgagc ctgatcttcg acggcaactg gaacggcgac cgcagcacct acggcctgag  | 1800 |
| catcatcctg cccagaccg agctgagctt ctacctgcc ctgcaccgcg tgtgctgga     | 1860 |
| ccgctgacc cacatcatcc gcaaggccg catctggatg cacaaggagc gccaggagaa    | 1920 |
| cgtgcagaag atcatcctgg agggcaccca gcgcatggag gaccagggcc agagcatcat  | 1980 |
| ccccatgctg accggcgagg tgatccccgt gatggagctg ctgagcagca tgaagagcca  | 2040 |
| cagcgtgccc gaggagatcg acatcgccga caccgtgctg aacgacgacg acatcgccga  | 2100 |
| cagctgccac gagggttcc tgctgaacgc catcagcagc cacctgcaga cctgcggctg   | 2160 |
|                                                                    |      |
| cagcgtggtg gtgggcagca gcgccgagaa ggtgaacaag atcgtgcgca ccctgtgcct  | 2220 |
| gttctgacc cccgccgagc gcaagtgcag ccgctgtgc gaggccgaga gcagcttcaa    | 2280 |
| gtacgagagc ggctgttcg tgcaggccct gctgaaggac agcaccggca gcttcgtgct   | 2340 |
| gcccttccgc caggtgatgt acgcccccta cccaccacc cacatcgacg tggacgtgaa   | 2400 |
| caccgtgaag agatgcccc cctgccacga gcacatctac aaccagcgcc gctacatgcg   | 2460 |
| cagcgagctg accgccttct ggccgccac cagcgaggag gacatggccc aggacacat    | 2520 |
| catctacacc gacgagagct tcacccccga cctgaacatc ttccaggacg tgctgcaccg  | 2580 |
|                                                                    |      |
| cgacaccctg gtgaaggcct tcttggaaca ggtgttccag ctgaagcccg gcctgagcct  | 2640 |
| gcgcagcacc ttcttgccc agttcttctg ggtgctgcac cgcaaggccc tgacctgat    | 2700 |
| caagtacatc gaggacgaca cccagaaggg caagaagccc ttcaagagcc tgcgcaacct  | 2760 |
| gaagatcgac ctggacctga ccgccgagg cgacctgaac atcatcatgg ccctggccga   | 2820 |
| gaagatcaag cccggcctgc acagcttcat ctctggccgc ccttctaca ccagcgtgca   | 2880 |
| ggagcgcgac gtgctgatga ctttctaaca attgttaatt aagtttaaac cctcgaggcc  | 2940 |
| gcaagcttat cgataatcaa cctctggatt acaaaatttg tgaaagattg actggtattc  | 3000 |

ttaactatgt tgctcctttt acgctatgtg gatacgtgc tttaatgcct ttgtatcatg 3060  
ctattgcttc ccgatggctt ttcatcttct cctccttgta taaatcctgg ttgctgtctc 3120  
tttatgagga gttgtggccc gttgtcaggc aacgtggcgt ggtgtgcaact gtgtttgctg 3180  
acgcaacccc cactggttgg ggcattgcca ccacctgtca gctcctttcc gggactttcg 3240  
ctttccccc cctattgcc acggcggaac tcacgcgcgc ctgccttgcc cgctgctgga 3300  
caggggctcg gctgttgggc actgacaatt ccgtggtgtt gtcggggaaa tcacgtcctt 3360  
ttccttggct gctgcctgt gttgccacct ggattctgcg cgggacgtcc ttctgctacg 3420

tccttcggc cctcaatcca gcgaccttc ctccccggg cctgctgccg gctctcggc 3480  
ctcttcggc tcttcgctt cgccctcaga cgagtcggat ctccctttgg gccgcctccc 3540  
cgcatcgata ccgtcgacta gagctcgctg atcagcctcg actgtgcctt ctagtgtcca 3600  
gccatctgtt gtttccccct ccccgtgcc ttcccttgacc ctggaagggt ccaactccac 3660  
tgtcctttcc taataaaatg aggaaattgc atcgattgt ctgagtaggt gtcattctat 3720  
tctgggggtt ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca 3780  
tgctggggag agatccacga taacaacag cttttttggg gtgaacatat tgactgaatt 3840

ccctgcaggt tggccactcc ctctctgcgc gctcgtcgc tcaactgaggc cggccgggca 3900  
aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag tgagcgagcg agcgcgcaga 3960  
gagggagtgg ccaactccat cactaggggt tcct 3994

<210> 53  
<211> 3997  
<212> DNA  
<213> Artificial Sequence  
<220><223> Synthetic Polynucleotide  
<400> 53

ttggccactc cctctctgcg cgctcgtcg ctcaactgagg ccgggcgacc aaaggtcgcc 60  
cgacgcccg gctttgccc ggccgcctca gtgagcgagc gagcgcgag agagggagtg 120  
gccaaactca tcaactagggg ttcttgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240  
tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300  
cctagttatt aatagtaatc aattacgggg tcattagttc atagccata tatggagtcc 360  
cgcgttacat aacttacgtt aaatggcccg cctggctgac cgcccaacga ccccgccca 420  
ttgacgtcaa taatgacgta tgttccata gtaacgcaa tagggacttt ccattgacgt 480  
caatgggtgg actatttacg gtaactgcc cacttggcag tacatcaagt gtatcatatg 540

ccaagtacgc cccctattga cgtcaatgac ggtaaattggc ccgcctggca ttatgccag 600

tacatgacct tatgggactt tctacttgg cagtacatct acgtattagt catcgctatt 660

accatggtcg aggtgagccc caggttctgc ttactctcc ccatctcccc cccctcccca 720

cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg 780

gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg 840

agagggtcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg 900

cggcggcggc ggcgcccta taaaagcga agcgcgcggc gggcgggagt cgctgcgacg 960

ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cggccgcccc ggctctgact 1020

gaccgcgtta ctccacagg tgagcgggcg ggacggccct tctcctcagc gctgtaatta 1080

gcgcttgggt taatgacggc ttgttgagg cttgctgaag gctgtatgct gttgtcgggg 1140

ccggtttcgg ggcttagtg aagccacaga tgtacggccc cgaaaccggc cccaggacac 1200

aaggcctgtt actagcactc acatggaaca aatggccacc gtgggaggat gacaatttct 1260

gtggctgcgt gaaagccttg aggggctccg ggagctagag cctctgctaa ccatgttcat 1320

gccttcttct ttttctaca gctcctgggc aacgtgctgg ttattgtgct gtctcatcat 1380

tttggcaaag aattcctcga agatccgaag ggaaagtctt ccacgactgt gggatccgtt 1440

cgaagatac accggttgag ccacatgag caccctgtgc ccccccca gccccccgt 1500

ggccaagacc gagatcgccc tgagcggcaa gagccccctg ctggccgcca ctttcgccta 1560

ctgggacaac atcctgggcc cccgcgtgcg ccacatctgg gcccccaaga ccgagcaggt 1620

gctgctgagc gacggcgaga tcaccttctt ggccaaccac accctgaacg gcgagatcct 1680

gcgcaacgcc gagagcggcg ccatcgacgt gaagtcttct gtgctgagcg agaaggcgct 1740

gatcatcgtg agcctgatct tcgacggcaa ctggaacggc gaccgcagca cctacggcct 1800

gagcatcatc ctgccccaga ccgagctgag cttctacctg cccctgcacc gcgtgtgcgt 1860

ggaccgcctg acccacatca tccgaaggg ccgcatctgg atgcacaagg agcgcagga 1920

gaacgtgcag aagatcatcc tggagggcac cgagcgcagt gaggaccagg gccagagcat 1980

catcccatg ctgaccggcg aggtgatccc cgtgatggag ctgctgagca gcatgaagag 2040

ccacagcgtg cccgaggaga tcgacatcgc cgacaccgtg ctgaacgacg acgacatcgg 2100

cgacagctgc cagagggtt tctgctgaa cgccatcagc agccacctgc agacctgcgg 2160

ctgcagcgtg gtgtgggca gcagcgccga gaaggtgaac aagatcgtgc gcacctgtg 2220

cctgttctg acccccgccg agcgcaagtg cagccgcctg tgcgaggccg agagcagctt 2280

caagtacgag agcggcctgt tcgtgcaggg cctgctgaag gacagcacgc gcagcttcgt 2340  
 gctgcccttc cgccaggatg tgtacgcccc ctacccacc acccacatcg acgtggacgt 2400  
 gaacaccgtg aagcagatgc ccccttgcca cgagcacatc tacaaccagc gccgctacat 2460  
 gcgcagcgag ctgaccgcct tctggcgcgc caccagcgag gaggacatgg cccaggacac 2520  
 catcatctac accgacgaga gcttcacccc cgacctgaac atcttcagg acgtgctgca 2580  
 ccgcgacacc ctggtgaagg ccttcctgga ccaggtgttc cagctgaagc ccggcctgag 2640  
 cctgcgcagc accttcctgg ccagttcct gctggtgctg caccgcaagg ccctgaccct 2700  
  
