(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년09월02일
(11) 등록번호 10-2700963
(24) 등록일자 2024년08월27일

(51) 국제특허분류(Int. Cl.)
C12N 15/86 (2006.01) *A61K 48/00* (2006.01)
A61K 9/00 (2006.01) *C07K 14/435* (2006.01)
(52) CPC특허분류
C12N 15/86 (2013.01)
A61K 48/00 (2024.01)
(21) 출원번호 10-2022-7002135(분할)
(22) 출원일자(국제) 2018년10월03일
심사청구일자 2022년01월20일
(85) 번역문제출일자 2022년01월20일
(65) 공개번호 10-2022-0015499
(43) 공개일자 2022년02월08일
(62) 원출원 특허 10-2020-7012190
원출원일자(국제) 2018년10월03일
심사청구일자 2021년10월01일
(86) 국제출원번호 PCT/US2018/054227
(87) 국제공개번호 WO 2019/070894
국제공개일자 2019년04월11일
(30) 우선권주장
62/567,311 2017년10월03일 미국(US)
(뒷면에 계속)
(56) 선행기술조사문헌
JP2003506662 A*
(뒷면에 계속)

(73) 특허권자
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

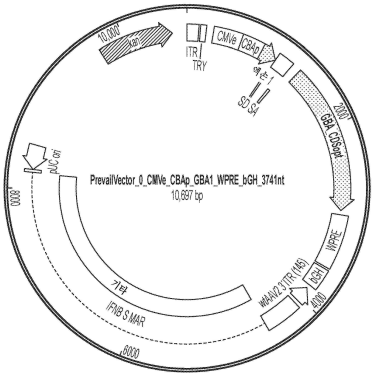
전체 청구항 수 : 총 45 항 심사관 : 손영희

(54) 발명의 명칭 **리소좀 장애를 위한 유전자 요법**

(57) 요 약

일부 측면에서, 본 개시내용은 비정상적인 리소좀 기능과 연관된 질환, 예를 들어, 파킨슨병 (PD) 및 고셰병 치료를 위한 조성물 및 방법에 관한 것이다. 일부 실시양태에서, 본 개시내용은 베타-글루코세레브로시다제 (GBA) 또는 그의 일부를 코딩하는 트랜스진을 단독으로, 또는 하나 이상의 PD 연관 유전자와 함께 조합하여 포함하는 발현 구축물을 제공한다. 일부 실시양태에서, 본 개시내용은 파킨슨병 방법을 필요로 하는 대상체에게 상기 발현 구축물을 투여함으로써 파킨슨병 방법을 제공한다.

대 표 도 - 도1



(52) CPC특허분류
A61K 9/0019 (2013.01)
C07K 14/435 (2013.01)
C12N 2710/14043 (2013.01)
C12N 2750/14143 (2013.01)
C12N 2800/22 (2013.01)

(56) 선행기술조사문헌
JP61082035 A*
US05490559 A*
US06691742 B1*
US20090178279 A1*
*는 심사관에 의하여 인용된 문헌

(30) 우선권주장
62/567,310 2017년10월03일 미국(US)
62/567,301 2017년10월03일 미국(US)
62/567,319 2017년10월03일 미국(US)
62/567,296 2017년10월03일 미국(US)

명 세 서

청구범위

청구항 1

(i) TREM2 단백질을 코딩하는 트랜스진을 포함하는 발현 구축물이며, 여기서 TREM2 단백질은 서열식별번호 58의 핵산 서열에 의해 코딩된 것인 발현 구축물; 및

(ii) 상기 발현 구축물을 플랭킹하는 2개의 아데노-연관 바이러스 역전된 말단 반복부 (ITR) 서열

을 포함하는 단리된 핵산.

청구항 2

제1항에 있어서, 트랜스진이 프로모터에 작동가능하게 연결된 것인 단리된 핵산.

청구항 3

제2항에 있어서, 프로모터가 닭 베타 액틴 (CBA) 프로모터인 단리된 핵산.

청구항 4

제2항에 있어서, CMV 인핸서를 추가로 포함하는 단리된 핵산.

청구항 5

제1항에 있어서, 우드척 간염 바이러스 전사후 조절 요소 (WPRE)를 추가로 포함하는 단리된 핵산.

청구항 6

제1항에 있어서, 소 성장 호르몬 폴리A 신호 테일을 추가로 포함하는 단리된 핵산.

청구항 7

제1항에 있어서, 각각의 ITR 서열이 야생형 AAV2 ITR 서열인 단리된 핵산.

청구항 8

제1항에 있어서, 각각의 ITR 서열이 트랜스진에 인접한 "D" 영역을 포함하고, 여기서 "D" 영역이 서열식별번호 27에 기재된 서열을 포함하는 것인 단리된 핵산.

청구항 9

제1항에 있어서, ITR 서열 중 하나 이상이 트랜스진 기준으로 ITR 서열의 외부에 위치한 "D" 영역을 포함하고, 여기서 "D" 영역이 서열식별번호 27에 기재된 서열을 포함하는 것인 단리된 핵산.

청구항 10

제1항에 있어서, 트랜스진 기준으로 5'에 위치한 ITR 서열이 트랜스진에 인접한 "D" 영역을 포함하고, 트랜스진 기준으로 3'에 위치한 ITR 서열이 트랜스진 기준으로 ITR 서열의 외부에 위치한 "D" 영역을 포함하며, 여기서 "D" 영역이 서열식별번호 27에 기재된 서열을 포함하는 것인 단리된 핵산.

청구항 11

제10항에 있어서, 5' ITR의 핵산 서열이 서열식별번호 1의 뉴클레오티드 1 내지 145이고, 3' ITR의 핵산 서열이 서열식별번호 1의 뉴클레오티드 3867 내지 4011인 단리된 핵산.

청구항 12

제11항에 있어서, 5' ITR과 트랜스진 사이에 TRY 영역을 추가로 포함하며, 여기서 TRY 영역은 서열식별번호 28

에 기재된 서열을 갖는 것인 단리된 핵산.

청구항 13

2개의 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR)에 의해 플랭킹된, TREM2 단백질을 코딩하는 트랜스진을 포함하는 발현 구축물을 포함하는 핵산을 포함하는 재조합 아데노-연관 바이러스 (rAAV) 벡터이며, 여기서 TREM2 단백질은 서열식별번호 58의 핵산 서열에 의해 코딩된 것인 rAAV 벡터.

청구항 14

제13항에 있어서, 트랜스진이 프로모터에 작동가능하게 연결된 것인 rAAV 벡터.

청구항 15

제14항에 있어서, 프로모터가 닭 베타 액틴 (CBA) 프로모터인 rAAV 벡터.

청구항 16

제14항에 있어서, CMV 인핸서를 추가로 포함하는 rAAV 벡터.

청구항 17

제13항에 있어서, 우드척 간염 바이러스 전사후 조절 요소 (WPRE)를 추가로 포함하는 rAAV 벡터.

청구항 18

제13항에 있어서, 소 성장 호르몬 폴리A 신호 테일을 추가로 포함하는 rAAV 벡터.

청구항 19

제13항에 있어서, 각각의 ITR 서열이 야생형 AAV2 ITR 서열인 rAAV 벡터.

청구항 20

제13항에 있어서, 각각의 ITR 서열이 트랜스진에 인접한 "D" 영역을 포함하고, 여기서 "D" 영역이 서열식별번호 27에 기재된 서열을 포함하는 것인 rAAV 벡터.

청구항 21

제13항에 있어서, ITR 서열 중 하나 이상이 트랜스진 기준으로 ITR 서열의 외부에 위치한 "D" 영역을 포함하고, 여기서 "D" 영역이 서열식별번호 27에 기재된 서열을 포함하는 것인 rAAV 벡터.

청구항 22

제13항에 있어서, 트랜스진 기준으로 5'에 위치한 ITR 서열이 트랜스진에 인접한 "D" 영역을 포함하고, 트랜스진 기준으로 3'에 위치한 ITR 서열이 트랜스진 기준으로 ITR 서열의 외부에 위치한 "D" 영역을 포함하며, 여기서 "D" 영역이 서열식별번호 27에 기재된 서열을 포함하는 것인 rAAV 벡터.

청구항 23

제22항에 있어서, 5' ITR의 핵산 서열이 서열식별번호 1의 뉴클레오티드 1 내지 145이고, 3' ITR의 핵산 서열이 서열식별번호 1의 뉴클레오티드 3867 내지 4011인 rAAV 벡터.

청구항 24

제23항에 있어서, 5' ITR과 트랜스진 사이에 TRY 영역을 추가로 포함하며, 여기서 TRY 영역은 서열식별번호 28에 기재된 서열을 갖는 것인 rAAV 벡터.

청구항 25

(i) AAV 캡시드 단백질; 및

(ii) 제13항의 rAAV 벡터

를 포함하는 재조합 아데노-연관 바이러스 (rAAV).

청구항 26

제25항에 있어서, AAV 캡시드 단백질이 AAV9 캡시드 단백질인 rAAV.

청구항 27

5'에서 3' 순서로:

(a) 5' AAV ITR;

(b) CMV 인핸서;

(c) CBA 프로모터;

(d) 서열식별번호 58의 핵산 서열에 의해 코딩된, TREM2 단백질을 코딩하는 트랜스진;

(e) WPRE;

(f) 소 성장 호르몬 폴리A 신호 테일; 및

(g) 3' AAV ITR

을 포함하는 핵산을 포함하는 재조합 아데노-연관 바이러스 (AAV) 벡터.

청구항 28

(i) AAV 캡시드 단백질; 및

(ii) 제27항의 rAAV 벡터

를 포함하는 재조합 아데노-연관 바이러스 (rAAV).

청구항 29

제28항에 있어서, AAV 캡시드 단백질이 AAV9 캡시드 단백질인 rAAV.

청구항 30

제13항의 rAAV 벡터를 포함하는 플라스미드.

청구항 31

서열식별번호 58에 기재된 핵산 서열을 포함하는 배큘로바이러스 벡터.

청구항 32

(i) 하나 이상의 아데노-연관 바이러스 rep 단백질 및/또는 하나 이상의 아데노-연관 바이러스 cap 단백질을 코딩하는 제1 벡터; 및

(ii) 서열식별번호 58에 기재된 핵산 서열을 포함하는 제2 벡터

를 포함하는 단리된 세포.

청구항 33

제32항에 있어서, 제1 벡터가 플라스미드이고 제2 벡터가 플라스미드인 단리된 세포.

청구항 34

제33항에 있어서, 단리된 세포가 포유동물 세포인 단리된 세포.

청구항 35

제34항에 있어서, 포유동물 세포가 HEK293 세포인 단리된 세포.

청구항 36

제32항에 있어서, 제1 벡터가 배큘로바이러스 벡터이고 제2 벡터가 배큘로바이러스 벡터인 단리된 세포.

청구항 37

제32항에 있어서, 단리된 세포가 곤충 세포인 단리된 세포.

청구항 38

제37항에 있어서, 곤충 세포가 SF9 세포인 단리된 세포

청구항 39

(i) 하나 이상의 아데노-연관 바이러스 rep 단백질 및/또는 하나 이상의 아데노-연관 바이러스 cap 단백질을 코딩하는 제1 벡터, 및 서열식별번호 58의 뉴클레오티드 서열을 포함하는 발현 카세트를 포함하는 재조합 AAV 벡터를 인간 체내 세포를 제외한 세포에 전달하는 단계;

(ii) rAAV를 패키징하도록 허용하는 조건 하에 세포를 배양하는 단계; 및

(iii) rAAV의 수집을 위해 배양된 세포 또는 배양 배지를 수확하는 단계

를 포함하는, 제25항 내지 제29항 중 어느 한 항의 rAAV를 생산하는 방법.

청구항 40

알츠하이머병을 앓거나 또는 앓을 것으로 의심되는 대상체를 치료하기 위한 제약 조성물이며, 제1항 내지 제12항 중 어느 한 항의 단리된 핵산, 제13항 내지 제24항 중 어느 한 항의 rAAV 벡터, 또는 제25항 내지 제29항 중 어느 한 항의 rAAV를 포함하고, 대상체에게 투여되는 제약 조성물.

청구항 41

제40항에 있어서, 투여가 대상체의 CNS에의 직접 주사를 포함하는 것인 제약 조성물.

청구항 42

제41항에 있어서, 직접 주사가 대뇌내 주사, 실질내 주사, 경막내 주사, 대수조내 (ICM) 주사, 또는 그의 임의의 조합인 제약 조성물.

청구항 43

제41항에 있어서, 대상체의 CNS에의 직접 주사가 컨벡션 강화 전달법 (CED)을 포함하는 것인 제약 조성물.

청구항 44

제40항에 있어서, 투여가 말초 주사를 포함하는 것인 제약 조성물.

청구항 45

제44항에 있어서, 말초 주사가 정맥내 주사인 제약 조성물.

발명의 설명

기 술 분 야

[0001] 관련 출원

[0002] 본 출원은 35 U.S.C. 119(e)하에 2017년 10월 3일 출원된 미국 가출원 일련 번호 62/567,296 (발명의 명칭: "GENE THERAPIES FOR LYSOSOMAL DISORDERS"), 2017년 10월 3일 출원된 미국 가출원 일련 번호 62/567,311 (발명의 명칭: "GENE THERAPIES FOR LYSOSOMAL DISORDERS"), 2017년 10월 3일 출원된 미국 가출원 일련 번호 62/567,319 (발명의 명칭: "GENE THERAPIES FOR LYSOSOMAL DISORDERS"), 2018년 10월 3일 출원된 미국 가출원

일련 번호 62/567,301 (발명의 명칭: "GENE THERAPIES FOR LYSOSOMAL DISORDERS"), 및 2017년 10월 3일 출원된 미국 가출원 일련 번호 62/567,310 (발명의 명칭: "GENE THERAPIES FOR LYSOSOMAL DISORDERS")의 출원일의 이익을 주장하고, 각 출원의 전체 내용이 본원에서 참조로 포함된다.

배 경 기 술

[0003] 고셰병(Gaucher disease)은 리소좀 산 β-글루코세레브로시다제 (Gcase, "GBA")의 결핍에 기인한, 선천성 글리코스핑고리피드 대사 이상 희귀 질환이다. 환자는 간비장비대, 범혈구감소증으로 이어지는 골수 기능 부전, 폐 장애 및 섬유증, 골 결손을 비롯한, 비-CNS 증상 및 소견을 앓는다. 추가로, 상당수의 환자들은 결함성 도약 안구 운동 및 시선, 발작, 인지 결손, 발달 지연, 및 파킨슨병을 비롯한 운동 장애를 포함하는, 신경학적 증상을 앓는다.

[0004] 말초 질환 및 조혈 골수 및 내장에서의 주요 임상 증상을 처리하는 여러 치료제들이 존재하며, 이는 하기 기술되는 바와 같은 효소 대체 요법, 결함성 Gcase에 결합하고, 안정성을 개선시키는 샤페론 유사 소분자 약물, 및 증상 및 소견으로 이어지는, 고셰병에서 축적되는 기질의 생산을 차단하는 기질 감소 요법을 포함한다. 그러나, 고셰병의 다른 측면 (특히, 골격 및 뇌에서 발병되는 것)은 치료에 대하여 난치성인 것으로 보인다.

발명의 내용

[0005] (*GBA1* 유전자의 두 염색체 대립유전자 모두에 돌연변이를 갖는) 고셰병 환자 이외에도, GBA1의 단 하나의 대립유전자에만 돌연변이를 갖는 환자는 파킨슨병 (PD)에 걸릴 위험이 매우 높다. 보행 장애, 휴식시 떨림, 경직, 및 흔히 우울증, 수면 장애, 및 인지 저하를 포함하는, PD 증상의 중증도는 효소 활성 감소 정도와 상관관계를 갖는다. 따라서, 고셰병 환자는 가장 중증의 과정을 갖는 반면, *GBA1* 중 단일의 가벼운 돌연변이를 갖는 환자는 전형적으로 더욱 양성의 과정을 갖는다. 돌연변이 보유자는 또한, 실행 기능 장애, 정신병, 및 PD 유사 운동 장애를 특징으로 하는 루이소체 치매, 및 특유의 운동 및 인지 장애를 갖는 다계통 위축을 비롯한, 다른 PD 관련 장애의 위험이 높다. 이러한 장애의 거침없이 계속되는 과정을 변경하는 요법은 현재 없다.

[0006] 예컨대, Gcase (예컨대, *GBA1* 유전자의 유전자 생성물)와 같은 효소 결핍 뿐만 아니라, 리소좀 기능 또는 리소좀으로의 거대분자 수송에 관여하는 다수의 유전자 (예컨대, 리소좀 막 단백질 1 (LIMP), 이는 또한 SCARB2로도 지칭)에서의 일반적인 변이체가 PD 위험 증가와 연관이 있다. 본 개시내용은 부분적으로 하나 이상의 PD 연관 유전자, 예를 들어, Gcase, GBA2, 프로사포신, 프로그래뉼린, LIMP2, GALC, CTSB, SMPD1, GCH1, RAB7, VPS35, IL-34, TREM2, TMEM106B, 또는 상기 중 임의의 것 (또는 그의 일부)의 조합을 코딩하는 발현 구축물 (예컨대, 벡터)에 기초한다. 일부 실시양태에서, 본원에 기술된 유전자 생성물의 조합은 대상체에서 발현되었을 때, 함께 (예컨대, 시너지적으로) 작용하여 PD의 하나 이상의 징후 및 증상을 감소시킨다.

[0007] 따라서, 일부 측면에서, 본 개시내용은 Gcase (예컨대, *GBA1* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) Gcase 코딩 서열을 포함한다. 일부 실시양태에서, Gcase를 코딩하는 핵산 서열은 서열식별번호(SEQ ID NO:) 14에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_000148.2에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 15에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, Gcase 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0008] 일부 측면에서, 본 개시내용은 프로사포신 (예컨대, *PSAP* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) 프로사포신 코딩 서열을 포함한다. 일부 실시양태에서, 프로사포신을 코딩하는 핵산 서열은 서열식별번호 16에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_002769.1에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 17에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, 프로사포신 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0009] 일부 측면에서, 본 개시내용은 LIMP2/SCARB2 (예컨대, *SCARB2* 유전자의 유전자 생성물)를 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세

포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) SCARB2 코딩 서열을 포함한다. 일부 실시양태에서, LIMP2/SCARB2를 코딩하는 핵산 서열은 서열식별번호 18에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_005497.1에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 29에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, SCARB2 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0010] 일부 측면에서, 본 개시내용은 GBA2 단백질 (예컨대, *GBA2* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) GBA2 코딩 서열을 포함한다. 일부 실시양태에서, GBA2를 코딩하는 핵산 서열은 서열식별번호 30에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_065995.1에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 31에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, GBA2 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0011] 일부 측면에서, 본 개시내용은 GALC 단백질 (예컨대, *GALC* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) GALC 코딩 서열을 포함한다. 일부 실시양태에서, GALC를 코딩하는 핵산 서열은 서열식별번호 33에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_000144.2에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 34에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, GALC 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0012] 일부 측면에서, 본 개시내용은 CTSB 단백질 (예컨대, *CTSB* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) CTSB 코딩 서열을 포함한다. 일부 실시양태에서, CTSB를 코딩하는 핵산 서열은 서열식별번호 35에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_001899.1에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 36에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, CTSB 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0013] 일부 측면에서, 본 개시내용은 SMPD1 단백질 (예컨대, *SMPD1* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) *SMPD1* 코딩 서열을 포함한다. 일부 실시양태에서, SMPD1을 코딩하는 핵산 서열은 서열식별번호 37에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_000534.3에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 38에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, SMPD1 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0014] 일부 측면에서, 본 개시내용은 GCH1 단백질 (예컨대, *GCH1* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) GCH1 코딩 서열을 포함한다. 일부 실시양태에서, GCH1을 코딩하는 핵산 서열은 서열식별번호 45에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_000534.3에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 46에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, GCH1 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0015] 일부 측면에서, 본 개시내용은 RAB7L 단백질 (예컨대, *RAB7L* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) RAB7L 코딩 서열을 포함한다. 일부 실시양태에서, RAB7L을 코딩하는 핵산 서열은 서열식별번호 47에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_003920.1에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 48에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관

바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, RAB7L 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0016] 일부 측면에서, 본 개시내용은 VPS35 단백질 (예컨대, *VPS35* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) VPS35 코딩 서열을 포함한다. 일부 실시양태에서, VPS35를 코딩하는 핵산 서열은 서열식별번호 49에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_060676.2에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 50에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, VPS35 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0017] 일부 측면에서, 본 개시내용은 IL-34 단백질 (예컨대, *IL34* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) IL-34 코딩 서열을 포함한다. 일부 실시양태에서, IL-34를 코딩하는 핵산 서열은 서열식별번호 55에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_689669.2에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 56에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, IL_34 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0018] 일부 측면에서, 본 개시내용은 TREM2 단백질 (예컨대, *TREM* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) TREM2 코딩 서열을 포함한다. 일부 실시양태에서, TREM2를 코딩하는 핵산 서열은 서열식별번호 57에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_061838.1에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 58에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, TREM2 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0019] 일부 측면에서, 본 개시내용은 TMEM106B 단백질 (예컨대, *TMEM106B* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) TMEM106B 코딩 서열을 포함한다. 일부 실시양태에서, TMEM106B를 코딩하는 핵산 서열은 서열식별번호 63에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_060844.2에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 64에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, TMEM106B 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0020] 일부 측면에서, 본 개시내용은 프로그래뉼린 (예컨대, *PGRN* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) 프로사포신 코딩 서열을 포함한다. 일부 실시양태에서, 프로그래뉼린 (PRGN)을 코딩하는 핵산 서열은 서열식별번호 67에 기재된 바와 같은 (예컨대, NCBI 참조 서열 NP_002078.1에 기재된 바와 같은) 아미노산 서열을 포함하는 단백질을 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 68에 기재된 바와 같은 서열을 포함한다. 일부 실시양태에서, 발현 구축물은 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR), 예를 들어, 프로사포신 단백질을 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.

[0021] 일부 측면에서, 본 개시내용은 제1 유전자 생성물 및 제2 유전자 생성물을 코딩하는 발현 구축물을 포함하는 단리된 핵산이며, 여기서 각 유전자 생성물은 독립적으로 표 1에 기재된 유전자 생성물 또는 그의 일부로부터 선택되는 것인, 단리된 핵산을 제공한다.

[0022] 일부 실시양태에서, 제1 유전자 생성물 또는 제2 유전자 생성물은 Gcase 단백질, 또는 그의 일부이다. 일부 실시양태에서, 제1 유전자 생성물은 Gcase 단백질이고, 제2 유전자 생성물은 GBA2, 프로사포신, 프로그래뉼린, LIMP2, GALC, CTSB, SMPD1, GCH1, RAB7, VPS35, IL-34, TREM2, 및 TMEM106B로부터 선택된다.

[0023] 일부 실시양태에서, 발현 구축물은 간섭 핵산 (예컨대, shRNA, miRNA, dsRNA 등)을 추가로 코딩한다. 일부 실

시양태에서, 간섭 핵산은 α-시누클레인 (α-시누클레인)의 발현을 억제시킨다. 일부 실시양태에서, α-시누클레인을 표적화하는 간섭 핵산은 서열식별번호 20-25 중 어느 하나에 기재된 서열을 포함한다. 일부 실시양태에서, α-시누클레인을 표적화하는 간섭 핵산은 서열식별번호 20-25 중 어느 하나에 기재된 서열에 결합한다 (예컨대, 그와 하이브리드화한다).

[0024] 일부 실시양태에서, 간섭 핵산은 TMEM106B의 발현을 억제시킨다. 일부 실시양태에서, TMEM106B를 표적화하는 간섭 핵산은 서열식별번호 64 또는 65에 기재된 서열을 포함한다. 일부 실시양태에서, TMEM106B를 표적화하는 간섭 핵산은 서열식별번호 64 또는 65에 기재된 서열에 결합한다 (예컨대, 그와 하이브리드화한다).

[0025] 일부 실시양태에서, 발현 구축물은 하나 이상의 프로모터를 추가로 포함한다. 일부 실시양태에서, 프로모터는 닭 베타 액틴 (CBA) 프로모터, CAG 프로모터, CD68 프로모터, 또는 JeT 프로모터이다. 일부 실시양태에서, 프로모터는 RNA pol II 프로모터 (예컨대, 또는 RNA pol III 프로모터 (예컨대, U6 등)이다.

[0026] 일부 실시양태에서, 발현 구축물은 내부 리보솜 진입 부위 (IRES)를 추가로 포함한다. 일부 실시양태에서, IRES는 제1 유전자 생성물과 제2 유전자 생성물 사이에 위치한다.

[0027] 일부 실시양태에서, 발현 구축물은 자기 절단 펩티드 코딩 서열을 추가로 포함한다. 일부 실시양태에서, 자기 절단 펩티드는 T2A 펩티드이다.

[0028] 일부 실시양태에서, 발현 구축물은 2개의 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR) 서열을 포함한다. 일부 실시양태에서, ITR 서열은 제1 유전자 생성물 및 제2 유전자 생성물에 플랭킹된다 (예컨대, 5' 단부에서 3' 단부 방향으로 하기와 같이 배열되어 있다: ITR-제1 유전자 생성물-제2 유전자 생성물-ITR). 일부 실시양태에서, 단리된 핵산의 ITR 서열 중 하나에는 기능성 말단 분해 부위 (trs)가 결손되어 있다. 예를 들어, 일부 실시양태에서, ITR 중 하나는 ΔITR이다.

[0029] 일부 측면에서, 본 개시내용은 변형된 "D" 영역 (예컨대, 야생형 AAV2 ITR 기준으로 변형된 D 서열, 서열식별번호 29)을 갖는 ITR을 포함하는 rAAV 벡터에 관한 것이다. 일부 실시양태에서, 변형된 D 영역을 갖는 ITR은 rAAV 벡터의 5' ITR이다. 일부 실시양태에서, 변형된 "D" 영역은 예를 들어, 서열식별번호 26에 기재되어 있는 바와 같은 "S" 서열을 포함한다. 일부 실시양태에서, 변형된 "D" 영역을 갖는 ITR은 rAAV 벡터의 3' ITR이다. 일부 실시양태에서, 일부 실시양태에서, 변형된 "D" 영역은 "D" 영역이 ITR의 3' 단부에 (예컨대, 벡터의 트랜스진 인서트 기준으로 ITR의 외부 또는 말단 단부 상에) 위치하는 3' ITR을 포함한다. 일부 실시양태에서, 변형된 "D" 영역은 서열식별번호 26 또는 27에 기재되어 있는 바와 같은 서열을 포함한다.

[0030] 일부 실시양태에서, 단리된 핵산 (예컨대, rAAV 벡터)는 TRY 영역을 포함한다. 일부 실시양태에서, TRY 영역은 서열식별번호 28에 기재되어 있는 바와 같은 서열을 포함한다.

[0031] 일부 실시양태에서, 본 개시내용에 의해 기술된 단리된 핵산은 서열식별번호 1-78 중 어느 하나에 기재된 서열을 포함하거나 또는 그로 이루어지거나 또는 상기 서열을 갖는 펩티드를 코딩한다.

[0032] 일부 측면에서, 본 개시내용은 본 개시내용에 의해 기술된 바와 같은 단리된 핵산을 포함하는 벡터를 제공한다. 일부 실시양태에서, 벡터는 플라스미드, 또는 바이러스 벡터이다. 일부 실시양태에서, 바이러스 벡터는 재조합 AAV (rAAV) 벡터또는 배큘로바이러스 벡터이다. 일부 실시양태에서, rAAV 벡터는 단일-가닥 (예컨대, 단일-가닥 DNA)이다.

[0033] 일부 측면에서, 본 개시내용은 본 개시내용에 의해 기술된 바와 같은 단리된 핵산, 또는 본 개시내용에 의해 기술된 바와 같은 벡터를 포함하는 숙주 세포를 제공한다.

[0034] 일부 측면에서, 본 개시내용은 본 개시내용에 의해 기술된 바와 같은 캡시드 단백질 및 단리된 핵산 또는 벡터를 포함하는 재조합 아데노-연관 바이러스 (rAAV)를 제공한다.

[0035] 일부 실시양태에서, 캡시드 단백질, 예를 들어, AAV9 캡시드 단백질 또는 AAVrh.10 캡시드 단백질은 혈액-뇌 장벽을 건널 수 있다. 일부 실시양태에서, rAAV는 중추 신경계 (CNS)의 뉴런 세포 및 비-뉴런 세포를 형질도입한다.

[0036] 일부 측면에서, 본 개시내용은 파킨슨병을 앓거나 또는 앓을 것으로 의심되는 대상체에게 본 개시내용에 의해 기술된 바와 같은 조성물 (예컨대, 단리된 핵산 또는 벡터 또는 rAAV를 포함하는 조성물)을 투여하는 단계를 포함하는, 파킨슨병을 앓거나 또는 앓을 것으로 의심되는 대상체를 치료하는 방법을 제공한다.

[0037] 일부 실시양태에서, 투여는 대상체의 CNS에의 직접 주사를 포함한다. 일부 실시양태에서, 직접 주사는 대뇌내

주사, 실질내 주사, 경막내 주사, 대수조내 주사, 또는 그의 임의의 조합이다. 일부 실시양태에서, 대상체의 CNS에의 직접 주사는 컨벡션 강화 전달법 (CED)을 포함한다.

[0038] 일부 실시양태에서, 투여는 말초 주사를 포함한다. 일부 실시양태에서, 말초 주사는 정맥내 주사이다.

도면의 간단한 설명

[0039] 도 1은 Gcase (예컨대, *GBA1* 또는 그의 일부)를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다.

도 2는 Gcase (예컨대, *GBA1* 또는 그의 일부) 및 LIMP2 (SCARB2) 또는 그의 일부를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 LIMP2의 코딩 서열은 내부 리보솜 진입 부위 (IRES)에 의해 이격되어 있다.

도 3은 Gcase (예컨대, *GBA1* 또는 그의 일부) 및 LIMP2 (SCARB2) 또는 그의 일부를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 LIMP2의 코딩 서열의 발현은 각각 별개의 프로모터에 의해 구동된다.

도 4는 Gcase (예컨대, *GBA1* 또는 그의 일부), LIMP2 (SCARB2) 또는 그의 일부, 및 α-Syn에 대한 간섭 RNA를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다.

도 5는 Gcase (예컨대, *GBA1* 또는 그의 일부), 프로사포신 (예컨대, *PSAP* 또는 그의 일부), 및 α-Syn에 대한 간섭 RNA를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다.

도 6은 Gcase (예컨대, *GBA1* 또는 그의 일부) 및 프로사포신 (예컨대, *PSAP* 또는 그의 일부)을 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 프로사포신의 코딩 서열은 내부 리보솜 진입 부위 (IRES)에 의해 이격되어 있다.

도 7은 Gcase (예컨대, *GBA1* 또는 그의 일부)를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. 본 실시양태에서, 벡터는 인간 *GBA1*의 코돈 최적화된 코딩 서열을 구성적으로 발현하기 위해, 4개 파트: CMV 인핸서 (CMVe), CBA 프로모터 (CBAp), 엑손 1, 및 인트론 (int)으로 이루어진, CBA 프로모터 요소 (CBA)를 포함한다. 3' 영역은 또한 WPRE 조절 요소, 이어서, bGH 폴리A 테일을 함유한다. 3개의 전사 조절 활성화 부위가 프로모터 영역의 5' 단부에 포함되어 있다: TATA, RBS, 및 YY1. 플랭킹 ITR을 통해 개재 서열은 정확하게 패키징될 수 있다. 5' ITR 서열의 두 변이체 (삽입 박스)를 평가하였다: 상기 변이체들은 야생형 AAV2 ITR의 20-뉴클레오티드 "D" 영역 내에서 수개의 뉴클레오티드 차이를 보인다. 일부 실시양태에서, rAAV 벡터는 상단 라인에 제시된 "D" 도메인 뉴클레오티드 서열을 함유한다. 일부 실시양태에서, rAAV 벡터 돌연변이체 "D" 도메인 (예컨대, "S" 도메인, 여기서 뉴클레오티드 변이는 하단 라인에 제시)을 보유한다.

도 8은 도 6에 기술된 벡터의 한 실시양태를 도시한 개략도이다.

도 9는 파킨슨병을 앓는 CBE 마우스 모델에서의 Gcase (예컨대, *GBA1* 또는 그의 일부)를 코딩하는 트랜스진을 포함하는 rAAV의 전달에 대한 대표적인 데이터를 보여주는 것이다. P8에 PBS 비히클, 25 mg/kg CBE, 37.5 mg/kg CBE, 또는 50 mg/kg CBE (좌측부터 우측으로)의 매일 IP 전달을 개시하였다. 매일 2회에 걸쳐 생존 (좌측 상단)을 체크하였고, 체중 (우측 상단)은 매일 체크하였다. 모든 군은 n = 8으로 시작하였다. P23에 오픈 필드에서의 총 이동 거리 (좌측 하단) 및 P24에 로타로드에서 떨어질 때까지의 시간 (하단의 중간)에 의해 행동을 사정하였다. CBE를 중단한 조건 (3일째) 및 CBE를 중단하지 않은 조건 (1일째), 둘 모두의 PBS 및 25 mg/kg CBE 처리군 중의 마우스의 피질에서 GCase 기질의 수준을 분석하였다. 응집체 GluSph 및 GalSph 수준 (우측 하단)은 조직 습식 중량 1 mg당 pmol로 제시되어 있다. 평균이 제시되어 있다. 오차 막대는 SEM이다. *p<0.05; **p<0.01; ***p<0.001, 선형 회귀에 의한 처리군에 대한 명목 p 값.

도 10은 CBE 마우스 모델에서의 최대 rAAV 용량에 대한 연구 디자인의 한 실시양태를 도시한 개략도이다. 간략하면, rAAV를 P3에 ICV 주사에 의해 전달하고, P8에 매일 CBE 처리를 개시하였다. P24-25에 오픈 필드 및 로타로드 검정법으로 행동을 사정하고, P36 및 P38에 기질 수준을 측정하였다.

도 11은 CBE 마우스 모델에서의 최대 rAAV 용량의 생전 사정에 대한 대표적인 데이터를 보여주는 것이다. P3에, 마우스를 ICV 전달을 통해 부형제 또는 8.8e9 vg rAAV-GBA1로 처리하였다. P8에 PBS 또는 25 mg/kg CBE의 매일 IP 전달을 개시하였다. 연구 종료시, 마우스 중 절반을 P36 (1일째)에 그의 최종 CBE 투약 후 다음날 희생시켰고, 나머지 절단은 P38 (3일째)에 희생시키기 전에 3일 동안 CBE를 중단하였다. 모든 처리군 (부형제 +

PBS의 경우, n = 8, rAAV-GBA1+ PBS의 경우, n = 7, 부형제 + CBE의 경우, n = 8, 및 변이체 + CBE의 경우, n = 9)의 체중을 매일 측정하였고 (좌측 상단), P36의 체중을 분석하였다 (우측 상단). P23에 오픈 필드에서의 총 이동 거리 (좌측 하단) 및 P24에 로타로드에서 떨어질 때까지의 시간 (우측 하단)에 의해 행동을 사정하고, 각 동물에 대하여 3회 시도에 대한 중간값으로서 평가하였다. 치사에 기인하여, 행동 검정법에서 부형제 + CBE 군의 경우, n = 7인 반면, 다른 군은 모두 n=8이었다. 동물들 간의 평균이 제시되어 있다. 오차 막대는 SEM이다. *p<0.05; ***p<0.001, CBE 처리된 동물에서 선형 회귀에 의한 처리군에 대한 명목 p 값.

도 12는 CBE 마우스 모델에서의 최대 rAAV 용량의 생화학적 사정에 대한 대표적인 데이터를 보여주는 것이다. CBE 중단 전 (1일째) 또는 CBE 중단 후 (3일째), 모든 처리군 (부형제 + PBS의 경우, n = 8, 변이체 + PBS의 경우, n = 7, 부형제 + CBE의 경우, n = 7, 및 변이체 + CBE의 경우, n = 9)의 피질을 사용하여 군에서의 GCase 활성 (좌측 상단), GluSph 수준 (우측 상단), GluCer 수준 (좌측 하단), 및 벡터 게놈 (우측 하단)을 측정하였다. 생체분포는 게놈 DNA 1 ㎍당 벡터 게놈으로 제시되어 있다. 평균이 제시되어 있다. 오차 막대는 SEM이다. (*)p<0.1; **p<0.01; ***p<0.001, CBE 처리된 동물에서 선형 회귀에 의한 처리군에 대한 명목 p 값, 여기서 수집일 및 성별은 공변량으로서 보정.

도 13은 부형제 + PBS, 부형제 + CBE, 및 변이체 + CBE 처리군의 투여 이후의 CBE 마우스 모델에서의 행동 및 생화학적 상관관계에 대한 대표적인 데이터를 보여주는 것이다. 처리군 간에, 로타로드에서의 수행능은 GluCer 축적과 음의 상관관계를 가졌고 (A, 선형 회귀에 의해 p=0.0012), GluSph 축적은 증가된 GCase 활성과 음의 상관관계를 가졌다 (B, 선형 회귀에 의해 p=0.0086).

도 14는 CBE 마우스 모델에서의 변이체의 생체분포에 대한 대표적인 데이터를 보여주는 것이다. 모든 처리군 (부형제 + PBS의 경우, n = 8, 변이체+ PBS의 경우, n = 7, 부형제 + CBE의 경우, n = 7, 및 변이체+ CBE의 경우, n = 9)에 대해 간, 비장, 신장, 및 생식선에서 벡터 게놈의 존재를 사정하였다. 생체분포는 게놈 DNA 1 ㎍당 벡터 게놈으로 제시되어 있다. 벡터 참조 표준 곡선을 이용하여 정량적 PCR에 의해 벡터 게놈 존재를 정량화하였고; A260 광학 밀도 측정에 의해 게놈 DNA 농도를 평가하였다. 평균이 제시되어 있다. 오차 막대는 SEM이다. *p<0.05; **p<0.01; ***p<0.001, CBE 처리된 동물에서 선형 회귀에 의한 처리군에 대한 명목 p 값, 여기서 수집일 및 성별은 공변량으로서 보정.

도 15는 CBE 마우스 모델에서의 rAAV 용량 범위의 생전 사정에 대한 대표적인 데이터를 보여주는 것이다. 마우스는 P3에 ICV 전달에 의해 부형제 또는 3개의 상이한 용량의 rAAV-GBA1: 3.2e9 vg, 1.0e10 vg, 또는 3.2e10 vg 중 하나를 받았다. P8에, 25 mg/kg CBE의 매일 IP 처리를 개시하였다. 부형제 및 CBE 또는 부형제 및 PBS를 받은 마우스는 대조군으로서의 역할을 하였다. 모든 처리군은 군당 n = 10 (5M/5F)으로 시작하였다. 모든 마우스를 그의 최종 CBE 투약 후 다음날 (P38-P40) 희생시켰다. 모든 처리군을 매일 체중을 측정하고, P36에 그의 체중을 분석하였다. P24에 로타로드에서 떨어질 때까지의 시간, 및 P30에 테이퍼 빔(Tapered Beam)을 건너는 데 소요되는 시간에 의해 운동 수행능을 사정하였다. 조기 치사에 기인하여, 행동 검정법에 참여한 마우스의 마리수는 부형제 + PBS의 경우, n = 10, 부형제 + CBE의 경우, n = 9, 및 3.2e9 vg rAAV-GBA1+ CBE의 경우, n = 6, 1.0e10 vg rAAV-GBA1+ CBE의 경우, n = 10, 3.2e10 vg rAAV-GBA1+ CBE의 경우, n = 7이었다. 평균이 제시되어 있다. 오차 막대는 SEM이고; * p<0.05; **p<0.01, CBE 처리군에서 선형 회귀에 의한 명목 p 값, 여기서 성별은 공변량으로서 보정.

도 16은 CBE 마우스 모델에서의 rAAV 용량 범위의 생화학적 사정에 대한 대표적인 데이터를 보여주는 것이다. 모든 처리군 (부형제 + PBS의 경우, n = 10, 부형제 + CBE의 경우, n = 9, 및 3.2e9 vg rAAV-GBA1+ CBE의 경우, n = 6, 1.0e10 vg rAAV-GBA1+ CBE의 경우, n = 10, 3.2e10 vg rAAV-GBA1+ CBE의 경우, n = 7)의 피질을 이용하여 GCase 활성, GluSph 수준, GluCer 수준, 및 벡터 게놈을 측정하였다. GCase 활성은 총 단백질 1 mg당 GCase ng로 제시되어 있다. GluSph 및 GluCer 수준은 조직 습식 중량 1 mg당 pmol로 제시되어 있다. 생체분포는 게놈 DNA 1 ㎍당 벡터 게놈으로 제시되어 있다. 벡터 참조 표준 곡선을 이용하여 정량적 PCR에 의해 벡터 게놈 존재를 정량화하였고; A260 광학 밀도 측정에 의해 게놈 DNA 농도를 평가하였다. 벡터 게놈 존재를 또한 간에서도 측정하였다 (E). 평균이 제시되어 있다. 오차 막대는 SEM이다. **p<0.01; ***p<0.001, CBE 처리군에서 선형 회귀에 의한 명목 p 값, 여기서 성별은 공변량으로서 보정.

도 17은 유전적 마우스 모델에서 최대 용량 rAAV-GBA1에서의 테이퍼 빔 분석에 대한 대표적인 데이터를 보여주는 것이다. rAAV-GBA1 투여 후 4주째에 빔 워크(Beam Walk)에 의해 처리군 (WT + 부형제 (n = 5), 4L/PS-NA + 부형제 (n = 6), 및 4L/PS-NA + rAAV-GBA1 (n = 5))의 운동 수행능을 검정하였다. 총 슬립수 및 활동 시간은 상이한 빔에서의 5회 시도에 대한 총합으로 제시되어 있다. 속도 및 속도당 슬립수는 상이한 빔에서의 5회 시

도에 대한 평균으로 제시되어 있다. 평균이 제시되어 있다. 오차 막대는 SEM이다.

도 18은 프로그래뉼린 (PGRN) 단백질을 코딩하는 rAAV 구축물의 시험관내 발현에 대한 대표적인 데이터를 보여주는 것이다. 좌측 패널은 프로그래뉼린 (PGRN) ELISA 검정법의 표준 곡선을 보여주는 것이다. 하단 패널은 rAAV가 형질도입된 HEK293T 세포의 세포 용해물 중에서 ELISA 검정법에 의해 측정된 PGRN 발현의 용량-반응을 보여주는 것이다. MOI = 감염 다중도 (세포당 벡터 게놈).

도 19는 프로사포신 (*PSAP*), *SCARB2*, 및/또는 하나 이상의 억제성 핵산과 함께 조합된, *GBA1*을 코딩하는 rAAV 구축물의 시험관내 발현에 대한 대표적인 데이터를 보여주는 것이다. 데이터는, HEK293 세포를 각 구축물로 형질감염시켰을 때, 모의 형질감염된 세포와 비교하여 관심 트랜스진이 과다발현되었다는 것을 나타낸다.

도 20은 ITR의 "외부"에 (예컨대, 트랜스진 인서트 또는 발현 구축물 기준으로 ITR의 말단에 인접하게) 위치하는 "D" 영역을 포함하는 rAAV 벡터 (상단), 및 벡터의 "내부"에 (예컨대, 벡터의 트랜스진 인서트에 인접한 위치에) ITR을 갖는 야생형 rAAV 벡터를 도시한 개략도이다.

도 21은 GBA2 또는 그의 일부, 및 α-Syn에 대한 간섭 RNA를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다.

도 22는 Gcase (예컨대, *GBA1* 또는 그의 일부) 및 갈락토실세라미다제 (예컨대, *GALC* 또는 그의 일부)를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 갈락토실세라미다제의 코딩 서열의 발현은 T2A 자기 절단 펩티드 서열에 의해 이격되어 있다.

도 23은 Gcase (예컨대, *GBA1* 또는 그의 일부) 및 갈락토실세라미다제 (예컨대, *GALC* 또는 그의 일부)를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 갈락토실세라미다제의 코딩 서열의 발현은 T2A 자기 절단 펩티드 서열에 의해 이격되어 있다.

도 24는 Gcase (예컨대, *GBA1* 또는 그의 일부), 카텝신 B (예컨대, *CTSB* 또는 그의 일부), 및 α-Syn에 대한 간섭 RNA를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 카텝신 B의 코딩 서열의 발현은 T2A 자기 절단 펩티드 서열에 의해 이격되어 있다.

도 25는 Gcase (예컨대, *GBA1* 또는 그의 일부), 스핑고마이엘린 포스포디에스테라제 1 (예컨대, *SMPD1* 그의 일부, 및 α-Syn에 대한 간섭 RNA를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다.

도 26은 Gcase (예컨대, *GBA1* 또는 그의 일부) 및 갈락토실세라미다제 (예컨대, *GALC* 또는 그의 일부)를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 갈락토실세라미다제의 코딩 서열은 내부 리보솜 진입 부위 (IRES)에 의해 이격되어 있다.

도 27은 Gcase (예컨대, *GBA1* 또는 그의 일부) 및 카텝신 B (예컨대, *CTSB* 또는 그의 일부)를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 카텝신 B의 코딩 서열의 발현은 각각 별개의 프로모터에 의해 구동된다.

도 28은 Gcase (예컨대, *GBA1* 또는 그의 일부), GCH1 (예컨대, *GCH1* 또는 그의 일부), 및 α-Syn에 대한 간섭 RNA를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 GCH1의 코딩 서열은 T2A 자기 절단 펩티드 서열에 의해 이격되어 있다.

도 29는 Gcase (예컨대, *GBA1* 또는 그의 일부), RAB7L1 (예컨대, *RAB7L1* 또는 그의 일부), 및 α-Syn에 대한 간섭 RNA를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 RAB7L1의 코딩 서열은 T2A 자기 절단 펩티드 서열에 의해 이격되어 있다.

도 30은 Gcase (예컨대, *GBA1* 또는 그의 일부), GCH1 (예컨대, *GCH1* 또는 그의 일부), 및 α-Syn에 대한 간섭 RNA를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 GCH1의 코딩 서열의 발현은 내부 리보솜 진입 부위 (IRES)이다.

도 31은 VPS35 (예컨대, *VPS35* 또는 그의 일부) 및 α-Syn 및 TMEM106B에 대한 간섭 RNA를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다.

도 32는 Gcase (예컨대, *GBA1* 또는 그의 일부), IL-34 (예컨대, *IL34* 또는 그의 일부), 및 α-Syn에 대한 간섭 RNA를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 IL-34의 코딩 서열은 T2A 자기 절단 펩티드 서열에 의해 이격되어 있다.

도 33은 Gcase (예컨대, *GBA1* 또는 그의 일부) 및 IL-34 (예컨대, *IL34* 또는 그의 일부)를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 IL-34의 코딩 서열은 내부 리보솜 진입 부위 (IRES)에 의해 이격되어 있다.

도 34는 Gcase (예컨대, *GBA1* 또는 그의 일부) 및 TREM2 (예컨대, *TREM2* 또는 그의 일부)를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 TREM2의 코딩 서열의 발현은 각각 별개의 프로모터에 의해 구동된다.

도 35는 Gcase (예컨대, *GBA1* 또는 그의 일부) 및 IL-34 (예컨대, *IL34* 또는 그의 일부)를 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. Gcase 및 IL-34의 코딩 서열의 발현은 각각 별개의 프로모터에 의해 구동된다.

도 36a-36b는 qPCR 및 ELISA에 의해 측정된 바, 대조군 형질도입된 세포 대비 HEK293 세포에서의 TREM2 및 GBA1의 과다발현에 대한 대표적인 데이터를 보여주는 것이다. 도 36a는 TREM2의 과다발현에 대한 데이터를 보여주는 것이다. 도 36b는 동일 구축물로부터의 GBA1의 과다발현에 대한 데이터를 보여주는 것이다.

도 37은 GFP 리포터 검정법 (상단) 및 α-Syn 검정법 (하단)에 의해, 시험관내에서 *SCNA*의 침묵화가 성공적으로 이루어졌다는 것을 나타내는 대표적인 데이터를 보여주는 것이다.

도 38은 GFP 리포터 검정법 (상단) 및 α-Syn 검정법 (하단)에 의해, 시험관내에서 *TMEM106B*의 침묵화가 성공적으로 이루어졌다는 것을 나타내는 대표적인 데이터를 보여주는 것이다.

도 39는 PGRN을 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다.

도 40은 "D" 서열이 야생형으로 (동그라미 표시) 또는 대안적으로 (예컨대, "외부에"; 사각형 표시) 배치되어 있는 ITR을 갖는 rAAV를 이용한 HEK293 세포의 형질도입에 대한 데이터를 보여주는 것이다. "외부에" 배치된 ITR을 갖는 rAAV는 야생형 ITR을 갖는 rAAV만큼 효율적으로 세포를 형질도입시킬 수 있었다.

발명을 실시하기 위한 구체적인 내용

[0040] 본 개시내용은 부분적으로 대상체에서 PD 연관 유전자 생성물의 조합의 발현을 위한 조성물 및 방법에 기초한다. 유전자 생성물은 단백질, 단백질의 단편 (예컨대, 일부), PD 연관 유전자를 억제시키는 간섭 핵산 등일 수 있다. 일부 실시양태에서, 유전자 생성물은 PD 연관 유전자에 의해 코딩되는 단백질 또는 단백질 단편이다. 일부 실시양태에서, 유전자 생성물은 PD 연관 유전자를 억제시키는 간섭 핵산 (예컨대, shRNA, siRNA, miRNA, amiRNA 등)이다.

[0041] PD 연관 유전자란, PD와 유전적으로, 생화학적으로 또는 기능적으로 연관이 있는 유전자 생성물을 코딩하는 유전자를 지칭한다. 예를 들어, (단백질 Gcase를 코딩하는) *GBA1* 유전자 중 돌연변이를 갖는 개체는 *GBA1* 중에 돌연변이를 가지지 않는 것과 비교하였을 때, PD가 발생할 위험이 증가되어 있는 것으로 관찰되었다. 또 다른 예에서, PD는 α-시누클레인 (α-Syn)을 포함하는 단백질 응집체의 축적과 연관이 있고; 따라서, (α-Syn을 코딩하는) *SCNA*가 PD 연관 유전자이다. 일부 실시양태에서, 본원에 기술된 발현 카세트는 야생형 또는 비-돌연변이체 형태의 PD 연관 유전자 (또는 그의 코딩 서열)이다. PD 연관 유전자의 예는 표 1에 열거되어 있다.

[0042] <표 1>

[0043] PD 연관 유전자의 예

명칭	유전자	기능	NCBI 수탁 번호
리소좀 막 단백질 2	*SCARB2/LIMP2*	글루코실세라미다제에 대한 리소좀 수용체 (GBA 표적화)	NP_005497.1 (이소형 1), NP_001191184.1 (이소형 2)
프로사포신	*PSAP*	리소좀 구획으로 국재화되고, 짧은 올리고당 기를 갖는 글리코스핑고리피드의 이화작용을 촉진시키는 사포신 A, B, C, 및 D에 대한 전구체	AAH01503.1, AAH07612.1, AAH04275.1, AAA60303.1
베타-글루코세레브로시다제	*GBA1*	글루코세레브로시드의 베타-글리코시드 결합을 절단한다	NP_001005742.1 (이소형 1), NP_001165282.1 (이소형 2), NP_001165283.1 (이소형 3)
비-리소좀 글루코실세라미다제	*GBA2*	글루코실세라미드의 유리 글루코스 및 세라미드로의 전환을 촉진시킨다	NP_065995.1 (이소형 1), NP_001317589.1 (이소형 2)
갈락토실세라미다제	*GALC*	세라미드 유도체로부터 갈락토스를 제거한다	EAW81359.1 (이소형 CRA_a), EAW81360.1 (이소형 CRA_b), EAW81362.1 (이소형 CRA_c)
스핑고마이엘린 포스포디에스테라제 1	*SMPD1*	스핑고마이엘린을 세라미드로 전환시킨다	EAW68726.1 (이소형 CRA_a), EAW68727.1 (이소형 CRA_b), EAW68728.1 (이소형 CRA_c), EAW68729.1 (이소형 CRA_d)

[0044]

카텝신 B	*CTSB*	단백질의 세포내 분해 및 전환에 관여하는 것으로 간주되는 티올 프로테아제; 이는 또한 종양 침윤 및 전이에도 연루	AAC37547.1, AAH95408.1, AAH10240.1
RAB7, 구성원 RAS 온코진 패밀리-유사 1	*RAB7L1*	소포 수송을 조절한다	AAH02585.1
액포 단백질 분류 연관 단백질 35	*VPS35*	레트로머 카그로-선택적 복합체의 성분	NP_060676.2
GTP 시클로히드롤라제 1	*GCH1*	구아노신 트리포스페이트를 가수분해하여 7,8-디히드로네오프테린 트리포스페이트를 형성하는 것을 담당	AAH25415.1
인터류킨 34	*IL34*	단핵구의 성장 또는 생존을 증가시키고; 콜로니 자극 인자 1 수용체에 결합함으로써 활성을 유도한다	AAH29804.1
골수 세포 상에서 발현되는 트리거링(triggering) 수용체 2	*TREM2*	TYRO 단백질 티로신 키아제 결합 단백질과 수용체 신호전달 복합체를 형성하고; 면역 반응에서 작용하고, 만성 염증에 관여할 수 있다	AAF69824.1
프로그래뉼린	*PGRN*	발생, 염증, 세포 증식 및 단백질 항상성에서 중요한 역할을 한다	NP_002087.1

[0045]

[0046] *단리된 핵산 및 벡터*

[0047] 단리된 핵산은 DNA 또는 RNA일 수 있다. 일부 측면에서, 본 개시내용은 하나 이상의 PD 연관 유전자, 예를 들어, Gcase (예컨대, *GBA1* 유전자의 유전자 생성물) 또는 그의 일부를 코딩하는 발현 구축물을 포함하는 단리된 핵산 (예컨대, rAAV 벡터)을 제공한다. β-글루코세레브로시다제 또는 GBA로도 지칭되는 Gcase란, 당지질 대사에서 중간체인 화학물질 글루코세레브로시드의 베타-글리코시드 결합을 절단하는 리소좀 단백질을 지칭한다. 인간에서, Gcase는 염색체 1에 위치하는, *GBA1* 유전자에 의해 코딩된다. 일부 실시양태에서, *GBA1*은 NCBI 참조 서열 NCBI 참조 서열 NP_000148.2 (서열식별번호 14)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대, 포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) Gcase 코딩 서열, 예컨대, 서열식별번호 15에 기재된 서열을 포함한다.

[0048] 일부 측면에서, 본 개시내용은 프로사포신 (예컨대, *PSAP* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. 프로사포신은 짧은 올리고당 기를 갖는 글리코스핑고리피드의 이화작용을 촉진시키는 스핑고리피드 활성인자 단백질 (사포신) A, B, C, 및 D에 대한 전구체 당단백질이다. 인간에서, *PSAP* 유전자는 염색체 10에 위치한다. 일부 실시양태에서, *PSAP*는 NCBI 참조 서열 NP_002769.1 (예컨대, 서열식별번호 16)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 (예컨대,

포유동물 세포, 예를 들어, 인간 세포에서의 발현을 위해 코돈 최적화된) 프로사포신 코딩 서열, 예컨대, 서열식별번호 17에 기재된 서열을 포함한다.

[0049] 본 개시내용의 측면은 LIMP2/SCARB2 (예컨대, *SCARB2* 유전자의 유전자 생성물)를 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. SCARB2는 세포 내에서의 리소좀 및 엔도솜 수송을 조절하는 막 단백질을 지칭한다. 인간에서, *SCARB2* 유전자는 염색체 4에 위치한다. 일부 실시양태에서, *SCARB2* 유전자는 NCBI 참조 서열 NP_005497.1 (서열식별번호 18)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 19에 기재된 서열을 포함한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 SCARB2 코딩 서열을 포함한다.

[0050] 본 개시내용의 측면은 GBA2 단백질 (예컨대, *GBA2* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. GBA2 단백질은 비-리소좀 글루코실세라미다제를 지칭한다. 인간에서, *GBA2* 유전자는 염색체 9에 위치한다. 일부 실시양태에서, *GBA2* 유전자는 NCBI 참조 서열 NP_065995.1 (서열식별번호 30)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 31에 기재된 서열을 포함한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 GBA2 코딩 서열을 포함한다.

[0051] 본 개시내용의 측면은 GALC 단백질 (예컨대, *GALC* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. GALC 단백질은, 갈락토세레브로시드, 갈락토실스핑고신, 락토실세라미드, 및 모노갈락토실디글리세리드의 갈락토스 에스테르 결합을 가수분해하는 효소인 갈락토실세라미다제 (또는 갈락토세레브로시다제)를 지칭한다. 인간에서, *GALC* 유전자는 염색체 14에 위치한다. 일부 실시양태에서, *GALC* 유전자는 NCBI 참조 서열 NP_000144.2 (서열식별번호 33)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 34에 기재된 서열을 포함한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 GALC 코딩 서열을 포함한다.

[0052] 본 개시내용의 측면은 CTSB 단백질 (예컨대, *CTSB* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. CTSB 단백질은, 세포내 단백질분해에서 중요한 역할을 하는 리소좀 시스테인 프로테아제인 카텝신 B를 지칭한다. 인간에서, *CTSB* 유전자는 염색체 8에 위치한다. 일부 실시양태에서, *CTSB* 유전자는 NCBI 참조 서열 NP_001899.1 (서열식별번호 35)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 36에 기재된 서열을 포함한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 CTSB 코딩 서열을 포함한다.

[0053] 본 개시내용의 측면은 SMPD1 단백질 (예컨대, *SMPD1* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. SMPD1 단백질은, 스핑고리피드 대사에 관여하는 히드롤라제 효소인 스핑고마이엘린 포스포디에스테라제 1을 지칭한다. 인간에서, *SMPD1* 유전자는 염색체 11에 위치한다. 일부 실시양태에서, *SMPD1* 유전자는 NCBI 참조 서열 NP_000534.3 (서열식별번호 37)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 38에 기재된 서열을 포함한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 SMPD1 코딩 서열을 포함한다.

[0054] 본 개시내용의 측면은 GCH1 단백질 (예컨대, *GCH1* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. GCH1 단백질은, 폴레이트 및 바이오프테린 생합성 경로의 일부인 히드롤라제 효소인 GTP 시클로히드롤라제 I을 지칭한다. 인간에서, *GCH1* 유전자는 염색체 14에 위치한다. 일부 실시양태에서, *GCH1* 유전자는 NCBI 참조 서열 NP_000152.1 (서열식별번호 45)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 46에 기재된 서열을 포함한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 GCH1 코딩 서열을 포함한다.

[0055] 본 개시내용의 측면은 RAB7L 단백질 (예컨대, *RAB7L* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. RAB7L 단백질은, GTP 결합 단백질인, RAB7, 구성원 RAS 온코진 패밀리 유사 1을 지칭한다. 인간에서, *RAB7L* 유전자는 염색체 1에 위치한다. 일부 실시양태에서, *RAB7L* 유전자는 NCBI 참조 서열 NP_003920.1 (서열식별번호 47)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 48에 기재된 서열을 포함한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 RAB7L 코딩 서열을 포함한다.

[0056] 본 개시내용의 측면은 VPS35 단백질 (예컨대, *VPS35* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. VPS35 단백질은, 단백질의 엔도솜에서 트랜스-골지 네트워크로의 역핵 수송에 관여하는 단백질 복합체의 일부인, 액포 단백질 분류 연관 단백질 35를 지칭한다. 인간에서, *VPS35* 유전자는

염색체 16에 위치한다. 일부 실시양태에서, *VPS35* 유전자는 NCBI 참조 서열 NP_060676.2 (서열식별번호 49)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 50에 기재된 서열을 포함한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 VPS35 코딩 서열을 포함한다.

[0057] 본 개시내용의 측면은 IL-34 단백질 (예컨대, *IL34* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. IL-34 단백질은, 단핵구의 성장 및 생존을 증가시키는 시토카인인 인터류킨 34를 지칭한다. 인간에서, *IL34* 유전자는 염색체 16에 위치한다. 일부 실시양태에서, *IL34* 유전자는 NCBI 참조 서열 NP_689669.2 (서열식별번호 55)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 56에 기재된 서열을 포함한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 IL-34 코딩 서열을 포함한다.

[0058] 본 개시내용의 측면은 TREM2 단백질 (예컨대, *TREM2* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. TREM2 단백질은, 골수 세포 상에서 발견되는 면역글로불린 슈퍼패밀리 수용체인, 골수 세포 상세어 발현되는 트리거링(triggering) 수용체 2를 지칭한다. 인간에서, *TREM2* 유전자는 염색체 6에 위치한다. 일부 실시양태에서, *TREM2* 유전자는 NCBI 참조 서열 NP_061838.1 (서열식별번호 57)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 58에 기재된 서열을 포함한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 TREM2 코딩 서열을 포함한다.

[0059] 본 개시내용의 측면은 TMEM106B 단백질 (예컨대, *TMEM106B* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. TMEM106B 단백질은, 수상돌기 형태발생 및 리소좀 수송 조절에 관여하는 단백질인, 막횡단 단백질 106B를 지칭한다. 인간에서, *TMEM106B* 유전자는 염색체 7에 위치한다. 일부 실시양태에서, *TMEM106B* 유전자는 NCBI 참조 서열 NP_060844.2 (서열식별번호 62)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 63에 기재된 서열을 포함한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 TMEM106B 코딩 서열을 포함한다.

[0060] 본 개시내용의 측면은 프로그래뉼린 단백질 (예컨대, *PRGN* 유전자의 유전자 생성물)을 코딩하는 발현 구축물을 포함하는 단리된 핵산에 관한 것이다. PGRN 단백질은, 발생, 염증, 세포 증식 및 단백질 항상성에 관여하는 단백질인 프로그래뉼린을 지칭한다. 인간에서, *PGRN* 유전자는 염색체 17에 위치한다. 일부 실시양태에서, *PGRN* 유전자는 NCBI 참조 서열 NP_002078.1 (서열식별번호 67)에 의해 제시되는 펩티드를 코딩한다. 일부 실시양태에서, 단리된 핵산은 서열식별번호 68에 기재된 서열을 포함한다. 일부 실시양태에서, 단리된 핵산은 코돈 최적화된 PGRN 코딩 서열을 포함한다.

[0061] 일부 측면에서, 본 개시내용은 제1 유전자 생성물 및 제2 유전자 생성물을 코딩하는 발현 구축물을 포함하는 단리된 핵산이며, 여기서 각 유전자 생성물은 독립적으로 표 1에 기재된 유전자 생성물 또는 그의 일부로부터 선택되는 것인, 단리된 핵산을 제공한다.

[0062] 일부 실시양태에서, 유전자 생성물은 자연적으로 발생된 유전자의 코딩 부분 (예컨대, cDNA)에 의해 코딩된다. 일부 실시양태에서, 제1 유전자 생성물은 *GBA1* 유전자에 의해 코딩된 단백질 (또는 그의 단편)이다. 일부 실시양태에서, 유전자 생성물은 표 1에 열거된 또 다른 유전자, 예를 들어, *SCARB2/LIMP2* 유전자 또는 *PSAP* 유전자에 의해 코딩된 단백질 (또는 그의 단편)이다. 그러나, 통상의 기술자는 제1 유전자 생성물 (예컨대, Gcase) 및 제2 유전자 생성물 (예컨대, LIMP2 등)의 발현 순서가 일반적으로 역전될 수 있다는 것 (예컨대, LIMP2가 제1 유전자 생성물이고, Gcase가 제2 유전자 생성물이다)을 이해한다. 일부 실시양태에서, 유전자 생성물은 표 1에 열거된 유전자의 단편 (또는 일부)이다. 단백질 단편은 표 1에 열거된 유전자에 의해 코딩된 단백질의 약 50%, 약 60%, 약 70%, 약 80%, 약 90% 또는 약 99%를 포함할 수 있다. 일부 실시양태에서, 단백질 단편은 표 1에 열거된 유전자에 의해 코딩된 단백질의 50% 내지 99.9% (예컨대, 50% 내지 99.9% 사이의 임의의 값)를 포함한다.

[0063] 일부 실시양태에서, 발현 구축물은 모노시스트론성이다 (예컨대, 발현 구축물은 제1 유전자 생성물 및 제2 유전자 생성물을 포함하는 단일 융합 단백질을 코딩한다). 일부 실시양태에서, 발현 구축물은 폴리시스트론성이다 (예컨대, 발현 구축물은 2개의 전혀 다른 유전자 생성물, 예를 들어, 2개의 상이한 단백질 또는 단백질 단편을 코딩한다).

[0064] 폴리시스트론성 발현 벡터는 하나 이상의 (예컨대, 1, 2, 3, 4, 5개 또는 그 초과의) 프로모터를 포함할 수 있다. 임의의 적합한 프로모터, 예를 들어, 구성적 프로모터, 유도성 프로모터, 내인성 프로모터, 조직 특이적 프로모터 (예컨대, 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 의해 기술된 것과 같은 것), CD68 프로모터, 또는 JeT 프로모터 (예를 들어, 문헌 [Tornøe et al. (2002) *Gene* 297(1-2):21-32]에 의해 기술된 것과 같은 것)이다. 일부 실시양태에서, 프로모터는 제1 유전자 생성물, 제2 유전자 생성물, 또는 제1 유전자 생성물 및 제2 유전자 생성물을 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, 하나 이상의 추가의 조절 서열을 포함한다.

[0065] 일부 실시양태에서, 제1 유전자 생성물을 코딩하는 핵산 서열 및 제2 유전자 생성물을 코딩하는 핵산 서열은 내부 리보솜 진입 부위 (IRES)를 코딩하는 핵산 서열에 의해 이격된다. IRES 부위의 예는 예를 들어, 문헌 [Mokrejs et al. (2006) *Nucleic Acids Res.* 34(Database issue):D125-30]에 기술되어 있다. 일부 실시양태에서, 제1 유전자 생성물을 코딩하는 핵산 서열 및 제2 유전자 생성물을 코딩하는 핵산 서열은 자기 절단 펩티드를 코딩하는 핵산 서열에 의해 이격된다. 자기 절단 펩티드의 예는 T2A, P2A, E2A, F2A, BmCPV 2A, 및 BmIFV 2A, 및 문헌 [Liu et al. (2017) *Sci Rep.* 7: 2193]에 의해 기술된 것을 포함하나, 이에 제한되지 않는다. 일부 실시양태에서, 자기 절단 펩티드는 T2A 펩티드이다.

[0066] 병리학상, 장애, 예컨대, PD 및 고셰병은 대개는 α-시누클레인 (α-Syn) 단백질로 구성된 단백질 응집체의 축적과 연관이 있다. 따라서, 일부 실시양태에서, 본원에 기술된 단리된 핵산은 α-Syn 단백질의 발현을 감소시키거나, 또는 방해하는 억제성 핵산을 포함한다. 억제성 핵산을 코딩하는 서열은 발현 벡터의 비번역 영역 (예컨대, 인트론, 5'UTR, 3'UTR 등)에 배치될 수 있다.

[0067] 일부 실시양태에서, 억제성 핵산은 발현 구축물의 인트론에, 예를 들어, 제1 유전자 생성물을 코딩하는 서열의 상류쪽 인트론에 위치한다. 억제성 핵산은 이중 가닥 RNA (dsRNA), siRNA, 마이크로RNA (miRNA), 인공 miRNA (amiRNA), 또는 RNA 압타머일 수 있다. 일반적으로, 억제성 핵산은 표적 RNA (예컨대, mRNA)의 약 6개 내지 약 30개의 (예컨대, 6 내지 30 사이의 임의의 정수 (6 및 30 포함)) 연속적인 뉴클레오티드에 결합한다 (예컨대, 그와 하이브리드화한다). 일부 실시양태에서, 억제성 핵산 분자는 miRNA 또는 amiRNA, 예를 들어, *SNCA* (α-Syn 단백질을 코딩하는 유전자) 또는 *TMEM106B* (예컨대, TMEM106B 단백질을 코딩하는 유전자)를 표적화하는 miRNA이다. 일부 실시양태에서, miRNA는 하이브리드화 대상인 *SNCA* mRNA의 영역과의 미스매치는 포함하지 않는다 (예컨대, miRNA는 "완벽한 것이다"). 일부 실시양태에서, 억제성 핵산은 shRNA (예컨대, *SNCA* 또는 *TMEM106B*를 표적화하는 shRNA)이다. 일부 실시양태에서, 억제성 핵산은 miR-155 스캐폴드 및 *SCNA* 또는 *TMEM106B* 표적화 서열을 포함하는 인공 miRNA (amiRNA)이다.

[0068] 본원에 기술된 단리된 핵산은 그 자체 단독으로, 또는 벡터의 일부로서 존재할 수 있다. 일반적으로, 벡터는 플라스미드, 코스미드, 파지미드, 박테리아 인공 염색체 (BAC), 또는 바이러스 벡터 (예컨대, 아데노바이러스 벡터, 아데노-연관 바이러스 (AAV) 벡터, 레트로바이러스 벡터, 배큘로바이러스 벡터 등)일 수 있다. 일부 실시양태에서, 벡터는 플라스미드 (예컨대, 본원에 기술된 바와 같은 단리된 핵산을 포함하는 플라스미드)이다. 일부 실시양태에서, rAAV 벡터는 단일-가닥 (예컨대, 단일-가닥 DNA)이다. 일부 실시양태에서, 벡터는 재조합 AAV (rAAV) 벡터이다. 일부 실시양태에서, 벡터는 배큘로바이러스 벡터 (예컨대, 오토그라파 칼리포르니카 핵다각체병 (AcNPV) 벡터)이다.

[0069] 전형적으로, rAAV 벡터 (예컨대, rAAV 게놈)는 두 AAV 역전된 말단 반복부 (ITR) 서열에 의해 플랭킹된 트랜스진 (예컨대, 하기 각각의 것: 프로모터, 인트론, 인핸서 서열, 단백질 코딩 서열, 억제성 RNA 코딩 서열, 폴리A 테일 서열 등 중 하나 이상의 것을 포함하는 발현 구축물)을 포함한다. 일부 실시양태에서, rAAV 벡터의 트랜스진은 본 개시내용에 의해 기술된 바와 같은 단리된 핵산을 포함한다. 일부 실시양태에서, rAAV 벡터의 두 ITR 서열은 각각 전장의 ITR (예컨대, 길이 대략 145 bp, 및 기능성 *Rep* 결합 부위 (RBS) 및 말단 분해 부위 (trs)를 함유하는 것)이다. 일부 실시양태에서, rAAV 벡터의 ITR들 중 하나는 말단절단된 것이다 (예컨대, 단축된 것이거나, 또는 전장의 것이 아니다). 일부 실시양태에서, 말단절단된 ITR에는 기능성 말단 분해 부위 (trs)가 결여되어 있고, 자기 상보성 AAV 벡터 (scAAV 벡터) 제조에 사용된다. 일부 실시양태에서, 말단절단된 ITR은 예를 들어, 문헌 [McCarty et al. (2003) *Gene Ther.* 10(26):2112-8]에 기술된 바와 같이, ΔITR이다.

[0070] 본 개시내용의 측면은 야생형 AAV ITR 기준으로, 예를 들어, 야생형 AAV2 ITR (예컨대, 서열식별번호 29) 기준으로 하나 이상의 변형 (예컨대, 핵산 부가, 결실, 치환 등)을 갖는 ITR을 포함하는 단리된 핵산 (예컨대, rAAV 벡터)에 관한 것이다. 야생형 AAV2 ITR의 구조는 도 20에 제시되어 있다. 일반적으로, 야생형 ITR은, 자기 어닐링하여, 각각 B/B' 및 C/C'로 지칭되는 서열에 의해 형성된) 2개의 크로스 아암, (서열 A/A'에 의해 형성된)

더욱 긴 스템 영역, 및 "D" 영역으로 지칭되는 단일-가닥 말단 영역으로 이루어진 회문 이중 가닥 T 형상의 헤어핀 구조를 형성하는 125개의 뉴클레오티드 영역을 포함한다 (도 20). 일반적으로, ITR의 "D" 영역은 A/A' 서열에 의해 형성된 스템 영역과 rAAV 벡터의 트랜스진을 함유하는 인서트 사이에 위치한다 (예컨대, ITR의 말단 기준으로 ITR의 "내부"에 또는 rAAV 벡터의 트랜스진 인서트 또는 발현 구축물에 인접하게 위치한다). 일부 실시양태에서, "D" 영역은 서열식별번호 27에 기재된 서열을 포함한다. "D" 영역은 예를 들어, 문헌 [Ling et al. (2015) *J Mol Genet Med* 9(3)]에 개시되어 있는 바와 같이, 캡시드 단백질에 의한 rAAV 벡터의 캡시드화에서 중요한 역할을 하는 것으로 관찰되었다.

[0071] 본 개시내용은 부분적으로 ITR의 "외부"에 (예컨대, 트랜스진 인서트 또는 발현 구축물 기준으로 ITR의 말단에 인접하게) 위치하는 "D" 영역을 포함하는 rAAV 벡터가 비변형된 (예컨대, 야생형) ITR로 ITR을 갖는 rAAV 벡터보다 AAV 캡시드 단백질에 의해 효율적으로 캡시드화된다는 놀라운 발견에 기초한다. 일부 실시양태에서, 변형된 "D" 서열 (예컨대, "외부 위치에 있는 "D" 서열")을 갖는 rAAV 벡터가 야생형 ITR 서열을 갖는 rAAV 벡터 기준으로 감소된 독성을 갖는다.

[0072] 일부 실시양태에서, 변형된 "D" 서열은 야생형 "D" 서열 (예컨대, 서열식별번호 27) 기준으로 적어도 하나의 뉴클레오티드를 포함한다. 변형된 "D" 서열은 야생형 "D" 서열 (예컨대, 서열식별번호 27) 기준으로 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개, 또는 10개 초과의 뉴클레오티드 치환을 가질 수 있다. 일부 실시양태에서, 변형된 "D" 서열은 야생형 "D" 서열 (예컨대, 서열식별번호 27) 기준으로 적어도 10, 11, 12, 13, 14, 15, 16, 17, 18, 또는 19개의 핵산 치환을 포함한다. 일부 실시양태에서, 변형된 "D" 서열은 야생형 "D" 서열 (예컨대, 서열식별번호 27)과 약 10% 내지 약 99% (예컨대, 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 기술된 바와 같이 "S" 서열로도 지칭되는, 서열식별번호 26에 기재된 서열을 포함한다.

[0073] 본 개시내용에 의해 기술된 바와 같은 단리된 핵산 또는 rAAV 벡터는 예를 들어, 서열식별번호 28에 기재된 바와 같은, 또는 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 기술된 바와 같은 "TRY" 서열을 추가로 포함할 수 있다. 일부 실시양태에서, TRY 서열은 ITR (예컨대, 5' ITR)과 단리된 핵산 또는 rAAV 벡터의 발현 구축물 (예컨대, 트랜스진 코딩 인서트) 사이에 위치한다.

[0074] 일부 측면에서, 본 개시내용은 본 개시내용에 의해 기술된 바와 같은 단리된 핵산 또는 rAAV 벡터를 포함하는 배큘로바이러스 벡터에 관한 것이다. 일부 실시양태에서, 배큘로바이러스 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 의해 기술된 바와 같은, 오토그라파 칼리포르니카 핵 다각체병 (AcNPV) 벡터이다.

[0075] 일부 측면에서, 본 개시내용은 본원에 기술된 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, E. 콜라이(*E. coli*) 세포이다.

[0076] *rAAV*

[0077] 일부 측면에서, 본 개시내용은 본원에 기술된 바와 같은 핵산을 코딩하는 트랜스진을 포함하는 재조합 AAV (rAAV) (예컨대, 본원에 기술된 바와 같은 rAAV 벡터)에 관한 것이다. "rAAV"라는 용어는 일반적으로 하나 이상의 AAV 캡시드 단백질에 의해 캡시드화된 rAAV 벡터를 포함하는 바이러스 입자를 지칭한다. 본 개시내용에 의해 기술된 rAAV는 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, 및 AAV10으로부터 선택되는 혈청형을 갖는 캡시드 단백질을 포함할 수 있다. 일부 실시양태에서, rAAV는 비-인간 숙주로부터의 캡시드 단백질, 예를 들어, 레수스 AAV 캡시드 단백질, 예컨대, AAVrh.10, AAVrh.39 등을 포함한다. 일부 실시양태에서, 본 개시내용에 의해 기술된 rAAV는 야생형 캡시드 단백질의 변이체인 캡시드 단백질, 예컨대, 유래 기점이 되는 야생형 AAV 캡시드 단백질 기준으로 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개, 또는 10개 초과의 (예컨대, 15, 20 25, 50, 100개 등) 아미노산 치환 (예컨대, 돌연변이)을 포함하는 캡시드 단백질 변이체를 포함한다.

[0078] 일부 실시양태에서, 본 개시내용에 의해 기술된 rAAV는 특히 CSF 공간 내로, 또는 뇌 실질 내로 직접 도입되었을 때, CNS를 통해 쉽게 확산된다. 따라서, 일부 실시양태에서, 본 개시내용에 의해 기술된 rAAV는 혈액-뇌 장벽 (BBB)을 건널 수 있는 캡시드 단백질을 포함한다. 예를 들어, 일부 실시양태에서, rAAV는 AAV9 또는

AAVrh.10 혈청형을 갖는 캡시드 단백질 포함한다. rAAV 제조는 예를 들어, 문헌 [Samulski et al. (1989) *J Virol.* 63(9):3822-8] 및 [Wright (2009) *Hum Gene Ther.* 20(7): 698-706]에 의해 기술되어 있다.

[0079] 일부 실시양태에서, 본 개시내용에 의해 기술된 바와 같은 rAAV (예컨대, AAV 캡시드 단백질에 의해 캡시드화되어 rAAV 캡시드 입자를 형성하는 재조합 rAAV 게놈을 포함하는 것)는 배큘로바이러스 벡터 발현 시스템 (BEVS)에서 제조된다. BEVS를 이용하여 rAAV를 제조하는 것은 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 의해 기술되어 있다. 그러나, rAAV는 임의의 적합한 방법 (예컨대, 재조합 rep 및 cap 유전자를 이용하는 방법)을 이용하여 제조될 수 있다.

[0080] *제약 조성물*

[0081] 일부 측면에서, 본 개시내용은 본원에 기술된 바와 같은 단리된 핵산 또는 rAAV 및 제약상 허용되는 담체를 포함하는 제약 조성물을 제공한다. 본원에서 사용되는 바, "제약상 허용되는"이라는 용어는 물질, 예컨대, 담체 또는 희석제가 화합물의 생물학적 활성 또는 특성을 폐기하지 않고, 비교적 비독성이라는 것, 예컨대, 물질이 비바람직한 생물학적 효과를 유발하지 않거나, 또는 그와 함께 함유되어 있는 조성물의 성분 중 어느 것과도 유해한 방식으로 상호작용하지 않으면서, 개체에게 투여될 수 있다는 것을 지칭한다.

[0082] 본원에서 사용되는 바, "제약상 허용되는 담체"라는 용어는 본 발명 내에서 유용한 화합물이 그의 의도된 기능을 수행할 수 있도록 상기 화합물을 환자 내에 또는 환자에게 운반 또는 수송하는 데 관여하는, 제약상 허용되는 물질, 조성물 또는 담체, 예컨대, 액체 또는 고체 충전제, 안정제, 분산화제, 현탁화제, 희석제, 부형제, 증점제, 용매 또는 캡슐화 물질을 의미한다. 본 발명의 실시에서 사용되는 제약 조성물에 포함될 수 있는 추가 성분은 관련 기술분야에 공지되어 있고, 예를 들어, 문헌 [Remington's Pharmaceutical Sciences (Genaro, Ed., Mack Publishing Co., 1985, Easton, PA)] (상기 문헌은 본원에서 참조로 포함된다)에 기술되어 있다.

[0083] 본원에서 제공된 조성물 (예컨대, 제약 조성물)은 장관 (예컨대, 경구적), 비경구적, 정맥내, 근육내, 동맥내, 골수내, 경막내, 피하, 뇌실내, 경피, 피부간(interdermal), 직장, 질내, 복강내, 국소 (분제, 연고, 크림제, 및/또는 점적제에 의한 경우), 점막, 비강, 협착, 설하를 비롯한, 임의의 경로에 의해; 기관내 주입, 기관지 주입, 및/또는 흡입에 의해; 및/또는 경구용 스프레이, 비강용 스프레이, 및/또는 에어로졸로서 투여될 수 있다. 경구 투여, 정맥내 투여 (예컨대, 전신 정맥내 주사), 혈액 및/또는 림프 공급을 통한 국부 투여, 및/또는 환부로의 직접 투여가 구체적으로 고려되는 경로이다. 일반적으로, 가장 적절한 투여 경로는 작용제의 성질 (예컨대, 위장관 환경에서의 그의 안정성), 및/또는 대상체의 상태 (예컨대, 대상체가 경구 투여를 허용할 수 있는지 여부)를 비롯한, 다양한 인자에 의존할 것이다. 특정 실시양태에서, 본원에 기술된 화합물 또는 제약 조성물은 대상체의 안구로의 국소 투여에 적합하다.

[0084] *방법*

[0085] 본 개시내용은 부분적으로 함께 (예컨대, 시너지적으로) 작용하여 파킨슨병을 치료하는, 대상체에서의 PD 연관 유전자 생성물의 조합의 발현을 위한 조성물에 기초한다. 본원에서 사용되는 바, "치료하다" 또는 "치료하는"이라는 것은 (a) 파킨슨병의 발병을 예방하거나, 또는 지연시키고/거나; (b) 파킨슨병의 중증도를 감소시키고/거나; (c) 파킨슨병의 특징적인 증상의 발생을 감소시키거나, 또는 예방하고/거나; (d) 파킨슨병의 특징적인 증상의 악화를 예방하는 것을 지칭한다. 파킨슨병의 증상으로는 예를 들어, 운동 기능장애 (예컨대, 떨림, 경직, 몸동작이 느려지는 서동, 보행 곤란), 인지 기능장애 (예컨대, 치매, 우울증, 불안), 감정 및 행동 기능장애를 포함한다.

[0086] 따라서, 일부 측면에서, 본 개시내용은 파킨슨병을 앓거나 또는 앓을 것으로 의심되는 대상체에게 본 개시내용에 의해 기술된 바와 같은 조성물 (예컨대, 단리된 핵산 또는 벡터 또는 rAAV를 포함하는 조성물)을 투여하는 단계를 포함하는, 파킨슨병을 앓거나 또는 앓을 것으로 의심되는 대상체를 치료하는 방법을 제공한다.

[0087] 일부 실시양태에서, 조성물은 대상체의 CNS로 직접, 예를 들어, 대상체의 뇌 및/또는 척수 내로의 직접 주사에 의해 투여된다. CNS 직접 투여 방식의 예로는 대뇌내 주사, 뇌실내 주사, 수조내 주사, 실질내 주사, 경막내 주사, 및 상기의 임의의 조합을 포함하나, 이에 제한되지 않는다. 일부 실시양태에서, 대상체의 CNS 내로 직접 주사하면, 그 결과로 트랜스진은 대상체의 중뇌, 선조체 및/또는 대뇌 피질에서 발현된다 (예컨대, 제1 유전자 생성물, 제2 유전자 생성물, 및 적용가능할 경우, 제3 유전자 생성물은 발현된다). 일부 실시양태에서, CNS 내로 직접 주사하면, 그 결과로 트랜스진은 대상체의 척수 및/또는 CSF에서 발현된다 (예컨대, 제1 유전자 생성물, 제2 유전자 생성물, 및 적용가능할 경우, 제3 유전자 생성물은 발현된다).

[0088] 일부 실시양태에서, 대상체의 CNS로의 직접 주사는 컨벡션 강화 전달법 (CED)을 포함한다. 컨벡션 강화 전달법은 뇌를 외과적으로 노출시키고, 뇌의 표적 부위 내로 직접 소경 카테터를 배치한 후, 이어서, 치료제 (예컨대, 본원에 기술된 바와 같은 조성물 또는 rAAV)를 대상체의 뇌로 직접 주입하는 것을 포함하는 치료 전략법이다. CED는 예를 들어, 문헌 [Debinski et al. (2009) *Expert Rev Neurother.* 9(10):1519-27]에 의해 기술되어 있다.

[0089] 일부 실시양태에서, 조성물은 대상체에게 예를 들어, 말초 주사에 의해 말초적으로 투여된다. 말초 주사의 예로는 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.

[0090] 일부 실시양태에서, 본 개시내용에 의해 기술된 바와 같은 조성물 (예컨대, 단리된 핵산 또는 벡터 또는 rAAV를 포함하는 조성물)은 말초적으로 및 대상체의 CNS로 직접, 상기 두 방식 모두에 의해 투여된다. 예를 들어, 일부 실시양태에서, 대상체는 동맥내 주사에 의해 (예컨대, 경동맥 내로의 주사에 의해), 및 실질내 주사에 의해 (예컨대, CED에 의한 실질내 주사에 의해) 조성물을 투여받는다. 일부 실시양태에서, CNS로의 직접 주사 및 말초 주사는 동시에 이루어진다 (예컨대, 같은 시간에 수행된다). 일부 실시양태에서, 직접 주사는 말초 주사 이전에 (예컨대, 말초 주사 전 1분 내지 1주 사이에, 또는 그를 초과하는 시간에) 이루어진다. 일부 실시양태에서, 직접 주사는 말초 주사 이후에 (예컨대, 말초 주사 후 1분 내지 1주 사이에, 또는 그를 초과하는 시간에) 이루어진다.

[0091] 대상체에게 투여되는 본 개시내용에 의해 기술된 바와 같은 조성물 (예컨대, 단리된 핵산 또는 벡터 또는 rAAV를 포함하는 조성물)의 양은 투여 방법에 의존하여 달라질 것이다. 예를 들어, 일부 실시양태에서, 본원에 기술된 바와 같은 rAAV는 약 10^9 게놈 카피 (GC)/kg 내지 약 10^{14} GC/kg (예컨대, 약 10^9 GC/kg, 약 10^{10} GC/kg, 약 10^{11} GC/kg, 약 10^{12} GC/kg, 약 10^{12} GC/kg, 또는 약 10^{14} GC/kg)의 역가로 대상체에게 투여된다. 일부 실시양태에서, 대상체는 CSF 공간으로의 주사에 의해, 또는 실질내 주사에 의해 고역가로 (예컨대, $>10^{12}$ 게놈 카피 GC/kg (rAAV)) 투여받는다.

[0092] 본 개시내용에 의해 기술된 바와 같은 조성물 (예컨대, 단리된 핵산 또는 벡터 또는 rAAV를 포함하는 조성물)은 1회 또는 다회에 걸쳐 (예컨대, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20회 또는 그 초과) 대상체에게 투여될 수 있다. 일부 실시양태에서, 조성물은 대상체에게 연속하여 (예컨대, 만성적으로), 예를 들어, 주입 펌프를 통해 투여된다.

[0093] 실시예

[0094] *실시예 1: rAAV 벡터*

[0095] 삼중-플라스미드 형질감염을 위해 세포, 예컨대, HEK293 세포를 사용하여 AAV 벡터를 생성한다. ITR 서열은 각 관심 트랜스진에 대한 프로모터/인핸서 요소, 3' 폴리A 신호, 및 번역 후 신호, 예컨대, WPRE 요소를 포함하는 발현 구축물에 플랭킹된다. 단백질 서열의 융합에 의해; 또는 2A 펩티드 링커, 예컨대, T2A 또는 P2A 사용으로 펩티드 결합이 생성되지 못하게 방해됨에 따라 아미노산이 부가된 2개의 펩티드 단편이 유도됨으로써; IRES 요소를 사용하여; 또는 2개의 별개의 발현 카세트를 이용하는 발현에 의해 예컨대, *GBA1* 및 *LIMP2* 및/또는 프로사포신과 같은 다중 유전자 생성물이 동시에 발현될 수 있다. 발현 유전자 상류쪽의 효율적으로 스플라이싱된 짧은 인트론 서열의 존재가 발현 수준을 개선시킬 수 있다. shRNA 및 다른 조절 RNA가 잠재적으로 상기 서열 내에 포함될 수 있다. 본 개시내용에 의해 기술된 발현 구축물의 예는 도 1-8 및 21-35 및 하기 표 2에 제시되어 있다.

[0096] <표 2>

명칭	프로모터 1	shRNA	CDS1	폴리A1	비시스트론성 요소	프로모터 2	CDS2	폴리A2	ITR 사이의 길이
CMVe_CBAp_GBA1_WPRE_bGH	CBA		GBA1	WPRE-bGH					3741
LT1s_JetLong_mRNAiaSYn_SCARB2-T2A-GBA1_bGH	JetLong	aSyn	SCARB2	bGH	T2A		GBA1		4215
LI1_JetLong_SCARB2-IRES-GBA1_bGH	JetLong		SCARB2	bGH	IRES		GBA1		4399
FP1_JetLong_GBA1_bGH_JetLong_SCARB2_SV40L	JetLong		GBA1	bGH		JetLong	SCARB2	SV40L	4464
PrevailVector_LT2s_JetLong_mRNAiaSYn_PSAP-T2A-GBA1_bGH_4353nt	JetLong	aSyn	PSAP	bGH	T2A	-	GBA1	-	4353
PrevailVector_LI2_JetLong_PSAP_IRES_GBA1_SymtheticpolyA_4337nt	JetLong	-	PSAP	합성 pA	IRES	-	GBA1	-	4337
PrevailVector_10s_JetLong_mRNAiaSy_GBA2_WPRE_bGH_4308nt	JetLong	aSyn	GBA2	WPRE_bGH	-	-	-	-	4308
PrevailVector_FT4_JetLong_GBA1_T2A_GALC_SyntheticpolyA_4373nt	JetLong	-	GBA1	합성 pA	T2A	-	GALC	-	4373
PrevailVector_LT4_JetLong_GALC_T2A_GBA1_SyntheticpolyA_4373nt	JetLong	-	GALC	합성 pA	T2A	-	GBA1	-	4373
PrevailVector_LT5s_JetLong_mRNAiaSyn_CTSB-T2A-	JetLong	aSyn	CTSB	WPRE_bGH	T2A	-	GBA1	-	4392

[0097]

[0098]

GBA1_WPRE_bGH_4392nt									
PrevailVector_FT11t_JetLong_mRNAiaSyn_GBA1_T2S_SMPD1_SyntheticpolyA_4477nt	JetLong	aSyn	GBA1	합성 pA	T2A	-	SMPD1	-	4477
PrevailVector_LI4_JetLong_GALC_IRES_GBA1_SymtheticpolyA_4820nt	JetLong	-	GALC	합성 pA	IRES	-	GBA1	-	4820
PrevailVector_FP5_JetLong_GBA1_bGH_JetLong_CTSB_SV40l_4108nt	JetLong	-	GBA1	bGH	-	JetLong	CTSB	SV40L	4108
PrevailVector_FT6s_JetLong_mRNAiaSyn_GBA1-T2A-GCH1_WPRE_bGH_4125nt	JetLong	aSyn	GBA1	WPRE_bGH	T2A	-	GCH1	-	4125
PrevailVector_LT7s_JetLong_mRNAiaSyn_RAB7L1-T2A-GBA1_WPRE_bGH_3984nt	JetLong	aSyn	RAB7L1	WPRE_bGH	T2A	-	GBA1	-	3984
PrevailVector_FI6s_JetLong_mRNAiaSYn_GBA1-IRES-GCH1_bGH_3978nt	JetLong	aSyn	GBA1	bGH	IRES	-	GCH1	-	3978
PrevailVector_9st_JetLong_mRNAiaSyn_mRNAiTMEM106B_VPS35_WPRE_bGH_4182nt	JetLong	aSyn & TMEM 106B	VPS35	WPRE_bGH	-	-	-	-	4182
PrevailVector_FT12s_JetLong_mRNAiaSyn_GBA1-T2A-IL34_WPRE_bGH_4104nt	JetLong	aSyn	GBA1	WPRE_bGH	T2A	-	IL34	-	4104

[0099]

PrevailVector_FI12s_JetLong_mRNAiaSYn_GBA1-IRES-IL34_bGH_3957nt	JetLong	aSyn	GBA1	bGH	IRES	-	IL34	-	3957
PrevailVector_FP8_JetLong_GBA1_bGH_CD68_TREM2_SV40l_4253nt	JetLong	-	GBA1	bGH	-	CD68	TREM2	SV40L	4253
PrevailVector_FP12_CMVe_CBA_GBA1_bGH_JetLong_IL34_SV40l_4503nt	CBA		GBA1	bGH		JetLong	IL34	SV40L	4503

[0100] *실시예 2: GBA 결핍 세포로의 바이러스 형질도입에 관한 세포 기반 검정법*

[0101] 예를 들어, GD 환자 단핵구, 또는 hES 세포로부터의 섬유모세포, 또는 환자 유래의 유도 만능 줄기 세포 (iPSC)로서 *GBA1* 결핍 세포를 수득한다. 이들 세포는 기질, 예컨대, 글루코실세라미드 및 글루코실스핑고신 (GlcCer 및 GlcSph)을 축적한다. 야생형 또는 돌연변이체 배양 세포주를 Gcase 억제제, 예컨대, CBE로 처리하는 것 또한 GBA 결핍 세포를 수득하는 데 사용된다.

[0102] 상기 세포 모델을 이용하여, 해당 단백질에 대한 항체 또는 포스포-αSyn을 이용한 후, 형광 현미경법을 이용하여 영상화함으로써 단백질 응집체, 예컨대, α-시누클레인의 축적에 의해 리소좀 결함을 정량화한다. 단백질 마커, 예컨대, LAMP1, LAMP2, LIMP1, LIMP2에 대한 ICC에 의해, 또는 염료, 예컨대, 리소트래커(Lysotracker)

를 사용하여, 또는 형광성 덱스트란 또는 다른 마커의 세포내이입 구획을 통한 흡수에 의해 리소좀 이상에 대한 영상화 또한 수행한다. 리소좀과의 결함성 융합에 기인하는 자가포식 마커 축적에 대한, 예컨대 LC3에 대한 영상화 또한 수행될 수 있다. 웨스턴 블롯 및/또는 ELISA는 상기 마커의 비정상적인 축적을 정량화하는 데 사용된다. 또한, 당지질 기질 및 GBA1의 생성물의 축적은 표준 접근법을 이용하여 측정된다.

[0103] 치료 종점 (예컨대, PD 연관 병상의 감소)은 활성 및 기능을 확인하고, 정량화하기 위해 AAV 벡터의 형질도입의 발현과 관련하여 측정된다. Gcase는 또한 단백질 ELISA 측정을 사용하여, 또는 표준 Gcase 활성 검정법에 의해 정량화될 수 있다.

[0104] *실시예 3: 돌연변이체 마우스를 이용한 생체내 검정법*

[0105] 본 실시예는 돌연변이체 마우스를 이용한 AAV 벡터에 관한 생체내 검정법을 기술한다. 돌연변이체 마우스에서의 상기와 같은 AAV 벡터에 관한 생체내 연구는 예를 들어, 문헌 [Liou *et al.* (2006) *J. Biol. Chem.* 281(7):4242-4253], [Sun *et al.* (2005) *J. Lipid Res.* 46:2102-2113], 및 [Farfel-Becker *et al.* (2011) *Dis. Model Mech.* 4(6):746-752]에 의해 기술된 검정법을 이용하여 수행된다.

[0106] 예를 들어, 5-10 ㎕의 주사 부피로 진한 AAV 스톡을 사용하여 (예컨대, $2x10^{11}$ vg/마우스 용량의) 비히클 대조군 및 AAV 벡터의 경막내 또는 뇌실내 전달을 수행한다. 컨벡션 강화 전달법에 의한 실질내 전달을 수행한다.

[0107] 처리는 증상 발병 이전 또는 발병 이후에 개시한다. 측정되는 종점은 CNS 및 CSF 중 기질 축적, ELISA에 의한 Gcase 효소 축적 및 효소 활성의 축적, 운동 및 인지 종점, 리소좀 기능장애, 및 α-시누클레인 단량체, 원시섬유 또는 원섬유의 축적이다.

[0108] *실시예 4: 질환의 화학적 모델*

[0109] 본 실시예는 화학적으로 유도된 고셰병을 앓는 마우스 모델 (예컨대, CBE 마우스 모델)을 이용하는 AAV 벡터에 관한 생체내 검정법을 기술한다. 예를 들어, 문헌 [Vardi *et al.* (2016) *J Pathol.* 239(4):496-509]에 기술된 바와 같이, 화학적으로 유도된 고셰병을 앓는 마우스 모델에서 상기 AAV 벡터에 관한 생체내 연구를 수행한다.

[0110] 예를 들어, 5-10 ㎕의 주사 부피로 진한 AAV 스톡을 사용하여 (예컨대, $2x10^{11}$ vg/마우스의) 비히클 대조군 및 AAV 벡터의 경막내 또는 뇌실내 전달을 수행한다. 컨벡션 강화 전달법에 의한 실질내 전달을 수행한다. 꼬리정맥 주사에 의해 말초 전달을 달성한다.

[0111] 처리는 증상 발병 이전 또는 발병 이후에 개시한다. 측정되는 종점은 CNS 및 CSF 중 기질 축적, ELISA에 의한 Gcase 효소 축적 및 효소 활성의 축적, 운동 및 인지 종점, 리소좀 기능장애, 및 α-시누클레인 단량체, 원시섬유 또는 원섬유의 축적이다.

[0112] *실시예 5: PD, LBD, 고셰병 환자에서의 임상 시험*

[0113] 일부 실시양태에서, 특정 형태의 고셰병 (예컨대, GD1)을 앓는 환자는 파킨슨병 (PD) 또는 루이소체 치매 (LBD)가 발생할 위험이 증가되어 있다. 본 실시예는 고셰병, PD 및/또는 LBD를 앓는 환자에서 본 개시내용에 의해 기술된 바와 같은 rAAV의 안전성 및 효능을 사정하는 임상 시험을 기술한다.

[0114] 문헌 [Grabowski et al. (1995) *Ann. Intern. Med.* 122(1):33-39]에 기술된 것과 유사한 연구 디자인을 사용하여 고셰병, PD 및/또는 LBD 치료를 위한 상기 벡터의 임상 시험을 수행한다.