 gatcaagtac atcgaggacg acaccagaa gggcaagaag cccttcaaga gcctgcgcaa 2760  
 cctgaagatc gacctggacc tgaccgccga gggcgacctg aacatcatca tggccctggc 2820  
 cgagaagatc aagccccggc tgcacagctt catcttcggc cgcccttct acaccagcgt 2880  
 gcaggagcgc gacgtgctga tgaccttcta acaattgtta attaagtta aacctcgag 2940  
 gccgcaagct tatcgataat caacctctgg attacaaaat ttgtgaaaga ttgactggtg 3000  
 ttcttaacta tgttgcctct ttacgctat gtggatacgc tgctttaatg cctttgtatc 3060  
 atgctattgc ttcccgatg gctttcattt tctctcctt gtataaatcc tggttgctgt 3120  
  
 ctctttatga ggagtgtgg cccgttgtca ggcaacgtgg cgtggtgtgc actgtgtttg 3180  
 ctgacgcaac cccactggt tggggcattg ccaccacctg tcagctcctt tccgggactt 3240  
 tcgctttccc cctccctatt gccacggcgg aactcatgc cgctgcctt gcccgctgct 3300  
 ggacaggggc tcggtgttg ggcactgaca attccgtggt gttgtcgggg aatcatcgt 3360  
 cctttccttg gctgctcgc tgtgttgcca cctggattct gcgcgggacg tccttctgct 3420  
 acgtcccttc ggccctcaat ccagcggacc ttcttccc cggcctgctg ccggctctgc 3480  
 ggccctcttc gcgtcttcgc cttcgcctc agacgagtc gatctcctt tgggccgcct 3540  
  
 ccccgcatcg ataccgtga ctagagctcg ctgatcagcc tcgactgtgc cttctagtgt 3600  
 ccagccatct gtgtttgcc cctccccgt gccttccttg accctggaag gtgccactcc 3660  
 cactgtcctt tcctaataaa atgaggaaat tgcacgcat tgtctgagta ggtgtcattc 3720  
 tattctgggg ggtggggtgg ggcaggacag caagggggag gattgggaag acaatagcag 3780  
 gcatgctggg gagagatcca cgataacaaa cagctttttt ggggtgaaca tattgactga 3840  
 attccctgca ggttggccac tcctctctg cgcgctcgt cgctcactga ggccgccgg 3900  
 gcaaagcccg ggcgtcgggc gacctttgt cgccggcct cagtgagcga gcgagcgcgc 3960  
  
 agagaggag tggccaactc catcactagg ggttcct 3997

<210> 54

<211> 3994

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 54

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc   | 60   |
| cgacgcccg gctttgcccc ggccgcctca gtgagcgagc gagcgcgag agagggagt     | 120  |
| gccaaactcca tacttagggg ttcttgctag ctctgggtat ttaagcccga gtgagcacgc | 180  |
| aggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc   | 240  |
| tccctagggt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac  | 300  |
|                                                                    |      |
| cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc  | 360  |
| cgcgttacat aacttacggt aaatggcccc cctggctgac cgcccaacga cccccccca   | 420  |
| ttgacgtcaa taatgacgta tgttccata gtaacccaa tagggacttt ccattgacgt    | 480  |
| caatgggtgg actatttacg gtaaaactgcc cacttgccag tacatcaagt gtatcatatg | 540  |
| ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgctggca ttatgccag    | 600  |
| tacatgacct taigggactt tctacttgg cagtacatct acgtattagt catcgctatt   | 660  |
| accatggtcg aggtgagccc cagttctgc ttactctcc ccatctccc cccctccca      | 720  |
|                                                                    |      |
| cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggccggggggg | 780  |
| gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggccgggcgg gccgaggcgg  | 840  |
| agaggtgcgg ccgcagccaa tcagagcggc gcgctccgaa agtttcttt tatggcgagg   | 900  |
| cggcggcggc ggccggccca taaaaagcga agcgcgcggc ggccgggagt cgctgcgacg  | 960  |
| ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc ccccccccc ggctctgact   | 1020 |
| gaccgcgtta ctccacagg tgagcgggcg ggacggccct tctcctcagc gctgtaatta   | 1080 |
| gcgcttggtt taatgacggc ttgttggagg ctgtctgaag gctgtatgct gttgtcttat  | 1140 |
|                                                                    |      |
| ttactttctc tgcacagtga agccacagat ggtgcagaga aagtaataa aggacacaag   | 1200 |
| gcctgttact agcactcaca tggaacaaat ggccaccgtg ggaggatgac aatttctgtg  | 1260 |
| gctgcgtgaa agccttgagg ggctccggga gctagacct ctgctaacca tgttcatgcc   | 1320 |
| ttcttctttt tctacagct cctgggcaac gtgctgttta ttgtctgtc tcatcatttt    | 1380 |
| ggcaaagaat tctcgaaga tccgaaggga aagtcttcca cgactgtggg atccgttcga   | 1440 |
| agatatcacc gggttagcca ccatgagcac cctgtgcccc cccccagcc ccgccgtggc   | 1500 |
| caagaccgag atcgccctga gcggcaagag cccctgctg gccgccacct tcgctactg    | 1560 |
|                                                                    |      |
| ggacaacatc ctgggcccc gcgtgcgcca catctgggcc cccaagaccg agcaggtgct   | 1620 |



gctgagcgac ggcgagatca ccttcctggc caaccacacc ctgaacggcg agatcctgcg 1680  
 caacgccgag agcggcgcca tcgacgtgaa gttcttcgtg ctgagcgaga agggcgtgat 1740  
 catcgtgagc ctgatcttcg acggcaactg gaacggcgac cgcagcacct acggcctgag 1800  
 catcatcctg cccagaccg agctgagctt ctacctgccc ctgcaccgcg tgtgcgtgga 1860  
 ccgctgacc cacatcatcc gcaagggccg catctggatg cacaaggagc gccaggagaa 1920  
 cgtgcagaag atcatcctgg agggcaccga gcgcatggag gaccagggcc agagcatcat 1980

ccccatgctg accggcgagg tgatccccgt gatggagctg ctgagcagca tgaagagcca 2040  
 cagcgtgccc gaggagatcg acatcgccga caccgtgctg aacgacgacg acatcgccga 2100  
 cagctgccac gagggttcc tgctgaacgc catcagcagc cacctgcaga cctgcggctg 2160  
 cagcgtgggt gtgggcagca gcgccgagaa ggtgaacaag atcgtgcgca ccctgtgcct 2220  
 gttctgacc cccgccgagc gcaagtgcag ccgctgtgc gaggccgaga gcagcttcaa 2280  
 gtacgagagc ggctgttcg tgcagggcct gctgaaggac agcaccggca gcttcgtgct 2340  
 gcccttcgc caggtgatgt acgcccccta cccaccacc cacatcgacg tggacgtgaa 2400

caccgtgaag cagatgcccc cctgccacga gcacatctac aaccagcgcc gctacatgcg 2460  
 cagcgagctg accgccttct ggcgccac cagcgaggag gacatggccc aggacacat 2520  
 catctacacc gacgagagct tcacccccga cctgaacatc ttccaggacg tgctgcaccg 2580  
 cgacaccctg gtgaaggcct tcctggacca ggtgttccag ctgaagcccc gcctgagcct 2640  
 gcgcagcacc ttcttgccc agttctgct ggtgctgcac cgcaaggccc tgaccctgat 2700  
 caagtacatc gaggacgaca cccagaaggg caagaagccc ttcaagagcc tgcgcaacct 2760  
 gaagatcgac ctggacctga ccgccgaggg cgacctgaac atcatcatgg ccctggccga 2820

gaagatcaag cccggcctgc acagcttcat cttcggccgc ccttctaca ccagcgtgca 2880  
 ggagcgcgac gtgctgatga cttttaaca attgttaatt aagtttaaac cctcgaggcc 2940  
 gcaagcttat cgataatcaa cctctggatt acaaaatttg tgaaagattg actggtattc 3000  
 ttaactatgt tgctcctttt acgctatgtg gatacgtgc tttaatgcct ttgtatcatg 3060  
 ctattgcttc ccgtatggct ttcattttct cctccttgta taaatcctgg ttgctgtctc 3120  
 tttatgagga gttgtggccc gttgtcaggc aacgtggcgt ggtgtgcact gtgtttgctg 3180  
 acgcaacccc cactggttgg ggcattgcca ccacctgtca gctcctttcc gggactttcg 3240

ctttccccc cctattgcc acggcggaac tcacgcccgc ctgccttgcc cgctgctgga 3300  
 caggggctcg gctgttgggc actgacaatt ccgtggtgtt gtcggggaaa tcacgtcct 3360  
 ttcttggtg gctgcctgt gttgccacct ggattctgcg cgggacgtcc ttctgctacg 3420  
 tcccttcgga cctcaatcca gcggaccttc ctcccgcg cctgctgccg gctctgcggc 3480

ctcttccgcg tcttcgcctt cgccctcaga cgagtcggat ctccctttgg gccgcctccc 3540  
 cgcatcgata ccgtcgacta gagctcgctg atcagcctcg actgtgcctt ctagtgtcca 3600  
 gccatctgtt gtttgcacct ccccggtgcc ttcccttgacc ctggaaggtg ccaactccac 3660

tgtcctttcc taataaaatg aggaaattgc atcgattgt ctgagtaggt gtcattctat 3720  
 tctggggggt ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca 3780  
 tgctggggag agatccacga taacaacag cttttttggg gtgaacatat tgactgaatt 3840  
 ccctgcaggt tggccactcc ctctctgcgc gctcgctcgc tcaactgaggc cgcccgggca 3900  
 aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag tgagcgagcg agcgcgcaga 3960  
 gagggagtgg ccaactccat cactaggggt tcct 3994

<210> 55  
 <211> 902  
 <212> DNA  
 <213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 55

ttggccaatc cctctctgcg cgctcgctcg ctcaactgagg ccgggcgacc aaaggtcgcc 60  
 cgacgcccgg gctttgcccg ggcgcctca gtgagcgagc gagcgcgag agagggagtg 120  
 gccaaactca tcaactagggg ttcttgctag ctctgggtat ttaagcccga gtgagcacgc 180  
 agggctctca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240  
 aaaaaaattg tcatcctccc acggtggcca tttgttccat gtgagtgcta gtaacaggcc 300  
 ttgtgtcctg gggccggttt cggggccgta catctgtggc ttactaagg ccccgaaacc 360

ggccccgaca acagcataca gccttcagca agcctccagt ggtctcatc agaacttata 420  
 agattcccaa atccaaagac atttcacgtt tatggtgatt tcccagaaca catagcgaca 480  
 tgcaaatatt gcagggcgcc actccctgt ccttcacagc catcttctg ccagggcgca 540  
 cgcgcgctgg gtgttccgc ctagtacac tgggcccgcg attccttgga gcgggttgat 600  
 gacgtcagcg ttcccatgg tgaagcttg atctgatccc taggttctag aaccggtgac 660  
 caattgttaa ttaagttaa accctcgagg ccgaagcag atccagata acaaacagct 720  
 tttttgggt gaacatatg actgaattcc ctgcaggtt gccactcct ctctgcgcgc 780

tcgctcgctc actgaggccg cccgggcaaa gcccgggcgt cgggcgacct ttggtcgccc 840  
 ggcctcagtg agcgagcgag cgcgagaga gggagtggcc aactccatca ctaggggttc 900  
 ct 902

<210> 56

<211> 2550

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 56

|                                                                     |      |
|---------------------------------------------------------------------|------|
| ttggccactc ccctctctgcg cgctcgctcg ctactgagg ccgcccgggc aaagcccggg   | 60   |
| cgctcgggcga cctttggctc cccggcctca gtgagcgagc gagcgcgag agagggagt    | 120  |
| gccaaactcca tcaactagggg ttcttgctag ctctgggtat ttaagcccga gtgagcacgc | 180  |
| agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc   | 240  |
| tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtag   | 300  |
| ctagttatta atagtaatca attacggggt cattagtcca tagcccatat atggagtcc    | 360  |
| gcgttacata acttacggta aatggcccg cctggctgacc gccaacgac ccccgccat     | 420  |
| tgacgtcaat aatgacgtat gtcccatag taacgccaat agggactttc cattgacgtc    | 480  |
| aatgggtgga gtatttacgg taaactgcc acttggcagt acatcaagt tatcatatgc     | 540  |
| caagtacgcc ccctattgac gtcaatgac gtaaatggcc cgctggcat tatgcccagt     | 600  |
| acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgtatta    | 660  |
| ccatggtcga ggtgagcccc acgttctgct tcaactctcc catctcccc cctccccac     | 720  |
| ccccaatttt gtatttatat ttttttaat tttttgtgc agcgatgggg gcgggggggg     | 780  |
| ggggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga   | 840  |
| gaggtcggcg ggcagccaat cagagcggcg cgctccgaaa gtttcctttt atggcgaggc   | 900  |
| ggcggcgggc gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgacgc   | 960  |
| tgccctcgcc ccgtgccccg ctccgccgcc gcctcgcgcc gcccgcgccg gctctgactg   | 1020 |
| accgcgttac tcccacaggt gagcgggcgg gacggccctt ctctcagcg ctgtaattag    | 1080 |
| cgcttggttt aatgacggct tggttgaggc ttgctgaagg ctgtatgctg ttgtcggggc   | 1140 |
| cggtttcggg gccttagtga agccacagat gtacggcccc gaaaccggcc ccaggacaca   | 1200 |
| aggcctgtta ctagcactca catggaacaa atggccaccg tgggaggatg acaatttctg   | 1260 |
| tggtcgcgtg aaagccttga ggggctccgg gagctagagc ctctgctaac catgttcatt   | 1320 |
| ccttcttctt tttctacag ctctgggca acgtgctggt tattgtgctg tctcatcatt     | 1380 |
| ttggcaaaga attcctcgaa gatccgaagg gaaagtcttc cacgactgtg ggatccgttc   | 1440 |
| gaagatatca ccggttgagc cacccaattg ttaattaagt ttaaacctc gaggcgcaa     | 1500 |

gcttatcgat aatcaacctc tggattacaa aatttgtgaa agattgactg gtattcttaa 1560  
ctatgttgct ccttttacgc tatgtggata cgtgcttta atgcctttgt atcatgctat 1620  
tgcttcccgat atggctttca ttttctctc cttgtataaa tcttggttgc tgtctcttta 1680  
tgaggagtgt tggcccggtg tcaggcaacg tggcgtgggtg tgcactgtgt ttgctgacgc 1740  
aacccccact ggttggggca ttgccaccac ctgtcagctc ctttccggga ctttcgcttt 1800  
ccccctccct attgccacgg cggaactcat cgcgcctgc cttgcccgct gctggacagg 1860

ggctcggctg ttgggcactg acaattccgt ggtgtgtgctg gggaaatcat cgtcctttcc 1920  
ttggctgctc gccgtgtgtg ccacctggat tctgcgcggg acgtccttct gctacgtccc 1980  
ttcggccctc aatccagcgg accttccttc ccgcggcctg ctgcggctc tgcggcctct 2040  
tccgcgtctt cgccttcgcc ctacagcag tggatctcc ctttgggccg cctccccga 2100  
tcgataccgt cgactagagc tcgtgatca gcctcgactg tgccttctag ttgccagcca 2160  
tctgttgttt gccctcccc cgtgccttcc ttgacctgg aaggtgccac tcccactgtc 2220  
ctttcctaataaaatgagga aattgcatcg cattgtctga gtaggtgtca ttctattctg 2280

gggggtgggg tgggacagga cagcaagggg gaggattggg aagacaatag caggcatgct 2340  
ggggagagat ccacgataac aaacagcttt ttgggggtga acatattgac tgaattccct 2400  
gcaggttggc cactccctct ctgcgcctc gctcgtcac tgaggccgc cgggcaaagc 2460  
ccgggcgtcg ggcgaccttt ggtcgcccg cctcagtgag cgagcgagcg cgcagagagg 2520  
gagtggccaa ctccatcact aggggttcc 2550