[0115] *실시예 6: 말초 질환 치료*

[0116] 일부 실시양태에서, 예를 들어, 문헌 [Biegstraaten *et al.* (2010) *Brain* 133(10):2909-2919]에 기술된 바와 같이, 특정 형태의 고셰병을 앓는 환자는 말초 신경병증의 증상을 보인다.

[0117] 본 실시예는 고셰병 (예컨대, 타입 1 고셰병)과 연관된 말초 신경병증의 치료를 위한 본원에 기술된 바와 같은 AAV 벡터에 관한 생체내 검정법을 기술한다. 간략하면, 말초 신경병증의 징후 또는 증상을 앓는 것으로 확인된 타입 1 고셰병 환자는 본 개시내용에 의해 기술된 바와 같은 rAAV를 투여받는다. 일부 실시양태에서, rAAV의 투여 후에, 예를 들어, 문헌 [Biegstraaten *et al.*]에 기술된 방법을 사용하여 대상체의 말초 신경병증 징후 및 증상을 모니터링한다.

[0118] 환자에 (예컨대, 환자의 혈청에, 환자의 말초 조직에 (예컨대, 간 조직, 비장 조직 등에) 존재하는 본 개시내용에 의해 기술된 바와 같은 형질도입된 유전자 생성물의 수준을 예를 들어, 웨스턴 블롯 분석, 효소 기능성 검정

법, 또는 영상화 연구에 의해 검정한다.

[0119] *실시예 7: CNS 형태의 치료*

[0120] 본 실시예는 CNS 형태의 고셰병의 치료를 위한 본원에 기술된 바와 같은 rAAV 벡터에 관한 생체내 검정법을 기술한다. 간략하면, CNS 형태의 고셰병 (예컨대, 타입 2 또는 타입 3 고셰병)을 앓는 것으로 확인된 고셰병 환자는 본 개시내용에 의해 기술된 바와 같은 rAAV를 투여받는다. 환자의 CNS에 (예컨대, 환자의 CNS의 혈청에, 환자의 뇌척수액 (CSF)에, 또는 환자의 CNS 조직에) 존재하는 본 개시내용에 의해 기술된 바와 같은 형질도입된 유전자 생성물의 수준을 예를 들어, 웨스턴 블롯 분석, 효소 기능성 검정법, 또는 영상화 연구에 의해 검정한다.

[0121] *실시예 8: GBA1에 돌연변이를 갖는 대상체에서 파킨슨병의 유전자 요법*

[0122] 본 실시예는 *GBA1* 중의 돌연변이를 특징으로 하는 파킨슨병을 앓는 대상체에게 *GBA1*을 코딩하는 재조합 아데노-연관 바이러스 (rAAV)를 투여하는 것을 기술한다.

[0123] rAAV-GBA1 벡터 인서트는, 인간 GBA1 (마룬)의 코돈 최적화된 코딩 서열 (CDS)을 구성적으로 발현하기 위해, 4개 파트: CMV 인핸서 (CMVe), CBA 프로모터 (CBAp), 엑손 1, 및 인트론 (int)으로 이루어진, CBA 프로모터 요소 (CBA)를 함유한다. 3' 영역은 또한 우드척 간염 바이러스 전사후 조절 요소(Woodchuck hepatitis virus Posttranscriptional Regulatory Element: WPRE), 전사후 조절 요소, 이어서, 소 성장 호르몬 폴리A 신호 (bGH 폴리A) 테일을 함유한다. 플랭킹 ITR을 통해 개재 서열은 정확하게 패키징될 수 있다. 5' ITR 서열의 두 변이체 (도 7, 삽입 박스, 하단 서열)를 평가하였다: 상기 변이체들은 ITR의 20-뉴클레오티드 "D" 영역 내에서 수개의 뉴클레오티드 차이를 보이며, 이는 패키징 및 발현 효율에 영향을 미치는 것으로 간주된다. rAAV-GBA1 벡터 생성물은 도 7 (삽입 박스, 상단 서열)에 제시된 "D" 도메인 뉴클레오티드 서열을 함유한다. 변이체 벡터는 임상전 연구에서 유사하게 수행된, 돌연변이체 "D" 도메인 (본원에서 "S" 도메인으로 명명, 여기서 뉴클레오티드 변이는 음영 표시로 제시)을 보유한다. 백본은 카나마이신에 대한 내성을 부여하는 유전자 뿐만 아니라, 역 패키징을 막는 스터퍼(stuffer) 서열을 함유한다. rAAV-GBA1 벡터를 도시하는 개략도는 도 8에 제시되어 있다. rAAV-GBA1 벡터는 AAV9 혈청형 캡시드 단백질을 이용하여 rAAV로 패키징된다.

[0124] 대수조 내로의 (대수조내; ICM) 투시 조영 유도하의 후두하 주사를 통해 rAAV-GBA1을 단일 용량으로 대상체에게 투여한다. rAAV-GBA1 투약 요법 연구의 한 실시양태는 하기와 같다:

[0125] 단일 용량의 rAAV-GBA1을, 비임상 약리학 및 독성 연구의 결과에 기초하여 결정된 두 용량 수준 (3e13 vg (저용량); 1e14 vg (고용량) 등) 중 하나로 환자 (N=12)에게 투여한다.

[0126] 초기 연구는 rAAV-GBA1 벡터 및 rAAV-GBA1 S-변이체 구축물 (하기에서 추가로 기술)의 효능 및 안전성을 사정하기 위해, GCase의 억제제인 콘두리톨-b-에폭시드 (CBE)를 매일 전달하는 것을 포함하는 화학적 마우스 모델에서 수행하였다. 추가로, 동형접합성 *GBA1* 돌연변이를 보유하고, 사포신 (4L/PS-NA)이 부분적으로 결핍되어 있는 유전적 마우스 모델에서도 초기 연구를 수행하였다. 벡터 안전성 및 효능을 추가로 평가하기 위해 마우스 및 비인간 영장류 (NHP)에서 추가의 용량 범위 연구를 수행한다.

[0127] AAV 백본 중의 약간 다른 두 버전의 5' 역전된 말단 반복부 (ITR)를 시험하여 제조가능성 및 트랜스진 발현을 사정하였다 (도 7). 145 bp 5' ITR 내의 20 bp "D" 도메인은 최적의 바이러스 벡터 제조를 위해 필요한 것으로 간주되지만, "D" 도메인 내의 돌연변이 또한 일부 경우에 트랜스진 발현을 증가시키는 것으로 보고된 바 있다. 따라서, 무손상 "D" 도메인을 보유하는 바이러스 벡터 rAAV-GBA1 이외에, 돌연변이체 D 도메인 (본원에서 "S" 도메인으로 명명)을 포함하는 제2 벡터 형태 또한 평가하였다. rAAV-GBA1 및 변이체, 둘 모두 동일한 트랜스진을 발현한다. 하기에서 상세하게 설명되는 바와 같이, 상기 두 벡터 모두 생체내에서 효과적인 바이러스를 생산한 반면, 야생형 "D" 도메인을 함유하는 rAAV-GBA1은 추가 개발을 위해 선택하였다.

[0128] GCase 결핍인 CBE 모델을 확립하기 위해, 어린 마우스에 GCase의 특정 억제제인 CBE를 투약하였다. 생후 8일째 (P8)를 시작으로 매일 IP 주사에 의해 CBE를 마우스에 제공하였다. 행동 표현형을 보이는 모델을 확립하기 위해 3개의 상이한 CBE 용량 (25 mg/kg, 37.5 mg/kg, 50 mg/kg) 및 PBS를 시험하였다 (도 9). 더 높은 고용량의 CBE는 용량에 의존하는 방식으로 치사에 이르게 하였다. 50 mg/kg CBE로 처리된 마우스는 모두 P23까지 사망하였고, 37.5 mg/kg CBE로 처리된 8마리의 마우스 중 5마리는 P27까지 사망하였다. 25 mg/kg CBE로 처리된 마우스 중에서 치사는 없었다. CBE 주사맞은 마우스는 오픈 필드 검정법에서 일반적인 운동 장애를 보이지 않은 반면 (PBS를 받은 마우스와 동일한 속도로 동일한 거리를 이동), CBE로 처리된 마우스는 로타로드 검정법에 의해

측정된 바, 운동 협조 및 균형 장애를 보였다.

[0129] 연구 종료시까지 살아남은 마우스를 그의 최종 CBE 투약 후 그 다음날 (P27, "1일째"), 또는 CBE 중단 3일 후 (P29, "3일째") 희생시켰다. 1일째 및 3일째 코호트, 둘 모두에서 GCase 기질의 축적을 평가하기 위해 25 mg/kg CBE를 받은 마우스의 피질에서 지질 분석을 수행하였다. (본 실시예에서 응집체로 측정된) GluSph 및 GalSph 수준은, GCase 부전증과 일치하는 PBS로 처리된 대조군과 비교하였을 때, CBE로 처리된 마우스에서 유의적으로 축적되었다.

[0130] 상기 기술된 연구에 기초하여, 25 mg/kg CBE 용량은 생존에는 영향을 주지 않으면서, 행동 장애를 일으켰는 바, 이에, 상기 용량을 선택하였다. CBE 처리 동안 뇌 전역에 걸쳐 GBA1을 광범위하게 분포시키고, 트랜스진을 발현시키기 위해, 생후 3일째 (P3) 뇌실내 (ICV) 주사에 의해 rAAV-GBA1 또는 부형제를 전달한 후, 이어서, P8에 IP CBE 또는 PBS의 매일 처리를 개시하였다 (도 10).

[0131] rAAV-GBA1을 받은, CBE로 처리된 마우스는 로타로드에서 부형제를 받은 것보다 통계학상 유의적으로 더욱 우수한 수행능을 보였다 (도 11). 변이체 처리군의 마우스는 다른 행동 척도, 예컨대, 검사 동안 이동한 총 거리에 있어서는 부형제로 처리된 마우스와 다르지 않았다 (도 11).

[0132] 생전 연구 완료시, 생화학적 분석을 위해, 마우스 중 절반을 최종 CBE 투약 후 그 다음날 (P36, "1일째"), 또는 CBE 중단 3일 후 (P38, "3일째") 희생시켰다 (도 12). 생물학적으로 삼중으로 수행된 형광분석 효소 검정법을 이용하여, 피질에서 GCase 활성을 사정하였다. rAAV-GBA1로 처리된 마우스에서는 GCase 활성이 증가한 반면, CBE 처리는 GCase 활성을 감소시켰다. 추가로, CBE 및 rAAV-GBA1, 둘 모두를 받은 마우스는 PBS 처리군과 유사한 GCase 활성 수준을 보였는데, 이는 rAAV-GBA1 전달이 CBE 처리에 의해 유도된 GCase 활성 억제를 극복할 수 있다는 것을 시사하는 것이다. 기질 GluCer 및 GluSph의 수준을 조사하기 위해 마우스의 운동 피질에서 지질 분석을 수행하였다. 두 지질 모두 CBE를 받은 마우스의 뇌에 축적되었고, rAAV-GBA1 처리는 기질 축적을 유의적으로 감소시켰다.

[0133] 지질 수준은 처리군 간에 GCase 활성 및 로타로드에서의 수행능, 둘 모두와 음의 상관관계를 가졌다. rAAV-GBA1 투여 후 GCase 활성 증가는 기질 감소 및 운동 기능 증진과 연관이 있었다 (도 13). 도 14에 제시된 바와 같이, qPCR에 의해 측정된 바, 벡터 게놈 존재에 의해 예비 생체분포를 사정하였다 (여기서 게놈 DNA 1 ㎍당 >100 벡터 게놈인 것을 양성인 것으로 정의한다). CBE로 처리 및 비처리된, rAAV-GBA1을 받은 마우스, 둘 모두 피질에서 rAAV-GBA1 벡터 게놈에 대해 양성이었고, 이는 ICV 전달을 통해 rAAV-GBA1이 피질로 전달된다는 것을 시사하는 것이다. 추가로, 벡터 게놈이 간에서 검출되었고, 비장에서는 거의 검출되지 않았고, 심장, 신장 또는 생식선에서는 전혀 검출되지 않았다. 모든 척도에서 1일째 군과 3일째 군 사이에 통계학상 유의적인 차이는 없었다.

[0134] CBE 모델에서의 좀더 큰 대규모 연구를 통해 CBE 모델에서 rAAV-GBA1의 유효 용량을 추가로 조사하였다. 25 mg/kg CBE 용량 모델을 이용하여, 부형제 또는 rAAV-GBA1을 P3에 ICV를 통해 전달하고, P8에 IP PBS 또는 CBE의 매일 처리를 개시하였다. 이전 연구에서 CBE를 중단한 군과 CBE를 중단하지 않은 군 사이에 유사성이 관찰되었다는 점을 고려하여, 모든 마우스를 최종 CBE 투약 (P38-40) 후 다음날 희생시켰다. 군당 마우스 10마리씩 (5M/5F), 하기 5개 군으로 하여, 3개의 상이한 rAAV-GBA1 용량의 효과를 사정하였다:

[0135] 부형제 ICV + PBS IP

[0136] 부형제 ICV + 25 mg/kg CBE IP

[0137] 3.2e9 vg (2.13e10 vg/g 뇌) rAAV-GBA1 ICV + 25 mg/kg CBE IP

[0138] 1.0e10 vg (6.67e10 vg/g 뇌) rAAV-GBA1 ICV + 25 mg/kg CBE IP

[0139] 3.2e10 vg (2.13e11 vg/g 뇌) rAAV-GBA1 ICV + 25 mg/kg CBE IP.

[0140] 최고 용량의 rAAV-GBA1은 P37에 CBE 처리와 관련된 체중 증가 불능을 구제하였다. 추가로, 부형제 + CBE 처리군과 비교하였을 때, 상기 용량은 로타로드 및 테이퍼 빔에서의 수행능을 통계학상 유의적으로 증가시켰다 (도 15). 부형제 처리군 및 rAAV-GBA1 처리군, 둘 모두를 포함하는, 여러 군에서 치사가 관찰되었다 (부형제 + PBS: 0; 부형제 + 25 mg/kg CBE: 1; 3.2e9 vg rAAV-GBA1+ 25 mg/kg CBE: 4; 1.0e10 vg rAAV-GBA1+ 25 mg/kg CBE: 0; 3.2e10 vg rAAV-GBA1+ 25 mg/kg CBE: 3).

[0141] 생전 연구 완료시, 생화학적 분석을 위해, 마우스를 희생시켰다 (도 16). 형광분석 검정법에 의해 생물학적으

로 삼중으로 피질에서 GCase 활성을 사정하였다. CBE로 처리된 마우스는 감소된 GCase 활성을 보인 반면, 고용량의 rAAV-GBA1을 받은 마우스는 CBE 처리군과 비교하였을 때, GCase 활성의 통계학상 유의적인 증가를 보였다. CBE로 처리된 마우스는 또한 GluCer 및 GluSph의 축적을 보였으며, 이 둘 모두는 고용량의 rAAV-GBA1 투여에 의해 구제되었다.

[0142] 확립된 화학적 CBE 모델에 추가로, Gba1에서의 V394L GD 돌연변이에 대해 동형접합성이고, 또한 GCase 국재화 및 활성에 영향을 주는 사포신이 부분적으로 결핍되어 있는 4L/PS-NA 유전적 모델에서도 또한 rAAV-GBA1을 평가한다. 상기 마우스는 빔 워크, 로타로드, 및 와이어 행(wire hang) 검정법에서의 그의 수행능에 의해 입증되는 바와 같이, 운동 강도, 협조, 및 균형 장애를 보인다. 전형적으로, 이들 마우스의 수명은 22주 미만이다. 초기 연구에서, P23에 3 ㎕의 최대 역가 바이러스를 ICV에 의해 전달하였고, 여기서 최종 용량은 2.4e10 vg (6.0e10 vg/g 뇌)였다. 군당 마우스 6마리씩이며, 처리군은 하기와 같았다:

[0143] WT + 부형제 ICV

[0144] 4L/PS-NA + 부형제 ICV

[0145] 4L/PS-NA + 2.4e10 vg (6.0e10 vg/g 뇌) rAAV-GBA1 ICV.

[0146] rAAV-GBA1 전달 후 4주째에 빔 워크 검사에 의한 운동 수행능을 사정하였다. rAAV-GBA1을 받은 돌연변이체 마우스 군은, 운동 기능은 거의 WT 수준으로까지 회복되면서, 부형제로 처리된 돌연변이체 마우스와 비교하였을 때, 총 슬립수는 더 적고, 속도당 슬립수도 더 적은 경향을 보였다 (도 17). 운동 표현형은 상기 마우스가 노화됨에 따라 점점 더 심각해지기 때문에, 상기 행동 및 다른 행동 검사에서의 그의 수행능은 더 늦은 후속 시점에 사정한다. 생전 연구 완료시, 상기 마우스에서 지질 수준, GCase 활성, 및 생체분포를 사정한다.

[0147] 0.03x, 0.1x, 및 1x 제안된 1상의 임상 고용량에 상응하는, 추가의 더 낮은 용량의 rAAV-GBA1을 CBE 모델을 이용하여 현재 시험하고 있다. 각 군은 군당 마우스 10마리씩 (5M/5F) 포함한다:

[0148] 부형제 ICV

[0149] 부형제 ICV + 25 mg/kg CBE IP

[0150] 3.2e8 vg (2.13e9 vg/g 뇌) rAAV-GBA1 ICV + 25 mg/kg CBE IP

[0151] 1.0e9 vg (6.67e9 vg/g 뇌) rAAV-GBA1 ICV + 25 mg/kg CBE IP

[0152] 1.0e10 vg (6.67e10 vg/g 뇌) rAAV-GBA1 ICV + 25 mg/kg CBE IP.

[0153] 운동 표현형 이외에도, 피질에서 지질 수준 및 GCase 활성을 사정한다. 시간 경과에 따른 처리 및 분석 또한 수행한다.

[0154] 효능 및 안전성 데이터를 평가하기 위해 더 큰 용량 범위의 연구를 개시하였다. 10마리의 4L/PS-NA 마우스 (군당 5M/5F)에 10 ㎕의 rAAV-GBA1을 주사하였다. 상대 성장 측정 뇌 중량 계산을 이용하면, 용량은 0.15x, 1.5x, 4.4x, 및 14.5x 제안된 1상의 임상 고용량과 상관관계가 있다. 주사군은 하기 군으로 이루어진다:

[0155] WT + 부형제 ICV

[0156] 4L/PS-NA + 부형제 ICV

[0157] 4L/PS-NA + 4.3e9 vg (1.1e10 vg/g 뇌) rAAV-GBA1 ICV

[0158] 4L/PS-NA + 4.3e10 vg (1.1e11 vg/g/ 뇌) rAAV-GBA1 ICV

[0159] 4L/PS-NA + 1.3e11 vg (3.2e11 vg/g 뇌) rAAV-GBA1 ICV

[0160] 4L/PS-NA + 4.3e11 vg (1.1e12 vg/g 뇌) rAAV-GBA1 ICV.

[0161] CBE 모델에서의 비임상 연구에 관한 요약은 하기 표 3에 제시되어 있다.

[0162] <표 3>

[0163] CBE 마우스 모델에서의 결과 요약

시험 물질	연구 번호	용량 코호트	행동 변화			지질	효소	BD	
			로타로	데이퍼빔	오픈필드			뇌	간
rAAV-GBA1	CBE 모델에서의 PRV-2018-005 용량 범위 rAAV-GBA1	3.2e9 vg (2.13e10 vg/g 뇌)	NS	NS	NS	NS	NS	+	-
		1.10e10 vg (6.67e10 vg/g 뇌)	T	NS	NS	T/S	NS	+	+
		2.3e10vg (2.13e11 vg/g 뇌)	S	S	NS	S	S	+	+
변이체	CBE 모델에서의 PRV-2018-005 용량 범위 변이체	8.8e9 vg (5.9e10 vg/g 뇌)	S	N/A	NS	S	S	+	+

[0164]

[0165] 양성 생체분포는 >100 vg/1 ㎍ 게놈 DNA인 것으로 정의된다는 점에 주의한다.

[0166] 약어: BD = 생체분포; NS = 비유의적; T = 경향; S = 유의적; N/A = 적용불가능; + = 양성; - = 음성.

[0167] *실시예 9: rAAV 벡터의 시험관내 분석*

[0168] rAAV 구축물을 시험관내 및 생체내에서 시험하였다. 도 18은 프로그래뉼린 (PGRN) 단백질을 코딩하는 rAAV 구축물의 시험관내 발현에 대한 대표적인 데이터를 보여주는 것이다. 좌측 패널은 프로그래뉼린 (PGRN) ELISA 검정법의 표준 곡선을 보여주는 것이다. 하단 패널은 rAAV가 형질도입된 HEK293T 세포의 세포 용해물 중에서 ELISA 검정법에 의해 측정된 PGRN 발현의 용량-반응을 보여주는 것이다. MOI = 감염 다중도 (세포당 벡터 게놈).

[0169] 프로사포신 (*PSAP*) 및 *SCARB2*를 단독으로, 또는 *GBA1* 및/또는 하나 이상의 억제성 RNA와 함께 조합하여 코딩하는 rAAV 벡터의 시험관내 활성을 사정하는 예비 연구를 수행하였다. PSAP 및 프로그래뉼린 (PGRN)을 코딩하는 한 구축물도 시험하였다. 시험된 벡터는 표 4에 제시된 것을 포함한다. "Opt"는 포유동물 세포 (예컨대, 인간 세포)에서의 발현을 위해 코돈 최적화된 핵산 서열을 지칭한다. 도 19는 HEK293 세포를 각 구축물로 형질감염시켰을 때, 모의 형질감염된 세포와 비교하여 상응한 유전자 생성물이 과다발현되었다는 것을 나타내는 대표적인 데이터를 보여주는 것이다.

[0170] TREM2를 단독으로, 또는 하나 이상의 억제성 RNA와 함께 조합하여 코딩하는 rAAV 벡터의 시험관내 활성을 사정하는 예비 연구를 수행하였다. 시험된 벡터는 표 4에 제시된 것을 포함한다. "Opt"는 포유동물 세포 (예컨대, 인간 세포)에서의 발현을 위해 코돈 최적화된 핵산 서열을 지칭한다. 도 36a-36b는 HEK293 세포를 각 구축물로 형질감염시켰을 때, 모의 형질감염된 세포와 비교하여 상응한 유전자 생성물이 과다발현되었다는 것을 나타내는 대표적인 데이터를 보여주는 것이다.

[0171] <표 4>

ID	프로모터	억제성 RNA	프로모터	트랜스진
I00015	JL_인트론	*SCNA*	JetLong	Opt-PSAP_GBA1
I00039	-	-	JetLong	Opt-PSAP-GRN
I00046	-	-	CBA	Opt-PSAP
I00014	JetLong	*SCNA*	JetLong	Opt-SCARB2_GBA1
I00040			JL, CD68	opt-GBA1, TREM2

[0172]

[0173] *실시예 10: SCNA 및 TMEM106B shRNA 구축물 시험*

[0174] *HEK293 세포*

[0175] 본 연구에서는 인간 배아 신장 293 세포주 (HEK293) (#85120602, 시그마-알드리치(Sigma-Aldrich)를 사용하였다. HEK293 세포를, 100 유닛/ml 페니실린 및 100 ㎍/ml 스트렙토마이신 (#15140122, 써모 피셔 사이언티픽(Thermo Fisher Scientific))을 함유하는, 10% 우태아 혈청 [FBS] [#10082147, 써모 피셔 사이언티픽])으로 보충된 배양 배지 (D-MEM [#11995065, 써모 피셔 사이언티픽] 중에서 유지시켰다.

[0176] *플라스미드 형질감염*

[0177] 제조사의 설명서에 따라 리포펙타민(Lipofectamine) 2000 형질감염 시약 (#11668019, 써모 피셔 사이언티픽)을 이용하여 플라스미드 형질감염을 수행하였다. 간략하면, HEK293 세포 (#12022001, 시그마-알드리치)를 항생제가 없는 배양 배지 중에 $3x10^5$개의 세포/ml인 밀도로 플레이팅하였다. 다음날, 플라스미드 및 리포펙타민 2000 시약을 옵티-MEM(Opti-MEM) 용액 (#31985062, 써모 피셔 사이언티픽) 중에서 조합하였다. 5분 후, 혼합물을 HEK293 배양물에 첨가하였다. 72시간 후, RNA 또는 단백질 추출을 위해 세포를 수거하거나, 또는 세포에 대해 영상화 분석을 수행하였다. 영상화 분석을 위해, 세포 플레이팅 이전에 플레이트를 0.01% 폴리-L-리신 용액 (P8920, 시그마-알드리치)으로 미리 코팅하였다.

[0178] *정량적 실시간 PCR (qRT-PCR)에 의한 유전자 발현 분석*

[0179] 제조사의 설명서에 따라 파워 SYBR 그린 셀-투-CT 키트(Power SYBR Green Cells-to-CT Kit) (#4402955, 써모 피셔 사이언티픽)를 이용하여 정량적 실시간 PCR (qRT-PCR)에 의해 상대적인 유전자 발현 수준을 측정하였다. 리포펙타민 2000 형질감염 시약을 이용하여 48 웰 플레이트 상에 플레이팅된 HEK293 세포 ($7.5x10^4$개의 세포/웰) 내로 후보 플라스미드를 일시적으로 형질감염시켰다 (50 ㎕ 옵티-MEM 용액 중 0.5 ㎍ 플라스미드 및 1.5 ㎕ 시약). 72시간 후, 세포로부터 RNA를 추출하고, 역전사를 위해 사용하여 제조사의 설명서에 따라 cDNA를 합성하였다. 정량적 PCR 분석을 위해, 파워 SYBR 그린 PCR 마스터 믹스(Power SYBR Green PCR Master Mix) (#4367659, 써모 피셔 사이언티픽)와 함께 유전자 특이적 프라이머 쌍 (최종 농도 250 nM)을 이용하여 2~5 ㎕의 cDNA 생성물을 이중으로 증폭시켰다. *SNCA, TMEM106B*, 및 *GAPDH* 유전자에 대한 프라이머 서열은 5'- AAG AGG GTG TTC TCT ATG TAG GC -3' (서열식별번호 71), 5'- GCT CCT CCA ACA TTT GTC ACT T -3' (서열식별번호 72) (*SNCA*에 대한 것), 5'-ACA CAG TAC CTA CCG TTA TAG CA-3' (서열식별번호 73), 5'-TGT TGT CAC AGT AAC TTG CAT CA-3' (서열식별번호 74) (*TMEM106B*에 대한 것), 및 5'- CTG GGC TAC ACT GAG CAC C -3' (서열식별번호 75), 5'- AAG TGG TCG TTG AGG GCA ATG -3' (서열식별번호 76) (*GAPDH*에 대한 것)이었다. 콴트스튜디오 3 리얼-타임 PCR 시스템(QuantStudio 3 Real-Time PCR system) (써모 피셔 사이언티픽)에서 정량적 PCR을 수행하였다. 하우스키핑 유전자 *GAPDH*에 의해 발현 수준을 정규화하였고, 비교 CT 방법을 이용하여 계산하였다.

[0180] *형광 영상화 분석*

[0181] *SNCA* 및 *TMEM106B* 넉다운 플라스미드를 확인하는 데 EGFP 코딩 영역 하류의 인간 *SNCA* 유전자의 3'-UTR을 함유하는 EGFP 리포터 플라스미드를 사용하였다. EGFP 리포터 플라스미드 및 후보 넉다운 플라스미드를 리포펙타민 2000 형질감염 시약을 이용하여, 폴리-L-리신으로 코팅된 96 웰 플레이트 상에 플레이팅된 HEK293 세포 ($3.0x10^4$개의 세포/웰) 내로 동시에 형질감염시켰다 (10 ㎕ 옵티-MEM 용액 중 0.04 ㎍ 리포터 플라스미드, 0.06 ㎍

넉다운 플라스미드 및 0.3 ㎕ 시약). 72시간 후, 배리오스칸 LUX(Varioskan LUX) 다중모드 판독기 (써모 피셔 사이언티픽)를 이용하여 여기 488 nm/방출 512 nm에서 EGFP 신호의 형광 강도를 측정하였다. 세포를 RT에서 10 분 동안 4% PFA로 고정시키고, 40 ㎍/ml 7-아미노액티노마이신 D (7-AAD)를 함유하는 D-PBS와 함께 RT에서 30 min 동안 인큐베이션시켰다. D-PBS로 세척한 후, 배리오스칸 판독기를 이용하여 여기 546 nm/방출 647 nm에서 7-AAD 신호의 형광 강도를 측정하여 세포 개수를 정량화하였다. 7-AAD 신호 수준에 대해 정규화된 EGFP 신호를 대조군 넉다운 샘플과 비교하였다.

[0182] *효소-결합된 면역흡착 검정법 (ELISA)*

[0183] 단백질 수준으로 넉다운 플라스미드를 확인하는 데 *SNCA* 코딩 영역 하류의 인간 *SNCA* 유전자 또는 *TMEM106B* 유전자의 3'-UTR을 함유하는 α-시누클레인 리포터 플라스미드를 사용하였다. HEK293 세포로부터 추출된 용해물을 이용하여 ELISA (#KHB0061, 써모 피셔 사이언티픽)에 의해 α-시누클레인 단백질의 수준을 측정하였다. 후보 플라스미드를 리포펙타민 2000 형질감염 시약을 이용하여, 48 웰 플레이트 상에 플레이팅된 HEK293 세포 (7.5×10^4개의 세포/웰) 내로 일시적으로 형질감염시켰다 (25 ㎕ 옵티-MEM 용액 중 0.1 ㎍ 리포터 플라스미드, 0.15 ㎍ 넉다운 플라스미드 및 0.75 ㎕ 시약). 72시간 후, 프로테아제 억제제 칵테일 (#P8340, 시그마-알드리치)로 보충된 방사성면역침전 검정법 (RIPA) 완충제 (#89900, 써모 피셔 사이언티픽) 중에서 세포를 용해시키고, 몇 초 동안 초음파 처리하였다. 얼음 상에서 30 min 동안 인큐베이션시킨 후, 용해물을 4℃에서 15 min 동안 20,000 x g로 원심분리하고, 상청액을 수집하였다. 단백질 수준을 정량화하였다. 배리오스칸 플레이트 판독기 중 450 nm에서 플레이트를 판독하고, 소프트맥스 프로 5(SoftMax Pro 5) 소프트웨어를 이용하여 농도를 계산하였다. 측정된 단백질 농도를 비신코닌산 검정법 (#23225, 써모 피셔 사이언티픽)을 이용하여 측정된 전체 단백질 농도로 정규화하였다.

[0184] 도 37 및 하기 표 5는 GFP 리포터 검정법 (상단) 및 α-Syn 검정법 (하단)에 의해, 시험관내에서 *SCNA*의 침묵화가 성공적으로 이루어졌다는 것을 나타내는 대표적인 데이터를 보여주는 것이다. 도 38 및 하기 표 6은 GFP 리포터 검정법 (상단) 및 α-Syn 검정법 (하단)에 의해, 시험관내에서 *TMEM106B*의 침묵화가 성공적으로 이루어졌다는 것을 나타내는 대표적인 데이터를 보여주는 것이다.

[0185] <표 5>

ID	프로모터	넉다운	프로모터	과다발현
I00007	CMV_인트론	SNCA_mi	CMV	opt-GBA1
I00008	H1	SNCA_sh	CMV	opt-GBA1
I00009	H1	SNCA_Pubsh4	CMV	opt-GBA1
I00014	JL_인트론	SNCA_mi	JetLong	opt-SCARB2_GBA
I00015	JL_인트론	SNCA_mi	JetLong	opt-PSAP_GBA
I00016	JL_인트론	SNCA_mi	JetLong	opt-CTSB_GBA
I00019	JL_인트론	SNCA_TMEM_mi	JetLong	opt-VPS35
I00023	JL_인트론	SNCA_mi	JetLong	opt-GBA1_IL34
I00024	JL_인트론	SNCA_mi	JetLong	opt-GBA2
I00028	인트론	SNCA_Broadsh	CMV	opt-GBA1
I00029	인트론	SNCA_Pubsh4	CMV	opt-GBA1

[0186]

[0187] <표 6>

ID	프로모터	녹다운	프로모터	과다발현
I00010	H1	TMEM_Pubsh	CMV	opt-GRN
I00011	JL_인트론	TMEM_mi	JetLong	opt-GBA1_GRN
I00012	H1	TMEM_sh	CMV	opt-GRN
I00019	JL_인트론	SNCA_TMEM_mi	JetLong	opt-VPS35

[0188]

[0189] *실시예 11: ITR "D" 서열 배치 및 세포 형질도입*

[0190] ITR "D" 서열의 배치가 rAAV 벡터의 세포 형질도입에 미치는 효과를 조사하였다. 도 20에 제시된 바와 같은, 1) 야생형 ITR (예컨대, 트랜스진 인서트에 인접하게 위치하고, ITR의 말단에는 원위에 위치하는 "D" 서열) 또는 2) "D" 서열이 벡터의 "외부"에 위치하는 ITR (예컨대, ITR의 말단에 인접하게 위치하고, 트랜스진 인서트에는 원위에 위치하는 "D" 서열)을 갖는 Gcase 코딩 rAAV를 HEK293 세포에 형질도입시켰다. 놀랍게도, 데이터는 "외부" 위치에 위치하는 "D" 서열을 갖는 rAAV가 효율적으로 패키징될 수 있고, 세포를 형질도입시킬 수 있는 능력을 보유한다는 것을 나타낸다 (도 40).

[0191] *실시예 12: 프로그래뉼린 rAAV의 시험관내 시험*

[0192] 도 39는 PGRN을 코딩하는 발현 구축물을 포함하는 벡터의 한 실시양태를 도시한 개략도이다. 프로그래뉼린은 예컨대, 대수조 내로의 주사와 같은 실질내 또는 경막내 주사에 의한 PGRN (예컨대, 코돈 최적화된 PGRN)을 코딩하는 rAAV 벡터의 주사에 의해 GRN 결실에 대해 이형접합성 또는 동형접합성인, *GRN* 결핍인 설치류의 CNS에서 과다발현된다.

[0193] 마우스가 2개월 또는 6개월되었을 때, 그에 주사하고, 6개월 또는 12개월까지 노화시키고, 하기 중 하나 이상의 것에 대하여 분석하였다: RNA 및 단백질 수준에서의 GRN의 발현 수준, 행동 검정법 (예컨대, 운동 개선), 생존 검정법 (예컨대, 생존 개선), 소교세포 및 염증성 마커, 신경아교증, 뉴런 손실, 지방갈색소증, 및/또는 리소좀 마커 축적 구제, 예컨대, LAMP1. PGRN 결핍 마우스에 대한 검정법은 예를 들어, 문헌 [Arrant et al. (2017) *Brain* 140: 1477-1465]; [Arrant et al. (2018) *J. Neuroscience* 38(9):2341-2358]; 및 [Amado et al. (2018) doi:https://doi.org/10.1101/30869] (상기 문헌의 전체 내용이 본원에서 참조로 포함된다)에 의해 기술되어 있다.

[0194] 등가물

[0195] 본 출원은 하기 문헌의 내용, 그 전문을 참조로 포함한다: 2018년 10월 3일 출원된, 대리인 사건 번호 P1094.70002WO00으로 언급되는 국제 PCT 출원; 2018년 10월 3일 출원된, 대리인 사건 번호 P1094.70004WO00으로 언급되는 국제 PCT 출원; 2017년 10월 3일 출원된 가출원 일련 번호 62/567,296 (발명의 명칭: "GENE THERAPIES FOR LYSOSOMAL DISORDERS"); 2017년 10월 3일 출원된 가출원 일련 번호 62/567,311 (발명의 명칭: "GENE THERAPIES FOR LYSOSOMAL DISORDERS"); 2017년 10월 3일 출원된 가출원 일련 번호 62/567,319 (발명의 명칭: "GENE THERAPIES FOR LYSOSOMAL DISORDERS"); 2017년 10월 3일 출원된 가출원 일련 번호 62/567,301 (발명의 명칭: "GENE THERAPIES FOR LYSOSOMAL DISORDERS"); 2017년 10월 3일 출원된 가출원 일련 번호 62/567,310 (발명의 명칭: "GENE THERAPIES FOR LYSOSOMAL DISORDERS"); 2017년 10월 3일 출원된 가출원 일련 번호 가출원 일련 번호 62/567,303 (발명의 명칭: "GENE THERAPIES FOR LYSOSOMAL DISORDERS"); 및 2017년 10월 3일 출원된 가출원 일련 번호 62/567,305 (발명의 명칭: "GENE THERAPIES FOR LYSOSOMAL DISORDERS").

[0196] 따라서, 본 발명의 적어도 하나의 실시양태의 여러 측면을 기술하였는 바, 관련 기술분야의 통상의 기술자는 다양한 변경, 변형 및 개선을 쉽게 착안해 낼 수 있음을 이해하여야 한다. 상기 변경, 변형 및 개선은 본 개시내용의 일부가 되는 것으로 의도되고, 본 발명의 정신 및 범주 내에 포함되는 것으로 의도된다. 따라서, 상기 설명 및 도면은 단지 예일 뿐이다.

[0197] 본원에 본 발명의 여러 실시양태가 기술되고, 예시되었지만, 관련 기술분야의 통상의 기술자는 기능을 수행하고/거나, 본원에 기술된 결과 및/또는 이점들 중 하나 이상의 것을 수득하기 위해 각종의 다른 수단 및/또는 구조

를 쉽게 구상해 낼 수 있을 것이며, 상기 변이 및/또는 변형은 각각 본 발명의 범주 내에 포함되는 것으로 간주된다. 더욱 일반적으로, 관련 기술분야의 통상의 기술자는 본원에 기술된 모든 파라미터, 치수, 물질 및 구성은 예시적인 것임을 의미하며, 실제 파라미터, 치수, 물질 및/또는 구성은 본 발명의 교시(들)가 사용되는 구체적인 적용 또는 적용들에 의존하게 된다는 것을 쉽게 이해할 것이다. 관련 기술분야의 통상의 기술자는 단지 통상의 실험만을 사용하고도, 본원에 기술된 본 발명의 구체적인 실시양태에 대한 다수의 등가물을 쉽게 이해하거나, 또는 확인할 수 있을 것이다. 그러므로, 상기 실시양태는 단지 예시로서 제공되는 것이며, 첨부된 청구범위 및 그의 등가물의 범주 내에서 본 발명은 구체적으로 기술되고, 청구된 것과 다르게 실시될 수 있다는 것을 이해하여야 한다. 본 발명은 본원에 기술된 각각의 개별 특징, 시스템, 물품, 물질, 및/또는 방법에 관한 것이다. 추가로, 상기 특징, 시스템, 물품, 물질, 및/또는 방법이 서로 모순되지 않는다면, 둘 이상의 상기 특징, 시스템, 물품, 물질, 및/또는 방법의 임의의 조합이 본 발명의 범주 내에 포함된다.

[0198] 명세서 중 및 청구범위 중 본원에서 사용되는 바, "단수형태"는 달리 명확하게 그 반대로 명시되지 않는 한, "적어도 하나"를 의미하는 것으로 이해되어야 한다.

[0199] 명세서 중 및 청구범위 중 본원에서 사용되는 바, "및/또는"이라는 어구는 그렇게 결합되어 있는 요소 중 "어느 하나 또는 그 둘 모두," 즉, 일부 경우에서는 공동으로 존재하고, 다른 경우에는 분리되어 존재하는 요소를 의미하는 것으로 이해되어야 한다. 달리 명확하게 그 반대로 명시되지 않는 한, 구체적으로 확인된 요소와 관련이 있든 또는 관련이 없든 상관없이, "및/또는" 절에 의해 구체적으로 확인된 요소 이외의 다른 요소가 임의적으로 존재할 수 있다. 따라서, 비제한적인 예로서, 예컨대, "포함하는"이라는 제한을 두지 않은 개방형의 표현과 함께 결합하여 사용될 때, "A 및/또는 B"라고 언급하는 것은 한 실시양태에서는 (임의적으로 B 이외의 다른 요소를 포함하면서) B 없이 A를 지칭할 수 있고; 또 다른 실시양태에서는 (임의적으로 A 이외의 다른 요소를 포함하면서) A 없이 B를 지칭할 수 있고; 추가의 또 다른 실시양태에서는 (임의적으로 다른 요소를 포함하면서) A 및 B, 둘 모두를 지칭할 수 있으며; 기타의 경우도 존재할 수 있다.

[0200] 명세서 중 및 청구범위 중 본원에서 사용되는 바, "또는"은 상기 정의된 "및/또는"과 동일한 의미를 갖는 것으로 이해되어야 한다. 예를 들어, 목록에서 항목을 분리할 때, "또는" 또는 "및/또는"은 포괄적인 의미로, 즉, 요소의 개수 또는 요소의 목록 중 적어도 하나를 포함할 뿐만 아니라, 그 중 1개 초과의 것을 포함하고, 열거되지 않은 추가 항목을 포함하는 것으로 해석되어야 한다. 그 반대로 명확하게 명시된 용어, 예컨대, "~중 단 하나만" 또는 "~중 정확하게 하나," 또는 청구범위에서 사용될 때, "~으로 이루어진"이라는 용어만이 요소의 개수 또는 요소의 목록 중 정확히 하나의 요소를 포함하는 것을 의미할 것이다. 일반적으로, 본원에서 사용되는 바, "또는"이라는 용어는 오직, 배타적인 용어, 예컨대, "둘 중 어느 하나," "~중 하나," "~중 단 하나만" 또는 "~중 정확하게 하나"라는 용어가 선행될 때 배타적인 대안 (즉, "하나 또는 나머지 다른 하나이지만, 그 둘 모두는 아닌 것")을 나타내는 것으로 해석되어야한다. 청구범위에서 사용될 때, "본질적으로 ~으로 이루어진"이라는 것은 특허법 분야에서 사용되는 바와 같이, 그의 일반적인 의미를 가져야 한다.

[0201] 본원, 본 명세서 및 청구범위에서 사용되는 바, 하나 이상의 요소의 목록과 관련하여, "적어도 하나의"라는 어구는, 반드시 요소의 목록 내에 구체적으로 열거된 각각의 모든 요소 중 적어도 하나를 포함하여야 하는 것이 아니라, 요소의 목록에서 요소의 임의의 조합을 배제하는 것이 아니라, 적어도 하나의 요소가 요소의 목록에서 요소 중 임의의 하나 이상의 것으로부터 선택된다는 것을 의미하는 것으로 이해되어야 한다. 상기 정의는 또한, 구체적으로 확인된 요소와 관련이 있든 또는 없든 그에 상관없이, "적어도 하나의"라는 어구가 지칭하는 요소의 목록 내에 구체적으로 확인된 요소 이외의 요소가 임의적으로 존재할 수 있다는 것도 허용한다. 따라서, 비제한적인 예로서, "A 및 B 중 적어도 하나의" (또는 동일하게, "A 또는 B 중 적어도 하나의," 또는 동일하게, "A 및/또는 B 중 적어도 하나의")라는 것은 한 실시양태에서는 (임의적으로 B 이외의 다른 요소를 포함하면서) B는 존재하지 않고, 임의적으로 1개 초과의 A를 포함하는 적어도 하나의 A를 지칭할 수 있고; 또 다른 실시양태에서는, (임의적으로 A 이외의 다른 요소를 포함하면서) A는 존재하지 않고, 임의적으로 1개 초과의 B를 포함하는 적어도 하나의 B를 지칭할 수 있고; 추가의 또 다른 실시양태에서, 임의적으로 1개 초과의 A를 포함하는 적어도 하나의 A, 및 임의적으로 1개 초과의 B를 포함하는 적어도 하나의 B (및 다른 요소를 포함하는 것) 등을 칭할 수 있다.

[0202] 상기 명세서 뿐만 아니라, 청구항에서, 예컨대, "포함하는," "비롯한," "보유하는," "갖는," "함유하는," "수반하는," "보유하는" 등의 모든 전환 어구는 제한을 두지 않은 개방형인 것으로, 즉, ~을 포함하나, 이에 제한되지 않는다는 것을 의미하는 것으로 이해되어야 한다. 미국 특허청의 심사 절차에 관한 매뉴얼, 섹션 2111.03(United States Patent Office Manual of Patent Examining Procedures, Section 2111.03)에 기재되어 있는 바와 같이, 오직 "~으로 이루어진" 및 "본질적으로 ~으로 이루어진"이라는 전환 어구만이 각각 폐쇄형 또

는 반-폐쇄형 전환 어구여야 한다.

[0203] 청구항 중의 청구항을 수식하는 서수 용어, 예컨대, "제1," "제2," "제3" 등의 사용은 그 단독으로 한 청구항의 또 다른 청구항 대비의 임의의 우선함, 우위임, 또는 순서, 또는 방법의 조치가 수행되는 시간적 순서를 내포하는 것이 아니라, 이는 단지 청구항을 구별하기 위해 특정명을 갖는 한 청구항을 동일명을 갖는 또 다른 항을 구별하는 (서수 용어의 사용을 위한) 라벨로서 사용된다.

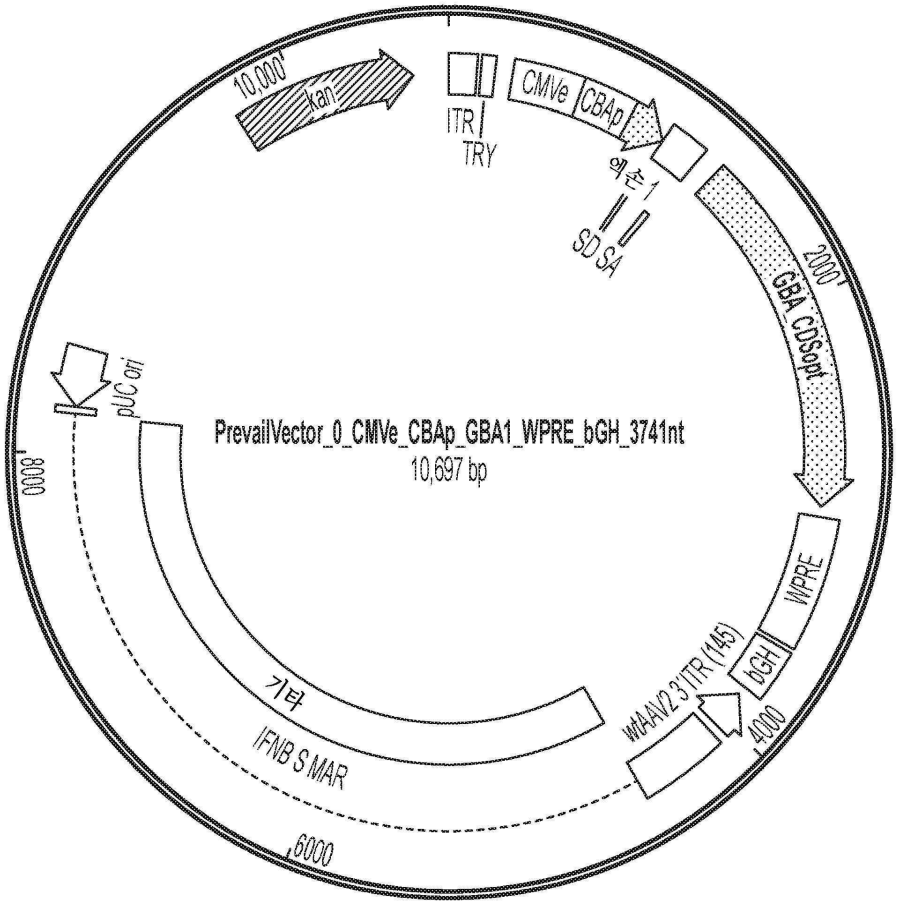
[0204] 명백하게 반대로 명시되지 않는 한, 1개 초과의 단계 또는 조치를 포함하는 본원에서 청구되는 임의의 방법에서, 본 방법의 단계 또는 조치의 순서가 반드시 본 방법의 단계 또는 조치가 언급된 순서로 제한되는 것은 아니라는 것도 또한 이해하여야 한다.