<210> 57

<211> 4142

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 57

ttggccactc cctctctgcg cgctcgtcgc ctactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg 120  
gccaactcca tctactagggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc 180  
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240  
tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300  
cctagtattt aatagtaatc aattacgggg tcattagttc atagccata tatggagttc 360  
cgcgttacat aacttacggt aaatggcccc cctggctgac cgcccaacga cccccccca 420

|                                                                   |      |
|-------------------------------------------------------------------|------|
| ttgacgtcaa taatgacgta tggtcccata gtaacgcaa tagggacttt ccattgacgt  | 480  |
|                                                                   |      |
| caatgggtgg actatttacg gtaaactgcc cacttggcag tacatcaagt gtatcatatg | 540  |
| ccaagtacgc cccctattga cgtcaatgac ggtaaattggc ccgcctggca ttatgccag | 600  |
| tacatgacct tatgggactt tctacttgg cagtacatct acgtattagt catcgctatt  | 660  |
| accatggtcg aggtgagccc cagttctgc ttactctcc ccatctcccc cccctccca    | 720  |
| cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg | 780  |
| gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg | 840  |
| agaggtgcgg cggcagccaa tcagagcggc gcgtccgaa agtttcttt tatggcgagg   | 900  |
|                                                                   |      |
| cggcggcggc ggcggcccta taaaaagcga agcgcgcggc ggcggggagt cgctgcgacg | 960  |
| ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cggccgcccc ggctctgact | 1020 |
| gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctctccgg gctgtaatta  | 1080 |
| gcgcttgggt taatgacggc ttgtctggag gcttgctttg ggctgtatgc tgcggttgcg | 1140 |
| gtgcctgcgc cttttggcct ctgactgagg cgcaggcacc gcaaccgcag gacacaaggc | 1200 |
| cctttatcag cactcacatg gaacaaatgg ccaccgtggg aggatgacaa ctggaggctt | 1260 |
| gctttgggct gtatgctgic tggttaatct ttatcagggt tttggcctct gactgaacct | 1320 |
|                                                                   |      |
| gataaagatt aaccagacag gacacaaggc cctttatcag cactcacatg gaacaaatgg | 1380 |
| ccaccgtggg aggatgacaa tttctgtggc tgcgtgaaag ccttgagggg ctccgggagc | 1440 |
| tagagcctct gtaaacatg ttcatgcctt cttcttttct ctacagctcc tgggcaactg  | 1500 |
| gctggttatt gtctgtctc atcattttgg caaagaattc ctgaagatc cgaagggaaa   | 1560 |
| gtcttcacg actgtgggat ccgttcgaag atatcaccgg ttgagccacc atgagcacc   | 1620 |
| tgtgcccccc cccagcccc gccgtggcca agaccgagat cgccctgagc ggcaagagcc  | 1680 |
| ccctgtctggc cgccaccttc gcctactggg acaacatcct gggccccgc gtgcgccaca | 1740 |
|                                                                   |      |
| tctgggcccc caagaccgag cagggtctgc tgagcgacgg cgagatcacc ttcctggcca | 1800 |
| accacacct gaacggcgag atcctgcgca acgccgagag cggcgccatc gacgtgaagt  | 1860 |
| tcttctgtct gagcgagaag ggcgtgatca tcgtgagcct gatcttcgac ggcaactgga | 1920 |
| acggcgaccg cagcacctac ggcctgagca tcactctgcc ccagaccgag ctgagcttct | 1980 |
| acctgcccc gcaccgcgtg tgcgtggacc gctgaccca catcatccgc aagggccgca   | 2040 |
| tctggatgca caaggagcgc caggagaacg tgcagaagat catcctggag ggcaccgagc | 2100 |
| gcatggagga ccagggccag agcatcatcc ccatgctgac cggcgagggt atccccgtga | 2160 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| tggagctgct gagcagcatg aagagccaca gcgtgcccga ggagatcgac atcgccgaca  | 2220 |
| ccgtgtctgaa cgacgacgac atcgccgaca gctgccacga gggcttcctg ctgaacgcca | 2280 |
| tcagcagcca cctgcagacc tgcggctgca gcgtggtggt gggcagcagc gccgagaagg  | 2340 |
| tgaacaagat cgtgcgcacc ctgtgcctgt tctgacccc cgccgagcgc aagtgcagcc   | 2400 |
| gcctgtgcga ggccgagagc agcttcaagt acgagagcgg cctgttcctg cagggcctgc  | 2460 |
| tgaaggacag caccggcagc ttctgtctgc cttccgcca ggtgatgtac gcccctacc    | 2520 |
| ccaccacca catcgacgtg gacgtgaaca ccgtgaagca gatgcccccc tgccacgagc   | 2580 |
|                                                                    |      |
| acatctacaa ccagcggcgc tacatgcgca gcgagctgac cgccttctgg cgcgccacca  | 2640 |
| gcgaggagga catggcccag gacaccatca tctacaccga cgagagcttc acccccgacc  | 2700 |
| tgaacatctt ccaggacgtg ctgcaccgcg acaccctggt gaaggccttc ctggaccagg  | 2760 |
| tgttccagct gaagcccggc ctgagcctgc gcagcacctt cctggcccag ttctgtctgg  | 2820 |
| tgctgcaccg caaggccctg accctgatca agtacctga ggacgacacc cagaagggca   | 2880 |
| agaagccctt caagagcctg cgcaacctga agatcgacct ggacctgacc gccgaggcgc  | 2940 |
| acctgaacat catcatggcc ctggccgaga agatcaagcc cggcctgcac agcttcatct  | 3000 |
|                                                                    |      |
| tcggccgccc ctctacacc agcgtgcagg agcgcgacgt gctgatgacc ttctaacaat   | 3060 |
| tgtaattaa gtttaaaccc tcgaggccgc aagcttatcg ataataacc tctggattac    | 3120 |
| aaaatttgtg aaagattgac tggatttctt aactatgttg ctctttttac gctatgtgga  | 3180 |
| tacgtgcttt taatgccttt gtatcatgct attgcttccc gtatggcttt cattttctcc  | 3240 |
| tccttgata aatcctggtt gctgtctctt tatgaggagt tgtggcccggt tgtcaggcaa  | 3300 |
| cgtggcgtgg tgtgactgt gtttctgac gcaaccccca ctggttgggg cattgccacc    | 3360 |
| acctgtcagc tcctttccgg gactttcgtt tccccctcc ctattgccac ggcggaactc   | 3420 |
|                                                                    |      |
| atcgccgctt gccttgcccg ctgctggaca ggggctcggc tgttgggcac tgacaattcc  | 3480 |
| gtggtgttgt cggggaaatc atcgtccttt ccttggtctg tcgctgtgt tgccacctgg   | 3540 |
| attctgcgcg ggacgtcctt ctgctacgtc ccttcggccc tcaatccagc ggaccttctt  | 3600 |
| tcccgggcc tgetgccgc tctgcggcct cttccgcgtc ttgccttcg cctcagacg      | 3660 |
| agtcggatct ccttttgggc cgctccccg catcgatacc gtcactaga gctcgtgat     | 3720 |
| cagcctgcac tgtgccttct agttgccagc catctgttgt ttgccctcc cccgtgcctt   | 3780 |
| ccttgacctt ggaaggtgcc actccactg tcctttccta ataaaatgag gaaattgcat   | 3840 |
|                                                                    |      |
| cgcattgtct gagiaggigt cattctattc tggggggtgg ggtggggcag gacagcaagg  | 3900 |
| gggaggattg ggaagacaat agcaggcatg ctggggagag atccacgata acaaacagct  | 3960 |
| tttttggggt gaacatatgt actgaattcc ctgcaggttg gccactccct ctctgcgcgc  | 4020 |

|                                                                   |      |
|-------------------------------------------------------------------|------|
| tcgctcgctc actgaggccg cccgggcaaa gcccgggcgt cgggcgacct ttggtcgccc | 4080 |
| ggcctcagtg agcgagcgag cgcgagaga gggagtggcc aactccatca ctaggggttc  | 4140 |
| ct                                                                | 4142 |
| <210> 58                                                          |      |
| <211> 4143                                                        |      |
| <212> DNA                                                         |      |
| <213> Artificial Sequence                                         |      |
| <220><223> Synthetic Polynucleotide                               |      |
| <400> 58                                                          |      |
| ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc  | 60   |
| cgacgcccg gctttgccc ggcgccctca gtgagcgagc gagcgcgag agagggagtg    | 120  |
| gccaaactca tctactaggg ttcttgctag ctctgggtat ttaagcccga gtgagcacgc | 180  |
| agggtctcca tttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc  | 240  |
| tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac | 300  |
| cctagttatt aatagtaatc aattacgggg tcattagttc atagccata tatggagttc  | 360  |
| cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga ccccgccca  | 420  |
| ttgacgtcaa taatgacgta tgttccata gtaacgcaa tagggacttt ccattgacgt   | 480  |
| caatgggtgg actatttacg gtaactgcc cacttggcag tacatcaagt gtatcatatg  | 540  |
| ccaagtacgc ccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgccag   | 600  |
| tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt | 660  |
| accatggtcg aggtgagccc cagttctgc ttactctcc ccatctccc cccctccca     | 720  |
| cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg | 780  |
| gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg | 840  |
| agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttcttt tatggcgagg  | 900  |
| cggcggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg | 960  |
| ctgccttcgc ccctgcccc gctccgccgc cgctcgcgc cggccgcccc ggctctgact   | 1020 |
| gaccgcgtta ctccacagg tgagcgggcg ggacggccct tctcctccgg gctgtaatta  | 1080 |
| gcgcttggtt taatgacggc ttgtctggag gcttgctttg ggctgtatgc tggctgcggt | 1140 |
| tgcggtgcct gcttttgcc tctgactgag caggcaccgc aaccgcagcc aggacacaag  | 1200 |
| gccctttatc agcactcaca tggaacaaat ggccaccgtg ggaggatgac aactggaggc | 1260 |
| ttgctttggg ctgtatgctg ggttgtttcc ctcttgttt attttggcct ctgactgaaa  | 1320 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| acaaggaggg aaacaacca ggacacaagg ccctttatca gcactcacat ggaacaaatg   | 1380 |
| gccaccgtgg gaggatgaca atttctgtgg ctgcgtgaaa gccttgagg gctccgggag   | 1440 |
| ctagagcctc tgctaaccat gttcatgcct tcttcttttt cctacagctc ctgggcaacg  | 1500 |
| tgctggttat tigtctgtct catcattttg gcaaagaatt cctcgaagat ccgaaggga   | 1560 |
| agtcttccac gactgtggga tccgttcgaa gatatcaccg gttgagccac catgagcacc  | 1620 |
| ctgtgcccc cccccagccc cgccgtggcc aagaccgaga tcgccctgag cggcaagagc   | 1680 |
| cccctgctgg ccgccacctt cgcctactgg gacaacatcc tgggcccccg cgtgcgccac  | 1740 |
| atctgggccc ccaagaccga gcaggtgctg ctgagcgacg gcgagatcac cttcctggcc  | 1800 |
| aaccacaccc tgaacggcga gatcctgcgc aacgccgaga gcggcgccat cgacgtgaag  | 1860 |
| ttcttctgtc tgagcgagaa gggcgtgatc atcgtgagcc tgatcttcga cggcaactgg  | 1920 |
| aacggcgacc gcagcaccta cggcctgagc atcatcctgc ccagaccga gctgagcttc   | 1980 |
| tacctgcccc tgcaccgcgt gtgcgtggac cgcctgaccc acatcatccg caagggccgc  | 2040 |
| atctggatgc acaaggagcg ccaggagaac gtgcagaaga tcacctgga gggcaccgag   | 2100 |
| cgcatggagg accagggccca gagcatcatc cccatgctga ccggcgaggt gatccccgtg | 2160 |
| atggagctgc tgagcagcat gaagagccac agcgtgcccg aggagatcga catcgccgac  | 2220 |
| accgtgctga acgacgacga catcgccgac agctgccacg agggcttcct gctgaacgcc  | 2280 |
| atcagcagcc acctgcagac ctgcggctgc agcgtggtgg tgggcagcag cgccgagaag  | 2340 |
| gtgaacaaga tcgtgcgcac cctgtgcctg ttctgaccc ccgccgagcg caagtgcagc   | 2400 |
| cgctgtgcg aggccgagag cagcttcaag tacgagagcg gcctgttcgt gcagggcctg   | 2460 |
| ctgaaggaca gcaccggcag cttcgtgctg cccttcgcc aggtgatgta cggccctac    | 2520 |
| cccaccaccc acatcgactt ggacgtgaac accgtgaagc agatgcccc ctgccacgag   | 2580 |
| cacatctaca accagcgccg ctacatgcgc agcgagctga ccgccttctg gcgcgccacc  | 2640 |
| agcgaggagg acatggccca ggacaccatc atctacaccg acgagagctt ccccccgac   | 2700 |
| ctgaacatct tccaggactt gctgcaccgc gacaccctgg tgaaggcctt cctggaccag  | 2760 |
| gtgttcacgc tgaagcccg cctgagcctg cgcagcacct tcctggccca gttcctgctg   | 2820 |
| gtgctgcacc gcaaggccct gaccctgatc aagtacatcg aggacgacac ccagaagggc  | 2880 |
| aagaagccct tcaagagcct gcgcaacctg aagatcgacc tggacctgac cgccgagggc  | 2940 |
| gacctgaaca tcacatggc cctggccgag aagatcaagc ccggcctgca cagcttcatc   | 3000 |
| ttcgccgcc cttctacac cagcgtgcag gagcgcgacg tgctgatgac cttctaaca     | 3060 |
| ttgttaatta agtttaaac ctcgaggccg caagcttata gataatcaac ctctggatta   | 3120 |
| caaaatttgt gaaagattga ctggtattct taactatgtt gctcctttta cgctatgtgg  | 3180 |



atacgtgct ttaatgcctt tgtatcatgc tattgcttcc cgtatggctt tcattttctc 3240  
ctccttgtat aaatcctggt tgctgtctct ttatgaggag ttgtggcccg ttgtcaggca 3300

acgtggcgtg gtgtgactg tgtttgctga cgcaaccccc actggttggg gcattgccac 3360  
cacctgtcag ctcttttccg ggactttcgc tttcccccctc cctattgccg cggcggaact 3420  
catcgccgcc tgccttgccc gctgctggac aggggctcgg ctgttgggca ctgacaattc 3480  
cgtggtgttg tgggggaaat catcgtcctt tccttggctg ctgcctgtg ttgccacctg 3540  
gattctgcgc gggacgtcct tctgtacgt cccttcggcc ctcaatccag cggaccttcc 3600  
ttcccgggc ctgctgccg ctctcgggcc tcttcgcgt cttgccttc gccctcagac 3660  
gagtcggatc tccctttggg ccgcctcccc gcatcgatac cgtcgactag agctcgtga 3720

tcagcctega ctgtgccttc tagttgccag ccatctgttg tttgcccctc ccccgtgcct 3780  
tccttgacc tggagggtgc cactccact gtcccttctt aataaaatga ggaaattgca 3840  
tcgcattgtc tgagtaggtg tcattctatt ctggggggtg ggggtgggca ggacagcaag 3900  
ggggaggatt ggaagacaa tagcaggcat gctggggaga gatccacgat aacaaacagc 3960  
ttttttggg tgaacatatt gactgaattc cctgcagggt ggccactccc tctctgcgcg 4020  
ctcgtcgtc cactgaggcc gcccgggcaa agcccggcg tcgggcgacc tttggtcgcc 4080  
cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc actaggggtt 4140

cct 4143

<210> 59  
<211> 4140  
<212> DNA  
<213> Artificial Sequence  
<220><223> Synthetic Polynucleotide  
<400> 59

ttggccactc cctctctgcg cgctcgtcgc ctactgagg ccgggcgacc aaaggtcgcc 60  
cgacgcccg gctttgccc ggccgcctca gtgagcgagc gagcgcgag agagggagtg 120  
gccaaactca tctactaggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc 180  
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240  
tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300

cctagttatt aatagtaac aattacgggg tcattagttc atagcccata tatggagtgc 360  
cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga cccccccca 420  
ttgacgtcaa taatgacgta tgttccata gtaacgcaa tagggacttt ccattgacgt 480

caatgggtgg aciattttacg gtaaaactgcc cacttggcag tacatcaagt gtatcatatg 540  
 ccaagtagcg cccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgccag 600  
 tacatgacct tatgggactt tectacttgg cagtacatct acgtattagt catcgctatt 660  
 accatggtcg aggtgagccc cacgttctgc ttcactctcc ccatctcccc cccctcccca 720  
  
 cccccaattt tgtattttatt tattttttta ttattttgtg cagcgatggg ggcggggggg 780  
 gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg 840  
 agaggtagcg cggcagccaa tcagagcggc gcgctccgaa agtttcttt tatggcgagg 900  
 cggcgggcgg ggcggcccta taaaaagcga agcgcgcggc gggcgggagt cgtgcgacg 960  
 ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cggccgcccc ggctctgact 1020  
 gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctcctccgg gctgtaatta 1080  
 gcgcttggtt taatgacggc ttgtctggag gcttgctttg ggctgtatgc tgcggtgcct 1140  
  
 gcgcccgcgg cttttggcct ctgactgagc gcggggcgca ggcaccgcag gacacaaggc 1200  
 cctttatcag cactcacatg gaacaaatgg ccaccgtggg aggatgacaa ctggaggctt 1260  
 gctttgggct gtatgctgtt tcttctgggt aatctttatt ttggcctctg actgataaag 1320  
 attaacaga agaacagga cacaaggccc ttatcagca ctacatgga acaaatggcc 1380  
 accgtgggag gatgacaatt tctgtggctg cgtgaaagcc ttgaggggct ccgggagcta 1440  
 gagcctctgc taaccatgtt catgccttct tcttttctt acagctcctg ggcaacgtgc 1500  
 tggttattgt gctgtctcat cattttggca aagaattcct cgaagatccg aagggaagt 1560  
  
 ctccacgac tgtgggatcc gttcgaagat atcaccggtt gagccacat gagcacctg 1620  
 tgccccccc ccagccccgc cgtggccaag accgagatcg cctgagcgg caagagcccc 1680  
 ctgctggcgg ccaccttcgc ctactgggac aacatcctgg gccccgcgt gcgccacatc 1740  
 tgggccccca agaccgagca ggtgctgctg agcgacggcg agatcacctt cctggccaac 1800  
 cacacctga acggcgagat cctgcgcaac gccgagagcg gcgccatcga cgtgaagtgc 1860  
 ttcgtgctga gcgagaaggg cgtgatcatc gtgagcctga tcttcgacgg caactggaac 1920  
 ggcgaccgca gcacctacgg cctgagcatc atcctgcccc agaccgagct gagcttctac 1980  
  