[0205] 서열

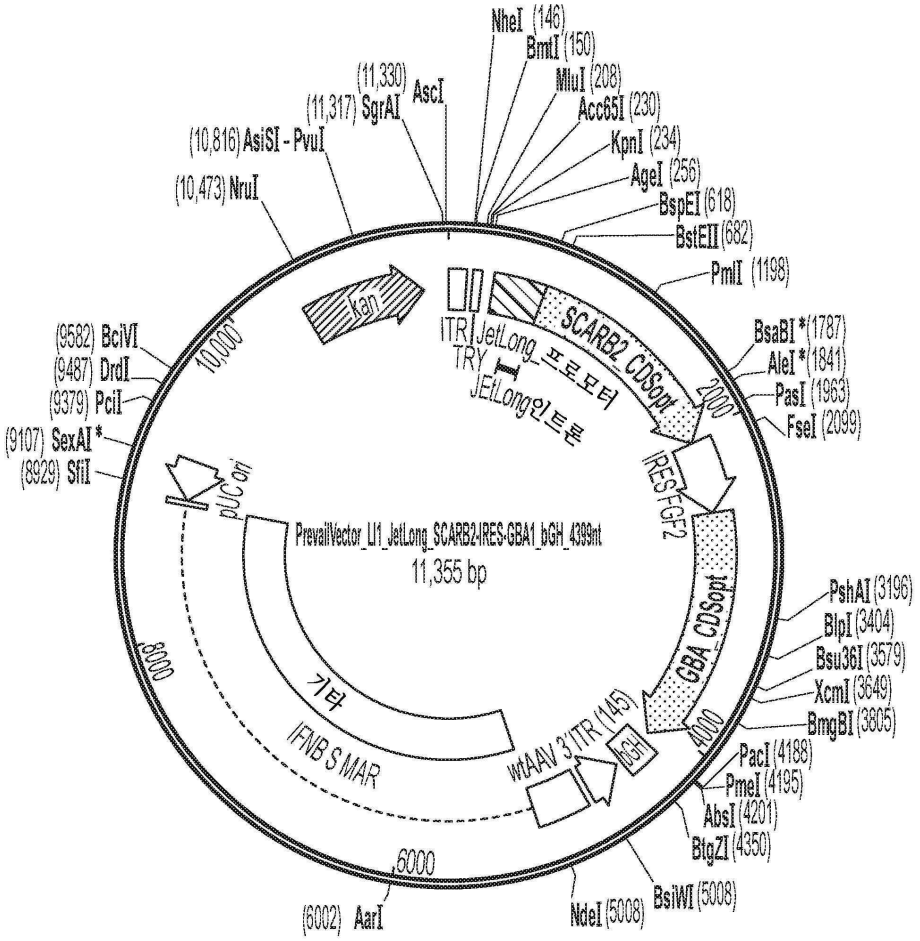
[0206] 일부 실시양태에서, 하나 이상의 유전자 생성물 (예컨대, 제1, 제2 및/또는 제3 유전자 생성물)을 코딩하는 발현 카세트는 서열식별번호 1-78 중 어느 하나에 기재된 서열을 포함하거나 또는 그로 이루어진다 (또는 상기 서열을 갖는 펩티드를 코딩한다). 일부 실시양태에서, 유전자 생성물은 서열식별번호 1-78 중 어느 하나의 일부 (예컨대, 단편)에 의해 코딩된다.

도면

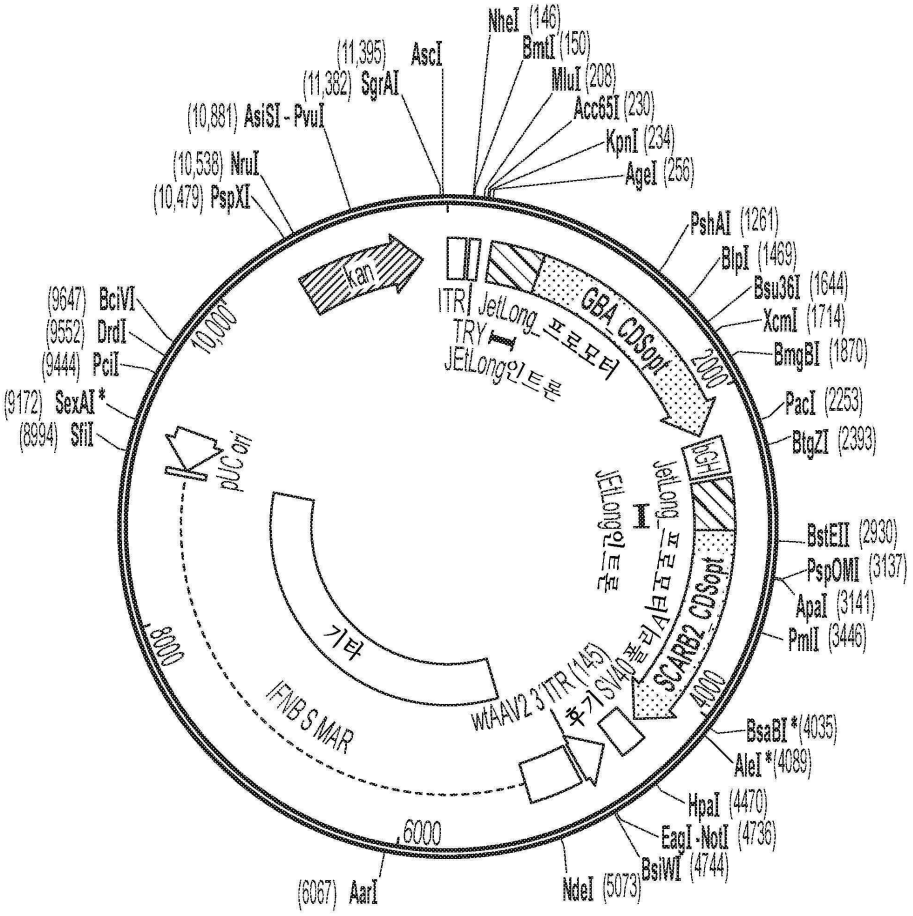
도면1



도면2

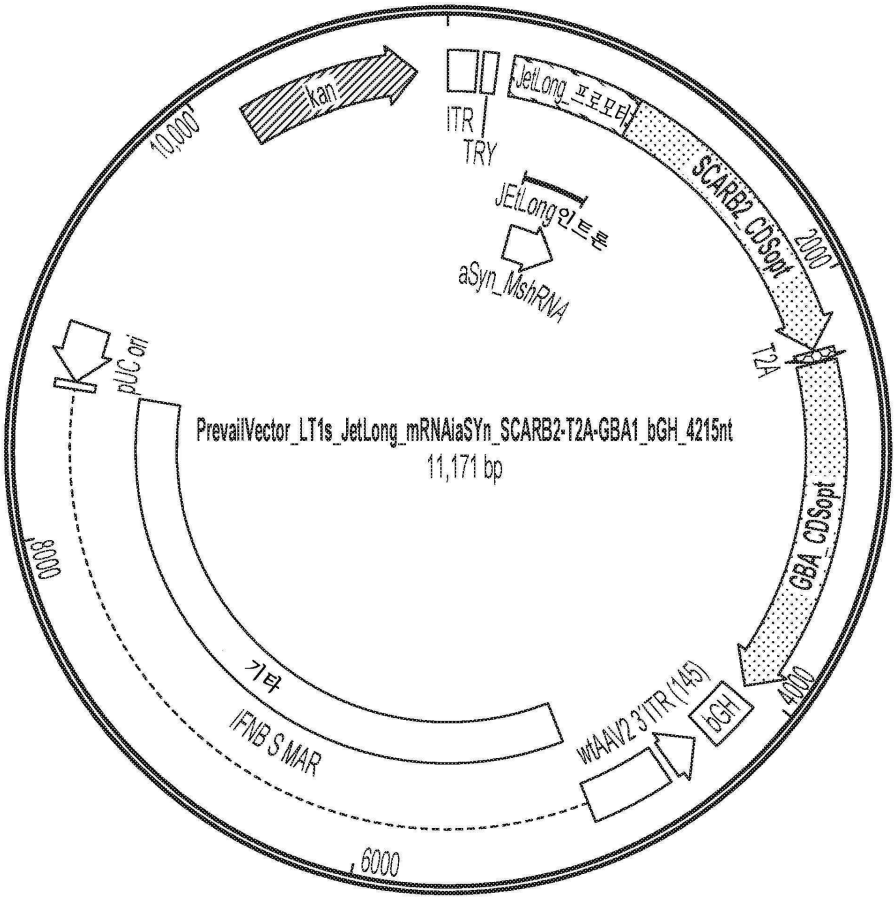

NheI (146)
BmtI (150)
MluI (208)
Acc65I (230)
KpnI (234)
AgeI (256)
BspEI (618)
BstEII (682)
PmlI (1198)
BsaBI* (1787)
AleI* (1841)
PasI (1963)
FseI (2099)
PshAI (3196)
BlpI (3404)
Bsu36I (3579)
XcmI (3649)
BmgBI (3805)
PacI (4188)
PmeI (4195)
AbsI (4201)
BtgZI (4350)
BsiWI (5008)
NdeI (5008)
(6002) AarI
(8929) SfiI
(9107) SexAI*
(9379) PciI
(9487) DrdI
(9582) BciVI
(10,473) NruI
(10,816) AsiSI - PvuI
(11,317) SgrAI
(11,330) AscI
kan
ITR
JetLong_프로모터
TRY
JEtLong인트론
SCARB2 CDSopt
IRES FGF2
GBA CDSopt
bGH
wtAAV 3'ITR (145)
pUC ori
PrevailVector_LI1_JetLong_SCARB2-IRES-GBA1_bGH_4399nt
11,355 bp
기타
IFNB S MAR
2000
4000
6000
8000
10000

도면3

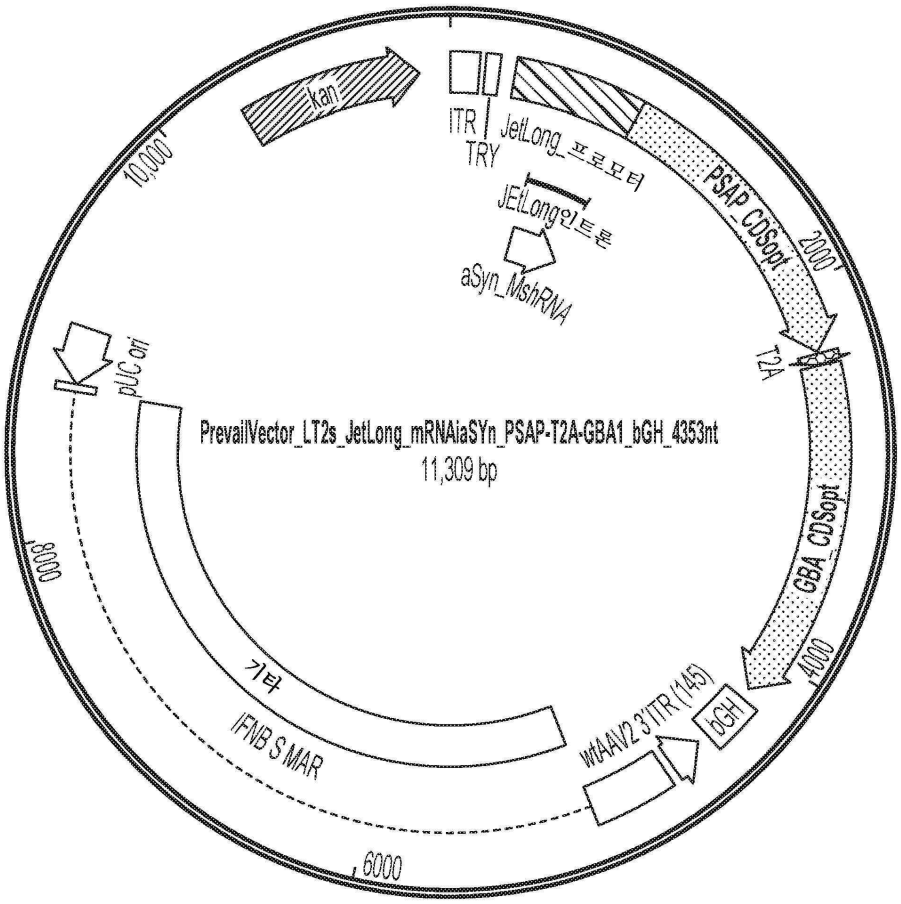


PrevailVector_FP1_JetLong_GBA1_bGH_JetLong_SCARB2_SV40L_4464nt
11,420 bp

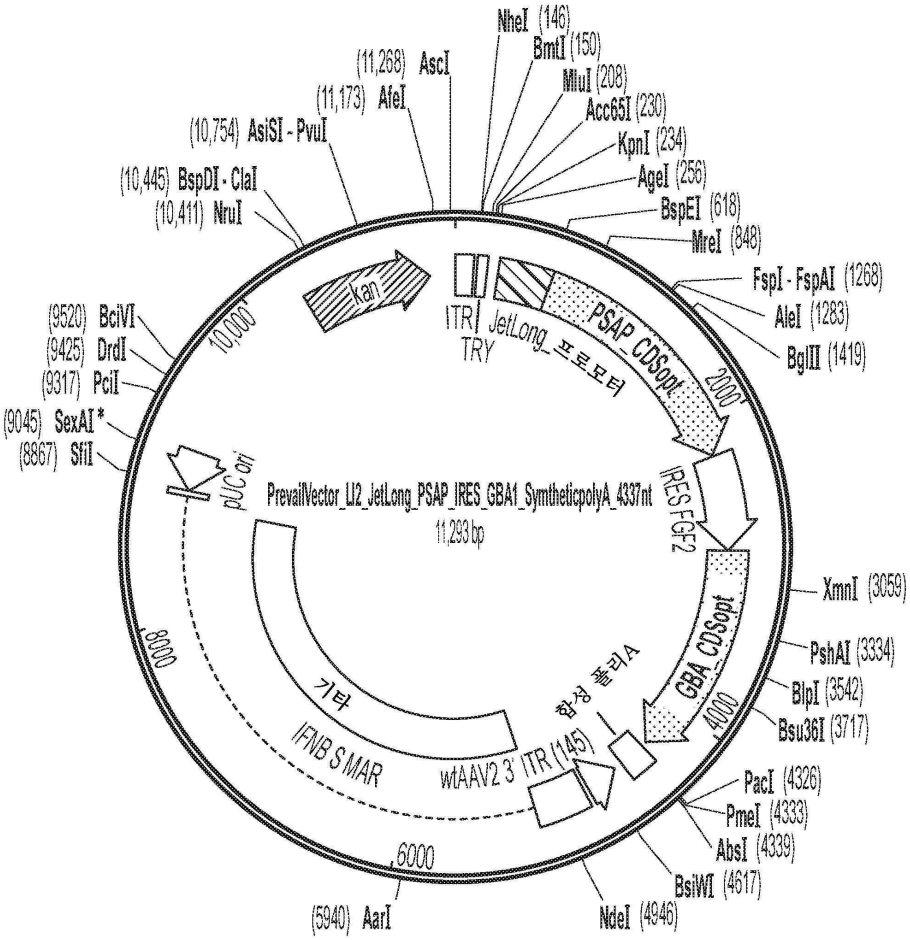
도면4


kan
ITR
TRY
JetLong_프로모터
SCARB2_CDSopt
JEtLong인트론
aSyn_MshRNA
2000
T2A
pUC ori
PrevailVector_LT1s_JetLong_mRNAiaSYn_SCARB2-T2A-GBA1_bGH_4215nt
11,171 bp
10,000
8000
GBA_CDSopt
기타
IFNB S MAR
wtAAV2 3'ITR (145)
bGH
4000
6000

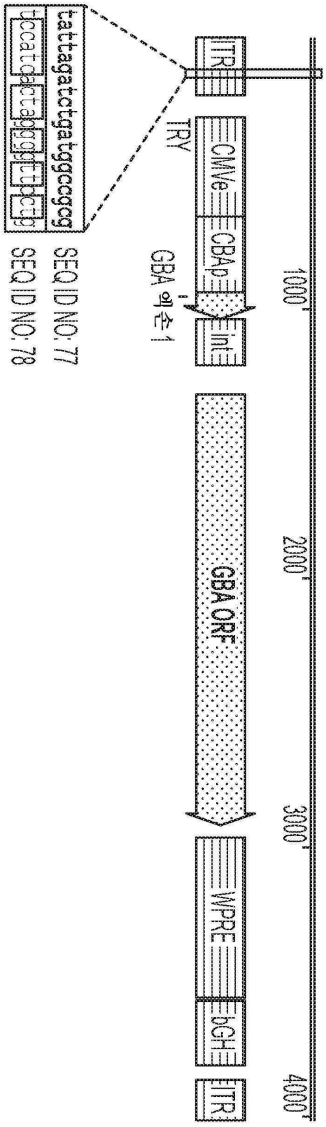
도면5


kan
ITR
TRY
JetLong_프로모터
JEtLong인트론
aSyn_MshRNA
PSAP CDSopt
2000
T2A
GBA CDSopt
4000
bGH
wtAAV2 3'ITR (145)
6000
IFNB S MAR
기타
8000
pUC ori
10,000
PrevailVector_LT2s_JetLong_mRNAiaSYn_PSAP-T2A-GBA1_bGH_4353nt
11,309 bp

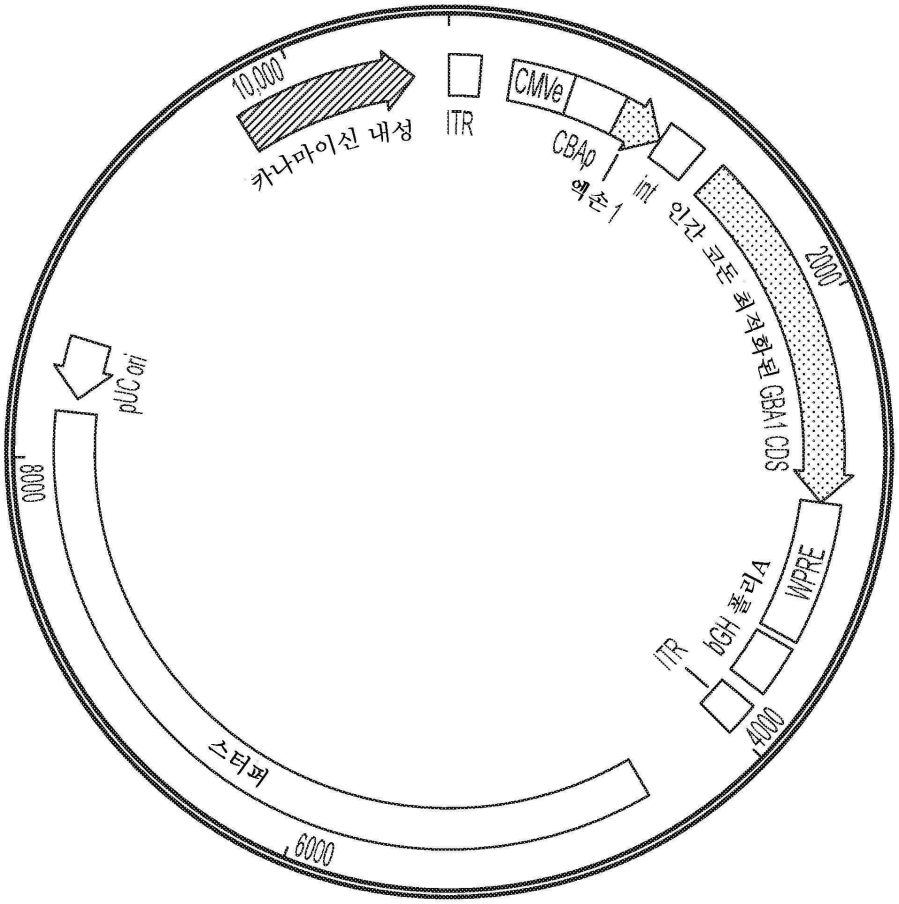
도면6


NheI (146)
BmtI (150)
MluI (208)
Acc65I (230)
KpnI (234)
AgeI (256)
BspEI (618)
MreI (848)
FspI - FspAI (1268)
AleI (1283)
BglII (1419)
(11,268) AscI
(11,173) AfeI
(10,754) AsiSI - PvuI
(10,445) BspDI - ClaI
(10,411) NruI
(9520) BciVI
(9425) DrdI
(9317) PciI
(9045) SexAI*
(8867) SfiI
kan
ITR
TRY
JetLong_프로모터
PSAP_CDSopt
IRES FGF2
GBA_CDSopt
PrevailVector_LI2_JetLong_PSAP_IRES_GBA1_SymtheticpolyA_4337nt
11,293 bp
pUC ori
기타
IFNB S MAR
wtAAV2 3' ITR (145)
합성 폴리A
XmnI (3059)
PshAI (3334)
BlpI (3542)
Bsu36I (3717)
PacI (4326)
PmeI (4333)
AbsI (4339)
BsiWI (4617)
NdeI (4946)
(5940) AarI
10,000
2000
4000
6000
8000

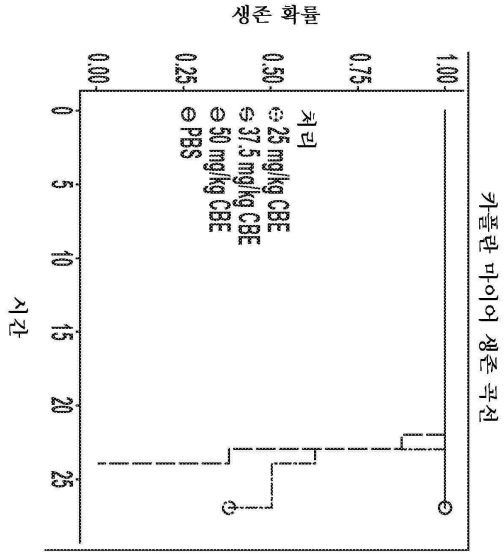
도면7

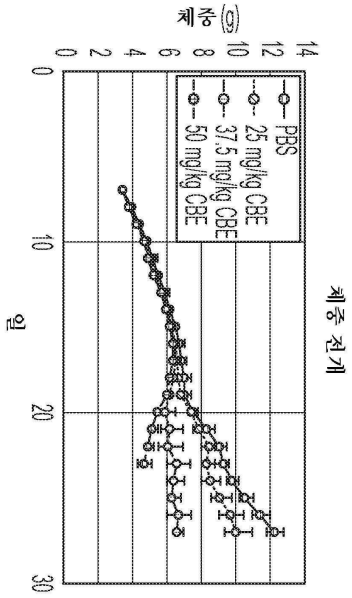

ITR
CMVe
CBAp
intE
GBA ORF
WPRE
bGH
ITR
1000
2000
3000
4000
TRY
GBA 엑손1
tattagatctgatggccggg
SEQ ID NO: 77
SEQ ID NO: 78

도면8

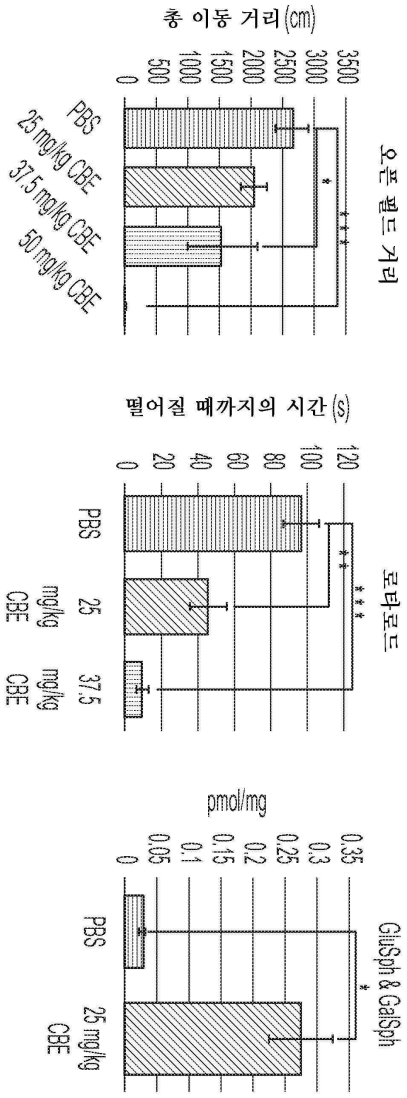

10,000
카나마이신 내성
ITR
CMVe
CBAp
엑손1
int
인간 코돈 최적화된 GBA1 CDS
2000
WPRE
bGH 폴리A
ITR
4000
스터퍼
6000
8000
pUC ori

도면9i

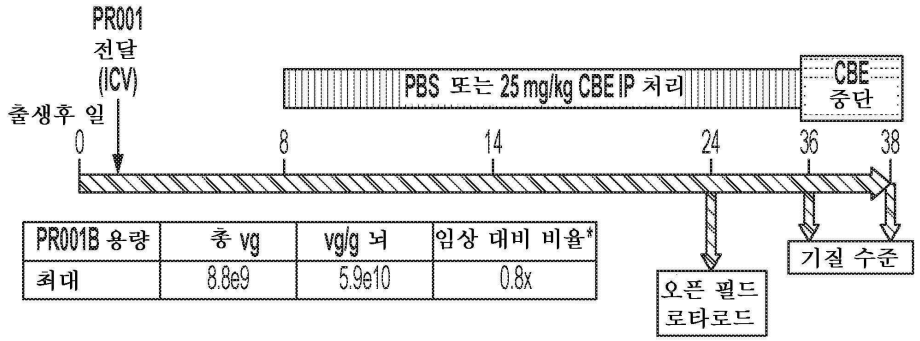

카플란 마이어 생존 곡선
생존 확률
1.00
0.75
0.50
0.25
0.00
처리
25 mg/kg CBE
37.5 mg/kg CBE
50 mg/kg CBE
PBS
0
5
10
15
20
25
시간


체중 진개
체중 (g)
PBS
25 mg/kg CBE
37.5 mg/kg CBE
50 mg/kg CBE
14
12
10
8
6
4
2
0
0
10
20
30
일

도면9ii

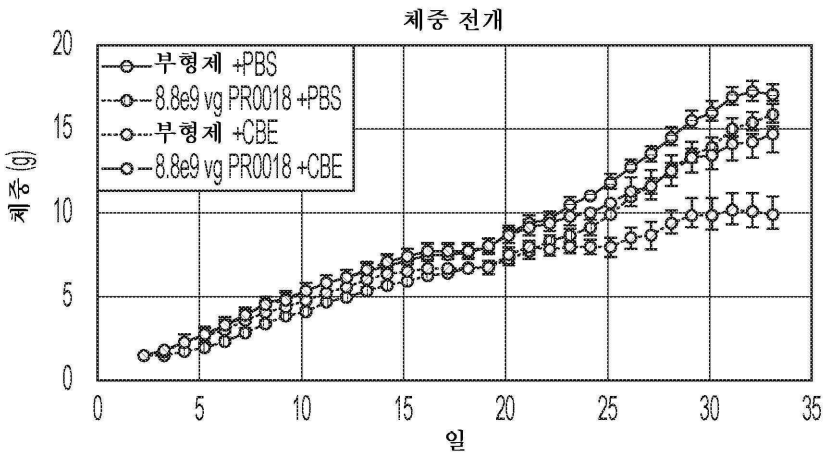


도면10

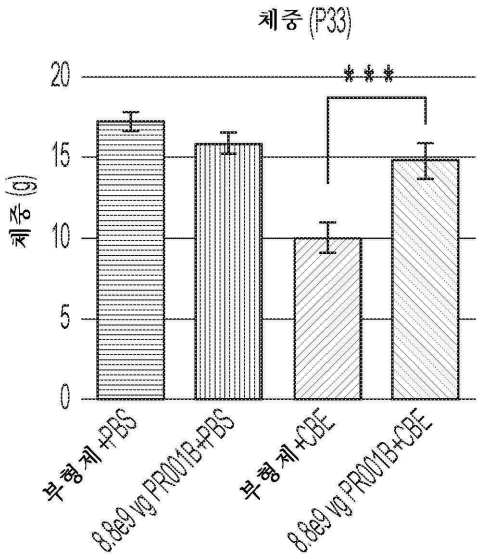


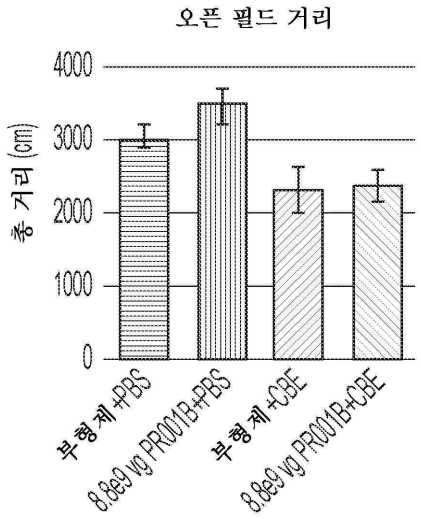
PR001B 용량	총 vg	vg/g 뇌	임상 대비 비율*
최대	8.8e9	5.9e10	0.8x

도면11i

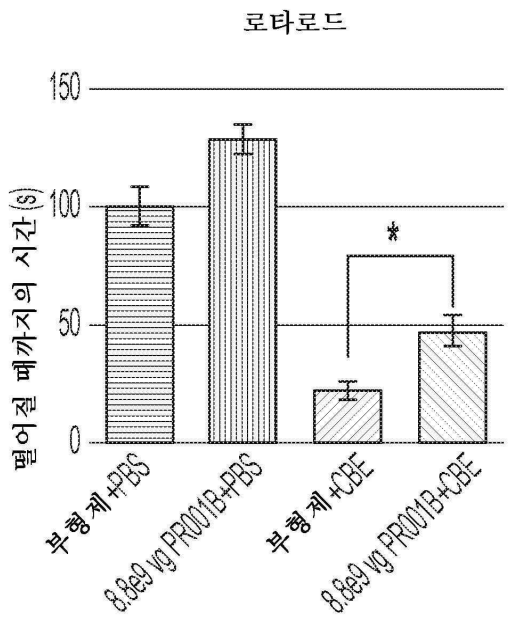


도면11ii

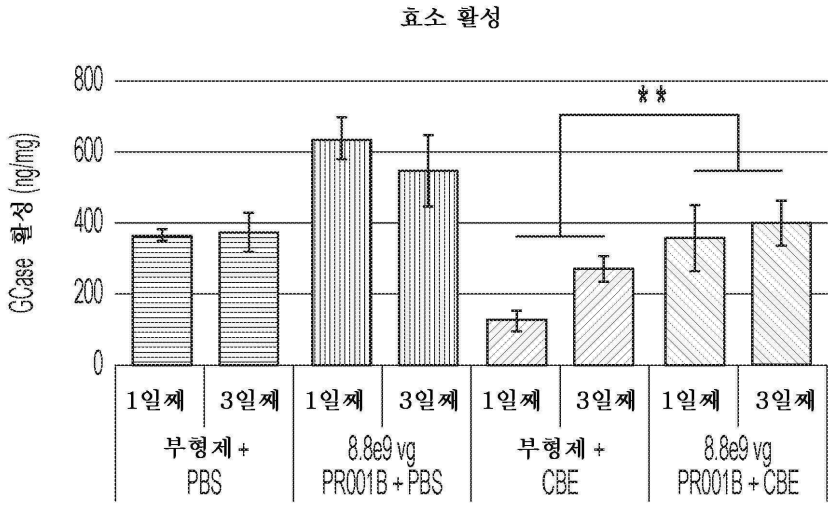




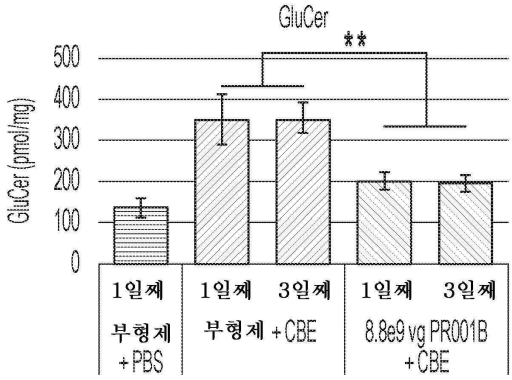
도면11iii

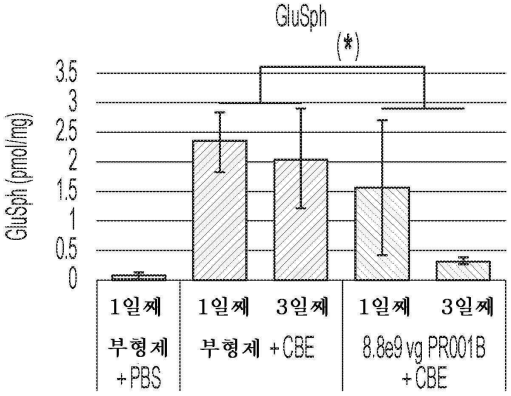


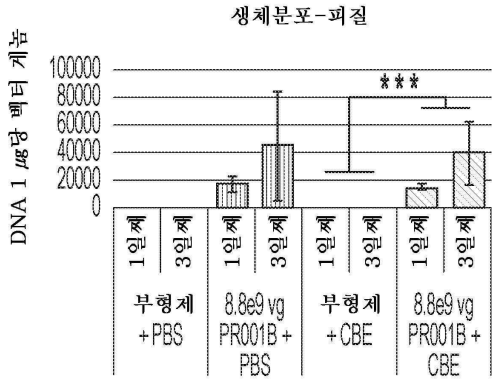
도면12i



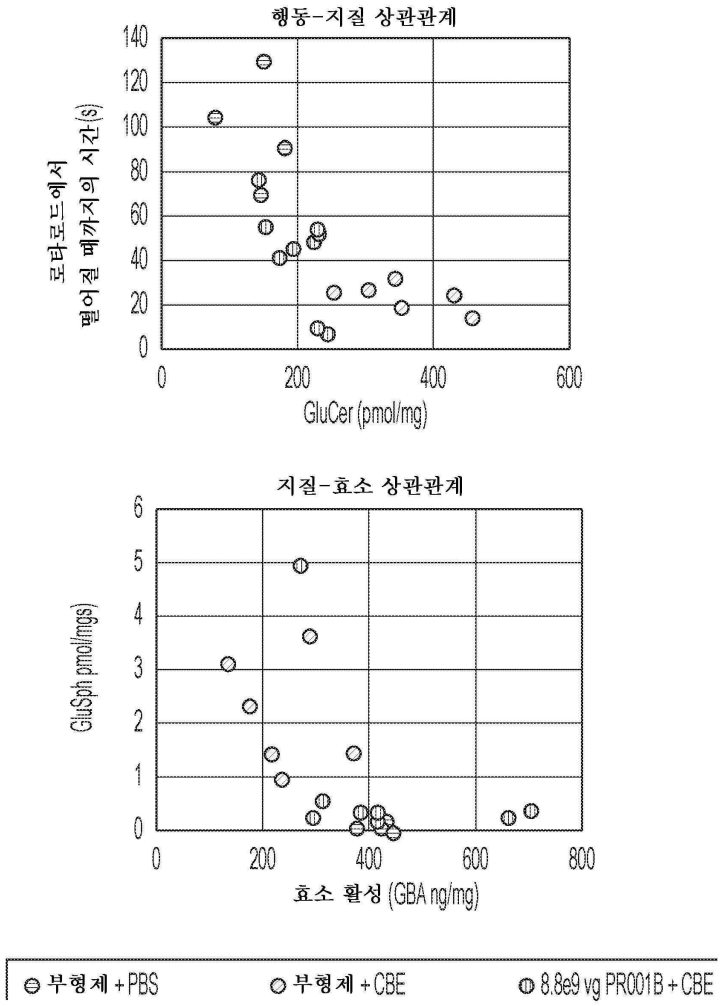
도면12ii


GluCer
**
GluCer (pmol/mg)
500
400
300
200
100
0
1일째
부형제 +PBS
1일째
3일째
부형제 +CBE
1일째
3일째
8.8e9 vg PR001B +CBE

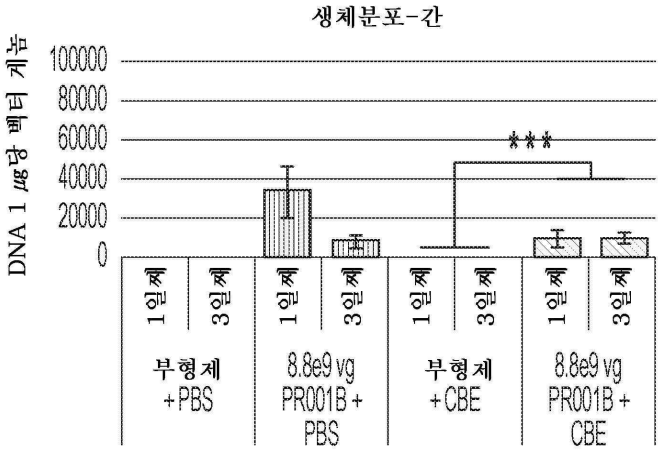

GluSph
(*)
GluSph (pmol/mg)
3.5
3
2.5
2
1.5
1
0.5
0
1일째
부형제 +PBS
1일째
3일째
부형제 +CBE
1일째
3일째
8.8e9 vg PR001B +CBE

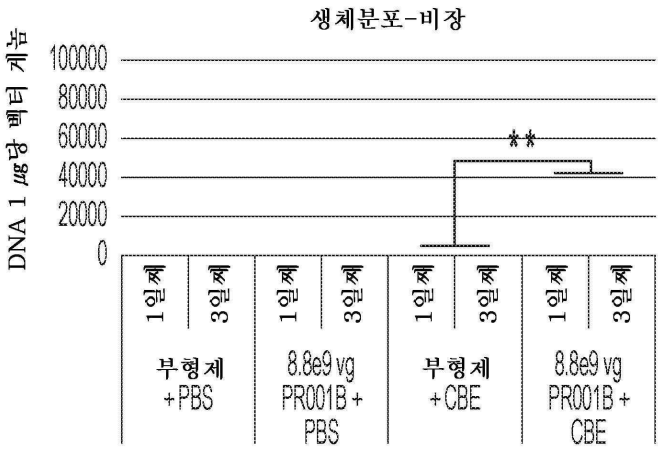

생체분포-피질
DNA 1 μg당 벡터 게놈
100000
80000
60000
40000
20000
0
***
1일째
3일째
부형제 +PBS
1일째
3일째
8.8e9 vg PR001B + PBS
1일째
3일째
부형제 +CBE
1일째
3일째
8.8e9 vg PR001B + CBE

도면13

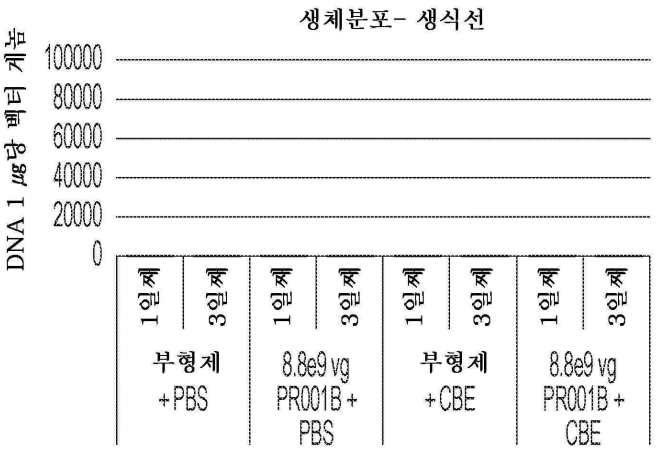

행동-지질 상관관계
로타로드에서
떨어질 때까지의 시간(s)
140
120
100
80
60
40
20
0
0
200
400
600
GluCer (pmol/mg)
지질-효소 상관관계
GluSph pmol/mgs)
6
5
4
3
2
1
0
0
200
400
600
800
효소 활성 (GBA ng/mg)
부형제 + PBS
부형제 + CBE
8.8e9 vg PR001B + CBE

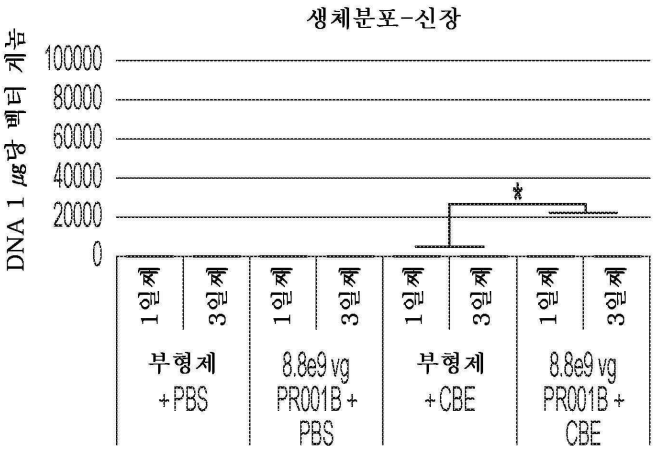
도면14i


생체분포-간
DNA 1 μg당 벡터 게놈
100000
80000
60000
40000
20000
0
***
1일째
3일째
1일째
3일째
1일째
3일째
1일째
3일째
부형제 +PBS
8.8e9 vg PR001B + PBS
부형제 +CBE
8.8e9 vg PR001B + CBE

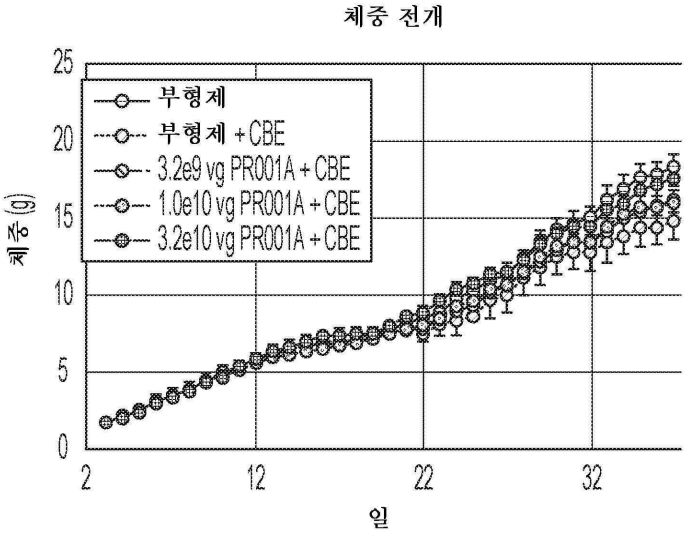

생체분포-비장
DNA 1 μg당 벡터 게놈
100000
80000
60000
40000
20000
0
**
1일째
3일째
1일째
3일째
1일째
3일째
1일째
3일째
부형제 +PBS
8.8e9 vg PR001B + PBS
부형제 +CBE
8.8e9 vg PR001B + CBE

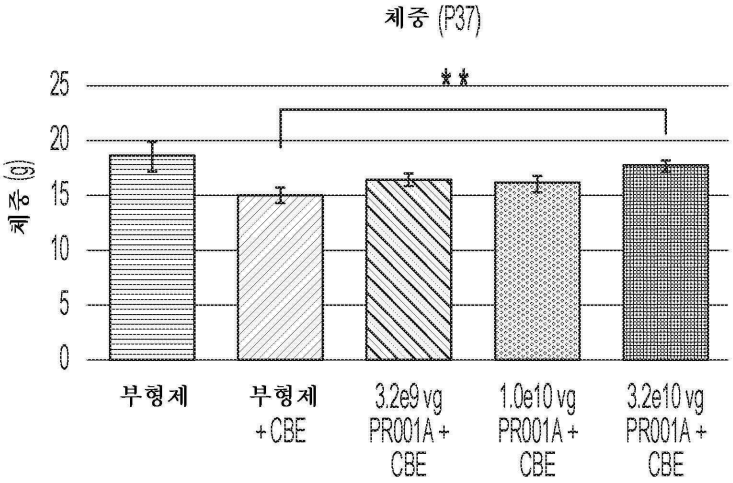
도면14ii


생체분포- 생식선
DNA 1 μg당 벡터 게놈
100000
80000
60000
40000
20000
0
1일째
3일째
1일째
3일째
1일째
3일째
1일째
3일째
부형제 +PBS
8.8e9 vg PR001B + PBS
부형제 +CBE
8.8e9 vg PR001B + CBE

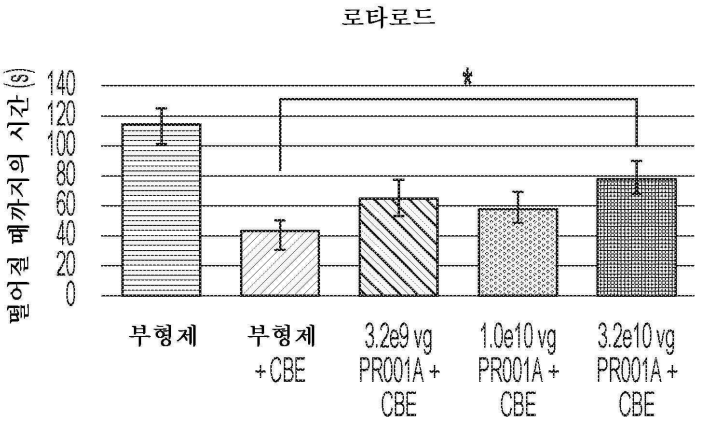

생체분포-신장
DNA 1 μg당 벡터 게놈
100000
80000
60000
40000
20000
0
*
1일째
3일째
1일째
3일째
1일째
3일째
1일째
3일째
부형제 +PBS
8.8e9 vg PR001B + PBS
부형제 +CBE
8.8e9 vg PR001B + CBE

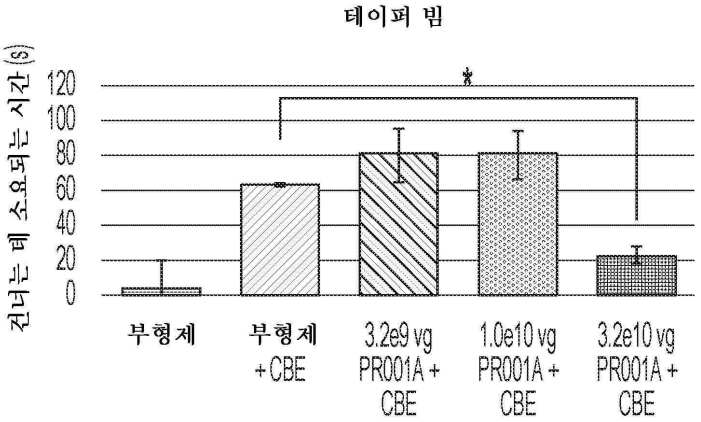
도면15i


체중 전개
부형제
부형제 +CBE
3.2e9 vg PR001A + CBE
1.0e10 vg PR001A + CBE
3.2e10 vg PR001A + CBE
체중 (g)
일
0
5
10
15
20
25
2
12
22
32

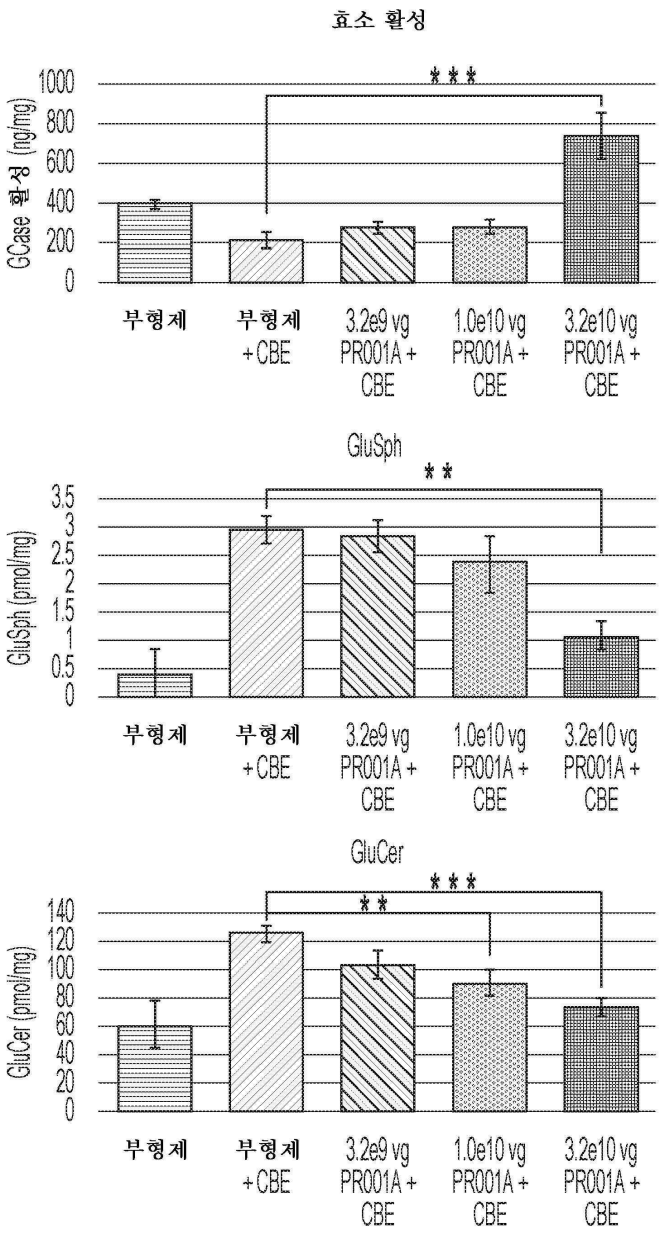

체중 (P37)
체중 (g)
**
0
5
10
15
20
25
부형제
부형제 +CBE
3.2e9 vg PR001A + CBE
1.0e10 vg PR001A + CBE
3.2e10 vg PR001A + CBE

도면15ii

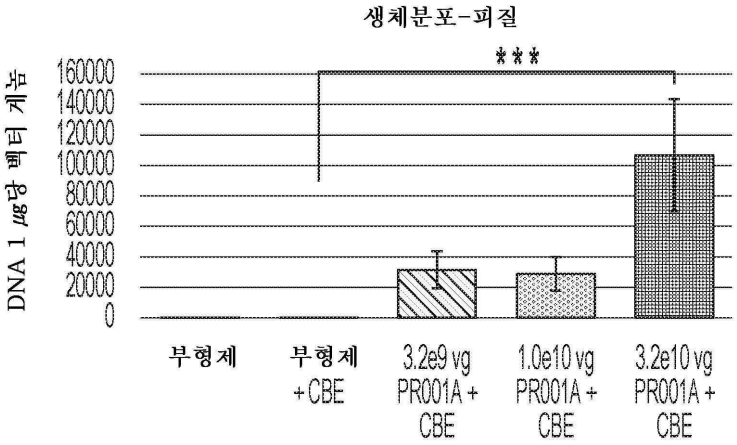

로타로드
떨어질 때까지의 시간 (s)
140
120
100
80
60
40
20
0
*
부형제
부형제 +CBE
3.2e9 vg PR001A + CBE
1.0e10 vg PR001A + CBE
3.2e10 vg PR001A + CBE