 ctgcccctgc accgcgtgtg cgtggaccgc ctgaccaca tcatccgaa gggccgcatc 2040  
 tggatgcaca aggagcgcca ggagaactg cagaagatca tcttgagggg caccgagcgc 2100  
 atggaggacc agggccagag catcatcccc atgctgaccg gcgaggtgat ccccgatg 2160  
 gagctgctga gcagcatgaa gagccacagc gtccccgagg agatcgacat cgccgacacc 2220  
 gtgctgaacg acgacgacat cggcgacagc tgccacgagg gcttctctgt gaacgccatc 2280  
 agcagccacc tgcagacctg cggctgcagc gtggtggtgg gcagcagcg cgagaaggtg 2340

aacaagatcg tgcgcacctt gtgcctgttc ctgacccccg ccgagcgcaa gtgcagccgc 2400

ctgtgcgagg ccgagagcag cttcaagtac gagagcggcc tgttcgtgca gggcctgctg 2460

aaggacagca cgggcagctt cgtgctgccc ttccgccagg tgatgtacgc cccctacccc 2520

accaccaca tcgacgtgga cgtgaacacc gtgaagcaga tgccccctg ccacgagcac 2580

atctacaacc agcgccgcta catgcgcagc gagctgaccg ccttctggcg cgccaccagc 2640

gaggaggaca tggcccagga caccatcatc tacaccgacg agagcttcac ccccgacctg 2700

aacatcttcc aggacgtgct gcaccgcgac accctgggtga aggccttcct ggaccaggtg 2760

ttccagctga agcccggcct gagcctgcgc agcaccttcc tggcccagtt cctgctggtg 2820

ctgcaccgca aggccctgac cctgatcaag tacatcgagg acgacacca gaagggaag 2880

aagcccttca agagcctgcg caacctgaag atcgacctgg acctgaccgc cgagggcgac 2940

ctgaacatca tcatggccct ggccgagaag atcaagcccc gcctgcacag cttcatcttc 3000

ggccgccccct tctacaccag cgtgcaggag cgcgacgtgc tgatgacctt ctaacaattg 3060

ttaattaagt ttaaaccctc gaggcgcaa gcttatcgat aatcaacctc tggattacaa 3120

aatttgtgaa agattgactg gtattcttaa ctatgttgct ccttttacgc tatgtggata 3180

cgctgcttta atgcctttgt atcatgctat tgcttcccgt atggctttca ttttctctc 3240

cttgtataaa tcttggttgc tgtctcttta tgaggagtgt tggcccggtg tcaggcaacg 3300

tggcgtggtg tgactgtgt ttgctgacgc aacccccact ggttggggca ttgccaccac 3360

ctgtcagctc ctttccggga ctttgccttt cccctccct attgccacgg cggaactcat 3420

cgccgcctgc cttgcccgct gctggacagg ggtcggctg ttgggcactg acaattccgt 3480

ggtgttgtcg gggaaatcat cgtcctttcc ttggctgctc gccgtgtgtg ccacctggat 3540

tctgcgcggg acgtccttct gctacgtccc ttggccctc aatccagcgg accttcctc 3600

ccgcggcctg ctgccgctc tgcggcctct tccgcgtctt cgccttcgcc ctcagacgag 3660

tcggatctcc ctttgggccg cctccccgca tcgataccgt cgactagagc tcgctgatca 3720

gcctcgaactg tgccttctag ttgccagca tctgttgttt gccctccccc cgtgccttcc 3780

ttgacctgg aaggtgccac tcccactgtc ctttccaat aaaatgagga aattgcatcg 3840

cattgtctga gtaggtgtca ttctattctg gggggtgggg tggggcagga cagcaagggg 3900

gaggattggg aagacaatag caggcatgct ggggagagat ccacgataac aaacagcttt 3960

tttggggtga acatattgac tgaattccct gcaggttggc cactccctct ctgcgcgctc 4020

gctcgtcac tgaggccgcc cgggcaaagc cggggtgctg ggcgacctt ggtcggccgg 4080

cctcagttag cgagcgagcg cgcagagagg gaggggccaa ctccatcact aggggttcct 4140

<210> 60

<211> 3793

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 60

tctcattaga agtgaggcgg ggccggccaa atcgaatgga caccgggtaa ttagcagggt 60

taccagata ctccagcacc ttttcccggt cggccgtgta cctgccattc acgtccatgc 120

cattgatggc cagcactgca tgaccactg cagaggtaga gctaaccgtc agcgaagggtg 180

cagccccggg attccgcga ggggacaagg gacccgacac aaccctttt cccccaaccc 240

cgcacctaca accagccac ttctacagca ctggggccct cccaccccg caccgcccac 300

gggcccagc ctagccacc tcggatgccg tcccgtggc cgaaagcaac caacacacgc 360

tcatcgtgta gcttagagcag cagatccagc ggataactga aagttttctc agcctcagcc 420

cgtggcgcgt agctgtccaa ctggtaaactc aagccgccag cttgtttcac cacatacaca 480

ctaaaaatcg ccatcgtgc cttgccgtc ggaaactggt attcagctc taccgacgg 540

cccctcccc gaaccgcac acagcacttg ccgccggccc caccacagcc tctctctct 600

cctctctctc ctccgcgcc ccccgtagc ccacctgtg cacttgcga ctgggagcga 660

cagctcggg cataagtagt gccgaaaagt tagctgccga gacctggtgg attgcttttc 720

gtttatcagt gcaggaaaac agcgtatag tactgcgtca caactagcgc agactccggc 780

agtatttagg cggtagcggt tgggaactag aatccacttc ctgtcttcg cctcaggcta 840

gagggcgagc gcttcgccgt gggacttctt ctgcctggct ccgcctcttg ccccggaagt 900

actcacagcg gacgggtggt tttgggccc tttctgagca gcgttctct tttgtccgac 960

atcttgacga ggctgcgggt tctgtgcta ttctccgagc ttcgcaatgg taagcttcag 1020

gggtgtgaag tcgccggcgt tcttgggttt gaggactcag tggggagagc cttcggcggg 1080

agcgtctctt ggctgcccg cctcggttgc agggcgggcg cggttattgc ttggcccatg 1140

tgctctggtg gtggagtttg cgggggctga gggcgagta ttaggggact ttggcgtat 1200

ttgaggacct ggttgcatc ccgctgccct cctacagccg cctaaggagc acaagaagaa 1260

gaaggacgct ggaaagtcgg ccaagaaaga caaagacca gtgaacaaat ccgggggcaa 1320

ggccaaaaag aaggtagaaa taagacctct ctgaaagaga ctagggttaa ctctctcgta 1380

atcctctagt aataggtaac ttgtatagta agtggttttt cagggttaga tttctagagt 1440

caaatgtga gattttatct tccgctacc actcgttctt tttccatta ggatcatgaa 1500

|                                                                    |      |
|--------------------------------------------------------------------|------|
| aatgggtctg ttgtgcgaag tgtctgccgc tgtgcctgct gtgttatttt taactgatct  | 1560 |
|                                                                    |      |
| agtggggctc ggccctctgt tgaaggccaa aaacgtgtcg gtgttttttt ttgttttttg  | 1620 |
| tttttagtaat gtgtaattta tccttgataa cggtggaaca gatttctctg acgcagatta | 1680 |
| ctcgagaggg aaaggtgtct tctgccagaa atactaactt gtttctgttt tgttttggtg  | 1740 |
| agcagaagtg gtccaaaggc aaagtctggg acaagctcaa taacttagtc ttgtttgaca  | 1800 |
| aagctaccta tgataaactc tgtaaggaag ttcccaacta taaacttata accccagctg  | 1860 |
| tggctctctga gagactgaag attcagggtc ccttgccag gccagccctt caggagctcc  | 1920 |
| ttagtaaagg tgaggggtgt atcctacatg tgtgtttttg taggttaaat tgtcttgacc  | 1980 |
|                                                                    |      |
| atgttaagca tcttcagtgg ttttctgga aaagcagaat taaaaaaaa aagcgtggct    | 2040 |
| tgaccattgg ctgttagtaa tgtaattctg acgtcttact cctgatcctg agatgaattc  | 2100 |
| tcagggttct tagccacttt tgtgccgtgg accctgtggc agtttagtga agcccaagga  | 2160 |
| tcttttatgt ttcgagtaaa tggatgcata gaattacagg gacaaccgtt ttgaaataa   | 2220 |
| ttagattact attttgaac aactttgaaa atgtttaaaa ctttatggt aaatatattg    | 2280 |
| ttgatgtatt aaattttaa accagaaatt tagtacggtc tactcagtag tatggctctga  | 2340 |
| ttaccataat tccacaataa taaggctcag ctaactatag tgactgaacg tctataattc  | 2400 |
|                                                                    |      |
| tagcactttg ggaggccaag gcgggtgaat caacggaggt caggagttaa agaccagcct  | 2460 |
| ggccaatatg gtgaaaacct gctctactga aagttagctg gacgtggggg cacacgtctg  | 2520 |
| taatcccagc tactcaggat gctgaggcat gaggatccct tgaaccagg agatggaggt   | 2580 |
| ggcagtgagc cgagatgaca ccaactgact ccagccttag tgacagcaaa agactgtctc  | 2640 |
| agaaaggggg ggggggtgga agataatgga gcctaattt aaaggaaaag taaggataga   | 2700 |
| tgatccgtta aaaacttgga ttctcggtta ccgaacgtca gattaagcaa ttctggagcc  | 2760 |
| aggtgcagtg gtacccttgt atttctagct acttgggagg ccaaagcagg aggatcattt  | 2820 |
|                                                                    |      |
| gagccaagga gtittaagac cattctgggc acctctgaga gaactctgtc tttttgtttt  | 2880 |
| cctttttttt aaatagagat gcggttttgc catgttgccc aggttggtct cctgggctca  | 2940 |
| agagatccac ctgtccaaag tgctgggatt acaggcatga gcctctgcac cggccaaaa   | 3000 |
| caaaccttac tagagtctca ttctgttgcc caggttggag tgcggagggg cagtcttgcc  | 3060 |
| tcaatgcaac caccaattcc tgggttcagg tggtctctac ctcagcttcc caagtagctg  | 3120 |
| gaattacaag catgtgccac catgcccagc taatttttgt atttttggtg gagatggggt  | 3180 |
| ttcaccttgt tggccaggtc ggtgtgcaac tccttacctc aagctatctg cccgtctcca  | 3240 |

cctcccaaag cagtgggatt ataagcatga gccaccgcgc ccagccaaaa accttactag 3300

tttctattgt agcatctgtt aagcatctca tctgtctatt ctctccccct aggacttate 3360

aaactggttt caaagcacag agctcaagta atttacacca gaaataccaa ggggtggagat 3420

gtccagctg ctgggaaga tgcatgaata ggtgagtagg aatgtgtggg ctcattggtg 3480

aggaggtaga tacaagctt tatggttctg attcttttaa tttttttta caggccaac 3540

cagctgtaca ttggaaaaa taaaacttta ttaaatcaaa tgaatgagta tgtctgttc 3600

ctaagaaaga caatgataaa gaatttgggtg gaaggtataa taggggtttg ttgactttgc 3660

ttttagcctc atggtagtgt gtagagagca tgattagctt ttttctgtat gtgactgctt 3720

cttcattgct gcagcttcag ttttgaattg atgtctgaaa ggaaataaag ggtaaacacg 3780

atgatgaagg gtg 3793

<210> 61

<211> 3996

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Polynucleotide

<400> 61

ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agaggagtg 120

gccaaactca tctactagggg ttctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag ctggatctg aattcggtac 300

cctagttatt aatagtaatc aattacgggg tcattagtgc atagccata tatggagtgc 360

cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga cccccccca 420

ttgacgtcaa taatgacgta tgttccata gtaacgcaa tagggacttt ccattgacgt 480

caatgggtgg actatttacg gtaaacgcc cacttggcag tacatcaagt gtatcatatg 540

ccaagtlacg cccctattga cgtcaatgac ggtaaattggc ccgcctggca ttatgccag 600

tacatgacct tatgggactt tctacttgg cagtacatct acgtattagt catcgctatt 660

accatggctg aggtgagccc cacgttctgc ttactctcc ccatctcccc cccctccca 720

cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg 780

gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg 840

agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg 900

cggcgggcgc ggcggcccta taaaagcga agcgcgcggc gggcgggagt cgctgcgacg 960

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cgcgcgcccc ggctctgact  | 1020 |
|                                                                    |      |
| gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctcctccgg gctgtaatta  | 1080 |
| gcgcttggtt taatgacggc ttgtctggag gcttgctttg ggctgtatgc tgataccttc  | 1140 |
| caccaaattc tttattttgg cctctgactg ataaagaatt gtggaaggta tcaggacaca  | 1200 |
| aggcccttta tcagcactca catggaacaa atggccaccg tgggaggatg acaatttctg  | 1260 |
| tggtcgcgtg aaagccttga ggggctccgg gagctagagc ctctgctaac catgttcattg | 1320 |
| ccttcttctt tttcctacag ctccctgggca acgtgctggg tattgtgctg tctcatcatt | 1380 |
| ttggcaaaga attcctcgaa gatccgaagg gaaagtcttc cacgactgtg ggatccgttc  | 1440 |
|                                                                    |      |
| gaagatatca ccggttgagc caccatgagc accctgtgcc cccccccag ccccgccgtg   | 1500 |
| gccaaagaccg agatcgccct gagcggcaag agccccctgc tggccgccac cttcgcctac | 1560 |
| tgggacaaca tcttgggccc ccgctgtgcgc cacatctggg cccccaagac cgagcaggtg | 1620 |
| ctgctgagcg acggcgagat caccctcttg gccaccaca cctgaacgg cgagatcctg    | 1680 |
| cgcaacgccg agagcggcgc catcgacgtg aagtctctcg tgctgagcga gaaggcgtg   | 1740 |
| atcatcgtga gctgatctt cgacggcaac tggaacggcg accgcagcac ctacggcctg   | 1800 |
| agcatcatcc tgccccagac cgagctgagc ttctacctgc cctgcaccg cgtgtgcgtg   | 1860 |
|                                                                    |      |
| gaccgcctga cccacatcat ccgcaagggc cgcacttgga tgcacaagga gcgccaggag  | 1920 |
| aacgtgcaga agatcatcct ggagggcacc gagcgcattg aggaccaggg ccagagcatc  | 1980 |
| atccccatgc tgaccggcga ggtgatcccc gtgatggagc tgctgagcag catgaagagc  | 2040 |
| cacagcgtgc ccgaggagat cgacatcgcc gacaccgtgc tgaacgacga cgacatcggc  | 2100 |
| gacagctgcc acgagggtt cctgctgaac gccatcagca gccacctgca gacctgcggc   | 2160 |
| tgcagcgtgg tggtaggcag cagcggcag aaggtgaaca agatcgtgcg caccctgtgc   | 2220 |
| ctgttctga cccccgcga gcgcaagtgc agccgcctgt gcgaggcca gagcagcttc     | 2280 |
|                                                                    |      |
| aagtacgaga gcggcctgtt cgtgcagggc ctgctgaagg acagcaccgg cagcttcgtg  | 2340 |
| ctgcccttc gccaggtgat gtacgcccc taccaccca cccacatcga cgtggacgtg     | 2400 |
| aacaccgtga agcagatgcc cccctgccac gagcacatct acaaccagcg ccgctacatg  | 2460 |
| cgcagcgagc tgaccgcctt ctggcgcgcc accagcgagg aggacatggc ccaggacacc  | 2520 |
| atcatctaca ccgacgagag cttaccccc gacctgaaca tcttcagga cgtgctgcac    | 2580 |
| cgcgacacc tggtaggc cttcctggac caggtgttc agctgaagcc cggcctgagc      | 2640 |
| ctgcgcagca cttcctggc ccagttctg ctggtgctgc accgcaagc cctgaccctg     | 2700 |

atcaagtaca tcgaggacga caccacagaag ggcaagaagc ccttcaagag cctgcgcaac 2760  
ctgaagatcg acctggacct gaccgccgag ggcgacctga acatcatcat ggccctggcc 2820  
gagaagatca agcccggcct gcacagcttc atcttcggcc gcccttcta caccagcgtg 2880  
caggagcgcg acgtgctgat gaccttctaa caattgttaa ttaagtttaa accctcgagg 2940  
ccgcaagctt atcgataatc aacctctgga ttacaaaatt tgtgaaagat tgactggtat 3000  
tcttaactat gttgctcctt ttacgctatg tggatacgtt gctttaatgc ctttgtatca 3060  
tgctattgct tcccgtatgg ctttcatttt ctctccttg tataaatcct ggttgctgtc 3120

tctttatgag gagtgtggc ccgttgtcag gcaacgtggc gtggtgtgca ctgtgtttgc 3180  
tgacgcaacc cccactggtt ggggcattgc caccacctgt cagctccttt ccgggacttt 3240  
cgctttcccc ctccctattg ccacggcgga actcatcgcc gcctgccttg cccgctgctg 3300  
gacaggggct cgctgtttg gcactgacaa ttccgtggtg ttgtcgggga aatcatcgtc 3360  
ctttccttgg ctgctcgctt gtgttgccac ctggattctg cgcgggacgt ccttctgcta 3420  
cgtccttctg gccctcaatc cagcggacct tccttccgc ggctgctgc cggtctgctg 3480  
gcctcttccg cgtcttcgcc ttccgctca gacgagtcgg atctcccttt gggccgcctc 3540