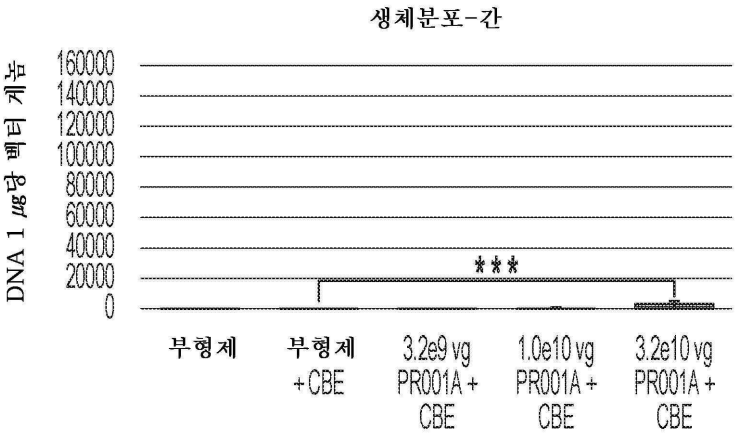

테이퍼 빔
건너는 데 소요되는 시간 (s)
120
100
80
60
40
20
0
*
부형제
부형제 +CBE
3.2e9 vg PR001A + CBE
1.0e10 vg PR001A + CBE
3.2e10 vg PR001A + CBE

도면16i

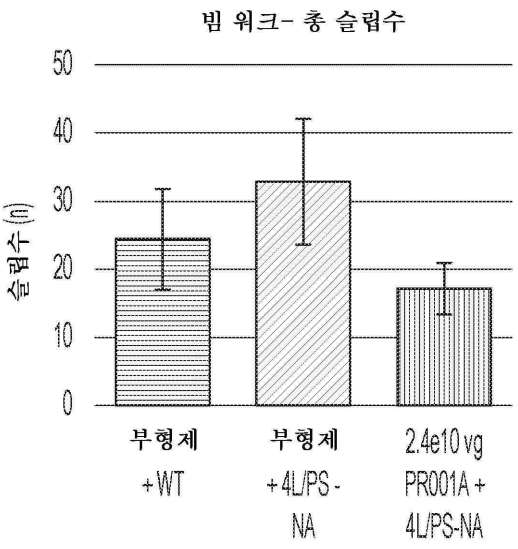

효소 활성
GCase 활성 (ng/mg)
1000
800
600
400
200
0
***
부형제
부형제 +CBE
3.2e9 vg PR001A + CBE
1.0e10 vg PR001A + CBE
3.2e10 vg PR001A + CBE
GluSph
GluSph (pmol/mg)
3.5
3
2.5
2
1.5
1
0.5
0
**
부형제
부형제 +CBE
3.2e9 vg PR001A + CBE
1.0e10 vg PR001A + CBE
3.2e10 vg PR001A + CBE
GluCer
GluCer (pmol/mg)
140
120
100
80
60
40
20
0
***
**
부형제
부형제 +CBE
3.2e9 vg PR001A + CBE
1.0e10 vg PR001A + CBE
3.2e10 vg PR001A + CBE

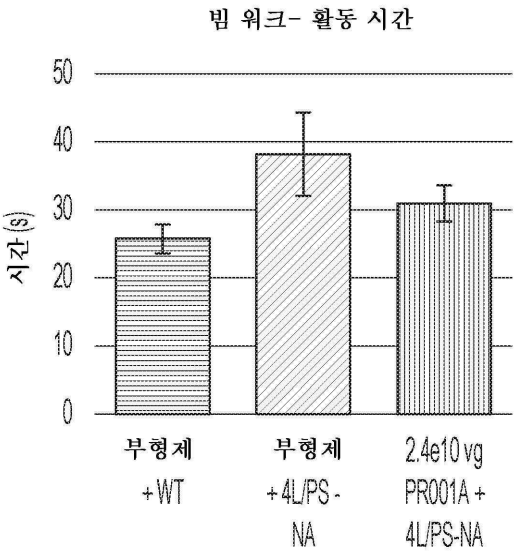
도면16ii


생체분포-피질
DNA 1 μg당 벡터 게놈
160000
140000
120000
100000
80000
60000
40000
20000
0
***
부형제
부형제 +CBE
3.2e9 vg PR001A + CBE
1.0e10 vg PR001A + CBE
3.2e10 vg PR001A + CBE

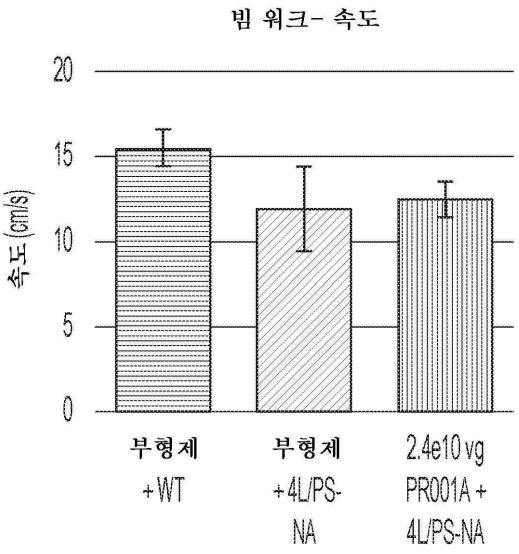

생체분포-간
DNA 1 μg당 벡터 게놈
160000
140000
120000
100000
80000
60000
40000
20000
0
***
부형제
부형제 +CBE
3.2e9 vg PR001A + CBE
1.0e10 vg PR001A + CBE
3.2e10 vg PR001A + CBE

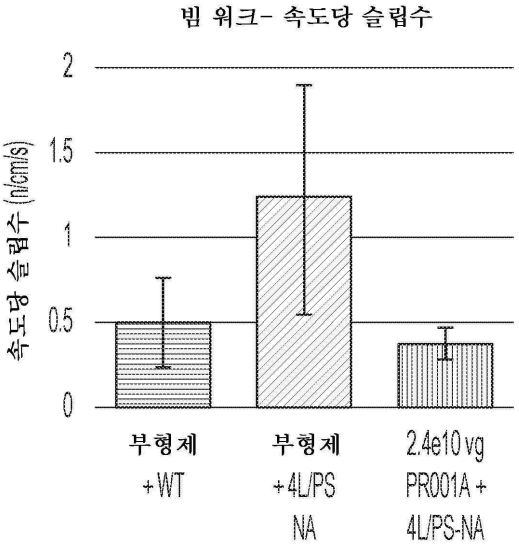
도면17i


빔 워크- 총 슬립수
슬립수 (n)
50
40
30
20
10
0
부형제 +WT
부형제 +4L/PS-NA
2.4e10 vg PR001A + 4L/PS-NA

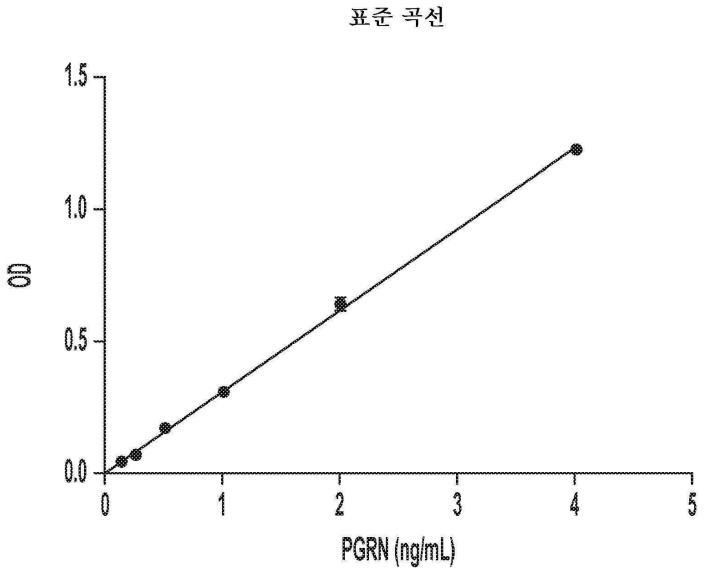

빔 워크- 활동 시간
시간 (s)
50
40
30
20
10
0
부형제 +WT
부형제 +4L/PS-NA
2.4e10 vg PR001A + 4L/PS-NA

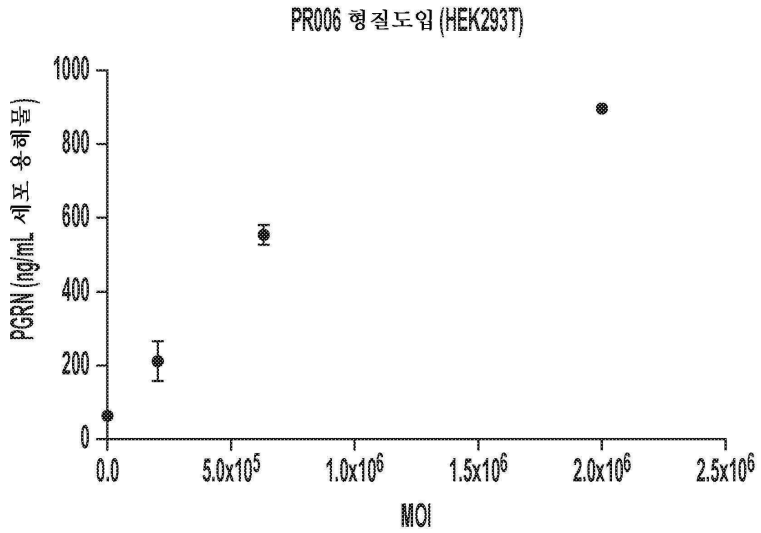
도면17ii


빔 워크- 속도
20
15
10
5
0
속도 (cm/s)
부형제 +WT
부형제 +4L/PS-NA
2.4e10 vg PR001A + 4L/PS-NA

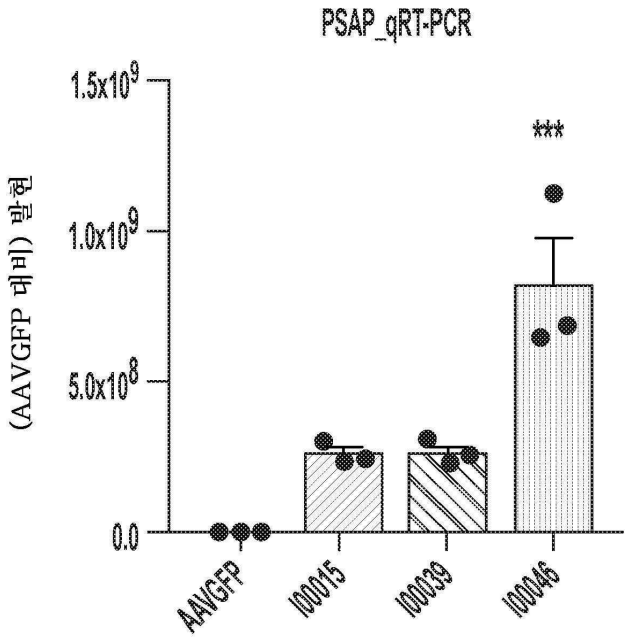

빔 워크- 속도당 슬립수
2
1.5
1
0.5
0
속도당 슬립수 (n/cm/s)
부형제 +WT
부형제 +4L/PS NA
2.4e10 vg PR001A + 4L/PS-NA

도면18

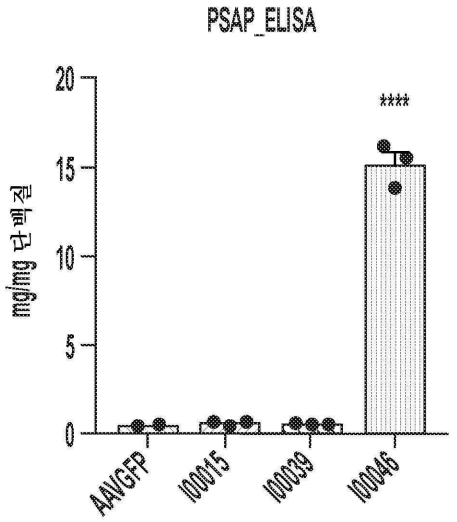

표준 곡선
OD
1.5
1.0
0.5
0.0
0
1
2
3
4
5
PGRN (ng/mL)

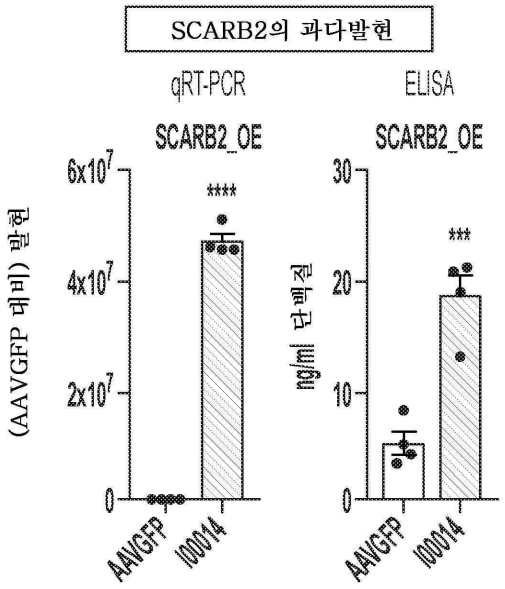

PR006 형질도입 (HEK293T)
PGRN (ng/mL 세포 용해물)
1000
800
600
400
200
0
0.0
5.0x10^5
1.0x10^6
1.5x10^6
2.0x10^6
2.5x10^6
MOI

도면19i

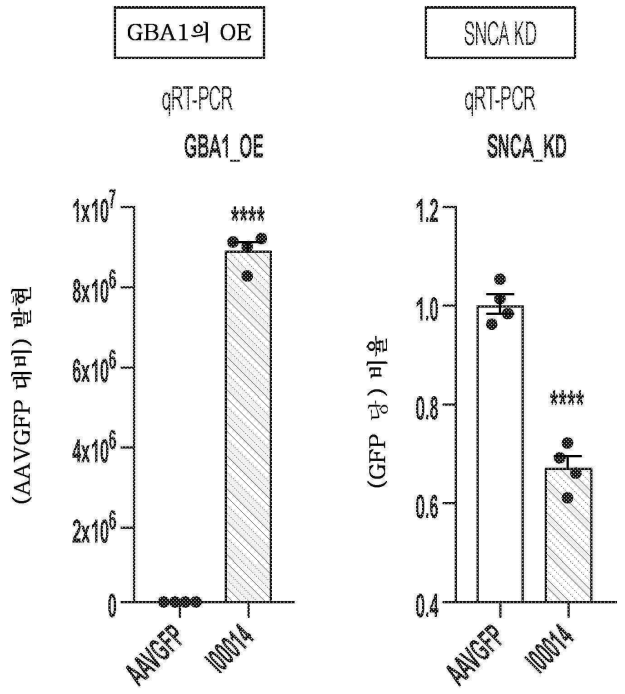

PSAP_qRT-PCR
1.5x10^9
1.0x10^9
5.0x10^8
0.0
(AAVGFP 대비) 발현
***
AAVGFP
I00015
I00039
I00046

도면19ii

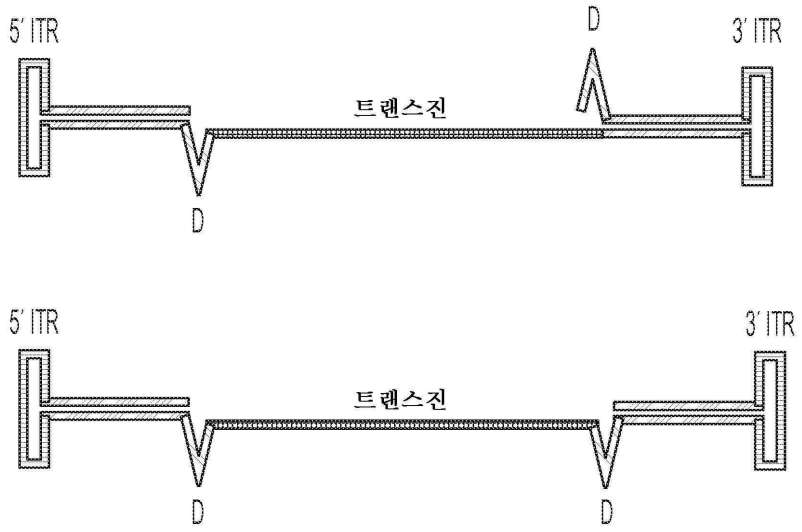

PSAP_ELISA
mg/mg 단백질
20
15
10
5
0
****
AAVGFP
I00015
I00039
I00046


SCARB2의 과다발현
qRT-PCR
ELISA
SCARB2_OE
SCARB2_OE
(AAVGFP 대비) 발현
6x10^7
4x10^7
2x10^7
0
****
ng/ml 단백질
30
20
10
0
***
AAVGFP
I00014
AAVGFP
I00014

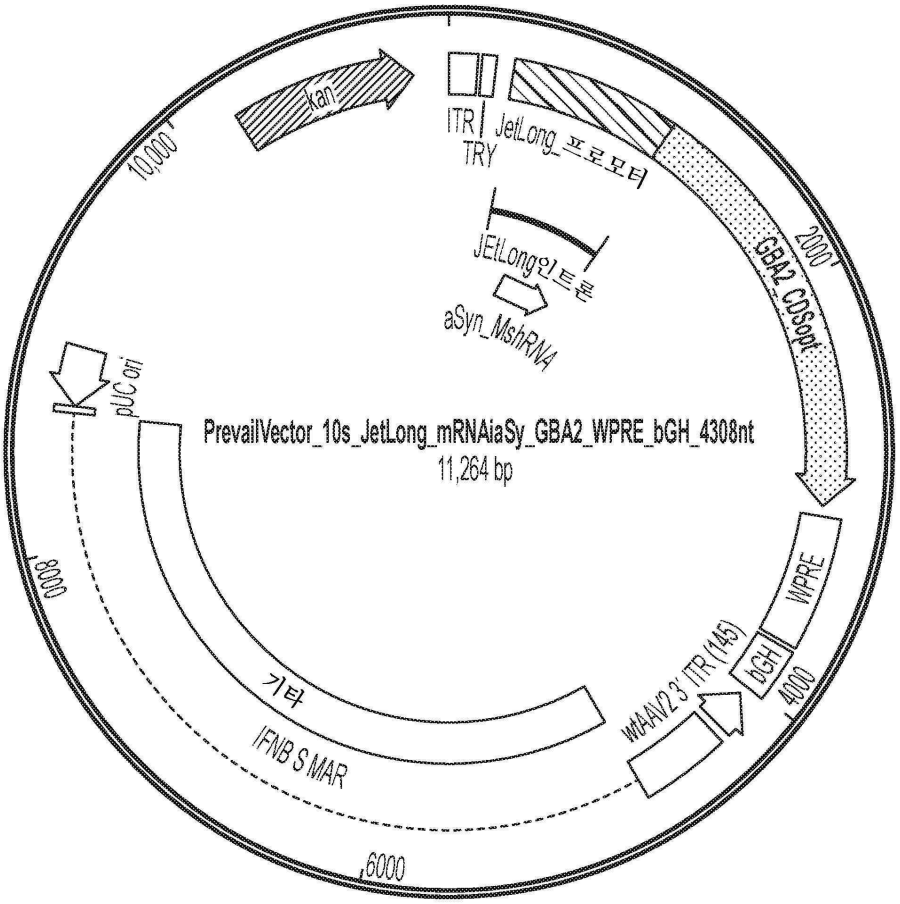
도면19iii



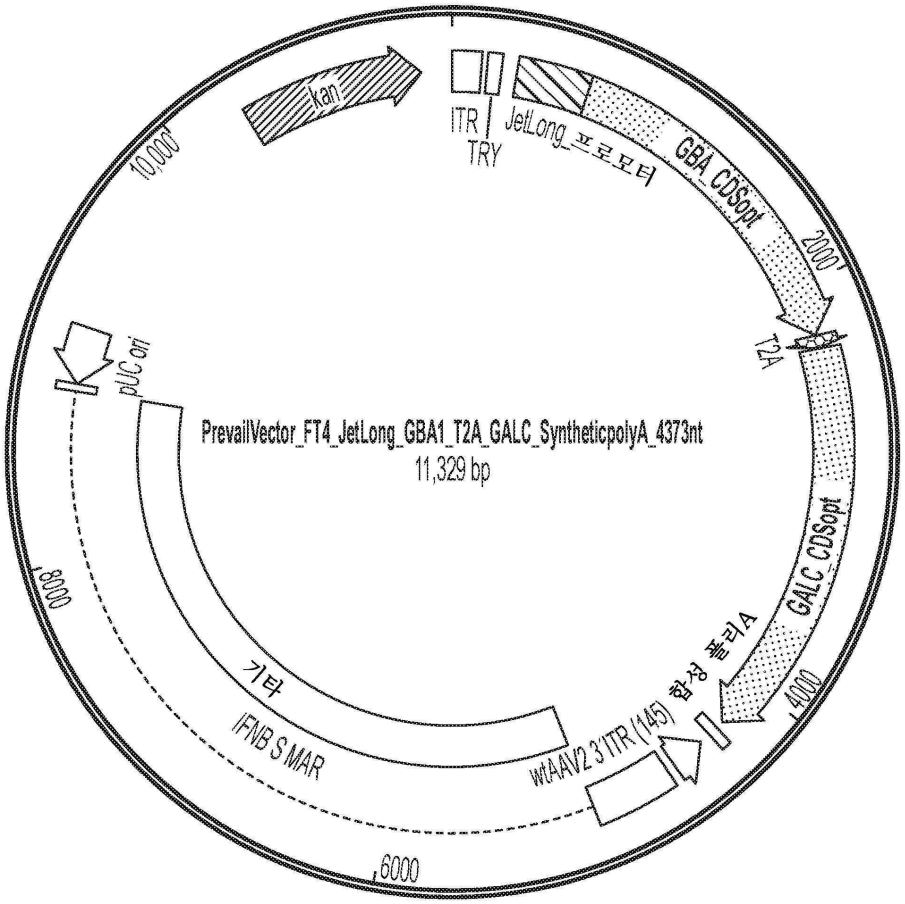
도면20



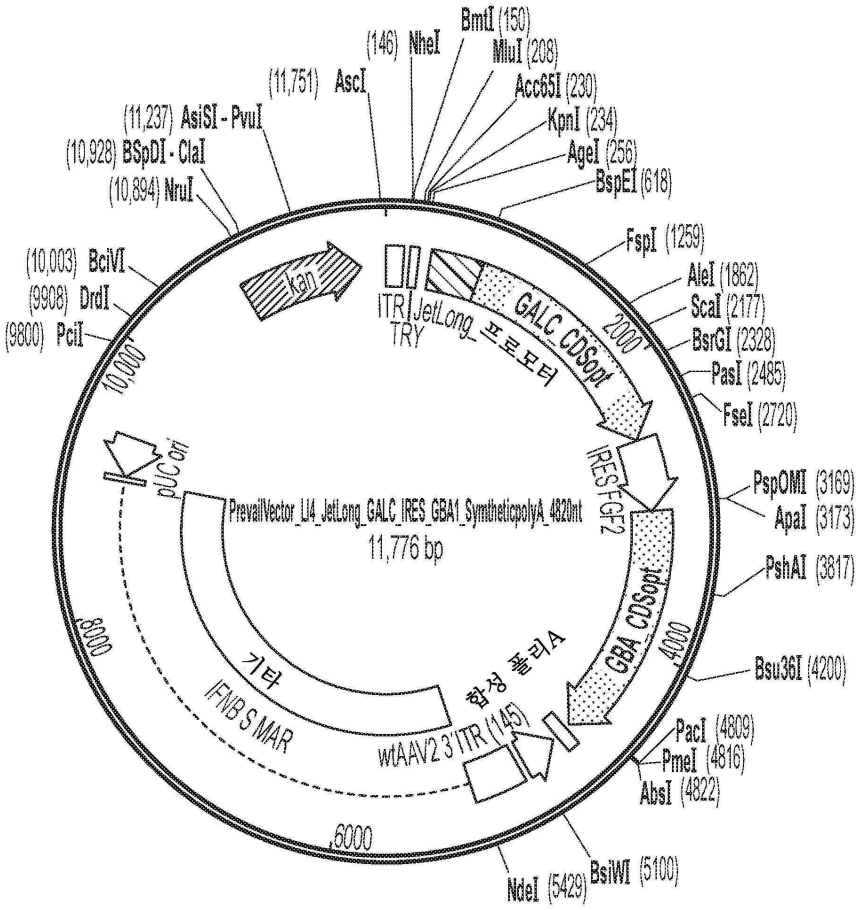
도면21


kan
ITR
TRY
JetLong_프로모터
JEtLong인트론
aSyn_MshRN4
GBA2 COSopt
2000
10,000
pUC ori
PrevailVector_10s_JetLong_mRNAiaSy_GBA2_WPRE_bGH_4308nt
11,264 bp
WPRE
bGH
4000
wtAAV2 3' ITR (145)
기타
IFNB S MAR
8000
6000

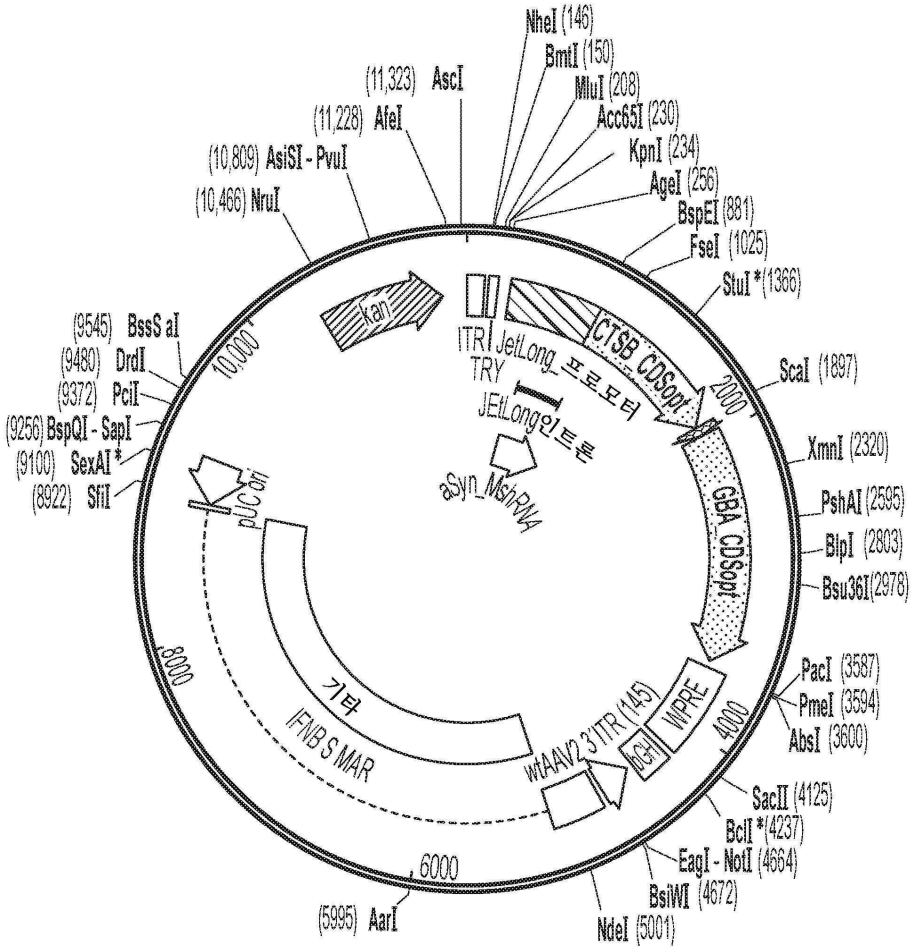
도면22


kan
ITR
TRY
JetLong_프로모터
GBA_CDSopt
2000
T2A
GALC_CDSopt
4000
합성 폴리A
wtAAV2 3'ITR (145)
6000
IFNB S MAR
기타
8000
pUC ori
10,000
PrevailVector_FT4_JetLong_GBA1_T2A_GALC_SyntheticpolyA_4373nt
11,329 bp

도면23

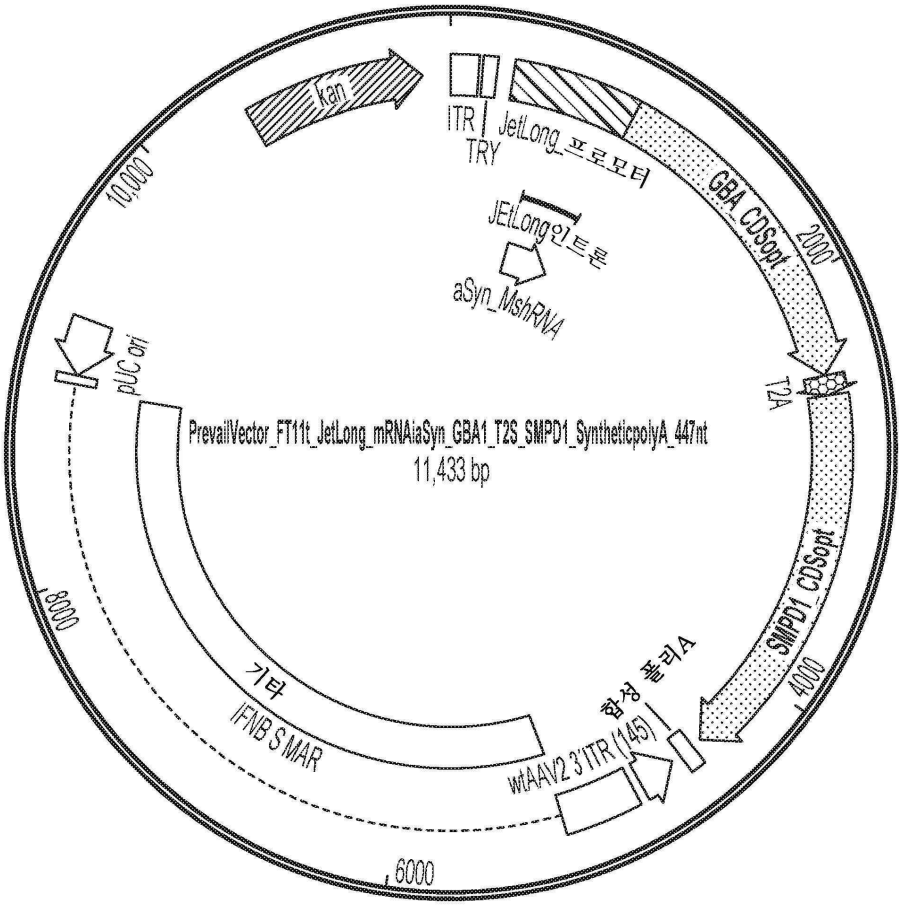

BmtI (150)
(146) NheI
MluI (208)
(11,751) AscI
Acc65I (230)
(11,237) AsiSI - PvuI
KpnI (234)
(10,928) BSpDI - ClaI
AgeI (256)
(10,894) NruI
BspEI (618)
FspI (1259)
(10,003) BciVI
AleI (1862)
(9908) DrdI
ScaI (2177)
(9800) PciI
BsrGI (2328)
PasI (2485)
FseI (2720)
PspOMI (3169)
ApaI (3173)
PshAI (3817)
Bsu36I (4200)
PacI (4809)
PmeI (4816)
AbsI (4822)
BsiWI (5100)
NdeI (5429)
kan
ITR
TRY
JetLong_프로모터
GALC_CDSopt
IRES FGF2
GBA_CDSopt
pUC ori
PrevailVector_LI4_JetLong_GALC_IRES_GBA1_SymtheticpolyA_4820int
11,776 bp
기타
IFNB S MAR
합성 폴리A
wtAAV2 3'ITR (145)
2000
4000
6000
8000
10,000

도면24

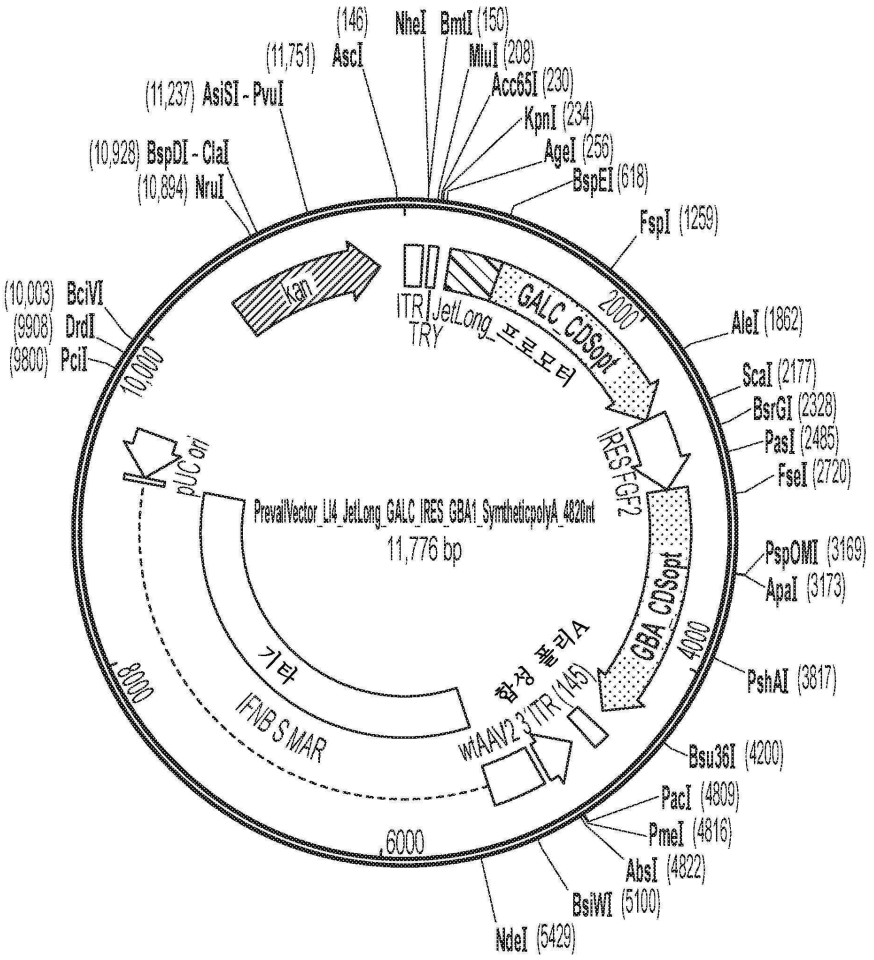


PrevailVector_LT5s_JetLong_mRNAiaSyn_CTSB-T2A-GBA1_WPRE_bGH_4392nt
11,348 bp

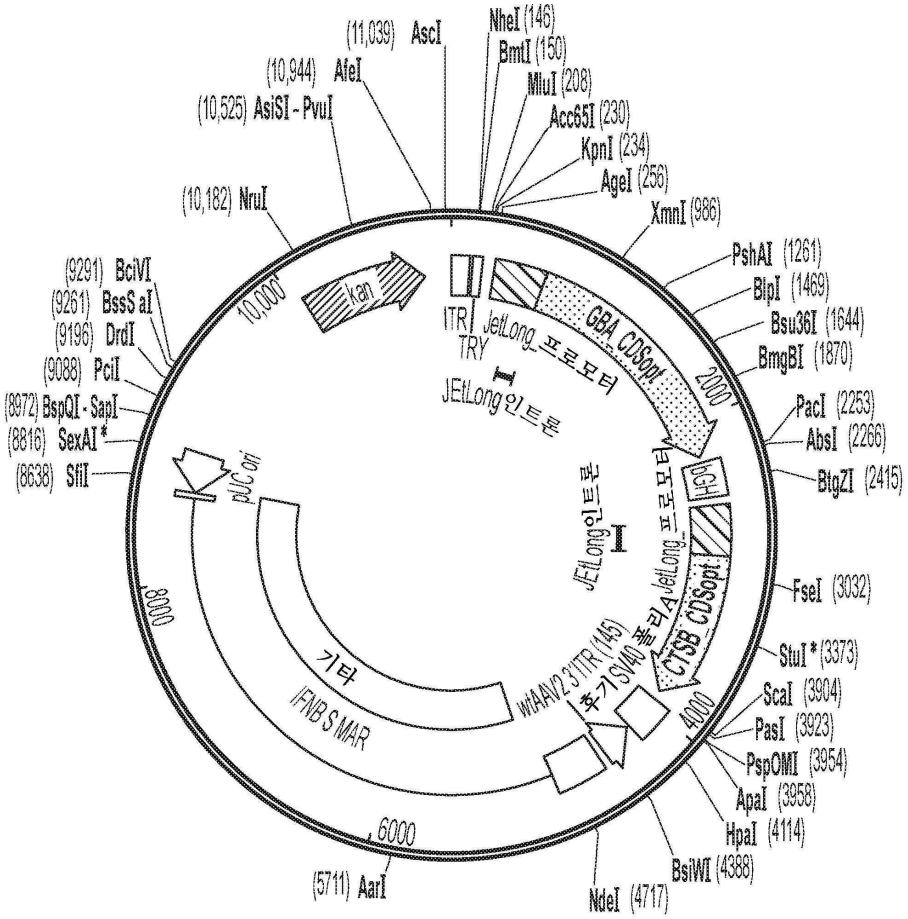
도면25


Kan
ITR
TRY
JetLong_프로모터
GBA_CDSopt
2000
JEtLong인트론
aSyn_MshRNA
pUC ori
T2A
PrevailVector_FT11t_JetLong_mRNAiaSyn_GBA1_T2S_SMPD1_SyntheticpolyA_447nt
11,433 bp
10,000
8000
SMPD1_CDSopt
4000
기타
IFNB S MAR
합성 폴리A
wtAAV2 3'ITR (145)
6000

도면26

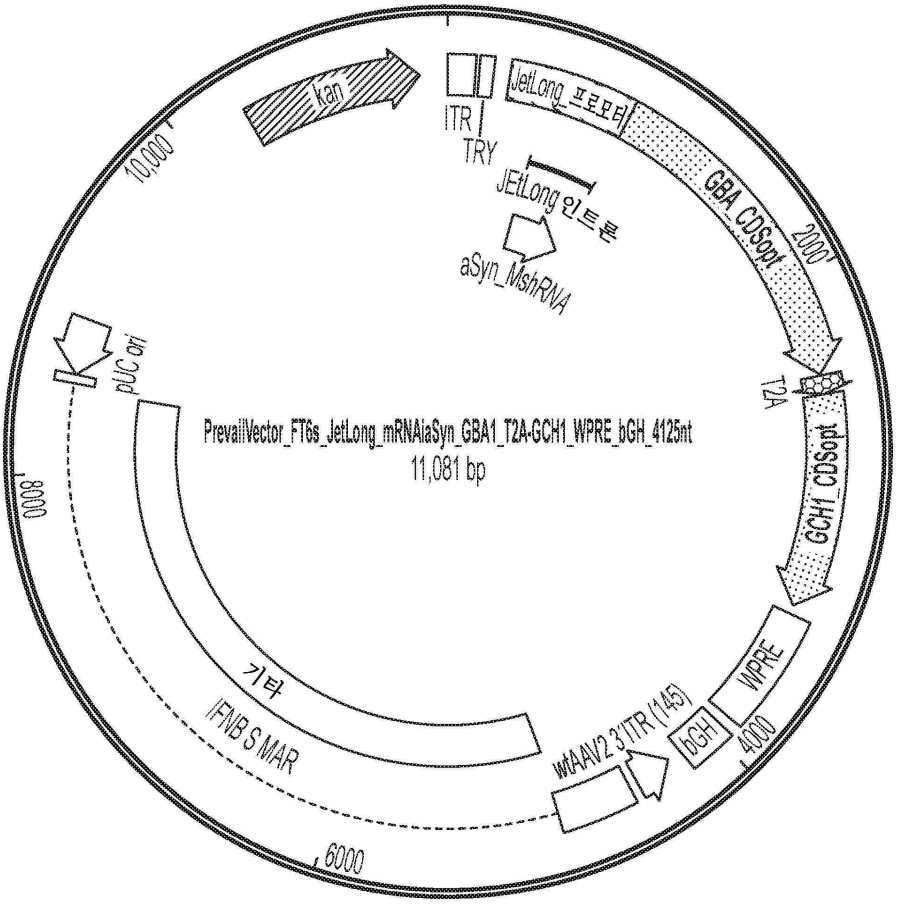

(146) NheI
BmtI (150)
(11,751) AscI
MluI (208)
(11,237) AsiSI - PvuI
Acc65I (230)
KpnI (234)
(10,928) BspDI - ClaI
AgeI (256)
(10,894) NruI
BspEI (618)
FspI (1259)
(10,003) BciVI
(9908) DrdI
(9800) PciI
kan
ITR
TRY
JetLong_프로모터
GALC CDSopt
2000
AleI (1862)
10,000
ScaI (2177)
BsrGI (2328)
PasI (2485)
FseI (2720)
pUC ori
IRES FGF2
PrevailVector_LI4_JetLong_GALC_IRES_GBA1_SyntheticpolyA_4820int
11,776 bp
PspOMI (3169)
ApaI (3173)
GBA CDSopt
4000
합성 폴리A
기타
PshAI (3817)
8000
IFNB S MAR
wtAAV2 3'ITR (145)
Bsu36I (4200)
PacI (4809)
PmeI (4816)
6000
AbsI (4822)
BsiWI (5100)
NdeI (5429)

도면27

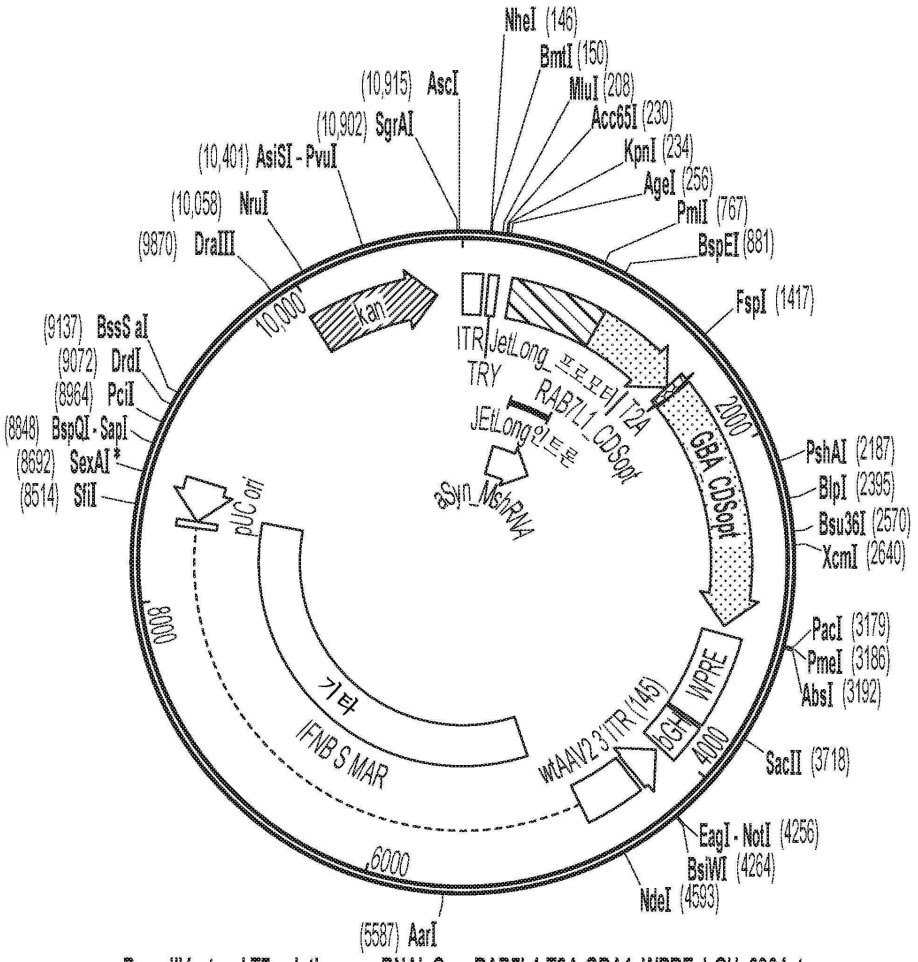


PrevailVector_FP5_JetLong_GBA1_bGH_JetLong_CTSB_SV40I_4108nt
11,064 bp

도면28

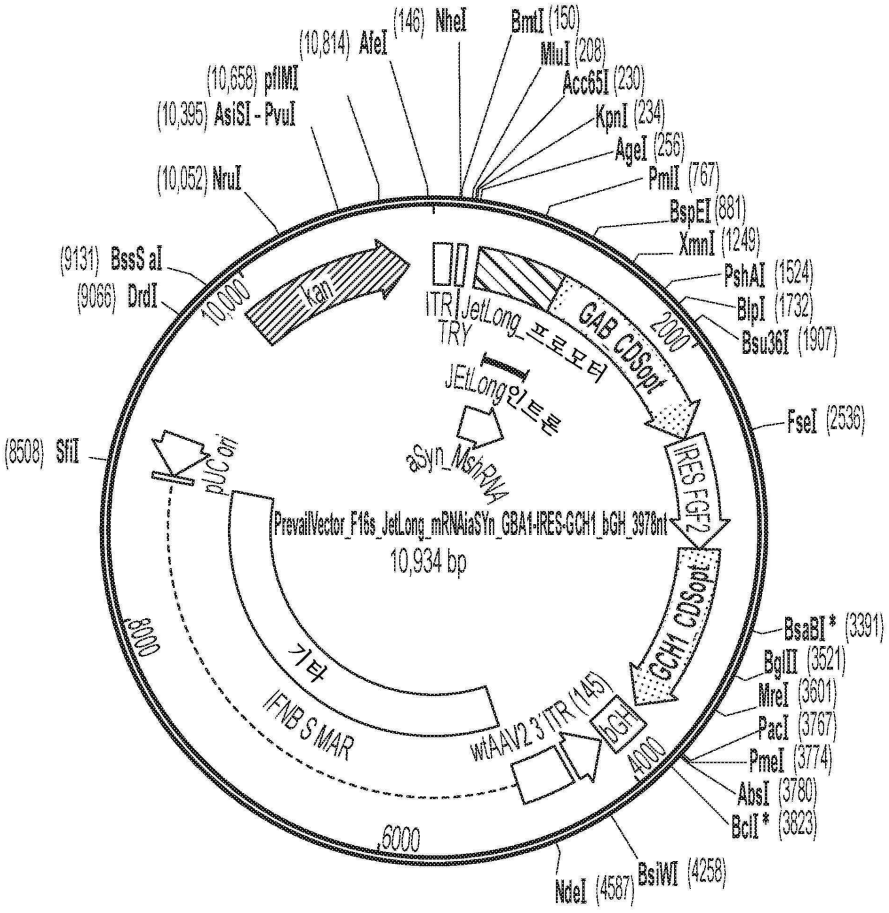

kan
ITR
TRY
JetLong_프로모터
JEtLong 인트론
aSyn_MshRNA
GBA_CDSopt
2000
T2A
GCH1_CDSopt
WPRE
4000
bGH
wtAAV2 3'ITR (145)
6000
기타
IFNB S MAR
8000
pUC ori
10,000
PrevailVector_FT6s_JetLong_mRNAiaSyn_GBA1_T2A-GCH1_WPRE_bGH_4125nt
11,081 bp

도면29

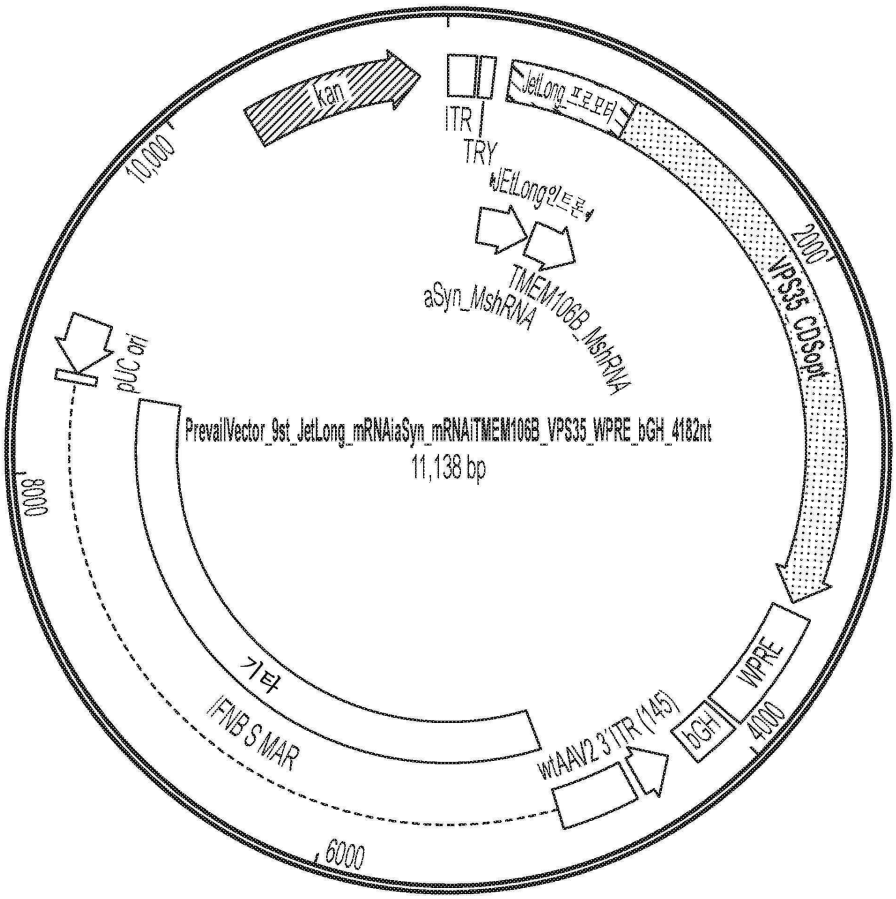


PrevailVector_LT7s_JetLong_mRNAiaSyn_RAB7L1-T2A-GBA1_WPRE_bGH_3984nt
10,940 bp

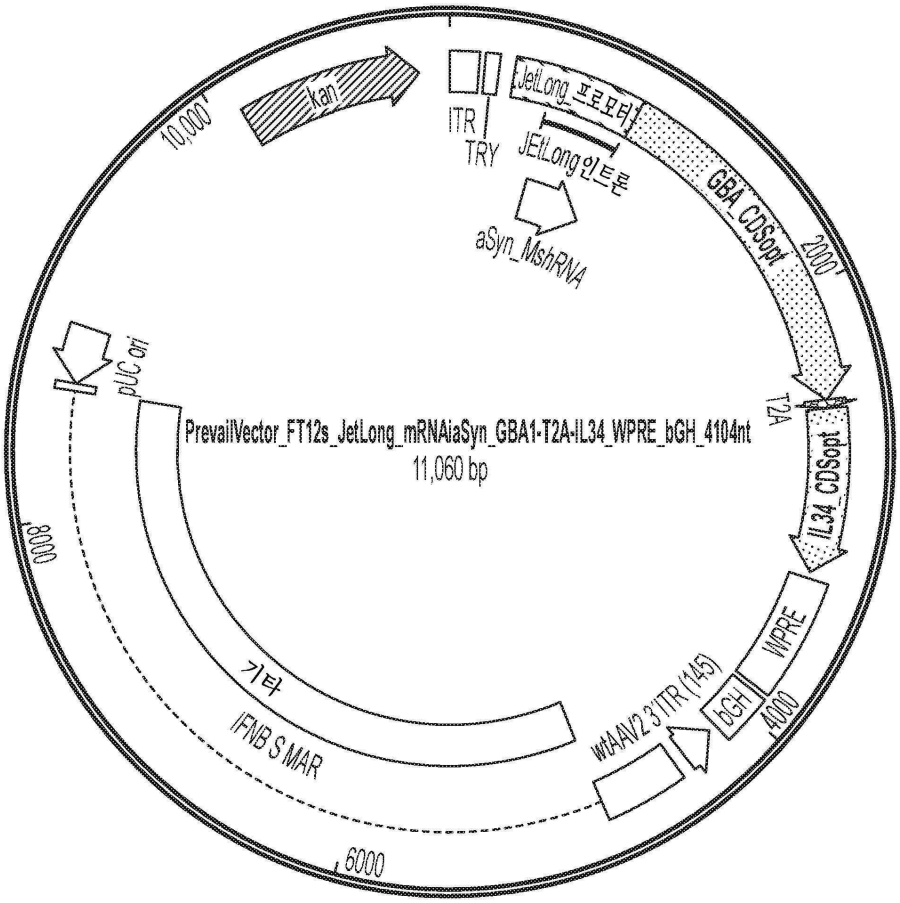
도면30


(146) NheI
BmtI (150)
MluI (208)
Acc65I (230)
KpnI (234)
AgeI (256)
PmlI (767)
BspEI (881)
XmnI (1249)
PshAI (1524)
BlpI (1732)
Bsu36I (1907)
FseI (2536)
BsaBI* (3391)
BglII (3521)
MreI (3601)
PacI (3767)
PmeI (3774)
AbsI (3780)
BclI* (3823)
BsiWI (4258)
NdeI (4587)
(8508) SfiI
(9131) BssS aI
(9066) DrdI
(10,052) NruI
(10,395) AsiSI - PvuI
(10,658) pflMI
(10,814) AfeI
Kan
ITR
TRY
JetLong_프로모터
GAB CDSopt
JETLong인트론
aSyn_MshRNA
IRES FGF2
GCH1 CDSopt
pUC ori
PrevailVector_F16s_JetLong_mRNAiaSYn_GBA1-IRES-GCH1_bGH_3978nt
10,934 bp
기타
IFNB S MAR
wtAAV2 3'ITR (145)
bGH
10,000
2000
4000
6000
8000

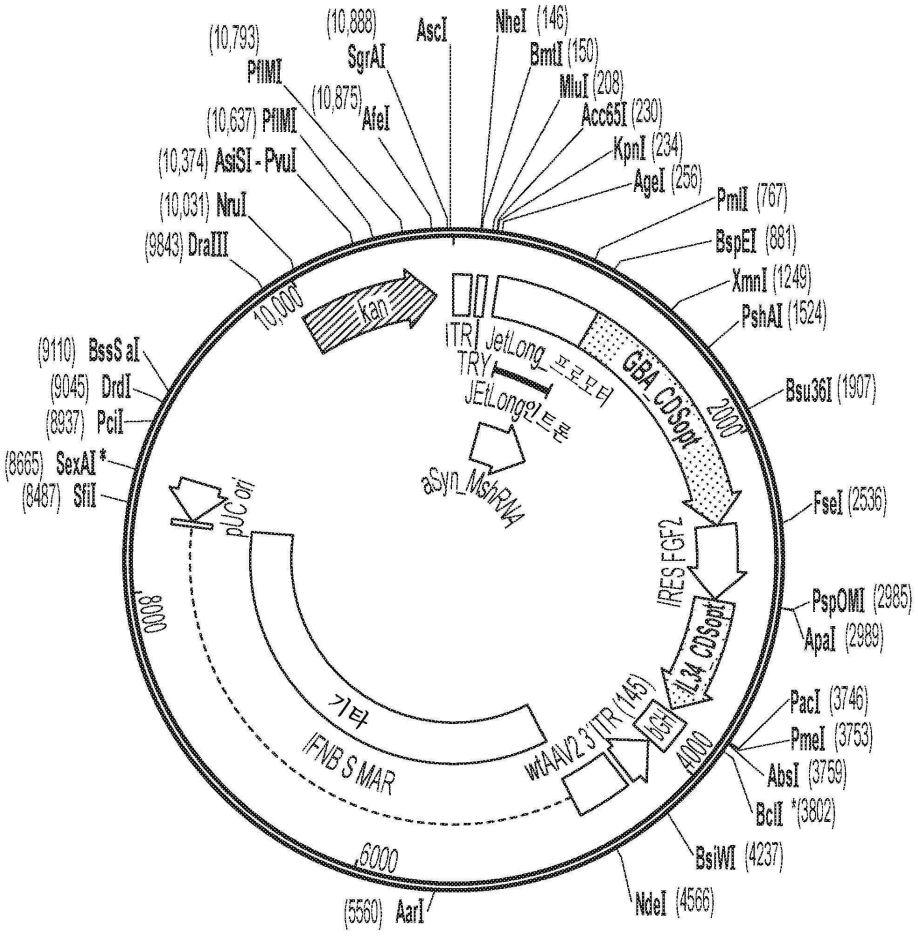
도면31


kan
ITR
TRY
JetLong_프로모터
JEtLong인트론
aSyn_MshRNA
TMEM106B_MshRNA
VPS35_CDSopt
2000
10,000
pUC ori
PrevailVector_9st_JetLong_mRNAiaSyn_mRNAiTMEM106B_VPS35_WPRE_bGH_4182nt
11,138 bp
8000
기타
IFNB S MAR
wtAAV2 3'ITR (145)
bGH
WPRE
4000
6000

도면32

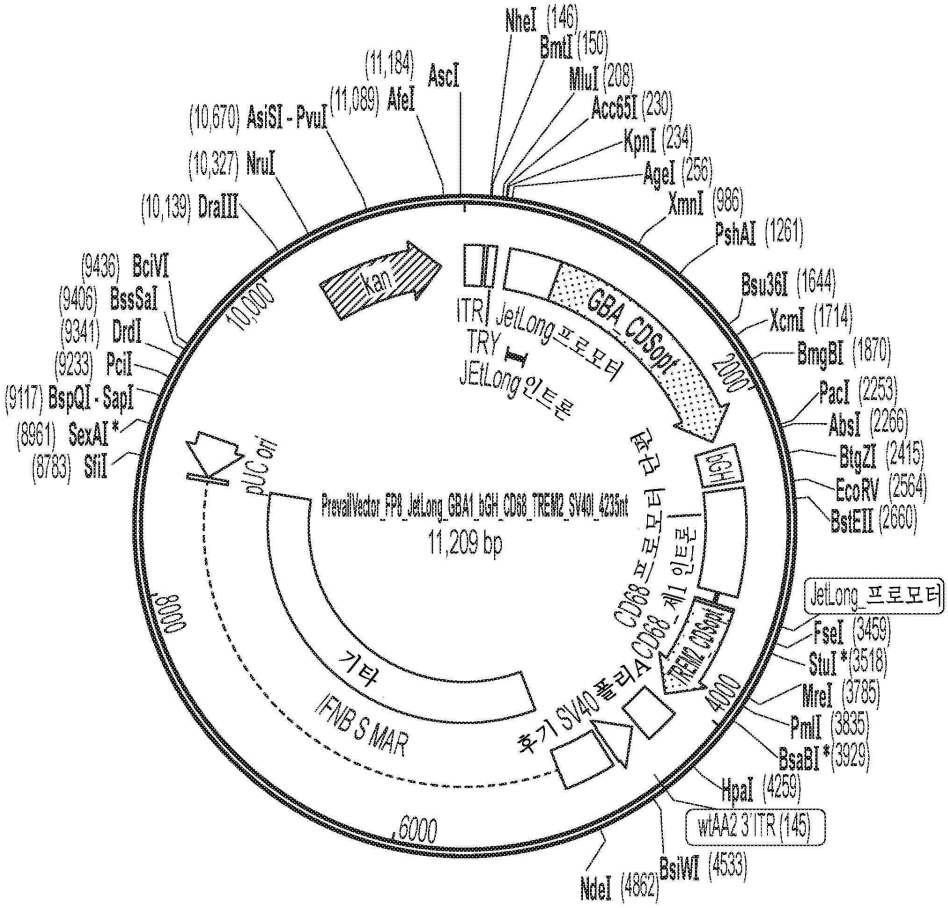

10,000
kan
ITR
TRY
JetLong_프로모터
JETLong인트론
aSyn_MshRNA
GBA_CDSopt
2000
T2A
IL34_CDSopt
WPRE
bGH
4000
wtAAV2 3'ITR (145)
6000
IFNB S MAR
기타
8000
pUC ori
PrevailVector_FT12s_JetLong_mRNAiaSyn_GBA1-T2A-IL34_WPRE_bGH_4104nt
11,060 bp

도면33

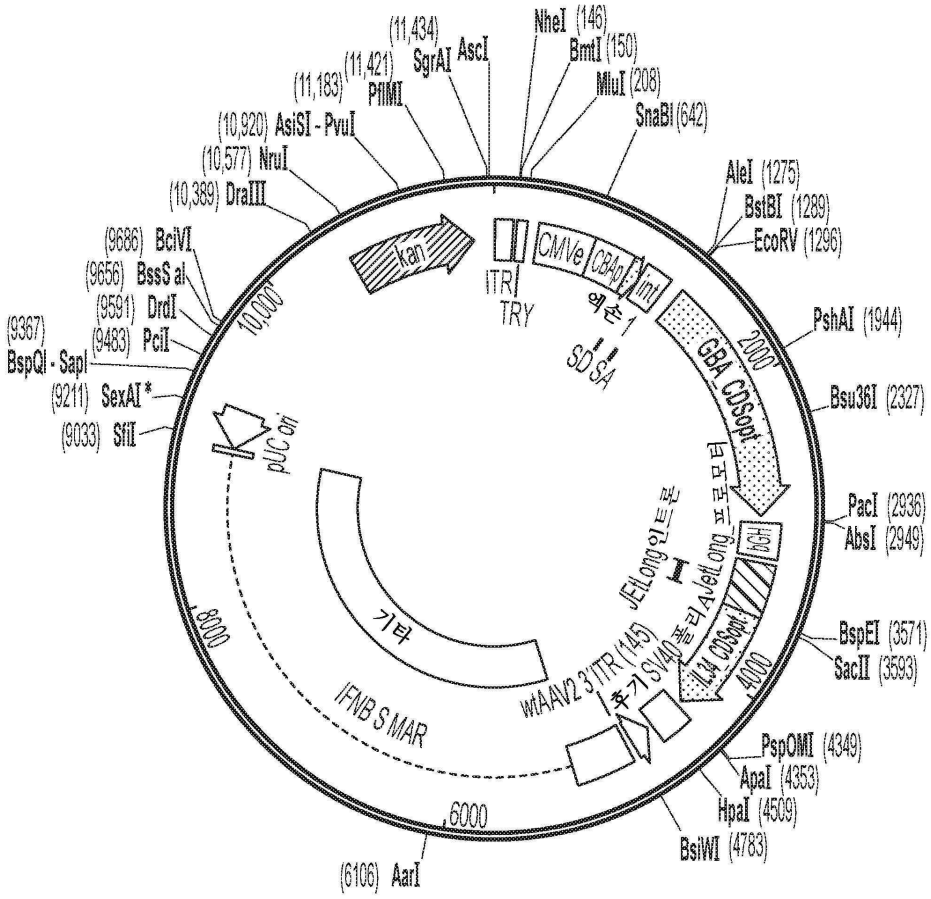


PrevailVector_FI12s_JetLong_mRNAiaSYn_GBA1-IRES-IL34_bGH_3957nt

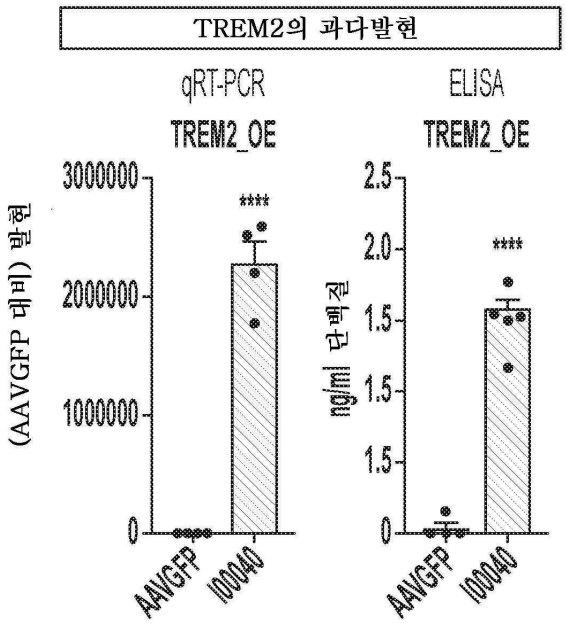
도면34


NheI (146)
BmtI (150)
MluI (208)
Acc65I (230)
KpnI (234)
AgeI (256)
XmnI (986)
PshAI (1261)
Bsu36I (1644)
XcmI (1714)
BmgBI (1870)
PacI (2253)
AbsI (2266)
BtgZI (2415)
EcoRV (2564)
BstEII (2660)
JetLong_프로모터
FseI (3459)
StuI* (3518)
MreI (3785)
PmlI (3835)
BsaBI* (3929)
HpaI (4259)
wIAA2 3'ITR (145)
BsiWI (4533)
NdeI (4862)
(11,184) AscI
(11,089) AfeI
(10,670) AsiSI - PvuI
(10,327) NruI
(10,139) DraIII
(9436) BciVI
(9406) BssSaI
(9341) DrdI
(9233) PciI
(9117) BspQI - SapI
(8961) SexAI*
(8783) SfiI
kan
ITR
TRY
JetLong_프로모터
JEtLong인트론
GBA CDSopt
2000
hGH
닭근 프로모터
CD68_제1 인트론
TREM2 CDSopt
4000
폴리A
SV40
후기
기타
IFNB S MAR
6000
8000
10,000
pUC ori
PrevailVector_FP8_JetLong_GBA1_hGH_CD68_TREM2_SV40l_4235nt
11,209 bp

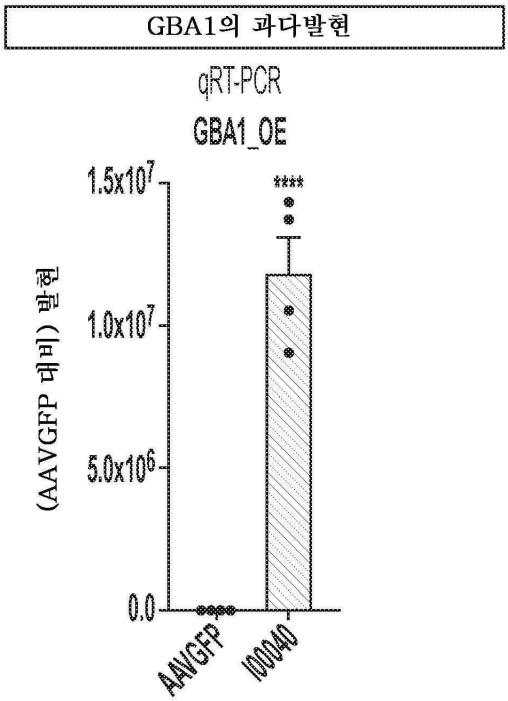
도면35


NheI (146)
BmtI (150)
MluI (208)
SnaBI (642)
AleI (1275)
BstBI (1289)
EcoRV (1296)
PshAI (1944)
Bsu36I (2327)
PacI (2936)
AbsI (2949)
BspEI (3571)
SacII (3593)
PspOMI (4349)
ApaI (4353)
HpaI (4509)
BsiWI (4783)
(6106) AarI
(9211) SexAI*
(9033) SfiI
(9367) BspQI - SapI
(9483) PciI
(9591) DrdI
(9656) BssS al
(9686) BciVI
(10,389) DraIII
(10,577) NruI
(10,920) AsiSI - PvuI
(11,183) PflMI
(11,421) SgrAI
(11,434) AscI
kan
ITR
TRY
CMVe
CBAo
Intr
엑손1
SD SA
GBA_CDSopt
2000
10,000
pUC ori
기타
JEtLong 인트론
JetLong 프로모터
bGH
IL34 CDSopt
4000
8000
IFNB S MAR
wtAAV2 3'ITR (145)
추가 SV40
6000
PrevailVector_FP12_CMVe_CBA_GBA1_bGH_JetLong_IL34_SV40I_4503nt
11,459 bp

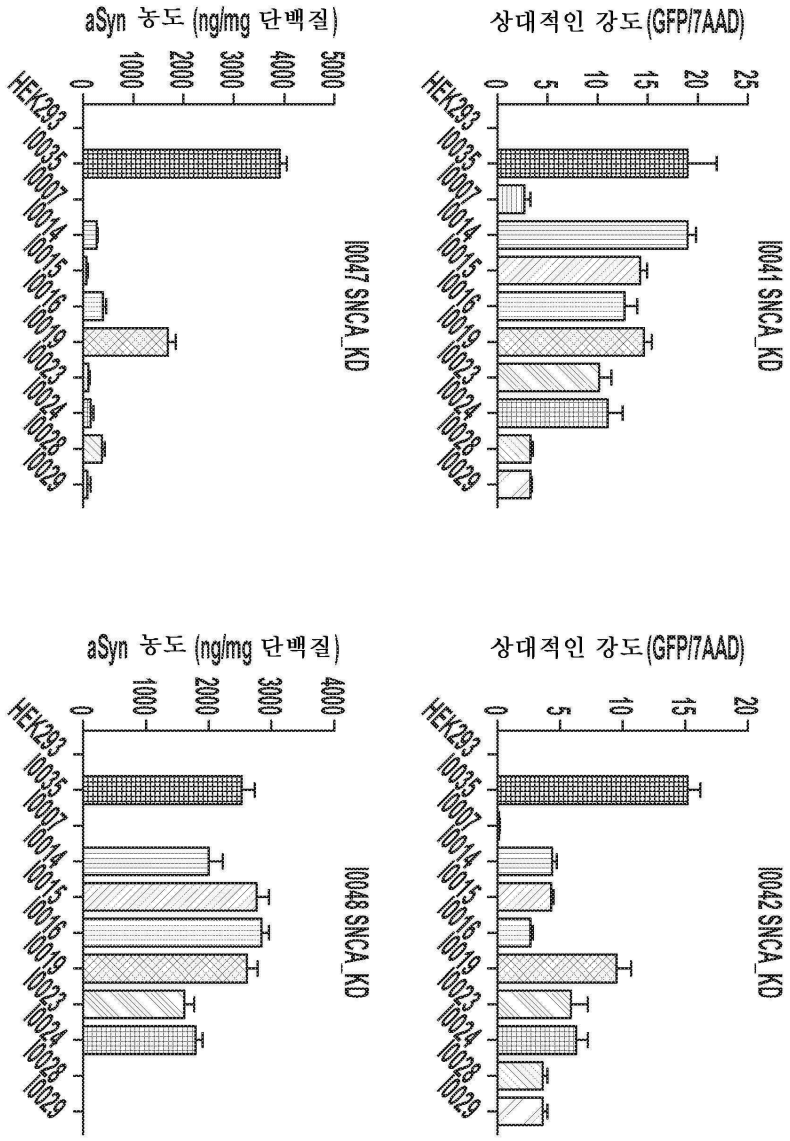
도면36a


TREM2의 과다발현
qRT-PCR
TREM2_OE
ELISA
TREM2_OE
(AAVGFP 대비) 발현
ng/ml 단백질
3000000
2000000
1000000
0
2.5
2.0
1.5
1.5
1.5
0
****
****
AAVGFP
100040
AAVGFP
100040

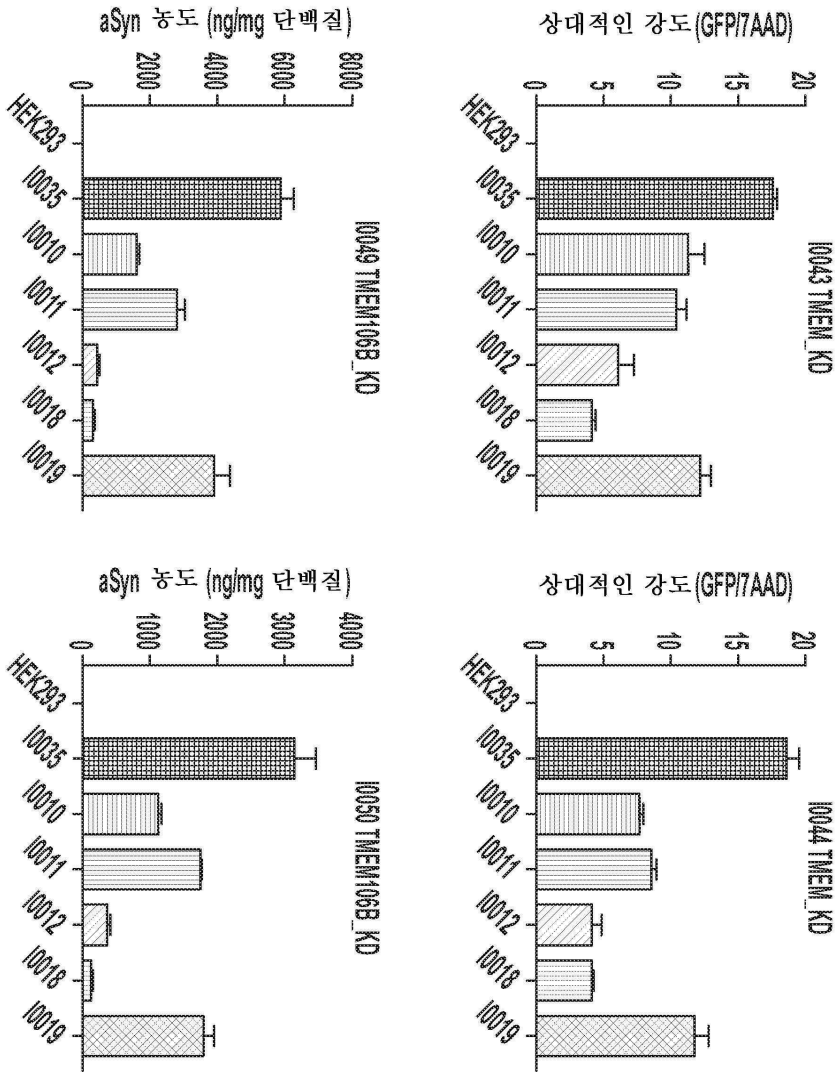
도면36b


GBA1의 과다발현
qRT-PCR
GBA1_OE
(AAVGFP 대비) 발현
1.5x10^7
1.0x10^7
5.0x10^6
0.0
****
AAVGFP
100040

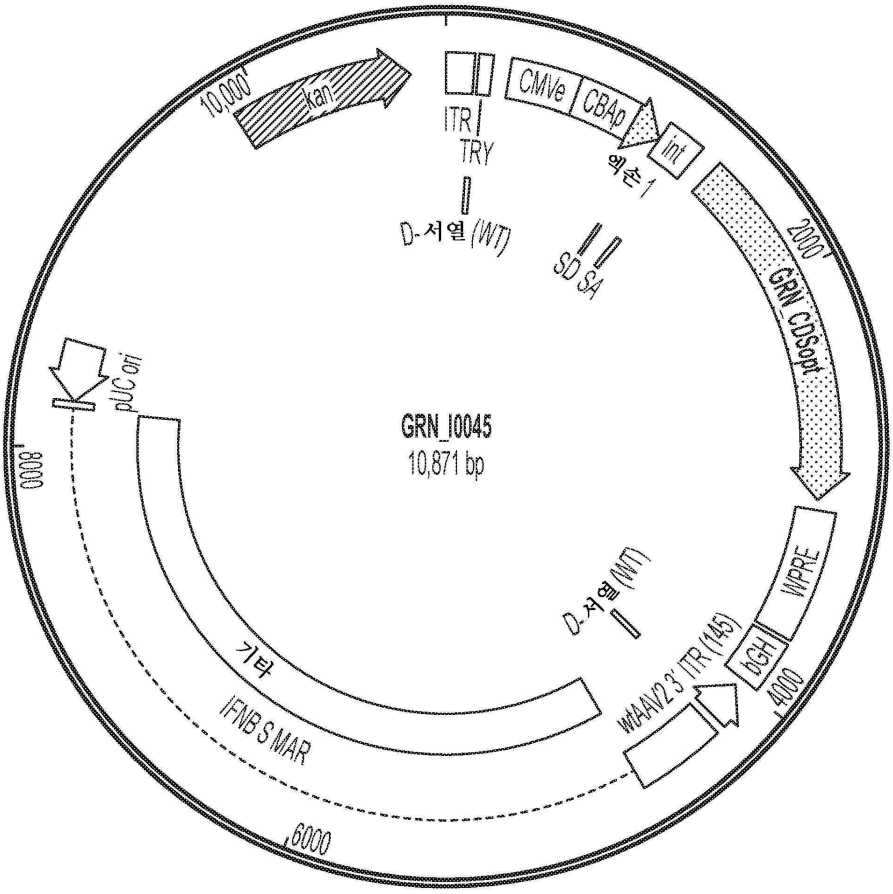
도면37


I0041 SNCA_KD
상대적인 강도(GFP/7AAD)
0
5
10
15
20
25
HEK293
I0035
I0007
I0014
I0015
I0016
I0019
I0023
I0024
I0028
I0029
I0047 SNCA_KD
aSyn 농도 (ng/mg 단백질)
0
1000
2000
3000
4000
5000
I0042 SNCA_KD
상대적인 강도(GFP/7AAD)
0
5
10
15
20
I0048 SNCA_KD
aSyn 농도 (ng/mg 단백질)
0
1000
2000
3000
4000

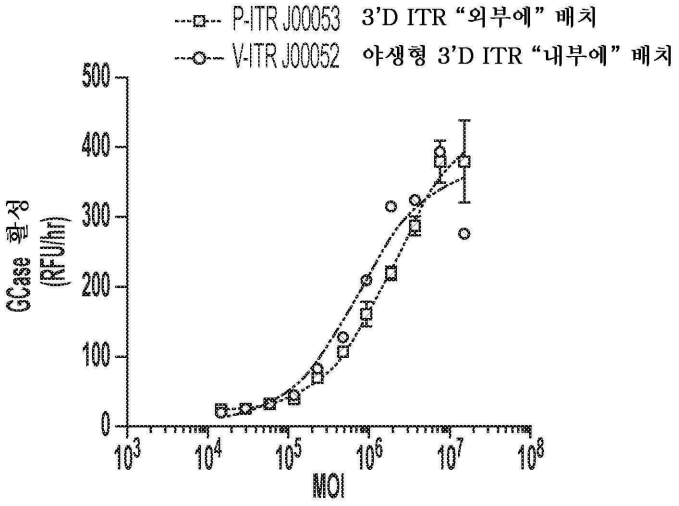
도면38


상대적인 강도(GFP/7AAD)
0
5
10
15
20
HEK293
I0035
I0010
I0011
I0012
I0018
I0019
I0043 TMEM_KD
aSyn 농도 (ng/mg 단백질)
0
2000
4000
6000
8000
I0049 TMEM106B_KD
상대적인 강도(GFP/7AAD)
I0044 TMEM_KD
aSyn 농도 (ng/mg 단백질)
0
1000
2000
3000
4000
I0050 TMEM106B_KD

도면39



도면40



서 열 목 록

SEQUENCE LISTING

<110> Prevail Therapeutics, Inc.

<120> GENE THERAPIES FOR LYSOSOMAL DISORDERS

<130> P1094.70003WO00

<140> Not Yet Assigned

<141> Concurrently herewith

<150> US 62/567,310

<151> 2017-10-03

<150> US 62/567,296

<151> 2017-10-03

<150> US 62/567,311

<151> 2017-10-03

<150> US 62/567,319

<151> 2017-10-03

<150> US 62/567,301

<151> 2017-10-03

<160> 78

<170> PatentIn version 3.5

<210> 1

<211> 10697

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 1

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300

cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc 360

cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga cccccgccca 420

ttgacgtcaa taatgacgta tgttcccata gtaacgccaa tagggacttt ccattgacgt 480

caatgggtgg actatttacg gtaaactgcc cacttggcag tacatcaagt gtatcatatg 540

ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgcccag 600

tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt 660

accatggtcg aggtgagccc cacgttctgc ttcactctcc ccatctcccc cccctcccca 720
cccccaattt tgtatttatt tatttttaa ttattttgtg cagcgatggg ggcgggggg 780
gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg 840

agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg 900
cggcggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg 960
ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cgcccgcccc ggctctgact 1020
gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctcctccgg gctgtaatta 1080
gcgcttggtt taatgacggc ttgttttctg tggctgcgtg aaagccttga ggggctccgg 1140
gagctagagc ctctgctaac catgttcatg ccttcttctt tttcctacag ctcctgggca 1200
acgtgctggt tattgtgctg tctcatcatt ttggcaaaga attcctcgaa gatccgaagg 1260

gaaagtcttc cacgactgtg ggatccgttc gaagatatca ccggttgagc caccatggaa 1320
ttcagcagcc ccagcagaga ggaatgcccc aagcctctga gccgggtgtc aatcatggcc 1380
ggatctctga caggactgct gctgcttcag gccgtgtctt gggcttctgg cgctagacct 1440
tgcatcccca agagcttcgg ctacagcagc gtcgtgtgcg tgtgcaatgc cacctactgc 1500
gacagcttcg accctcctac ctttcctgct ctgggcacct tcagcagata cgagagcacc 1560
agatccggca gacggatgga actgagcatg ggacccatcc aggccaatca cacaggcact 1620
ggcctgctgc tgacactgca gcctgagcag aaattccaga aagtgaaagg cttcggcgga 1680

gccatgacag atgccgccgc tctgaatatc ctggctctgt ctccaccagc tcagaacctg 1740
ctgctcaaga gctacttcag cgaggaaggc atcggctaca acatcatcag agtgcccatg 1800
gccagctgcg acttcagcat caggacctac acctacgccg acacacccga cgatttccag 1860
ctgcacaact tcagcctgcc tgaagaggac accaagctga agatccctct gatccacaga 1920
gccctgcagc tggcacaaag acccgtgtca ctgctggcct ctccatggac atctcccacc 1980
tggctgaaaa caaatggcgc cgtgaatggc aagggcagcc tgaaaggcca acctggcgac 2040
atctaccacc agacctgggc cagatacttc gtgaagttcc tggacgccta tgccgagcac 2100

aagctgcagt tttgggccgt gacagccgag aacgaacctt ctgctggact gctgagcggc 2160
taccccttc agtgcctggg ctttacaccc gagcaccagc gggactttat cgcccgtgat 2220
ctgggaccca cactggccaa tagcacccac cataatgtgc ggctgctgat gctggacgac 2280
cagagactgc ttctgcccca ctgggctaaa gtggtgctga cagatcctga ggccgccaaa 2340
tacgtgcacg gaatcgccgt gcactggtat ctggactttc tggcccctgc caaggccaca 2400
ctgggagaga cacacagact gttccccaac accatgctgt tcgccagcga agcctgtgtg 2460

ggcagcaagt tttgggaaca gagcgtgcgg ctcggcagct gggatagagg catgcagtac 2520

agccacagca tcatcaccaa cctgctgtac cacgtcgtcg gctggaccga ctggaatctg 2580
gccctgaatc ctgaaggcgg ccctaactgg gtccgaaact tcgtggacag ccccatcatc 2640
gtggacatca ccaaggacac cttctacaag cagcccatgt tctaccacct gggacacttc 2700
agcaagttca tccccgaggg ctctcagcgc gttggactgg tggcttccca gaagaacgat 2760
ctggacgccg tggctctgat gcaccctgat ggatctgctg tggtggtggt cctgaaccgc 2820
agcagcaaag atgtgcccct gaccatcaag gatcccgccg tgggattcct ggaaacaatc 2880
agccctggct actccatcca cacctacctg tggcgtagac agtgacaatt gttaattaag 2940

tttaaaccct cgaggccgca agcttatcga taatcaacct ctggattaca aaatttgtga 3000
aagattgact ggtattctta actatgttgc tccttttacg ctatgtggat acgctgcttt 3060
aatgcctttg tatcatgcta ttgcttcccg tatggctttc attttctcct ccttgtataa 3120
atcctggttg ctgtctcttt atgaggagtt gtggcccgtt gtcaggcaac gtggcgtggt 3180
gtgcactgtg tttgctgacg caacccccac tggttggggc attgccacca cctgtcagct 3240
cctttccggg actttcgctt tccccctccc tattgccacg gcggaactca tcgccgcctg 3300
ccttgcccgc tgctggacag gggctcggct gttgggcact gacaattccg tggtgttgtc 3360

ggggaaatca tcgtcctttc cttggctgct cgcctgtgtt gccacctgga ttctgcgcgg 3420
gacgtccttc tgctacgtcc cttcggccct caatccagcg gaccttcctt cccgcggcct 3480
gctgccggct ctgcggcctc ttccgcgtct tcgccttcgc cctcagacga gtcggatctc 3540
cctttgggcc gcctccccgc atcgataccg tcgactagag ctcgctgatc agcctcgact 3600
gtgccttcta gttgccagcc atctgttgtt tgcccctccc ccgtgccttc cttgaccctg 3660
gaaggtgcca ctcccactgt cctttcctaa taaaatgagg aaattgcatc gcattgtctg 3720
agtaggtgtc attctattct ggggggtggg gtggggcagg acagcaaggg ggaggattgg 3780

gaagacaata gcaggcatgc tggggagaga tccacgataa caaacagctt ttttggggtg 3840
aacatattga ctgaattccc tgcaggttgg ccactccctc tctgcgcgct cgctcgctca 3900
ctgaggccgc ccgggcaaag cccgggcgtc gggcgacctt tggtcgcccg gcctcagtga 3960
gcgagcgagc gcgcagagag ggagtggcca actccatcac taggggttcc tgcggccgct 4020
cgtacggtct cgaggaattc ctgcaggata acttgccaac ctcattctaa aatgtatata 4080
gaagcccaaa agacaataac aaaaatattc ttgtagaaca aaatgggaaa gaatgttcca 4140
ctaaatatca agatttagag caaagcatga gatgtgtggg gatagacagt gaggctgata 4200

aaatagagta gagctcagaa acagacccat tgatatatgt aagtgaccta tgaaaaaaat 4260
atggcatttt acaatgggaa aatgatggtc tttttctttt ttagaaaaac agggaaatat 4320
atttatatgt aaaaaataaa agggaaccca tatgtcatac catacacaca aaaaaattcc 4380
agtgaattat aagtctaaat ggagaaggca aaactttaaa tcttttagaa aataatatag 4440
aagcatgcag accagcctgg ccaacatgat gaaaccctct ctactaataa taaaatcagt 4500
agaactactc aggactactt tgagtgggaa gtccttttct atgaagactt ctttggccaa 4560
aattaggctc taaatgcaag gagatagtgc atcatgcctg gctgcactta ctgataaatg 4620

atgttatcac catctttaac caaatgcaca ggaacaagtt atggtactga tgtgctggat 4680
tgagaaggag ctctacttcc ttgacaggac acatttgtat caacttaaaa aagcagattt 4740
ttgccagcag aactattcat tcagaggtag gaaacttaga atagatgatg tcactgatta 4800
gcatggcttc cccatctcca cagctgcttc ccacccaggt tgcccacagt tgagtttgtc 4860
cagtgctcag ggctgcccac tctcagtaag aagccccaca ccagcccctc tccaaatatg 4920
ttggctgttc cttccattaa agtgacccca ctttagagca gcaagtggat ttctgtttct 4980
tacagttcag gaaggaggag tcagctgtga gaacctggag cctgagatgc ttctaagtcc 5040

cactgctact ggggtcaggg aagccagact ccagcatcag cagtcaggag cactaagccc 5100
ttgccaacat cctgtttctc agagaaactg cttccattat aatggttgtc ctttttaag 5160
ctatcaagcc aaacaaccag tgtctaccat tattctcatc acctgaagcc aagggttcta 5220
gcaaaagtca agctgtcttg taatggttga tgtgcctcca gcttctgtct tcagtcactc 5280
cactcttagc ctgctctgaa tcaactctga ccacagttcc ctggagcccc tgccacctgc 5340
tgcccctgcc accttctcca tctgcagtgc tgtgcagcct tctgcactct tgcagagcta 5400
ataggtggag acttgaagga agaggaggaa agtttctcat aatagccttg ctgcaagctc 5460

aaatgggagg tgggcactgt gcccaggagc cttggagcaa aggctgtgcc caacctctga 5520
ctgcatccag gtttggtctt gacagagata agaagccctg gcttttggag ccaaaatcta 5580
ggtcagactt aggcaggatt ctcaaagttt atcagcagaa catgaggcag aagacccttt 5640
ctgctccagc ttcttcaggc tcaaccttca tcagaataga tagaaagaga ggctgtgagg 5700
gttcttaaaa cagaagcaaa tctgactcag agaataaaca acctcctagt aaactacagc 5760
ttagacagag catctggtgg tgagtgtgct cagtgtccta ctcaactgtc tggtatcagc 5820
cctcatgagg acttctcttc tttccctcat agacctccat ctctgttttc cttagcctgc 5880

agaaatctgg atggctattc acagaatgcc tgtgctttca gagttgcatt ttttctctgg 5940
tattctggtt caagcatttg aaggtaggaa aggttctcca agtgcaagaa agccagccct 6000
gagcctcaac tgcctggcta gtgtggtcag taggatgcaa aggctgttga atgccacaag 6060

gccaaacttt aacctgtgta ccacaagcct agcagcagag gcagctctgc tcactggaac 6120
tctctgtctt ctttctcctg agccttttct tttcctgagt tttctagctc tcctcaacct 6180
tacctctgcc ctacccagga caaacccaag agccactgtt tctgtgatgt cctctccagc 6240
cctaattagg catcatgact tcagcctgac cttccatgct cagaagcagt gctaatccac 6300

ttcagatgag ctgctctatg caacacaggc agagcctaca aacctttgca ccagagccct 6360
ccacatatca gtgtttgttc atactcactt caacagcaaa tgtgactgct gagattaaga 6420
ttttacacaa gatggtctgt aatttcacag ttagttttat cccattaggt atgaaagaat 6480
tagcataatt ccccttaaac atgaatgaat cttagatttt ttaataaata gttttggaag 6540
taaagacaga gacatcagga gcacaaggaa tagcctgaga ggacaaacag aacaagaaag 6600
agtctggaaa tacacaggat gttcttggcc tcctcaaagc aagtgcaagc agatagtacc 6660
agcagcccca ggctatcaga gcccagtgaa gagaagtacc atgaaagcca cagctctaac 6720

caccctgttc cagagtgaca gacagtcccc aagacaagcc agcctgagcc agagagagaa 6780
ctgcaagaga aagtttctaa tttaggttct gttagattca gacaagtgca ggtcatcctc 6840
tctccacagc tactcacctc tccagcctaa caaagcctgc agtccacact ccaaccctgg 6900
tgtctcacct cctagcctct cccaacatcc tgctctctga ccatcttctg catctctcat 6960
ctcaccatct cccactgtct acagcctact cttgcaacta ccatctcatt ttctgacatc 7020
ctgtctacat cttctgccat actctgccat ctaccatacc acctcttacc atctaccaca 7080
ccatctttta tctccatccc tctcagaagc ctccaagctg aatcctgctt tatgtgttca 7140

tctcagcccc tgcatggaaa gctgacccca gaggcagaac tattcccaga gagcttggcc 7200
aagaaaaaca aaactaccag cctggccagg ctcaggagta gtaagctgca gtgtctgttg 7260
tgttctagct tcaacagctg caggagttcc actctcaaat gctccacatt tctcacatcc 7320
tcctgattct ggtcactacc catcttcaaa gaacagaata tctcacatca gcatactgtg 7380
aaggactagt catgggtgca gctgctcaga gctgcaaagt cattctggat ggtggagagc 7440
ttacaaacat ttcatgatgc tccccccgct ctgatggctg gagcccaatc cctacacaga 7500
ctcctgctgt atgtgttttc ctttcactct gagccacagc cagagggcag gcattcagtc 7560

tcctcttcag gctggggctg gggcactgag aactcaccca acaccttgct ctcactcctt 7620
ctgcaaaaca agaaagagct ttgtgctgca gtagccatga agaatgaaag gaaggcttta 7680
actaaaaaat gtcagagatt attttcaacc ccttactgtg gatcaccagc aaggaggaaa 7740
cacaacacag agacattttt tcccctcaaa ttatcaaaag aatcactgca tttgttaaag 7800
agagcaactg aatcaggaag cagagttttg aacatatcag aagttaggaa tctgcatcag 7860
agacaaatgc agtcatggtt gtttgctgca taccagccct aatcattaga agcctcatgg 7920

acttcaaaca tcattccctc tgacaagatg ctctagccta actccatgag ataaaataaa 7980

tctgcctttc agagccaaag aagagtccac cagcttcttc tcagtgtgaa caagagctcc 8040
agtcaggtta gtcagtccag tgcagtagag gagaccagtc tgcatcctct aattttcaaa 8100
ggcaagaaga tttgtttacc ctggacacca ggcacaagtg aggtcacaga gctcttagat 8160
atgcagtcct catgagtgag gagactaaag cgcatgccat caagacttca gtgtagagaa 8220
aacctccaaa aaagcctcct cactacttct ggaatagctc agaggccgag gcggcctcgg 8280
cctctgcata aataaaaaaa attagtcagc catggggcgg agaatgggcg gaactgggcg 8340
gagttagggg cgggatgggc ggagttaggg gcgggactat ggttgctgac taattgagat 8400

gcatgctttg catacttctg cctgctgggg agcctgggga ctttccacac ctggttgctg 8460
actaattgag atgcatgctt tgcatacttc tgcctgctgg ggagcctggg gactttccac 8520
accctaactg acacacattc cacagctgca ttaatgaatc ggccaacgcg cggggagagg 8580
cggtttgcgt attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt 8640
tcggctgcgg cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc 8700
aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa 8760
aaaggccgcg ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa 8820

tcgacgctca agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc 8880
ccctggaagc tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc 8940
cgcctttctc ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag 9000
ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga 9060
ccgctgcgcc ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc 9120
gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac 9180
agagttcttg aagtggtggc ctaactacgg ctacactaga agaacagtat ttggtatctg 9240

cgctctgctg aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca 9300
aaccaccgct ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa 9360
aggatctcaa gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa 9420
ctcacgttaa gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt 9480
aaattaaaaa tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag 9540
ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat 9600
agttgcctga ctcctgcaaa ccacgttgtg tctcaaaatc tctgatgtta cattgcacaa 9660

gataaaaata tatcatcatg aacaataaaa ctgtctgctt acataaacag taatacaagg 9720

ggtgttatga gccatattca acgggaaacg tcttgctcga ggccgcgatt aaattccaac 9780

atggatgctg atttatatgg gtataaatgg gctcgcgata atgtcgggca atcaggtgcg 9840

acaatctatc gattgtatgg gaagcccgat gcgccagagt tgtttctgaa acatggcaaa 9900

ggtagcgttg ccaatgatgt tacagatgag atggtcagac taaactggct gacggaattt 9960

atgcctcttc cgaccatcaa gcattttatc cgtactcctg atgatgcatg gttactcacc 10020

actgcgatcc ccgggaaaac agcattccag gtattagaag aatatcctga ttcaggtgaa 10080

aatattgttg atgcgctggc agtgttcctg cgccggttgc attcgattcc tgtttgtaat 10140

tgtcctttta acagcgatcg cgtatttcgt ctcgctcagg cgcaatcacg aatgaataac 10200

ggtttggttg atgcgagtga ttttgatgac gagcgtaatg gctggcctgt tgaacaagtc 10260

tggaaagaaa tgcataagct tttgccattc tcaccggatt cagtcgtcac tcatggtgat 10320

ttctcacttg ataaccttat ttttgacgag gggaaattaa taggttgtat tgatgttgga 10380

cgagtcggaa tcgcagaccg ataccaggat cttgccatcc tatggaactg cctcggtgag 10440

ttttctcctt cattacagaa acggcttttt caaaaatatg gtattgataa tcctgatatg 10500

aataaattgc agtttcattt gatgctcgat gagtttttct aagggcggcc tgccaccata 10560

cccacgccga aacaagcgct catgagcccg aagtggcgag cccgatcttc cccatcggtg 10620

atgtcggcga tataggcgcc agcaaccgca cctgtggcgc cggtgatgag ggcgcgccaa 10680

gtcgacgtcc ggcagtc 10697

<210> 2

<211> 11355

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 2

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420

tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctgcag ccctaggaat gcatctagac aattgtacta accttcttct 600
ctttcctctc ctgacagtcc ggaaagccac catgggccgc tgctgcttct acaccgccgg 660
caccctgagc ctgctgctgc tggtgaccag cgtgaccctg ctggtggccc gcgtgttcca 720
gaaggccgtg gaccagagca tcgagaagaa gatcgtgctg cgcaacggca ccgaggcctt 780
cgacagctgg gagaagcccc ccctgcccgt gtacacccag ttctacttct tcaacgtgac 840
caaccccgag gagatcctgc gcggcgagac cccccgcgtg gaggaggtgg gcccctacac 900
ctaccgcgag ctgcgcaaca aggccaacat ccagttcggc gacaacggca ccaccatcag 960

cgccgtgagc aacaaggcct acgtgttcga gcgcgaccag agcgtgggcg accccaagat 1020
cgacctgatc cgcaccctga acatccccgt gctgaccgtg atcgagtgga gccaggtgca 1080
cttcctgcgc gagatcatcg aggccatgct gaaggcctac cagcagaagc tgttcgtgac 1140
ccacaccgtg gacgagctgc tgtggggcta caaggacgag atcctgagcc tgatccacgt 1200
gttccgcccc gacatcagcc cctacttcgg cctgttctac gagaagaacg gcaccaacga 1260
cggcgactac gtgttcctga ccggcgagga cagctacctg aacttcacca agatcgtgga 1320
gtggaacggc aagaccagcc tggactggtg gatcaccgac aagtgcaaca tgatcaacgg 1380

caccgacggc gacagcttcc acccccctgat caccaaggac gaggtgctgt acgtgttccc 1440
cagcgacttc tgccgcagcg tgtacatcac cttcagcgac tacgagagcg tgcagggcct 1500
gcccgccttc cgctacaagg tgcccgccga gatcctggcc aacaccagcg acaacgccgg 1560
cttctgcatc cccgagggca actgcctggg cagcggcgtg ctgaacgtga gcatctgcaa 1620
gaacggcgcc cccatcatca tgagcttccc ccacttctac caggccgacg agcgcttcgt 1680
gagcgccatc gagggcatgc accccaacca ggaggaccac gagaccttcg tggacatcaa 1740
ccccctgacc ggcatcatcc tgaaggccgc caagcgcttc cagatcaaca tctacgtgaa 1800