cccgcacga taccgtcgac tagagctcgc tgatcacct cgactgtgcc ttctagtgc 3600  
cagccatctg ttgtttgccc ctccccgtg ccttcttga ccctggaagg tgccactccc 3660  
actgtccttt ctaataaaa tgaggaaatt gcatcgcat gtctgagtag gtgtcattct 3720  
attctggggg gtggggtggg gcaggacagc aagggggagg attgggaaga caatagcagg 3780  
catgctgggg agagatccac gataacaaac agcttttttg gggtgaacat attgactgaa 3840  
ttccctgcag gttggccact ccctctctgc gcgctcgtc gctcactgag gccgcccggg 3900  
caaagcccgg gcgtcgggag acctttggtc gcccggcctc agtgagcgag cgagcgcga 3960

gagagggagt ggccaactcc atcactaggg gttcct 3996

<210> 62  
<211> 3994  
<212> DNA  
<213> Artificial Sequence  
<220><223> Synthetic Polynucleotide  
<400> 62

ttggccactc cctctctgcg cgctcgtcgt ctcactgagg ccgggcgacc aaaggtcgcc 60  
cgacgcccg gctttgcccc ggccggcctca gtgagcgagc gagcgcgcag agagggagtg 120  
gccaaactca tcaactagggg ttccctgctag ctctgggtat ttaagcccga gtgagcacgc 180  
aggtcttcca tttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240



|                                                                   |      |
|-------------------------------------------------------------------|------|
| tccctagggt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtag | 300  |
|                                                                   |      |
| cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc | 360  |
| cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga ccccgccca  | 420  |
| ttgacgtcaa taatgacgta tgttccata gtaacgcaa tagggacttt ccattgacgt   | 480  |
| caatgggtgg actatttacg gtaaacgcc cacttggcag tacatcaagt gtatcatatg  | 540  |
| ccaagtacgc ccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgccag   | 600  |
| tacatgacct tatgggactt tctacttgg cagtacatct acgtattagt catcgctatt  | 660  |
| accatggtcg aggtgagccc cagttctgc ttactctcc ccctctccc cccctccca     | 720  |
|                                                                   |      |
| cccccaattt tgiatttatt tatttttaa ttatttttg cagcgatggg ggcggggggg   | 780  |
| gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg | 840  |
| agagggtcgg cgccagccaa tcagagcggc gcgctccgaa agtttcttt tatggcgagg  | 900  |
| cggcgcgggc ggcggcccta taaaagcga agcgcgcggc ggcggggagt cgctgcgacg  | 960  |
| ctgccttcgc ccgtgcccc gctccgccgc cgcctcgcgc ccccccccc ggccttgact   | 1020 |
| gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctctccgg gctgtaatta  | 1080 |
| gcgcttgggt taatgacggc ttgtctggag gcttgctttg ggctgtatgc tgataagtcc | 1140 |
|                                                                   |      |
| tttactaagg agcttttggc ctctgactga gctccttgta aggacttatc aggacacaag | 1200 |
| gccctttatc agcactcaca tggaacaaat ggccaccgtg ggaggatgac aatttctgtg | 1260 |
| gctgcgtgaa agccttgagg ggctccggga gctagagcct ctgctaacca tgttcatgcc | 1320 |
| ttcttctttt tctacagct cctgggcaac gtgctgttta ttgtctgtc tcatcatttt   | 1380 |
| ggcaaagaat tctcgaaga tccgaaggga aagtcttcca cgactgtggg atccgttcga  | 1440 |
| agatatcacc ggttgagcca ccatgagcac cctgtgcccc cccccagcc ccgccgtggc  | 1500 |
| caagaccgag atcgccctga gcggcaagag cccctgctg gccgccacct tcgcctactg  | 1560 |
|                                                                   |      |
| ggacaacatc ctggggcccc gcgtgcgcca catctgggcc cccaagaccg agcaggtgct | 1620 |
| gctgagcgac ggcgagatca ctttcttggc caaccacacc ctgaacggcg agatctgcg  | 1680 |
| caacgccgag agcggcgcca tcgacgtgaa gttcttcgtg ctgagcgaga agggcgtgat | 1740 |
| catcgtgagc ctgatcttcg acggcaactg gaacggcgac cgcagcacct acggcctgag | 1800 |
| catcatcctg cccagaccg agctgagctt ctacctgccc ctgcaccgcg tgtgcgtgga  | 1860 |
| ccgcctgacc cacatcatcc gcaagggccg catctggatg cacaaggagc gccaggagaa | 1920 |
| ctgtcagaag atcatcctgg agggcaccga gcgcatggag gaccagggcc agagcatcat | 1980 |

ccccatgctg accggcgagg tgatccccgt gatggagctg ctgagcagca tgaagagcca 2040  
 cagcgtgccc gaggagatcg acatcgccga caccgtgctg aacgacgacg acatcgccga 2100  
 cagctgccac gagggcttcc tgctgaacgc catcagcagc cacctgcaga cctgcggctg 2160  
 cagcgtggtg gtgggcagca gcgccgagaa ggtgaacaag atcgtgcgca ccctgtgcct 2220  
 gttctgacc cccgccgagc gcaagtgcag ccgctgtgc gaggccgaga gcagcttcaa 2280  
 gtacgagagc ggctgttccg tgcagggcct gctgaaggac agcaccggca gcttcgtgct 2340  
 gcccttccgc caggtgatgt acgccccctia cccaccacc cacatcgacg tggacgtgaa 2400  
  
 caccgtgaag cagatgcccc cctgccacga gcacatctac aaccagcgcc gctacatgcg 2460  
 cagcgagctg accgccttct ggccgcgccac cagcgaggag gacatggccc aggacacat 2520  
 catctacacc gacgagagct tcacccccga cctgaacatc ttccaggacg tgctgcaccg 2580  
 cgacaccctg gtgaaggcct tcttggaaca ggtgttccag ctgaagcccg gcctgagcct 2640  
 gcgcagcacc ttcttgcccc agttctctgct ggtgtgcac cgcaaggccc tgaccctgat 2700  
 caagtacatc gaggacgaca cccagaaggg caagaagccc ttcaagagcc tgcgcaacct 2760  
 gaagatcgac ctggacctga ccgccgaggg cgacctgaac atcatcatgg ccctggccga 2820  
  
 gaagatcaag cccggcctgc acagcttcat cttcgccgc ccttctaca ccagcgtgca 2880  
 ggagcgcgac gtgctgatga ctttctaaca attgttaatt aagtttaaac cctcgaggcc 2940  
 gcaagcttat cgataatcaa cctctggatt acaaaatttg tgaaagattg actggtattc 3000  
 ttaactatgt tgctcctttt acgctatgtg gatacgtgc tttaatgcct ttgtatcatg 3060  
 ctattgcttc ccgtatggct ttcatcttct cctccttgta taaatcctgg ttgctgtctc 3120  
 tttatgagga gttgtggccc gttgtcaggc aacgtggcgt ggtgtgact gtgtttgctg 3180  
 acgcaacccc cactggttgg ggcattgcca ccacctgtca gctcctttcc gggactttcg 3240  
  
 ctttccccct cctattgcc agggcggaac tcacgcgcg ctgccttgcc cgctgctgga 3300  
 caggggctcg gctgttgggc actgacaatt ccgtggtgtt gtcggggaaa tcacgtcct 3360  
 ttcttggct gctgcctgt gttgccacct ggattctgcg cgggacgtcc ttctgctacg 3420  
 tcccttcggc cctcaatcca gcggaccttc cttcccgcgg cctgctgccg gctctcggc 3480  
 ctcttcgcg tcttcgctt cgccctcaga cgagtcggat ctccctttgg gccgcctccc 3540  
 cgcacgata ccgtcgacta gagctcgctg atcagcctcg actgtgctt ctagttgcca 3600  
 gccatctgtt gtttgcctt ccccgtgcc ttcttgacc ctggaagggt ccaactccac 3660  
  
 tgtcctttcc taataaaatg aggaaattgc atcgattgt ctgagtaggt gtcattctat 3720  
 tctggggggt ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca 3780  
 tgctggggag agatccagca taacaaacag cttttttggg gtgaacatat tgactgaatt 3840

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ccctgcaggt tggccactcc ctctctgcgc gctcgctcgc tctactgaggc cgcccgggca | 3900 |
| aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag tgagcgagcg agcgcgcaga  | 3960 |
| gaggagtggt ccaactccat cactaggggt tcct                              | 3994 |