gaagctggac gacttcgtgg agaccggcga catccgcacc atggtgttcc ccgtgatgta 1860
cctgaacgag agcgtgcaca tcgacaagga gaccgccagc cgcctgaaga gcatgatcaa 1920
caccaccctg atcatcacca acatccccta catcatcatg gccctgggcg tgttcttcgg 1980
cctggtgttc acctggctgg cctgcaaggg ccagggcagc atggacgagg gcaccgccga 2040
cgagcgcgcc cccctgatcc gcacctgatt gtggccgaac cgccgaactc agaggccggc 2100
cccagaaaac ccgagcgagt agggggcggc gcgcaggagg gaggagaact gggggcgcgg 2160
gaggctggtg ggtgtggggg gtggagatgt agaagatgtg acgccgcggc ccggcgggtg 2220

ccagattagc ggacgcggtg cccgcggttg caacgggatc ccgggcgctg cagcttggga 2280
ggcggctctc cccaggcggc gtccgcggag acacccatcc gtgaacccca ggtcccgggc 2340
cgccggctcg ccgcgcacca ggggccggcg gacagaagag cggccgagcg gctcgaggct 2400
gggggaccgc gggcgcggcc gcgcgctgcc gggcgggagg ctgggggcc ggggccgggg 2460
ccgtgccccg gagcgggtcg gaggccgggg ccggggccgg gggacggcgg ctccccgcgc 2520
ggctccagcg gctcggggat cccggccggg ccccgcaggg accatgatgg aattcagcag 2580
ccccagcaga gaggaatgcc ccaagcctct gagccgggtg tcaatcatgg ccggatctct 2640

gacaggactg ctgctgcttc aggccgtgtc ttgggcttct ggcgctagac cttgcatccc 2700
caagagcttc ggctacagca gcgtcgtgtg cgtgtgcaat gccacctact gcgacagctt 2760
cgaccctcct acctttcctg ctctgggcac cttcagcaga tacgagagca ccagatccgg 2820
cagacggatg gaactgagca tgggacccat ccaggccaat cacacaggca ctggcctgct 2880
gctgacactg cagcctgagc agaaattcca gaaagtgaaa ggcttcggcg gagccatgac 2940
agatgccgcc gctctgaata tcctggctct gtctccacca gctcagaacc tgctgctcaa 3000
gagctacttc agcgaggaag gcatcggcta caacatcatc agagtgccca tggccagctg 3060

cgacttcagc atcaggacct acacctacgc cgacacaccc gacgatttcc agctgcacaa 3120
cttcagcctg cctgaagagg acaccaagct gaagatccct ctgatccaca gagccctgca 3180
gctggcacaa agacccgtgt cactgctggc ctctccatgg acatctccca cctggctgaa 3240
aacaaatggc gccgtgaatg gcaagggcag cctgaaaggc caacctggcg acatctacca 3300
ccagacctgg gccagatact tcgtgaagtt cctggacgcc tatgccgagc acaagctgca 3360
gttttgggcc gtgacagccg agaacgaacc ttctgctgga ctgctgagcg gctacccctt 3420
tcagtgcctg ggctttacac ccgagcacca gcgggacttt atcgcccgtg atctgggacc 3480

cacactggcc aatagcaccc accataatgt gcggctgctg atgctggacg accagagact 3540
gcttctgccc cactgggcta aagtggtgct gacagatcct gaggccgcca aatacgtgca 3600
cggaatcgcc gtgcactggt atctggactt tctggcccct gccaaggcca cactgggaga 3660
gacacacaga ctgttcccca acaccatgct gttcgccagc gaagcctgtg tgggcagcaa 3720
gttttgggaa cagagcgtgc ggctcggcag ctgggataga ggcatgcagt acagccacag 3780
catcatcacc aacctgctgt accacgtcgt cggctggacc gactggaatc tggccctgaa 3840
tcctgaaggc ggccctaact gggtccgaaa cttcgtggac agccccatca tcgtggacat 3900

caccaaggac accttctaca agcagcccat gttctaccac ctgggacact tcagcaagtt 3960
catccccgag ggctctcagc gcgttggact ggtggcttcc cagaagaacg atctggacgc 4020
cgtggctctg atgcaccctg atggatctgc tgtggtggtg gtcctgaacc gcagcagcaa 4080

agatgtgccc ctgaccatca aggatcccgc cgtgggattc ctggaaacaa tcagccctgg 4140
ctactccatc cacacctacc tgtggcgtag acagtgacaa ttgttaatta agtttaaacc 4200
ctcgaggccg caagccgcat cgataccgtc gactagagct cgctgatcag cctcgactgt 4260
gccttctagt tgccagccat ctgttgtttg cccctccccc gtgccttcct tgaccctgga 4320

aggtgccact cccactgtcc tttcctaata aaatgaggaa attgcatcgc attgtctgag 4380
taggtgtcat tctattctgg ggggtggggt ggggcaggac agcaaggggg aggattggga 4440
agacaatagc aggcatgctg gggagagatc cacgataaca aacagctttt ttggggtgaa 4500
catattgact gaattccctg caggttggcc actccctctc tgcgcgctcg ctcgctcact 4560
gaggccgccc gggcaaagcc cgggcgtcgg gcgacctttg gtcgcccggc ctcagtgagc 4620
gagcgagcgc gcagagaggg agtggccaac tccatcacta ggggttcctg cggccgctcg 4680
tacggtctcg aggaattcct gcaggataac ttgccaacct cattctaaaa tgtatataga 4740

agcccaaaag acaataacaa aaatattctt gtagaacaaa atgggaaaga atgttccact 4800
aaatatcaag atttagagca aagcatgaga tgtgtgggga tagacagtga ggctgataaa 4860
atagagtaga gctcagaaac agacccattg atatatgtaa gtgacctatg aaaaaaatat 4920
ggcattttac aatgggaaaa tgatggtctt tttctttttt agaaaaacag ggaaatatat 4980
ttatatgtaa aaaataaaag ggaacccata tgtcatacca tacacacaaa aaaattccag 5040
tgaattataa gtctaaatgg agaaggcaaa actttaaatc ttttagaaaa taatatagaa 5100
gcatgcagac cagcctggcc aacatgatga aaccctctct actaataata aaatcagtag 5160

aactactcag gactactttg agtgggaagt ccttttctat gaagacttct ttggccaaaa 5220
ttaggctcta aatgcaagga gatagtgcat catgcctggc tgcacttact gataaatgat 5280
gttatcacca tctttaacca aatgcacagg aacaagttat ggtactgatg tgctggattg 5340
agaaggagct ctacttcctt gacaggacac atttgtatca acttaaaaaa gcagattttt 5400
gccagcagaa ctattcattc agaggtagga aacttagaat agatgatgtc actgattagc 5460
atggcttccc catctccaca gctgcttccc acccaggttg cccacagttg agtttgtcca 5520
gtgctcaggg ctgcccactc tcagtaagaa gccccacacc agcccctctc caaatatgtt 5580

ggctgttcct tccattaaag tgaccccact ttagagcagc aagtggattt ctgtttctta 5640
cagttcagga aggaggagtc agctgtgaga acctggagcc tgagatgctt ctaagtccca 5700
ctgctactgg ggtcagggaa gccagactcc agcatcagca gtcaggagca ctaagccctt 5760
gccaacatcc tgtttctcag agaaactgct tccattataa tggttgtcct tttttaagct 5820
atcaagccaa acaaccagtg tctaccatta ttctcatcac ctgaagccaa gggttctagc 5880
aaaagtcaag ctgtcttgta atggttgatg tgcctccagc ttctgtcttc agtcactcca 5940

ctcttagcct gctctgaatc aactctgacc acagttccct ggagcccctg ccacctgctg 6000

cccctgccac cttctccatc tgcagtgctg tgcagccttc tgcactcttg cagagctaat 6060
aggtggagac ttgaaggaag aggaggaaag tttctataa tagccttgct gcaagctcaa 6120
atgggaggtg ggcactgtgc ccaggagcct tggagcaaag gctgtgccca acctctgact 6180
gcatccaggt ttggtcttga cagagataag aagccctggc ttttggagcc aaaatctagg 6240
tcagacttag gcaggattct caaagtttat cagcagaaca tgaggcagaa gaccctttct 6300
gctccagctt cttcaggctc aaccttcatc agaatagata gaaagagagg ctgtgagggt 6360
tcttaaaaca gaagcaaatc tgactcagag aataaacaac ctcctagtaa actacagctt 6420

agacagagca tctggtggtg agtgtgctca gtgtcctact caactgtctg gtatcagccc 6480
tcatgaggac ttctcttctt tccctcatag acctccatct ctgttttcct tagcctgcag 6540
aaatctggat ggctattcac agaatgcctg tgctttcaga gttgcatttt ttctctggta 6600
ttctggttca agcatttgaa ggtaggaaag gttctccaag tgcaagaaag ccagccctga 6660
gcctcaactg cctggctagt gtggtcagta ggatgcaaag gctgttgaat gccacaaggc 6720
caaactttaa cctgtgtacc acaagcctag cagcagaggc agctctgctc actggaactc 6780
tctgtcttct ttctcctgag ccttttcttt tcctgagttt tctagctctc ctcaacctta 6840

cctctgccct acccaggaca aacccaagag ccactgtttc tgtgatgtcc tctccagccc 6900
taattaggca tcatgacttc agcctgacct tccatgctca gaagcagtgc taatccactt 6960
cagatgagct gctctatgca acacaggcag agcctacaaa cctttgcacc agagccctcc 7020
acatatcagt gtttgttcat actcacttca acagcaaatg tgactgctga gattaagatt 7080
ttacacaaga tggtctgtaa tttcacagtt agttttatcc cattaggtat gaaagaatta 7140
gcataattcc ccttaaacat gaatgaatct tagatttttt aataaatagt tttggaagta 7200
aagacagaga catcaggagc acaaggaata gcctgagagg acaaacagaa caagaaagag 7260

tctggaaata cacaggatgt tcttggcctc ctcaaagcaa gtgcaagcag atagtaccag 7320
cagccccagg ctatcagagc ccagtgaaga gaagtaccat gaaagccaca gctctaacca 7380
ccctgttcca gagtgacaga cagtccccaa gacaagccag cctgagccag agagagaact 7440
gcaagagaaa gtttctaatt taggttctgt tagattcaga caagtgcagg tcatcctctc 7500
tccacagcta ctcacctctc cagcctaaca aagcctgcag tccacactcc aaccctggtg 7560
tctcacctcc tagcctctcc caacatcctg ctctctgacc atcttctgca tctctcatct 7620
caccatctcc cactgtctac agcctactct tgcaactacc atctcatttt ctgacatcct 7680

gtctacatct tctgccatac tctgccatct accataccac ctcttaccat ctaccacacc 7740
atcttttatc tccatccctc tcagaagcct ccaagctgaa tcctgcttta tgtgttcatc 7800
tcagcccctg catggaaagc tgaccccaga ggcagaacta ttcccagaga gcttggccaa 7860
gaaaaacaaa actaccagcc tggccaggct caggagtagt aagctgcagt gtctgttgtg 7920
ttctagcttc aacagctgca ggagttccac tctcaaatgc tccacatttc tcacatcctc 7980
ctgattctgg tcactaccca tcttcaaaga acagaatatc tcacatcagc atactgtgaa 8040
ggactagtca tgggtgcagc tgctcagagc tgcaaagtca ttctggatgg tggagagctt 8100

acaaacattt catgatgctc cccccgctct gatggctgga gcccaatccc tacacagact 8160
cctgctgtat gtgttttcct ttcactctga gccacagcca gagggcaggc attcagtctc 8220
ctcttcaggc tggggctggg gcactgagaa ctcacccaac accttgctct cactccttct 8280
gcaaaacaag aaagagcttt gtgctgcagt agccatgaag aatgaaagga aggctttaac 8340
taaaaaatgt cagagattat tttcaacccc ttactgtgga tcaccagcaa ggaggaaaca 8400
caacacagag acatttttc ccctcaaatt atcaaaagaa tcactgcatt tgttaaagag 8460
agcaactgaa tcaggaagca gagttttgaa catatcagaa gttaggaatc tgcatcagag 8520

acaaatgcag tcatggttgt ttgctgcata ccagccctaa tcattagaag cctcatggac 8580
ttcaaacatc attccctctg acaagatgct ctagcctaac tccatgagat aaaataaatc 8640
tgcctttcag agccaaagaa gagtccacca gcttcttctc agtgtgaaca agagctccag 8700
tcaggttagt cagtccagtg cagtagagga gaccagtctg catcctctaa ttttcaaagg 8760
caagaagatt tgtttaccct ggacaccagg cacaagtgag gtcacagagc tcttagatat 8820
gcagtcctca tgagtgagga gactaaagcg catgccatca agacttcagt gtagagaaaa 8880
cctccaaaaa agcctcctca ctacttctgg aatagctcag aggccgaggc ggcctcggcc 8940

tctgcataaa taaaaaaaat tagtcagcca tggggcggag aatgggcgga actgggcgga 9000
gttaggggcg ggatgggcgg agttaggggc gggactatgg ttgctgacta attgagatgc 9060
atgctttgca tacttctgcc tgctggggag cctggggact ttccacacct ggttgctgac 9120
taattgagat gcatgctttg catacttctg cctgctgggg agcctgggga ctttccacac 9180
cctaactgac acacattcca cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg 9240
gtttgcgtat tgggcgctct tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc 9300
ggctgcggcg agcggtatca gctcactcaa aggcggtaat acggttatcc acagaatcag 9360

gggataacgc aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa 9420
aggccgcgtt gctggcgttt ttccataggc tccgcccccc tgacgagcat cacaaaaatc 9480
gacgctcaag tcagaggtgg cgaaacccga caggactata aagataccag gcgtttcccc 9540

ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga tacctgtccg 9600
cctttctccc ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt 9660
cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga accccccgtt cagcccgacc 9720
gctgcgcctt atccggtaac tatcgtcttg agtccaaccc ggtaagacac gacttatcgc 9780

cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag 9840
agttcttgaa gtggtggcct aactacggct acactagaag aacagtattt ggtatctgcg 9900
ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa 9960
ccaccgctgg tagcggtggt ttttttgttt gcaagcagca gattacgcgc agaaaaaaag 10020
gatctcaaga agatcctttg atcttttcta cggggtctga cgctcagtgg aacgaaaact 10080
cacgttaagg gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa 10140
attaaaaatg aagttttaaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt 10200

accaatgctt aatcagtgag gcacctatct cagcgatctg tctatttcgt tcatccatag 10260
ttgcctgact cctgcaaacc acgttgtgtc tcaaaatctc tgatgttaca ttgcacaaga 10320
taaaaatata tcatcatgaa caataaaact gtctgcttac ataaacagta atacaagggg 10380
tgttatgagc catattcaac gggaaacgtc ttgctcgagg ccgcgattaa attccaacat 10440
ggatgctgat ttatatgggt ataaatgggc tcgcgataat gtcgggcaat caggtgcgac 10500
aatctatcga ttgtatggga agcccgatgc gccagagttg tttctgaaac atggcaaagg 10560
tagcgttgcc aatgatgtta cagatgagat ggtcagacta aactggctga cggaatttat 10620

gcctcttccg accatcaagc attttatccg tactcctgat gatgcatggt tactcaccac 10680
tgcgatcccc gggaaaacag cattccaggt attagaagaa tatcctgatt caggtgaaaa 10740
tattgttgat gcgctggcag tgttcctgcg ccggttgcat tcgattcctg tttgtaattg 10800
tccttttaac agcgatcgcg tatttcgtct cgctcaggcg caatcacgaa tgaataacgg 10860
tttggttgat gcgagtgatt ttgatgacga gcgtaatggc tggcctgttg aacaagtctg 10920
gaaagaaatg cataagcttt tgccattctc accggattca gtcgtcactc atggtgattt 10980
ctcacttgat aaccttattt ttgacgaggg gaaattaata ggttgtattg atgttggacg 11040

agtcggaatc gcagaccgat accaggatct tgccatccta tggaactgcc tcggtgagtt 11100
ttctccttca ttacagaaac ggctttttca aaaatatggt attgataatc ctgatatgaa 11160
taaattgcag tttcatttga tgctcgatga gtttttctaa gggcggcctg ccaccatacc 11220
cacgccgaaa caagcgctca tgagcccgaa gtggcgagcc cgatcttccc catcggtgat 11280
gtcggcgata taggcgccag caaccgcacc tgtggcgccg gtgatgaggg cgcgccaagt 11340
cgacgtccgg cagtc 11355

<210> 3

<211> 11420

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 3

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420

tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctgcag ccctaggaat gcatctagac aattgtacta accttcttct 600

ctttcctctc ctgacagtcc ggaaagccac catggaattc agcagcccca gcagagagga 660

atgccccaag cctctgagcc gggtgtcaat catggccgga tctctgacag gactgctgct 720

gcttcaggcc gtgtcttggg cttctggcgc tagaccttgc atccccaaga gcttcggcta 780

cagcagcgtc gtgtgcgtgt gcaatgccac ctactgcgac agcttcgacc ctcctacctt 840

tcctgctctg ggcaccttca gcagatacga gagcaccaga tccggcagac ggatggaact 900

gagcatggga cccatccagg ccaatcacac aggcactggc ctgctgctga cactgcagcc 960

tgagcagaaa ttccagaaag tgaaaggctt cggcggagcc atgacagatg ccgccgctct 1020

gaatatcctg gctctgtctc caccagctca gaacctgctg ctcaagagct acttcagcga 1080

ggaaggcatc ggctacaaca tcatcagagt gcccatggcc agctgcgact tcagcatcag 1140

gacctacacc tacgccgaca cacccgacga tttccagctg cacaacttca gcctgcctga 1200

agaggacacc aagctgaaga tccctctgat ccacagagcc ctgcagctgg cacaaagacc 1260

cgtgtcactg ctggcctctc catggacatc tcccacctgg ctgaaaacaa atggcgccgt 1320

gaatggcaag ggcagcctga aaggccaacc tggcgacatc taccaccaga cctgggccag 1380

atacttcgtg aagttcctgg acgcctatgc cgagcacaag ctgcagtttt gggccgtgac 1440

agccgagaac gaaccttctg ctggactgct gagcggctac ccctttcagt gcctgggctt 1500

tacacccgag caccagcggg actttatcgc ccgtgatctg ggacccacac tggccaatag 1560
cacccaccat aatgtgcggc tgctgatgct ggacgaccag agactgcttc tgccccactg 1620
ggctaaagtg gtgctgacag atcctgaggc cgccaaatac gtgcacggaa tcgccgtgca 1680

ctggtatctg gactttctgg cccctgccaa ggccacactg ggagagacac acagactgtt 1740
ccccaacacc atgctgttcg ccagcgaagc ctgtgtgggc agcaagtttt gggaacagag 1800
cgtgcggctc ggcagctggg atagaggcat gcagtacagc cacagcatca tcaccaacct 1860
gctgtaccac gtcgtcggct ggaccgactg gaatctggcc ctgaatcctg aaggcggccc 1920
taactgggtc cgaaacttcg tggacagccc catcatcgtg gacatcacca aggacacctt 1980
ctacaagcag cccatgttct accacctggg acacttcagc aagttcatcc ccgagggctc 2040
tcagcgcgtt ggactggtgg cttcccagaa gaacgatctg gacgccgtgg ctctgatgca 2100

ccctgatgga tctgctgtgg tggtggtcct gaaccgcagc agcaaagatg tgcccctgac 2160
catcaaggat cccgccgtgg gattcctgga aacaatcagc cctggctact ccatccacac 2220
ctacctgtgg cgtagacagt gacaattgtt aattaagttt catcgatacc gtcgactaga 2280
gctcgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt ttgcccctcc 2340
cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta ataaaatgag 2400
gaaattgcat cgcattgtct gagtaggtgt cattctattc tggggggtgg ggtggggcag 2460
gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggagag atccacgata 2520

acaaacagct tttttggggg ggcggagtta gggcggagcc aatcagcgtg cgccgttccg 2580
aaagttgcct tttatggctg ggcggagaat gggcggtgaa cgccgatgat tatataagga 2640
cgcgccgggt gtggcacagc tagttccgtc gcagccggga tttgggtcgc ggttcttgtt 2700
tgtggatccc tgtgatcgtc acttggtaag tcactgactg tctatgcctg ggaaagggtg 2760
ggcaggagat ggggcagtgc aggaaaagtg gcactatgaa ccctgcagcc ctaggaatgc 2820
atctagacaa ttgtactaac cttcttctct ttcctctcct gacagtccgg aaagccacca 2880
tgggccgctg ctgcttctac accgccggca ccctgagcct gctgctgctg gtgaccagcg 2940

tgaccctgct ggtggcccgc gtgttccaga aggccgtgga ccagagcatc gagaagaaga 3000
tcgtgctgcg caacggcacc gaggccttcg acagctggga gaagcccccc ctgcccgtgt 3060
acacccagtt ctacttcttc aacgtgacca accccgagga gatcctgcgc ggcgagaccc 3120
cccgcgtgga ggaggtgggc ccctacacct accgcgagct gcgcaacaag gccaacatcc 3180
agttcggcga caacggcacc accatcagcg ccgtgagcaa caaggcctac gtgttcgagc 3240
gcgaccagag cgtgggcgac cccaagatcg acctgatccg caccctgaac atccccgtgc 3300

tgaccgtgat cgagtggagc caggtgcact tcctgcgcga gatcatcgag gccatgctga 3360

aggcctacca gcagaagctg ttcgtgaccc acaccgtgga cgagctgctg tggggctaca 3420
aggacgagat cctgagcctg atccacgtgt tccgccccga catcagcccc tacttcggcc 3480
tgttctacga gaagaacggc accaacgacg gcgactacgt gttcctgacc ggcgaggaca 3540
gctacctgaa cttcaccaag atcgtggagt ggaacggcaa gaccagcctg gactggtgga 3600
tcaccgacaa gtgcaacatg atcaacggca ccgacggcga cagcttccac cccctgatca 3660
ccaaggacga ggtgctgtac gtgttcccca gcgacttctg ccgcagcgtg tacatcacct 3720
tcagcgacta cgagagcgtg cagggcctgc ccgccttccg ctacaaggtg cccgccgaga 3780

tcctggccaa caccagcgac aacgccggct tctgcatccc cgagggcaac tgcctgggca 3840
gcggcgtgct gaacgtgagc atctgcaaga acggcgcccc catcatcatg agcttccccc 3900
acttctacca ggccgacgag cgcttcgtga gcgccatcga gggcatgcac cccaaccagg 3960
aggaccacga gaccttcgtg gacatcaacc ccctgaccgg catcatcctg aaggccgcca 4020
agcgcttcca gatcaacatc tacgtgaaga agctggacga cttcgtggag accggcgaca 4080
tccgcaccat ggtgttcccc gtgatgtacc tgaacgagag cgtgcacatc gacaaggaga 4140
ccgccagccg cctgaagagc atgatcaaca ccaccctgat catcaccaac atcccctaca 4200

tcatcatggc cctgggcgtg ttcttcggcc tggtgttcac ctggctggcc tgcaagggcc 4260
agggcagcat ggacgagggc accgccgacg agcgcgcccc cctgatccgc acctgaccca 4320
ggggactcaa tcagcctcga agacatgata agatacattg atgagtttgg acaaaccaca 4380
acaagaatgc agtgaaaaaa atgctttatt tgtgaaattt gtgatgctat tgctttattt 4440
gtaaccatta taagctgcaa taaacaagtt aacaacaaca attgcattca ttttatgttt 4500
caggttcagg gggagatgtg ggaggttttt taaagcaagt aaaacctcta caaatgtggt 4560
atgaacatat tgactgaatt ccctgcaggt tggccactcc ctctctgcgc gctcgctcgc 4620

tcactgaggc cgcccgggca aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag 4680
tgagcgagcg agcgcgcaga gagggagtgg ccaactccat cactaggggt tcctgcggcc 4740
gctcgtacgg tctcgaggaa ttcctgcagg ataacttgcc aacctcattc taaaatgtat 4800
atagaagccc aaaagacaat aacaaaaata ttcttgtaga acaaaatggg aaagaatgtt 4860
ccactaaata tcaagattta gagcaaagca tgagatgtgt ggggatagac agtgaggctg 4920
ataaaataga gtagagctca gaaacagacc cattgatata tgtaagtgac ctatgaaaaa 4980
aatatggcat tttacaatgg gaaaatgatg gtctttttct tttttagaaa aacagggaaa 5040

tatatttata tgtaaaaaat aaaagggaac ccatatgtca taccatacac acaaaaaaat 5100

tccagtgaat tataagtcta aatggagaag gcaaaacttt aaatctttta gaaaataata 5160

tagaagcatg cagaccagcc tggccaacat gatgaaaccc tctctactaa taataaaatc 5220

agtagaacta ctcaggacta ctttgagtgg gaagtccttt tctatgaaga cttctttggc 5280

caaaattagg ctctaaatgc aaggagatag tgcatcatgc ctggctgcac ttactgataa 5340

atgatgttat caccatcttt aaccaaatgc acaggaacaa gttatggtac tgatgtgctg 5400

gattgagaag gagctctact tccttgacag gacacatttg tatcaactta aaaaagcaga 5460

tttttgccag cagaactatt cattcagagg taggaaactt agaatagatg atgtcactga 5520

ttagcatggc ttccccatct ccacagctgc ttcccaccca ggttgcccac agttgagttt 5580

gtccagtgct cagggctgcc cactctcagt aagaagcccc acaccagccc ctctccaaat 5640

atgttggctg ttccttccat taaagtgacc ccactttaga gcagcaagtg gatttctgtt 5700

tcttacagtt caggaaggag gagtcagctg tgagaacctg gagcctgaga tgcttctaag 5760

tcccactgct actggggtca gggaagccag actccagcat cagcagtcag gagcactaag 5820

cccttgccaa catcctgttt ctcagagaaa ctgcttccat tataatggtt gtccttttt 5880

aagctatcaa gccaaacaac cagtgtctac cattattctc atcacctgaa gccaagggtt 5940

ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct ccagcttctg tcttcagtca 6000

ctccactctt agcctgctct gaatcaactc tgaccacagt tccctggagc ccctgccacc 6060

tgctgcccct gccaccttct ccatctgcag tgctgtgcag ccttctgcac tcttgcagag 6120

ctaataggtg gagacttgaa ggaagaggag gaaagtttct cataatagcc ttgctgcaag 6180

ctcaaatggg aggtgggcac tgtgcccagg agccttggag caaaggctgt gcccaacctc 6240

tgactgcatc caggtttggt cttgacagag ataagaagcc ctggcttttg gagccaaaat 6300

ctaggtcaga cttaggcagg attctcaaag tttatcagca gaacatgagg cagaagaccc 6360

tttctgctcc agcttcttca ggctcaacct tcatcagaat agatagaaag agaggctgtg 6420

agggttctta aaacagaagc aaatctgact cagagaataa acaacctcct agtaaactac 6480

agcttagaca gagcatctgg tggtgagtgt gctcagtgtc ctactcaact gtctggtatc 6540

agccctcatg aggacttctc ttctttccct catagacctc catctctgtt ttccttagcc 6600

tgcagaaatc tggatggcta ttcacagaat gcctgtgctt tcagagttgc atttttctc 6660

tggtattctg gttcaagcat ttgaaggtag gaaaggttct ccaagtgcaa gaaagccagc 6720

cctgagcctc aactgcctgg ctagtgtggt cagtaggatg caaaggctgt tgaatgccac 6780

aaggccaaac tttaacctgt gtaccacaag cctagcagca gaggcagctc tgctcactgg 6840

aactctctgt cttctttctc ctgagccttt tcttttcctg agttttctag ctctcctcaa 6900

ccttacctct gccctaccca ggacaaaccc aagagccact gtttctgtga tgtcctctcc 6960
agccctaatt aggcatcatg acttcagcct gaccttccat gctcagaagc agtgctaatc 7020
cacttcagat gagctgctct atgcaacaca ggcagagcct acaaaccttt gcaccagagc 7080
cctccacata tcagtgtttg ttcatactca cttcaacagc aaatgtgact gctgagatta 7140

agattttaca caagatggtc tgtaatttca cagttagttt tatcccatta ggtatgaaag 7200
aattagcata attcccctta aacatgaatg aatcttagat tttttaataa atagttttgg 7260
aagtaaagac agagacatca ggagcacaag gaatagcctg agaggacaaa cagaacaaga 7320
aagagtctgg aaatacacag gatgttcttg gcctcctcaa agcaagtgca agcagatagt 7380
accagcagcc ccaggctatc agagcccagt gaagagaagt accatgaaag ccacagctct 7440
aaccaccctg ttccagagtg acagacagtc cccaagacaa gccagcctga gccagagaga 7500
gaactgcaag agaaagtttc taatttaggt tctgttagat tcagacaagt gcaggtcatc 7560

ctctctccac agctactcac ctctccagcc taacaaagcc tgcagtccac actccaaccc 7620
tggtgtctca cctcctagcc tctcccaaca tcctgctctc tgaccatctt ctgcatctct 7680
catctcacca tctcccactg tctacagcct actcttgcaa ctaccatctc attttctgac 7740
atcctgtcta catcttctgc catactctgc catctaccat accacctctt accatctacc 7800
acaccatctt ttatctccat ccctctcaga agcctccaag ctgaatcctg ctttatgtgt 7860
tcatctcagc ccctgcatgg aaagctgacc ccagaggcag aactattccc agagagcttg 7920
gccaagaaaa acaaaactac cagcctggcc aggctcagga gtagtaagct gcagtgtctg 7980

ttgtgttcta gcttcaacag ctgcaggagt tccactctca aatgctccac atttctcaca 8040
tcctcctgat tctggtcact acccatcttc aaagaacaga atatctcaca tcagcatact 8100
gtgaaggact agtcatgggt gcagctgctc agagctgcaa agtcattctg gatggtggag 8160
agcttacaaa catttcatga tgctcccccc gctctgatgg ctggagccca atccctacac 8220
agactcctgc tgtatgtgtt ttcctttcac tctgagccac agccagaggg caggcattca 8280
gtctcctctt caggctgggg ctggggcact gagaactcac ccaacacctt gctctcactc 8340
cttctgcaaa acaagaaaga gctttgtgct gcagtagcca tgaagaatga aaggaaggct 8400

ttaactaaaa aatgtcagag attattttca acccccttact gtggatcacc agcaaggagg 8460
aaacacaaca cagagacatt ttttcccctc aaattatcaa aagaatcact gcatttgtta 8520
aagagagcaa ctgaatcagg aagcagagtt ttgaacatat cagaagttag gaatctgcat 8580
cagagacaaa tgcagtcatg gttgtttgct gcataccagc cctaatcatt agaagcctca 8640
tggacttcaa acatcattcc ctctgacaag atgctctagc ctaactccat gagataaaat 8700
aaatctgcct ttcagagcca aagaagagtc caccagcttc ttctcagtgt gaacaagagc 8760

tccagtcagg ttagtcagtc cagtgcagta gaggagacca gtctgcatcc tctaattttc 8820

aaaggcaaga agatttgttt accctggaca ccaggcacaa gtgaggtcac agagctctta 8880
gatatgcagt cctcatgagt gaggagacta aagcgcatgc catcaagact tcagtgtaga 8940
gaaaacctcc aaaaaagcct cctcactact tctggaatag ctcagaggcc gaggcggcct 9000
cggcctctgc ataaataaaa aaaattagtc agccatgggg cggagaatgg gcggaactgg 9060
gcggagttag gggcgggatg ggcggagtta ggggcgggac tatggttgct gactaattga 9120
gatgcatgct ttgcatactt ctgcctgctg gggagcctgg ggactttcca cacctggttg 9180
ctgactaatt gagatgcatg ctttgcatac ttctgcctgc tggggagcct ggggactttc 9240

cacaccctaa ctgacacaca ttccacagct gcattaatga atcggccaac gcgcggggag 9300
aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt 9360
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga 9420
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 9480
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc ccccctgacg agcatcacaa 9540
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 9600
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 9660

gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 9720
cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 9780
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 9840
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 9900
tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat 9960
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 10020
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 10080

aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 10140
aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct 10200
tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga 10260
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc 10320
catagttgcc tgactcctgc aaaccacgtt gtgtctcaaa atctctgatg ttacattgca 10380
caagataaaa atatatcatc atgaacaata aaactgtctg cttacataaa cagtaataca 10440
aggggtgtta tgagccatat tcaacgggaa acgtcttgct cgaggccgcg attaaattcc 10500

aacatggatg ctgatttata tgggtataaa tgggctcgcg ataatgtcgg gcaatcaggt 10560

gcgacaatct atcgattgta tgggaagccc gatgcgccag agttgtttct gaaacatggc 10620

aaaggtagcg ttgccaatga tgttacagat gagatggtca gactaaactg gctgacggaa 10680

tttatgcctc ttccgaccat caagcatttt atccgtactc ctgatgatgc atggttactc 10740

accactgcga tccccgggaa aacagcattc caggtattag aagaatatcc tgattcaggt 10800

gaaaatattg ttgatgcgct ggcagtgttc ctgcgccggt tgcattcgat tcctgtttgt 10860

aattgtcctt ttaacagcga tcgcgtattt cgtctcgctc aggcgcaatc acgaatgaat 10920

aacggtttgg ttgatgcgag tgattttgat gacgagcgta atggctggcc tgttgaacaa 10980

gtctggaaag aaatgcataa gcttttgcca ttctcaccgg attcagtcgt cactcatggt 11040

gatttctcac ttgataacct tatttttgac gaggggaaat taataggttg tattgatgtt 11100

ggacgagtcg gaatcgcaga ccgataccag gatcttgcca tcctatggaa ctgcctcggt 11160

gagttttctc cttcattaca gaaacggctt tttcaaaaat atggtattga taatcctgat 11220

atgaataaat tgcagtttca tttgatgctc gatgagtttt tctaagggcg gcctgccacc 11280

atacccacgc cgaaacaagc gctcatgagc ccgaagtggc gagcccgatc ttccccatcg 11340

gtgatgtcgg cgatataggc gccagcaacc gcacctgtgg cgccggtgat gagggcgcgc 11400

caagtcgacg tccggcagtc 11420

<210> 4

<211> 11171

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 4

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420

tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctcctg gtggcgaggg gagggggtg gtcctcgaac gccttgcaga 600

actggcctgg atacagagtg gaccggctgg ccccatctgg aagacttcga gatacactgt 660

tgtcttactg cgctcaacag tgtatctcga agtcttccaa atggtgccag ccatcgcagc 720
ggggtgcagg aaatgggggc agcccccctt tttggctatc cttccacgtg ttctttttg 780
tatcttttgt gtttcctaga aaacatctca gtcaccaccg cagccctagg aatgcatcta 840
gacaattgta ctaaccttct tctctttcct ctcctgacag tccggaaagc caccatgggc 900
cgctgctgct tctacaccgc cggcaccctg agcctgctgc tgctggtgac cagcgtgacc 960
ctgctggtgg cccgcgtgtt ccagaaggcc gtggaccaga gcatcgagaa gaagatcgtg 1020
ctgcgcaacg gcaccgaggc cttcgacagc tgggagaagc ccccccctgcc cgtgtacacc 1080

cagttctact tcttcaacgt gaccaacccc gaggagatcc tgcgcggcga gaccccccgc 1140
gtggaggagg tgggccccta cacctaccgc gagctgcgca acaaggccaa catccagttc 1200
ggcgacaacg gcaccaccat cagcgccgtg agcaacaagg cctacgtgtt cgagcgcgac 1260
cagagcgtgg gcgaccccaa gatcgacctg atccgcaccc tgaacatccc cgtgctgacc 1320
gtgatcgagt ggagccaggt gcacttcctg cgcgagatca tcgaggccat gctgaaggcc 1380
taccagcaga agctgttcgt gacccacacc gtggacgagc tgctgtgggg ctacaaggac 1440
gagatcctga gcctgatcca cgtgttccgc cccgacatca gcccctactt cggcctgttc 1500

tacgagaaga acggcaccaa cgacggcgac tacgtgttcc tgaccggcga ggacagctac 1560
ctgaacttca ccaagatcgt ggagtggaac ggcaagacca gcctggactg gtggatcacc 1620
gacaagtgca acatgatcaa cggcaccgac ggcgacagct tccacccccct gatcaccaag 1680
gacgaggtgc tgtacgtgtt ccccagcgac ttctgccgca gcgtgtacat caccttcagc 1740
gactacgaga gcgtgcaggg cctgcccgcc ttccgctaca aggtgcccgc cgagatcctg 1800
gccaacacca gcgacaacgc cggcttctgc atccccgagg gcaactgcct gggcagcggc 1860
gtgctgaacg tgagcatctg caagaacggc gcccccatca tcatgagctt cccccacttc 1920

taccaggccg acgagcgctt cgtgagcgcc atcgagggca tgcaccccaa ccaggaggac 1980
cacgagacct tcgtggacat caaccccctg accggcatca tcctgaaggc cgccaagcgc 2040
ttccagatca acatctacgt gaagaagctg gacgacttcg tggagaccgg cgacatccgc 2100
accatggtgt tccccgtgat gtacctgaac gagagcgtgc acatcgacaa ggagaccgcc 2160
agccgcctga agagcatgat caacaccacc ctgatcatca ccaacatccc ctacatcatc 2220
atggccctgg gcgtgttctt cggcctggtg ttcacctggc tggcctgcaa gggccagggc 2280
agcatggacg agggcaccgc cgacgagcgc gcccccctga tccgcaccga gggcagagga 2340

agtcttctga catgcggaga cgtggaagag aatcccggcc ctatggaatt cagcagcccc 2400
agcagagagg aatgccccaa gcctctgagc cgggtgtcaa tcatggccgg atctctgaca 2460
ggactgctgc tgcttcaggc cgtgtcttgg gcttctggcg ctagaccttg catccccaag 2520
agcttcggct acagcagcgt cgtgtgcgtg tgcaatgcca cctactgcga cagcttcgac 2580
cctcctacct ttcctgctct gggcaccttc agcagatacg agagcaccag atccggcaga 2640
cggatggaac tgagcatggg acccatccag gccaatcaca caggcactgg cctgctgctg 2700
acactgcagc ctgagcagaa attccagaaa gtgaaaggct tcggcggagc catgacagat 2760

gccgccgctc tgaatatcct ggctctgtct ccaccagctc agaacctgct gctcaagagc 2820
tacttcagcg aggaaggcat cggctacaac atcatcagag tgcccatggc cagctgcgac 2880
ttcagcatca ggacctacac ctacgccgac acacccgacg atttccagct gcacaacttc 2940
agcctgcctg aagaggacac caagctgaag atccctctga tccacagagc cctgcagctg 3000
gcacaaagac ccgtgtcact gctggcctct ccatggacat ctcccacctg gctgaaaaca 3060
aatggcgccg tgaatggcaa gggcagcctg aaaggccaac ctggcgacat ctaccaccag 3120
acctgggcca gatacttcgt gaagttcctg gacgcctatg ccgagcacaa gctgcagttt 3180

tgggccgtga cagccgagaa cgaaccttct gctggactgc tgagcggcta ccccttcag 3240
tgcctgggct ttacacccga gcaccagcgg gactttatcg cccgtgatct gggacccaca 3300
ctggccaata gcacccacca taatgtgcgg ctgctgatgc tggacgacca gagactgctt 3360
ctgccccact gggctaaagt ggtgctgaca gatcctgagg ccgccaaata cgtgcacgga 3420
atcgccgtgc actggtatct ggactttctg gcccctgcca aggccacact gggagagaca 3480
cacagactgt tccccaacac catgctgttc gccagcgaag cctgtgtggg cagcaagttt 3540
tgggaacaga gcgtgcggct cggcagctgg gatagaggca tgcagtacag ccacagcatc 3600

atcaccaacc tgctgtacca cgtcgtcggc tggaccgact ggaatctggc cctgaatcct 3660
gaaggcggcc ctaactgggt ccgaaacttc gtggacagcc ccatcatcgt ggacatcacc 3720
aaggacacct tctacaagca gcccatgttc taccacctgg gacacttcag caagttcatc 3780
cccgagggct ctcagcgcgt tggactggtg gcttcccaga agaacgatct ggacgccgtg 3840
gctctgatgc accctgatgg atctgctgtg gtggtggtcc tgaaccgcag cagcaaagat 3900
gtgcccctga ccatcaagga tcccgccgtg ggattcctgg aaacaatcag ccctggctac 3960
tccatccaca cctacctgtg gcgtagacag tgacaattgt taattaagtt taaaccctcg 4020

aggccgcaag ccgcatcgat accgtcgact agagctcgct gatcagcctc gactgtgcct 4080
tctagttgcc agccatctgt tgtttgcccc tcccccgtgc cttccttgac cctggaaggt 4140
gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgcattg tctgagtagg 4200

tgtcattcta ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac 4260
aatagcaggc atgctgggga gagatccacg ataacaaaca gctttttgg ggtgaacata 4320
ttgactgaat tccctgcagg ttggccactc cctctctgcg cgctcgctcg ctcactgagg 4380
ccgcccgggc aaagcccggg cgtcgggcga cctttggtcg cccggcctca gtgagcgagc 4440

gagcgcgcag agagggagtg gccaactcca tcactagggg ttcctgcggc cgctcgtacg 4500
gtctcgagga attcctgcag gataacttgc caacctcatt ctaaaatgta tatagaagcc 4560
caaaagacaa taacaaaaat attcttgtag aacaaaatgg gaaagaatgt tccactaaat 4620
atcaagattt agagcaaagc atgagatgtg tggggataga cagtgaggct gataaaatag 4680
agtagagctc agaaacagac ccattgatat atgtaagtga cctatgaaaa aaatatggca 4740
ttttacaatg ggaaaatgat ggtctttttc ttttttagaa aaacagggaa atatatttat 4800
atgtaaaaaa taaaagggaa cccatatgtc ataccataca cacaaaaaaa ttccagtgaa 4860

ttataagtct aaatggagaa ggcaaaactt taaatctttt agaaaataat atagaagcat 4920
gcagaccagc ctggccaaca tgatgaaacc ctctctacta ataataaaat cagtagaact 4980
actcaggact actttgagtg ggaagtcctt ttctatgaag acttctttgg ccaaaattag 5040
gctctaaatg caaggagata gtgcatcatg cctggctgca cttactgata aatgatgtta 5100
tcaccatctt taaccaaatg cacaggaaca agttatggta ctgatgtgct ggattgagaa 5160
ggagctctac ttccttgaca ggacacattt gtatcaactt aaaaaagcag atttttgcca 5220
gcagaactat tcattcagag gtaggaaact tagaatagat gatgtcactg attagcatgg 5280

cttccccatc tccacagctg cttcccaccc aggttgccca cagttgagtt tgtccagtgc 5340
tcagggctgc ccactctcag taagaagccc cacaccagcc cctctccaaa tatgttggct 5400
gttccttcca ttaaagtgac cccactttag agcagcaagt ggatttctgt ttcttacagt 5460
tcaggaagga ggagtcagct gtgagaacct ggagcctgag atgcttctaa gtcccactgc 5520
tactggggtc agggaagcca gactccagca tcagcagtca ggagcactaa gcccttgcca 5580
acatcctgtt tctcagagaa actgcttcca ttataatggt tgtccttttt taagctatca 5640
agccaaacaa ccagtgtcta ccattattct catcacctga agccaagggt tctagcaaaa 5700

gtcaagctgt cttgtaatgg ttgatgtgcc tccagcttct gtcttcagtc actccactct 5760
tagcctgctc tgaatcaact ctgaccacag ttccctggag cccctgccac ctgctgcccc 5820
tgccaccttc tccatctgca gtgctgtgca gccttctgca ctcttgcaga gctaataggt 5880
ggagacttga aggaagagga ggaaagtttc tcataatagc cttgctgcaa gctcaaatgg 5940
gaggtgggca ctgtgcccag gagccttgga gcaaaggctg tgcccaacct ctgactgcat 6000
ccaggtttgg tcttgacaga gataagaagc cctggctttt ggagccaaaa tctaggtcag 6060

acttaggcag gattctcaaa gtttatcagc agaacatgag gcagaagacc ctttctgctc 6120

cagcttcttc aggctcaacc ttcatcagaa tagatagaaa gagaggctgt gagggttctt 6180
aaaacagaag caaatctgac tcagagaata aacaacctcc tagtaaacta cagcttagac 6240
agagcatctg gtggtgagtg tgctcagtgt cctactcaac tgtctggtat cagccctcat 6300
gaggacttct cttctttccc tcatagacct ccatctctgt tttccttagc ctgcagaaat 6360
ctggatggct attcacagaa tgcctgtgct ttcagagttg catttttct ctggtattct 6420
ggttcaagca tttgaaggta ggaaaggttc tccaagtgca agaaagccag ccctgagcct 6480
caactgcctg gctagtgtgg tcagtaggat gcaaaggctg ttgaatgcca caaggccaaa 6540

ctttaacctg tgtaccacaa gcctagcagc agaggcagct ctgctcactg gaactctctg 6600
tcttctttct cctgagcctt ttcttttcct gagttttcta gctctcctca accttacctc 6660
tgccctaccc aggacaaacc caagagccac tgtttctgtg atgtcctctc cagccctaat 6720
taggcatcat gacttcagcc tgaccttcca tgctcagaag cagtgctaat ccacttcaga 6780
tgagctgctc tatgcaacac aggcagagcc tacaaacctt tgcaccagag ccctccacat 6840
atcagtgttt gttcatactc acttcaacag caaatgtgac tgctgagatt aagattttac 6900
acaagatggt ctgtaatttc acagttagtt ttatcccatt aggtatgaaa gaattagcat 6960

aattcccctt aaacatgaat gaatcttaga ttttttaata aatagttttg gaagtaaaga 7020
cagagacatc aggagcacaa ggaatagcct gagaggacaa acagaacaag aaagagtctg 7080
gaaatacaca ggatgttctt ggcctcctca aagcaagtgc aagcagatag taccagcagc 7140
cccaggctat cagagcccag tgaagagaag taccatgaaa gccacagctc taaccaccct 7200
gttccagagt gacagacagt ccccaagaca agccagcctg agccagagag agaactgcaa 7260
gagaaagttt ctaatttagg ttctgttaga ttcagacaag tgcaggtcat cctctctcca 7320
cagctactca cctctccagc ctaacaaagc ctgcagtcca cactccaacc ctggtgtctc 7380

acctcctagc ctctcccaac atcctgctct ctgaccatct tctgcatctc tcatctcacc 7440
atctcccact gtctacagcc tactcttgca actaccatct cattttctga catcctgtct 7500
acatcttctg ccatactctg ccatctacca taccacctct taccatctac cacaccatct 7560
tttatctcca tccctctcag aagcctccaa gctgaatcct gctttatgtg ttcatctcag 7620
cccctgcatg gaaagctgac cccagaggca gaactattcc cagagagctt ggccaagaaa 7680
aacaaaacta ccagcctggc caggctcagg agtagtaagc tgcagtgtct gttgtgttct 7740
agcttcaaca gctgcaggag ttccactctc aaatgctcca catttctcac atcctcctga 7800

ttctggtcac tacccatctt caaagaacag aatatctcac atcagcatac tgtgaaggac 7860
tagtcatggg tgcagctgct cagagctgca aagtcattct ggatggtgga gagcttacaa 7920
acatttcatg atgctccccc cgctctgatg gctggagccc aatccctaca cagactcctg 7980
ctgtatgtgt tttcctttca ctctgagcca cagccagagg gcaggcattc agtctcctct 8040
tcaggctggg gctggggcac tgagaactca cccaacacct tgctctcact ccttctgcaa 8100
aacaagaaag agctttgtgc tgcagtagcc atgaagaatg aaaggaaggc tttaactaaa 8160
aaatgtcaga gattattttc aacccccttac tgtggatcac cagcaaggag gaaacacaac 8220

acagagacat tttttcccct caaattatca aaagaatcac tgcatttgtt aaagagagca 8280
actgaatcag gaagcagagt tttgaacata tcagaagtta ggaatctgca tcagagacaa 8340
atgcagtcat ggttgtttgc tgcataccag ccctaatcat tagaagcctc atggacttca 8400
aacatcattc cctctgacaa gatgctctag cctaactcca tgagataaaa taaatctgcc 8460
tttcagagcc aaagaagagt ccaccagctt cttctcagtg tgaacaagag ctccagtcag 8520
gttagtcagt ccagtgcagt agaggagacc agtctgcatc ctctaatttt caaaggcaag 8580
aagatttgtt taccctggac accaggcaca agtgaggtca cagagctctt agatatgcag 8640

tcctcatgag tgaggagact aaagcgcatg ccatcaagac ttcagtgtag agaaaacctc 8700
caaaaaagcc tcctcactac ttctggaata gctcagaggc cgaggcggcc tcggcctctg 8760
cataaataaa aaaaattagt cagccatggg gcggagaatg ggcggaactg ggcggagtta 8820
ggggcgggat gggcggagtt aggggcggga ctatggttgc tgactaattg agatgcatgc 8880
tttgcatact tctgcctgct ggggagcctg gggactttcc acacctggtt gctgactaat 8940
tgagatgcat gctttgcata cttctgcctg ctggggagcc tggggacttt ccacacccta 9000
actgacacac attccacagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt 9060

gcgtattggg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct 9120
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 9180
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 9240
cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg 9300
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 9360
aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 9420
tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt 9480

gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg 9540
cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact 9600
ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt 9660

cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggta tctgcgctct 9720

gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 9780

cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 9840

tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg 9900

ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 9960

aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca 10020

atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc 10080

ctgactcctg caaaccacgt tgtgtctcaa aatctctgat gttacattgc acaagataaa 10140

aatatatcat catgaacaat aaaactgtct gcttacataa acagtaatac aaggggtgtt 10200

atgagccata ttcaacggga aacgtcttgc tcgaggccgc gattaaattc caacatggat 10260

gctgatttat atgggtataa atgggctcgc gataatgtcg ggcaatcagg tgcgacaatc 10320

tatcgattgt atgggaagcc cgatgcgcca gagttgtttc tgaaacatgg caaaggtagc 10380

gttgccaatg atgttacaga tgagatggtc agactaaact ggctgacgga atttatgcct 10440

cttccgacca tcaagcattt tatccgtact cctgatgatg catggttact caccactgcg 10500

atccccggga aaacagcatt ccaggtatta gaagaatatc ctgattcagg tgaaaatatt 10560

gttgatgcgc tggcagtgtt cctgcgccgg ttgcattcga ttcctgtttg taattgtcct 10620

tttaacagcg atcgcgtatt tcgtctcgct caggcgcaat cacgaatgaa taacggtttg 10680

gttgatgcga gtgattttga tgacgagcgt aatggctggc ctgttgaaca agtctggaaa 10740

gaaatgcata agcttttgcc attctcaccg gattcagtcg tcactcatgg tgatttctca 10800

cttgataacc ttatttttga cgaggggaaa ttaataggtt gtattgatgt tggacgagtc 10860

ggaatcgcag accgatacca ggatcttgcc atcctatgga actgcctcgg tgagttttct 10920

ccttcattac agaaacggct ttttcaaaaa tatggtattg ataatcctga tatgaataaa 10980

ttgcagtttc atttgatgct cgatgagttt ttctaagggc ggcctgccac catacccacg 11040

ccgaaacaag cgctcatgag cccgaagtgg cgagcccgat cttccccatc ggtgatgtcg 11100

gcgatatagg cgccagcaac cgcacctgtg gcgccggtga tgagggcgcg ccaagtcgac 11160

gtccggcagt c 11171

<210> 5

<211> 11309

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 5

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420

tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctcctg gtggcgaggg gagggggtg gtcctcgaac gccttgcaga 600

actggcctgg atacagagtg gaccggctgg ccccatctgg aagacttcga gatacactgt 660

tgtcttactg cgctcaacag tgtatctcga agtcttccaa atggtgccag ccatcgcagc 720

ggggtgcagg aaatgggggc agcccccctt tttggctatc cttccacgtg ttctttttg 780

tatcttttgt gtttcctaga aaacatctca gtcaccaccg cagccctagg aatgcatcta 840

gacaattgta ctaaccttct tctctttcct ctcctgacag tccggaaagc caccatgtac 900

gccctgttcc tgctggccag cctgctgggc gccgccctgg ccggccccgt gctgggcctg 960

aaggagtgca cccgcggcag cgccgtgtgg tgccagaacg tgaagaccgc cagcgactgc 1020

ggcgccgtga agcactgcct gcagaccgtg tggaacaagc ccaccgtgaa gagcctgccc 1080

tgcgacatct gcaaggacgt ggtgaccgcc gccggcgaca tgctgaagga caacgccacc 1140

gaggaggaga tcctggtgta cctggagaag acctgcgact ggctgcccaa gcccaacatg 1200

agcgccagct gcaaggagat cgtggacagc tacctgcccg tgatcctgga catcatcaag 1260

ggcgagatga gccgccccgg cgaggtgtgc agcgccctga acctgtgcga gagcctgcag 1320

aagcacctgg ccgagctgaa ccaccagaag cagctggaga gcaacaagat ccccgagctg 1380

gacatgaccg aggtggtggc ccccttcatg gccaacatcc ccctgctgct gtacccccag 1440

gacggccccc gcagcaagcc ccagcccaag gacaacggcg acgtgtgcca ggactgcatc 1500

cagatggtga ccgacatcca gaccgccgtg cgcaccaaca gcaccttcgt gcaggccctg 1560

gtggagcacg tgaaggagga gtgcgaccgc ctgggccccg gcatggccga catctgcaag 1620

aactacatca gccagtacag cgagatcgcc atccagatga tgatgcacat gcagcccaag 1680

gagatctgcg ccctggtggg cttctgcgac gaggtgaagg agatgcccat gcagaccctg 1740

gtgcccgcca aggtggccag caagaacgtg atccccgccc tggagctggt ggagcccatc 1800

aagaagcacg aggtgcccgc caagagcgac gtgtactgcg aggtgtgcga gttcctggtg 1860
aaggaggtga ccaagctgat cgacaacaac aagaccgaga aggagatcct ggacgccttc 1920
gacaagatgt gcagcaagct gcccaagagc ctgagcgagg agtgccagga ggtggtggac 1980

acctacggca gcagcatcct gagcatcctg ctggaggagg tgagccccga gctggtgtgc 2040
agcatgctgc acctgtgcag cggcacccgc ctgcccgccc tgaccgtgca cgtgacccag 2100
cccaaggacg gcggcttctg cgaggtgtgc aagaagctgg tgggctacct ggaccgcaac 2160
ctggagaaga acagcaccaa gcaggagatc ctggccgccc tggagaaggg ctgcagcttc 2220
ctgcccgacc cctaccagaa gcagtgcgac cagttcgtgg ccgagtacga gcccgtgctg 2280
atcgagatcc tggtggaggt gatggacccc agcttcgtgt gcctgaagat cggcgcctgc 2340
cccagcgccc acaagcccct gctgggcacc gagaagtgca tctggggccc cagctactgg 2400

tgccagaaca ccgagaccgc cgcccagtgc aacgccgtgg agcactgcaa gcgccacgtg 2460
tggaacgagg gcagaggaag tcttctgaca tgcggagacg tggaagagaa tcccggccct 2520
atggaattca gcagccccag cagagaggaa tgccccaagc ctctgagccg ggtgtcaatc 2580
atggccggat ctctgacagg actgctgctg cttcaggccg tgtcttgggc ttctggcgct 2640
agaccttgca tccccaagag cttcggctac agcagcgtcg tgtgcgtgtg caatgccacc 2700
tactgcgaca gcttcgaccc tcctaccttt cctgctctgg gcaccttcag cagatacgag 2760
agcaccagat ccggcagacg gatggaactg agcatgggac ccatccaggc caatcacaca 2820

ggcactggcc tgctgctgac actgcagcct gagcagaaat tccagaaagt gaaaggcttc 2880
ggcggagcca tgacagatgc cgccgctctg aatatcctgg ctctgtctcc accagctcag 2940
aacctgctgc tcaagagcta cttcagcgag gaaggcatcg gctacaacat catcagagtg 3000
cccatggcca gctgcgactt cagcatcagg acctacacct acgccgacac acccgacgat 3060
ttccagctgc acaacttcag cctgcctgaa gaggacacca agctgaagat ccctctgatc 3120
cacagagccc tgcagctggc acaaagaccc gtgtcactgc tggcctctcc atggacatct 3180
cccacctggc tgaaaacaaa tggcgccgtg aatggcaagg gcagcctgaa aggccaacct 3240

ggcgacatct accaccagac ctgggccaga tacttcgtga agttcctgga cgcctatgcc 3300
gagcacaagc tgcagttttg ggccgtgaca gccgagaacg aaccttctgc tggactgctg 3360
agcggctacc cctttcagtg cctgggcttt acacccgagc accagcggga ctttatcgcc 3420
cgtgatctgg gacccacact ggccaatagc acccaccata atgtgcggct gctgatgctg 3480
gacgaccaga gactgcttct gccccactgg gctaaagtgg tgctgacaga tcctgaggcc 3540
gccaaatacg tgcacggaat cgccgtgcac tggtatctgg actttctggc ccctgccaag 3600

gccacactgg gagagacaca cagactgttc cccaacacca tgctgttcgc cagcgaagcc 3660

tgtgtgggca gcaagttttg ggaacagagc gtgcggctcg gcagctggga tagaggcatg 3720
cagtacagcc acagcatcat caccaacctg ctgtaccacg tcgtcggctg gaccgactgg 3780
aatctggccc tgaatcctga aggcggccct aactgggtcc gaaacttcgt ggacagcccc 3840
atcatcgtgg acatcaccaa ggacaccttc tacaagcagc ccatgttcta ccacctggga 3900
cacttcagca agttcatccc cgagggctct cagcgcgttg gactggtggc ttcccagaag 3960
aacgatctgg acgccgtggc tctgatgcac cctgatggat ctgctgtggt ggtggtcctg 4020
aaccgcagca gcaaagatgt gcccctgacc atcaaggatc ccgccgtggg attcctggaa 4080

acaatcagcc ctggctactc catccacacc tacctgtggc gtagacagtg acaattgtta 4140
attaagttta aaccctcgag gccgcaagcc gcatcgatac cgtcgactag agctcgctga 4200
tcagcctcga ctgtgccttc tagttgccag ccatctgttg tttgcccctc ccccgtgcct 4260
tccttgaccc tggaaggtgc cactcccact gtcctttcct aataaaatga ggaaattgca 4320
tcgcattgtc tgagtaggtg tcattctatt ctggggggtg gggtggggca ggacagcaag 4380
ggggaggatt gggaagacaa tagcaggcat gctggggaga gatccacgat aacaaacagc 4440
ttttttgggg tgaacatatt gactgaattc cctgcaggtt ggccactccc tctctgcgcg 4500

ctcgctcgct cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc tttggtcgcc 4560
cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc actaggggtt 4620
cctgcggccg ctcgtacggt ctcgaggaat tcctgcagga taacttgcca acctcattct 4680
aaaatgtata tagaagccca aaagacaata acaaaaatat tcttgtagaa caaaatggga 4740
aagaatgttc cactaaatat caagatttag agcaaagcat gagatgtgtg gggatagaca 4800
gtgaggctga taaaatagag tagagctcag aaacagaccc attgatatat gtaagtgacc 4860
tatgaaaaaa atatggcatt ttacaatggg aaaatgatgg tctttttctt ttttagaaaa 4920

acagggaaat atatttatat gtaaaaaata aaagggaacc catatgtcat accatacaca 4980
caaaaaaatt ccagtgaatt ataagtctaa atggagaagg caaaacttta aatcttttag 5040
aaaataatat agaagcatgc agaccagcct ggccaacatg atgaaaccct ctctactaat 5100
aataaaatca gtagaactac tcaggactac tttgagtggg aagtcctttt ctatgaagac 5160
ttctttggcc aaaattaggc tctaaatgca aggagatagt gcatcatgcc tggctgcact 5220
tactgataaa tgatgttatc accatcttta accaaatgca caggaacaag ttatggtact 5280
gatgtgctgg attgagaagg agctctactt ccttgacagg acacatttgt atcaacttaa 5340

aaaagcagat ttttgccagc agaactattc attcagaggt aggaaactta gaatagatga 5400
tgtcactgat tagcatggct tccccatctc cacagctgct tcccacccag gttgcccaca 5460
gttgagtttg tccagtgctc agggctgccc actctcagta agaagcccca caccagcccc 5520
tctccaaata tgttggctgt tccttccatt aaagtgaccc cactttagag cagcaagtgg 5580
atttctgttt cttacagttc aggaaggagg agtcagctgt gagaacctgg agcctgagat 5640
gcttctaagt cccactgcta ctggggtcag ggaagccaga ctccagcatc agcagtcagg 5700
agcactaagc ccttgccaac atcctgtttc tcagagaaac tgcttccatt ataatggttg 5760

tccttttta agctatcaag ccaaacaacc agtgtctacc attattctca tcacctgaag 5820
ccaagggttc tagcaaaagt caagctgtct tgtaatggtt gatgtgcctc cagcttctgt 5880
cttcagtcac tccactctta gcctgctctg aatcaactct gaccacagtt ccctggagcc 5940
cctgccacct gctgcccctg ccaccttctc catctgcagt gctgtgcagc cttctgcact 6000
cttgcagagc taataggtgg agacttgaag gaagaggagg aaagtttctc ataatagcct 6060
tgctgcaagc tcaaatggga ggtgggcact gtgcccagga gccttggagc aaaggctgtg 6120
cccaacctct gactgcatcc aggtttggtc ttgacagaga taagaagccc tggcttttgg 6180

agccaaaatc taggtcagac ttaggcagga ttctcaaagt ttatcagcag aacatgaggc 6240
agaagaccct ttctgctcca gcttcttcag gctcaacctt catcagaata gatagaaaga 6300
gaggctgtga gggttcttaa aacagaagca aatctgactc agagaataaa caacctccta 6360
gtaaactaca gcttagacag agcatctggt ggtgagtgtg ctcagtgtcc tactcaactg 6420
tctggtatca gccctcatga ggacttctct tctttccctc atagacctcc atctctgttt 6480
tccttagcct gcagaaatct ggatggctat tcacagaatg cctgtgcttt cagagttgca 6540
ttttttctct ggtattctgg ttcaagcatt tgaaggtagg aaaggttctc caagtgcaag 6600

aaagccagcc ctgagcctca actgcctggc tagtgtggtc agtaggatgc aaaggctgtt 6660
gaatgccaca aggccaaact ttaacctgtg taccacaagc ctagcagcag aggcagctct 6720
gctcactgga actctctgtc ttctttctcc tgagcctttt cttttcctga gttttctagc 6780
tctcctcaac cttacctctg ccctacccag gacaaaccca agagccactg tttctgtgat 6840
gtcctctcca gccctaatta ggcatcatga cttcagcctg accttccatg ctcagaagca 6900
gtgctaatcc acttcagatg agctgctcta tgcaacacag gcagagccta caaacctttg 6960
caccagagcc ctccacatat cagtgtttgt tcatactcac ttcaacagca aatgtgactg 7020

ctgagattaa gattttacac aagatggtct gtaatttcac agttagtttt atcccattag 7080
gtatgaaaga attagcataa ttccccttaa acatgaatga atcttagatt ttttaataaa 7140
tagttttgga agtaaagaca gagacatcag gagcacaagg aatagcctga gaggacaaac 7200

agaacaagaa agagtctgga aatacacagg atgttcttgg cctcctcaaa gcaagtgcaa 7260
gcagatagta ccagcagccc caggctatca gagcccagtg aagagaagta ccatgaaagc 7320
cacagctcta accaccctgt tccagagtga cagacagtcc ccaagacaag ccagcctgag 7380
ccagagagag aactgcaaga gaaagtttct aatttaggtt ctgttagatt cagacaagtg 7440

caggtcatcc tctctccaca gctactcacc tctccagcct aacaaagcct gcagtccaca 7500
ctccaaccct ggtgtctcac ctcctagcct ctcccaacat cctgctctct gaccatcttc 7560
tgcatctctc atctcaccat ctcccactgt ctacagccta ctcttgcaac taccatctca 7620
ttttctgaca tcctgtctac atcttctgcc atactctgcc atctaccata ccacctctta 7680
ccatctacca caccatcttt tatctccatc cctctcagaa gcctccaagc tgaatcctgc 7740
tttatgtgtt catctcagcc cctgcatgga aagctgaccc cagaggcaga actattccca 7800
gagagcttgg ccaagaaaaa caaaactacc agcctggcca ggctcaggag tagtaagctg 7860

cagtgtctgt tgtgttctag cttcaacagc tgcaggagtt ccactctcaa atgctccaca 7920
tttctcacat cctcctgatt ctggtcacta cccatcttca aagaacagaa tatctcacat 7980
cagcatactg tgaaggacta gtcatgggtg cagctgctca gagctgcaaa gtcattctgg 8040
atggtggaga gcttacaaac atttcatgat gctccccccg ctctgatggc tggagcccaa 8100
tccctacaca gactcctgct gtatgtgttt tcctttcact ctgagccaca gccagagggc 8160
aggcattcag tctcctcttc aggctggggc tggggcactg agaactcacc caacaccttg 8220
ctctcactcc ttctgcaaaa caagaaagag ctttgtgctg cagtagccat gaagaatgaa 8280

aggaaggctt taactaaaaa atgtcagaga ttattttcaa ccccttactg tggatcacca 8340
gcaaggagga aacacaacac agagacattt tttcccctca aattatcaaa agaatcactg 8400
catttgttaa agagagcaac tgaatcagga agcagagttt tgaacatatc agaagttagg 8460
aatctgcatc agagacaaat gcagtcatgg ttgtttgctg cataccagcc ctaatcatta 8520
gaagcctcat ggacttcaaa catcattccc tctgacaaga tgctctagcc taactccatg 8580
agataaaata aatctgcctt tcagagccaa agaagagtcc accagcttct tctcagtgtg 8640
aacaagagct ccagtcaggt tagtcagtcc agtgcagtag aggagaccag tctgcatcct 8700

ctaattttca aaggcaagaa gatttgttta ccctggacac caggcacaag tgaggtcaca 8760
gagctcttag atatgcagtc ctcatgagtg aggagactaa agcgcatgcc atcaagactt 8820
cagtgtagag aaaacctcca aaaaagcctc ctcactactt ctggaatagc tcagaggccg 8880
aggcggcctc ggcctctgca taaataaaaa aaattagtca gccatggggc ggagaatggg 8940
cggaactggg cggagttagg ggcgggatgg gcggagttag gggcgggact atggttgctg 9000
actaattgag atgcatgctt tgcatacttc tgcctgctgg ggagcctggg gactttccac 9060

acctggttgc tgactaattg agatgcatgc tttgcatact tctgcctgct ggggagcctg 9120

gggactttcc acaccctaac tgacacacat tccacagctg cattaatgaa tcggccaacg 9180
cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct 9240
gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt 9300
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 9360
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga 9420
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 9480
ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac 9540

cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 9600
taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 9660
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag 9720
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt 9780
aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaagaacagt 9840
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 9900
atccggcaaa caaaccaccg ctggtagcgg tggtttttt gtttgcaagc agcagattac 9960

gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca 10020
gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 10080
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 10140
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 10200
tcgttcatcc atagttgcct gactcctgca aaccacgttg tgtctcaaaa tctctgatgt 10260
tacattgcac aagataaaaa tatatcatca tgaacaataa aactgtctgc ttacataaac 10320
agtaatacaa ggggtgttat gagccatatt caacgggaaa cgtcttgctc gaggccgcga 10380

ttaaattcca acatggatgc tgatttatat gggtataaat gggctcgcga taatgtcggg 10440
caatcaggtg cgacaatcta tcgattgtat gggaagcccg atgcgccaga gttgtttctg 10500
aaacatggca aaggtagcgt tgccaatgat gttacagatg agatggtcag actaaactgg 10560
ctgacggaat ttatgcctct tccgaccatc aagcatttta tccgtactcc tgatgatgca 10620
tggttactca ccactgcgat ccccgggaaa acagcattcc aggtattaga agaatatcct 10680
gattcaggtg aaaatattgt tgatgcgctg gcagtgttcc tgcgccggtt gcattcgatt 10740
cctgtttgta attgtccttt taacagcgat cgcgtatttc gtctcgctca ggcgcaatca 10800

cgaatgaata acggtttggt tgatgcgagt gattttgatg acgagcgtaa tggctggcct 10860

gttgaacaag tctggaaaga aatgcataag cttttgccat tctcaccgga ttcagtcgtc 10920

actcatggtg atttctcact tgataacctt atttttgacg aggggaaatt aataggttgt 10980

attgatgttg gacgagtcgg aatcgcagac cgataccagg atcttgccat cctatggaac 11040

tgcctcggtg agttttctcc ttcattacag aaacggcttt ttcaaaaata tggtattgat 11100

aatcctgata tgaataaatt gcagtttcat ttgatgctcg atgagttttt ctaagggcgg 11160

cctgccacca tacccacgcc gaaacaagcg ctcatgagcc cgaagtggcg agcccgatct 11220

tccccatcgg tgatgtcggc gatataggcg ccagcaaccg cacctgtggc gccggtgatg 11280

agggcgcgcc aagtcgacgt ccggcagtc 11309

<210> 6

<211> 11293

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 6

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420

tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctgcag ccctaggaat gcatctagac aattgtacta accttcttct 600

ctttcctctc ctgacagtcc ggaaagccac catgtacgcc ctgttcctgc tggccagcct 660

gctgggcgcc gccctggccg gccccgtgct gggcctgaag gagtgcaccc gcggcagcgc 720

cgtgtggtgc cagaacgtga agaccgccag cgactgcggc gccgtgaagc actgcctgca 780

gaccgtgtgg aacaagccca ccgtgaagag cctgccctgc gacatctgca aggacgtggt 840

gaccgccgcc ggcgacatgc tgaaggacaa cgccaccgag gaggagatcc tggtgtacct 900

ggagaagacc tgcgactggc tgcccaagcc caacatgagc gccagctgca aggagatcgt 960

ggacagctac ctgcccgtga tcctggacat catcaagggc gagatgagcc gccccggcga 1020

ggtgtgcagc gccctgaacc tgtgcgagag cctgcagaag cacctggccg agctgaacca 1080

ccagaagcag ctggagagca acaagatccc cgagctggac atgaccgagg tggtggcccc 1140
cttcatggcc aacatccccc tgctgctgta cccccaggac ggcccccgca gcaagcccca 1200
gcccaaggac aacggcgacg tgtgccagga ctgcatccag atggtgaccg acatccagac 1260
cgccgtgcgc accaacagca ccttcgtgca ggccctggtg gagcacgtga aggaggagtg 1320
cgaccgcctg ggccccggca tggccgacat ctgcaagaac tacatcagcc agtacagcga 1380
gatcgccatc cagatgatga tgcacatgca gcccaaggag atctgcgccc tggtgggctt 1440
ctgcgacgag gtgaaggaga tgcccatgca gaccctggtg cccgccaagg tggccagcaa 1500

gaacgtgatc cccgccctgg agctggtgga gcccatcaag aagcacgagg tgcccgccaa 1560
gagcgacgtg tactgcgagg tgtgcgagtt cctggtgaag gaggtgacca agctgatcga 1620
caacaacaag accgagaagg agatcctgga cgccttcgac aagatgtgca gcaagctgcc 1680
caagagcctg agcgaggagt gccaggaggt ggtggacacc tacggcagca gcatcctgag 1740
catcctgctg gaggaggtga gccccgagct ggtgtgcagc atgctgcacc tgtgcagcgg 1800
cacccgcctg cccgccctga ccgtgcacgt gacccagccc aaggacggcg gcttctgcga 1860
ggtgtgcaag aagctggtgg gctacctgga ccgcaacctg gagaagaaca gcaccaagca 1920

ggagatcctg gccgccctgg agaagggctg cagcttcctg cccgacccct accagaagca 1980
gtgcgaccag ttcgtggccg agtacgagcc cgtgctgatc gagatcctgg tggaggtgat 2040
ggaccccagc ttcgtgtgcc tgaagatcgg cgcctgcccc agcgcccaca agcccctgct 2100
gggcaccgag aagtgcatct ggggccccag ctactggtgc cagaacaccg agaccgccgc 2160
ccagtgcaac gccgtggagc actgcaagcg ccacgtgtgg aactgattgt ggccgaaccg 2220
ccgaactcag aggccggccc cagaaaaccc gagcgagtag ggggcggcgc gcaggaggga 2280
ggagaactgg gggcgcggga ggctggtggg tgtgggggt ggagatgtag aagatgtgac 2340

gccgcggccc ggcgggtgcc agattagcgg acgcggtgcc cgcggttgca acgggatccc 2400
gggcgctgca gcttgggagg cggctctccc caggcggcgt ccgcggagac acccatccgt 2460
gaaccccagg tcccgggccg ccggctcgcc gcgcaccagg ggccggcgga cagaagagcg 2520
gccgagcggc tcgaggctgg gggaccgcgg gcgcggccgc gcgctgccgg gcgggaggct 2580
gggggccgg ggccggggcc gtgccccgga gcgggtcgga ggccggggcc ggggccgggg 2640
gacggcggct ccccgcgcgg ctccagcggc tcggggatcc cggccgggcc ccgcagggac 2700
catgatggaa ttcagcagcc ccagcagaga ggaatgcccc aagcctctga gccgggtgtc 2760

aatcatggcc ggatctctga caggactgct gctgcttcag gccgtgtctt gggcttctgg 2820
cgctagacct tgcatcccca agagcttcgg ctacagcagc gtcgtgtgcg tgtgcaatgc 2880
cacctactgc gacagcttcg accctcctac ctttcctgct ctgggcacct tcagcagata 2940
cgagagcacc agatccggca gacggatgga actgagcatg ggacccatcc aggccaatca 3000
cacaggcact ggcctgctgc tgacactgca gcctgagcag aaattccaga aagtgaaagg 3060
cttcggcgga gccatgacag atgccgccgc tctgaatatc ctggctctgt ctccaccagc 3120
tcagaacctg ctgctcaaga gctacttcag cgaggaaggc atcggctaca acatcatcag 3180

agtgcccatg gccagctgcg acttcagcat caggacctac acctacgccg acacacccga 3240
cgatttccag ctgcacaact tcagcctgcc tgaagaggac accaagctga agatccctct 3300
gatccacaga gccctgcagc tggcacaaag acccgtgtca ctgctggcct ctccatggac 3360
atctcccacc tggctgaaaa caaatggcgc cgtgaatggc aagggcagcc tgaaaggcca 3420
acctggcgac atctaccacc agacctgggc cagatacttc gtgaagttcc tggacgccta 3480
tgccgagcac aagctgcagt tttgggccgt gacagccgag aacgaacctt ctgctggact 3540
gctgagcggc tacccctttc agtgcctggg ctttacaccc gagcaccagc gggactttat 3600

cgcccgtgat ctgggaccca cactggccaa tagcacccac cataatgtgc ggctgctgat 3660
gctggacgac cagagactgc ttctgcccca ctgggctaaa gtggtgctga cagatcctga 3720
ggccgccaaa tacgtgcacg gaatcgccgt gcactggtat ctggactttc tggcccctgc 3780
caaggccaca ctgggagaga cacacagact gttccccaac accatgctgt tcgccagcga 3840
agcctgtgtg ggcagcaagt tttgggaaca gagcgtgcgg ctcggcagct gggatagagg 3900
catgcagtac agccacagca tcatcaccaa cctgctgtac cacgtcgtcg gctggaccga 3960
ctggaatctg gccctgaatc ctgaaggcgg ccctaactgg gtccgaaact tcgtggacag 4020

ccccatcatc gtggacatca ccaaggacac cttctacaag cagcccatgt tctaccacct 4080
gggacacttc agcaagttca tccccgaggg ctctcagcgc gttggactgg tggcttccca 4140
gaagaacgat ctggacgccg tggctctgat gcaccctgat ggatctgctg tggtggtggt 4200
cctgaaccgc agcagcaaag atgtgcccct gaccatcaag gatcccgccg tgggattcct 4260
ggaaacaatc agccctggct actccatcca cacctacctg tggcgtagac agtgacaatt 4320
gttaattaag tttaaaccct cgaggccgca agcaataaaa tatctttatt ttcattacat 4380
ctgtgtgttg gtttttgtg tggagatcca cgataacaaa cagctttttt ggggtgaaca 4440

tattgactga attccctgca ggttggccac tccctctctg cgcgctcgct cgctcactga 4500
ggccgcccgg gcaaagcccg ggcgtcgggc gacctttggt cgcccggcct cagtgagcga 4560
gcgagcgcgc agagagggag tggccaactc catcactagg ggttcctgcg gccgctcgta 4620

cggtctcgag gaattcctgc aggataactt gccaacctca ttctaaaatg tatatagaag 4680
cccaaaagac aataacaaaa atattcttgt agaacaaaat gggaaagaat gttccactaa 4740
atatcaagat ttagagcaaa gcatgagatg tgtggggata gacagtgagg ctgataaaat 4800
agagtagagc tcagaaacag acccattgat atatgtaagt gacctatgaa aaaaatatgg 4860

cattttacaa tgggaaaatg atggtctttt tctttttag aaaaacaggg aaatatattt 4920
atatgtaaaa aataaaaggg aacccatatg tcataccata cacacaaaaa aattccagtg 4980
aattataagt ctaaatggag aaggcaaaac tttaaatctt ttagaaaata atatagaagc 5040
atgcagacca gcctggccaa catgatgaaa ccctctctac taataataaa atcagtagaa 5100
ctactcagga ctactttgag tgggaagtcc ttttctatga agacttcttt ggccaaaatt 5160
aggctctaaa tgcaaggaga tagtgcatca tgcctggctg cacttactga taaatgatgt 5220
tatcaccatc tttaaccaaa tgcacaggaa caagttatgg tactgatgtg ctggattgag 5280

aaggagctct acttccttga caggacacat ttgtatcaac ttaaaaaagc agatttttgc 5340
cagcagaact attcattcag aggtaggaaa cttagaatag atgatgtcac tgattagcat 5400
ggcttcccca tctccacagc tgcttcccac ccaggttgcc cacagttgag tttgtccagt 5460
gctcagggct gcccactctc agtaagaagc cccacaccag cccctctcca aatatgttgg 5520
ctgttccttc cattaaagtg accccacttt agagcagcaa gtggatttct gtttcttaca 5580
gttcaggaag gaggagtcag ctgtgagaac ctggagcctg agatgcttct aagtcccact 5640
gctactgggg tcagggaagc cagactccag catcagcagt caggagcact aagcccttgc 5700

caacatcctg tttctcagag aaactgcttc cattataatg gttgtccttt tttaagctat 5760
caagccaaac aaccagtgtc taccattatt ctcatcacct gaagccaagg gttctagcaa 5820
aagtcaagct gtcttgtaat ggttgatgtg cctccagctt ctgtcttcag tcactccact 5880
cttagcctgc tctgaatcaa ctctgaccac agttccctgg agcccctgcc acctgctgcc 5940
cctgccacct tctccatctg cagtgctgtg cagccttctg cactcttgca gagctaatag 6000
gtggagactt gaaggaagag gaggaaagtt tctcataata gccttgctgc aagctcaaat 6060
gggaggtggg cactgtgccc aggagccttg gagcaaaggc tgtgcccaac ctctgactgc 6120

atccaggttt ggtcttgaca gagataagaa gccctggctt ttggagccaa aatctaggtc 6180
agacttaggc aggattctca aagtttatca gcagaacatg aggcagaaga ccctttctgc 6240
tccagcttct tcaggctcaa ccttcatcag aatagataga aagagaggct gtgagggttc 6300
ttaaaacaga agcaaatctg actcagagaa taaacaacct cctagtaaac tacagcttag 6360
acagagcatc tggtggtgag tgtgctcagt gtcctactca actgtctggt atcagccctc 6420
atgaggactt ctcttctttc cctcatagac ctccatctct gttttcctta gcctgcagaa 6480

atctggatgg ctattcacag aatgcctgtg ctttcagagt tgcatttttt ctctggtatt 6540

ctggttcaag catttgaagg taggaaaggt tctccaagtg caagaaagcc agccctgagc 6600
ctcaactgcc tggctagtgt ggtcagtagg atgcaaaggc tgttgaatgc cacaaggcca 6660
aactttaacc tgtgtaccac aagcctagca gcagaggcag ctctgctcac tggaactctc 6720
tgtcttcttt ctcctgagcc ttttcttttc ctgagttttc tagctctcct caaccttacc 6780
tctgccctac ccaggacaaa cccaagagcc actgtttctg tgatgtcctc tccagcccta 6840
attaggcatc atgacttcag cctgaccttc catgctcaga agcagtgcta atccacttca 6900
gatgagctgc tctatgcaac acaggcagag cctacaaacc tttgcaccag agccctccac 6960

atatcagtgt ttgttcatac tcacttcaac agcaaatgtg actgctgaga ttaagatttt 7020
acacaagatg gtctgtaatt tcacagttag ttttatccca ttaggtatga aagaattagc 7080
ataattcccc ttaaacatga atgaatctta gatttttaa taaatagttt tggaagtaaa 7140
gacagagaca tcaggagcac aaggaatagc ctgagaggac aaacagaaca agaaagagtc 7200
tggaaataca caggatgttc ttggcctcct caaagcaagt gcaagcagat agtaccagca 7260
gccccaggct atcagagccc agtgaagaga agtaccatga aagccacagc tctaaccacc 7320
ctgttccaga gtgacagaca gtccccaaga caagccagcc tgagccagag agagaactgc 7380

aagagaaagt ttctaattta ggttctgtta gattcagaca agtgcaggtc atcctctctc 7440
cacagctact cacctctcca gcctaacaaa gcctgcagtc cacactccaa ccctggtgtc 7500
tcacctccta gcctctccca acatcctgct ctctgaccat cttctgcatc tctcatctca 7560
ccatctccca ctgtctacag cctactcttg caactaccat ctcattttct gacatcctgt 7620
ctacatcttc tgccatactc tgccatctac cataccacct cttaccatct accacaccat 7680
cttttatctc catccctctc agaagcctcc aagctgaatc ctgctttatg tgttcatctc 7740
agcccctgca tggaaagctg accccagagg cagaactatt cccagagagc ttggccaaga 7800

aaaacaaaac taccagcctg gccaggctca ggagtagtaa gctgcagtgt ctgttgtgtt 7860
ctagcttcaa cagctgcagg agttccactc tcaaatgctc cacatttctc acatcctcct 7920
gattctggtc actacccatc ttcaaagaac agaatatctc acatcagcat actgtgaagg 7980
actagtcatg ggtgcagctg ctcagagctg caaagtcatt ctggatggtg gagagcttac 8040
aaacatttca tgatgctccc cccgctctga tggctggagc ccaatcccta cacagactcc 8100
tgctgtatgt gttttccttt cactctgagc cacagccaga gggcaggcat tcagtctcct 8160
cttcaggctg gggctggggc actgagaact cacccaacac cttgctctca ctccttctgc 8220

aaaacaagaa agagctttgt gctgcagtag ccatgaagaa tgaaaggaag gctttaacta 8280
aaaaatgtca gagattattt tcaacccctt actgtggatc accagcaagg aggaaacaca 8340
acacagagac atttttttccc ctcaaattat caaaagaatc actgcatttg ttaaagagag 8400
caactgaatc aggaagcaga gttttgaaca tatcagaagt taggaatctg catcagagac 8460
aaatgcagtc atggttgttt gctgcatacc agccctaatc attagaagcc tcatggactt 8520
caaacatcat tccctctgac aagatgctct agcctaactc catgagataa aataaatctg 8580
cctttcagag ccaaagaaga gtccaccagc ttcttctcag tgtgaacaag agctccagtc 8640

aggttagtca gtccagtgca gtagaggaga ccagtctgca tcctctaatt ttcaaaggca 8700
agaagatttg tttaccctgg acaccaggca caagtgaggt cacagagctc ttagatatgc 8760
agtcctcatg agtgaggaga ctaaagcgca tgccatcaag acttcagtgt agagaaaacc 8820
tccaaaaaag cctcctcact acttctggaa tagctcagag gccgaggcgg cctcggcctc 8880
tgcataaata aaaaaaatta gtcagccatg gggcggagaa tgggcggaac tgggcggagt 8940
taggggcggg atgggcggag ttaggggcgg gactatggtt gctgactaat tgagatgcat 9000
gctttgcata cttctgcctg ctggggagcc tggggacttt ccacacctgg ttgctgacta 9060

attgagatgc atgctttgca tacttctgcc tgctggggag cctggggact ttccacaccc 9120
taactgacac acattccaca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt 9180
ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg 9240
ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg 9300
gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag 9360
gccgcgttgc tggcgttttt ccataggctc cgcccccctg acgagcatca caaaaatcga 9420
cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct 9480

ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc 9540
tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg 9600
gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc 9660
tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca 9720
ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag 9780
ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg tatctgcgct 9840
ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaaacc 9900

accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga 9960
tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa cgaaaactca 10020
cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat ccttttaaat 10080

taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac 10140

caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt 10200

gcctgactcc tgcaaaccac gttgtgtctc aaaatctctg atgttacatt gcacaagata 10260

aaaatatatc atcatgaaca ataaaactgt ctgcttacat aaacagtaat acaaggggtg 10320

ttatgagcca tattcaacgg gaaacgtctt gctcgaggcc gcgattaaat tccaacatgg 10380

atgctgattt atatgggtat aaatgggctc gcgataatgt cgggcaatca ggtgcgacaa 10440

tctatcgatt gtatgggaag cccgatgcgc cagagttgtt tctgaaacat ggcaaaggta 10500

gcgttgccaa tgatgttaca gatgagatgg tcagactaaa ctggctgacg gaatttatgc 10560

ctcttccgac catcaagcat tttatccgta ctcctgatga tgcatggtta ctcaccactg 10620

cgatccccgg gaaaacagca ttccaggtat tagaagaata tcctgattca ggtgaaaata 10680

ttgttgatgc gctggcagtg ttcctgcgcc ggttgcattc gattcctgtt tgtaattgtc 10740

cttttaacag cgatcgcgta tttcgtctcg ctcaggcgca atcacgaatg aataacggtt 10800

tggttgatgc gagtgatttt gatgacgagc gtaatggctg gcctgttgaa caagtctgga 10860

aagaaatgca taagcttttg ccattctcac cggattcagt cgtcactcat ggtgatttct 10920

cacttgataa ccttattttt gacgagggga aattaatagg ttgtattgat gttggacgag 10980

tcggaatcgc agaccgatac caggatcttg ccatcctatg gaactgcctc ggtgagtttt 11040

ctccttcatt acagaaacgg ctttttcaaa aatatggtat tgataatcct gatatgaata 11100

aattgcagtt tcatttgatg ctcgatgagt ttttctaagg gcggcctgcc accataccca 11160

cgccgaaaca agcgctcatg agcccgaagt ggcgagcccg atcttcccca tcggtgatgt 11220

cggcgatata ggcgccagca accgcacctg tggcgccggt gatgagggcg cgccaagtcg 11280

acgtccggca gtc 11293

<210> 7

<211> 10700

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 7

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgcccgggc aaagcccggg 60

cgtcgggcga cctttggtcg cccggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300

ctagttatta atagtaatca attacggggt cattagttca tagcccatat atggagttcc 360

gcgttacata acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat 420

tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc 480

aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc 540

caagtacgcc ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt 600

acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta 660

ccatggtcga ggtgagcccc acgttctgct tcactctccc catctccccc ccctccccac 720

ccccaatttt gtatttattt atttttaat tattttgtgc agcgatgggg gcgggggggg 780

gggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga 840

gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa gtttcctttt atggcgaggc 900

ggcggcggcg gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgacgc 960

tgccttcgcc ccgtgccccg ctccgccgcc gcctcgcgcc gcccgccccg gctctgactg 1020

accgcgttac tcccacaggt gagcgggcgg gacggccctt ctcctccggg ctgtaattag 1080

cgcttggttt aatgacggct tgtttctttt ctgtggctgc gtgaaagcct tgaggggctc 1140

cgggagctag agcctctgct aaccatgttc atgccttctt ctttttccta cagctcctgg 1200

gcaacgtgct ggttattgtg ctgtctcatc attttggcaa agaattcctc gaagatccga 1260

agggaaagtc ttccacgact gtgggatccg ttcgaagata tcaccggttg agccaccatg 1320

gaattcagca gccccagcag agaggaatgc cccaagcctc tgagccgggt gtcaatcatg 1380

gccggatctc tgacaggact gctgctgctt caggccgtgt cttgggcttc tggcgctaga 1440

ccttgcatcc ccaagagctt cggctacagc agcgtcgtgt gcgtgtgcaa tgccacctac 1500

tgcgacagct tcgaccctcc tacctttcct gctctgggca ccttcagcag atacgagagc 1560

accagatccg gcagacggat ggaactgagc atgggaccca tccaggccaa tcacacaggc 1620

actggcctgc tgctgacact gcagcctgag cagaaattcc agaaagtgaa aggcttcggc 1680

ggagccatga cagatgccgc cgctctgaat atcctggctc tgtctccacc agctcagaac 1740

ctgctgctca agagctactt cagcgaggaa ggcatcggct acaacatcat cagagtgccc 1800

atggccagct gcgacttcag catcaggacc tacacctacg ccgacacacc cgacgatttc 1860

cagctgcaca acttcagcct gcctgaagag gacaccaagc tgaagatccc tctgatccac 1920

agagccctgc agctggcaca aagacccgtg tcactgctgg cctctccatg gacatctccc 1980

acctggctga aaacaaatgg cgccgtgaat ggcaagggca gcctgaaagg ccaacctggc 2040

gacatctacc accagacctg ggccagatac ttcgtgaagt tcctggacgc ctatgccgag 2100

cacaagctgc agttttgggc cgtgacagcc gagaacgaac cttctgctgg actgctgagc 2160
ggctacccct ttcagtgcct gggctttaca cccgagcacc agcgggactt tatcgcccgt 2220
gatctgggac ccacactggc caatagcacc caccataatg tgcggctgct gatgctggac 2280

gaccagagac tgcttctgcc ccactgggct aaagtggtgc tgacagatcc tgaggccgcc 2340
aaatacgtgc acggaatcgc cgtgcactgg tatctggact ttctggcccc tgccaaggcc 2400
acactgggag agacacacag actgttcccc aacaccatgc tgttcgccag cgaagcctgt 2460
gtgggcagca agttttggga acagagcgtg cggctcggca gctgggatag aggcatgcag 2520
tacagccaca gcatcatcac caacctgctg taccacgtcg tcggctggac cgactggaat 2580
ctggccctga atcctgaagg cggccctaac tgggtccgaa acttcgtgga cagccccatc 2640
atcgtggaca tcaccaagga caccttctac aagcagccca tgttctacca cctgggacac 2700

ttcagcaagt tcatccccga gggctctcag cgcgttggac tggtggcttc ccagaagaac 2760
gatctggacg ccgtggctct gatgcaccct gatggatctg ctgtggtggt ggtcctgaac 2820
cgcagcagca aagatgtgcc cctgaccatc aaggatcccg ccgtgggatt cctggaaaca 2880
atcagccctg gctactccat ccacacctac ctgtggcgta gacagtgaca attgttaatt 2940
aagtttaaac cctcgaggcc gcaagcttat cgataatcaa cctctggatt acaaaatttg 3000
tgaaagattg actggtattc ttaactatgt tgctcctttt acgctatgtg gatacgctgc 3060
tttaatgcct ttgtatcatg ctattgcttc ccgtatggct ttcattttct cctccttgta 3120

taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtcaggc aacgtggcgt 3180
ggtgtgcact gtgtttgctg acgcaacccc cactggttgg ggcattgcca ccacctgtca 3240
gctcctttcc gggactttcg ctttccccct ccctattgcc acggcggaac tcatcgccgc 3300
ctgccttgcc cgctgctgga caggggctcg gctgttgggc actgacaatt ccgtggtgtt 3360
gtcggggaaa tcatcgtcct ttccttggct gctcgcctgt gttgccacct ggattctgcg 3420
cgggacgtcc ttctgctacg tcccttcggc cctcaatcca gcggaccttc cttcccgcgg 3480
cctgctgccg gctctgcggc ctcttccgcg tcttcgcctt cgccctcaga cgagtcggat 3540

ctccctttgg gccgcctccc cgcatcgata ccgtcgacta gagctcgctg atcagcctcg 3600
actgtgcctt ctagttgcca gccatctgtt gtttgcccct cccccgtgcc ttccttgacc 3660
ctggaaggtg ccactcccac tgtcctttcc taataaaatg aggaaattgc atcgcattgt 3720
ctgagtaggt gtcattctat tctggggggt ggggtggggc aggacagcaa gggggaggat 3780
tgggaagaca atagcaggca tgctggggag agatccacga taacaaacag ctttttttggg 3840
gtgaacatat tgactgaatt ccctgcaggt tggccactcc ctctctgcgc gctcgctcgc 3900

tcactgaggc cgcccgggca aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag 3960

tgagcgagcg agcgcgcaga gagggagtgg ccaactccat cactaggggt tcctgcggcc 4020
gctcgtacgg tctcgaggaa ttcctgcagg ataacttgcc aacctcattc taaaatgtat 4080
atagaagccc aaaagacaat aacaaaaata ttcttgtaga acaaaatggg aaagaatgtt 4140
ccactaaata tcaagattta gagcaaagca tgagatgtgt ggggatagac agtgaggctg 4200
ataaaataga gtagagctca gaaacagacc cattgatata tgtaagtgac ctatgaaaaa 4260
aatatggcat tttacaatgg gaaaatgatg gtctttttct tttttagaaa aacagggaaa 4320
tatatttata tgtaaaaaat aaaagggaac ccatatgtca taccatacac acaaaaaaat 4380

tccagtgaat tataagtcta aatggagaag gcaaaacttt aaatctttta gaaaataata 4440
tagaagcatg cagaccagcc tggccaacat gatgaaaccc tctctactaa taataaaatc 4500
agtagaacta ctcaggacta ctttgagtgg gaagtccttt tctatgaaga cttctttggc 4560
caaaattagg ctctaaatgc aaggagatag tgcatcatgc ctggctgcac ttactgataa 4620
atgatgttat caccatcttt aaccaaatgc acaggaacaa gttatggtac tgatgtgctg 4680
gattgagaag gagctctact tccttgacag gacacatttg tatcaactta aaaaagcaga 4740
tttttgccag cagaactatt cattcagagg taggaaactt agaatagatg atgtcactga 4800

ttagcatggc ttccccatct ccacagctgc ttcccaccca ggttgcccac agttgagttt 4860
gtccagtgct cagggctgcc cactctcagt aagaagcccc acaccagccc ctctccaaat 4920
atgttggctg ttccttccat taaagtgacc ccactttaga gcagcaagtg gatttctgtt 4980
tcttacagtt caggaaggag gagtcagctg tgagaacctg gagcctgaga tgcttctaag 5040
tcccactgct actggggtca gggaagccag actccagcat cagcagtcag gagcactaag 5100
cccttgccaa catcctgttt ctcagagaaa ctgcttccat tataatggtt gtccttttt 5160
aagctatcaa gccaaacaac cagtgtctac cattattctc atcacctgaa gccaagggtt 5220

ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct ccagcttctg tcttcagtca 5280
ctccactctt agcctgctct gaatcaactc tgaccacagt tccctggagc ccctgccacc 5340
tgctgcccct gccaccttct ccatctgcag tgctgtgcag ccttctgcac tcttgcagag 5400
ctaataggtg gagacttgaa ggaagaggag gaaagtttct cataatagcc ttgctgcaag 5460
ctcaaatggg aggtgggcac tgtgcccagg agccttggag caaaggctgt gcccaacctc 5520
tgactgcatc caggtttggt cttgacagag ataagaagcc ctggcttttg gagccaaaat 5580
ctaggtcaga cttaggcagg attctcaaag tttatcagca gaacatgagg cagaagaccc 5640

tttctgctcc agcttcttca ggctcaacct tcatcagaat agatagaaag agaggctgtg 5700
agggttctta aaacagaagc aaatctgact cagagaataa acaacctcct agtaaactac 5760
agcttagaca gagcatctgg tggtgagtgt gctcagtgtc ctactcaact gtctggtatc 5820
agccctcatg aggacttctc ttctttccct catagacctc catctctgtt ttccttagcc 5880
tgcagaaatc tggatggcta ttcacagaat gcctgtgctt tcagagttgc atttttctc 5940
tggtattctg gttcaagcat ttgaaggtag gaaaggttct ccaagtgcaa gaaagccagc 6000
cctgagcctc aactgcctgg ctagtgtggt cagtaggatg caaaggctgt tgaatgccac 6060

aaggccaaac tttaacctgt gtaccacaag cctagcagca gaggcagctc tgctcactgg 6120
aactctctgt cttctttctc ctgagccttt tcttttcctg agttttctag ctctcctcaa 6180
ccttacctct gccctaccca ggacaaaccc aagagccact gtttctgtga tgtcctctcc 6240
agccctaatt aggcatcatg acttcagcct gaccttccat gctcagaagc agtgctaatc 6300
cacttcagat gagctgctct atgcaacaca ggcagagcct acaaaccttt gcaccagagc 6360
cctccacata tcagtgtttg ttcatactca cttcaacagc aaatgtgact gctgagatta 6420
agattttaca caagatggtc tgtaatttca cagttagttt tatcccatta ggtatgaaag 6480

aattagcata attcccctta aacatgaatg aatcttagat tttttaataa atagttttgg 6540
aagtaaagac agagacatca ggagcacaag gaatagcctg agaggacaaa cagaacaaga 6600
aagagtctgg aaatacacag gatgttcttg gcctcctcaa agcaagtgca agcagatagt 6660
accagcagcc ccaggctatc agagcccagt gaagagaagt accatgaaag ccacagctct 6720
aaccaccctg ttccagagtg acagacagtc cccaagacaa gccagcctga gccagagaga 6780
gaactgcaag agaaagtttc taatttaggt tctgttagat tcagacaagt gcaggtcatc 6840
ctctctccac agctactcac ctctccagcc taacaaagcc tgcagtccac actccaaccc 6900

tggtgtctca cctcctagcc tctcccaaca tcctgctctc tgaccatctt ctgcatctct 6960
catctcacca tctcccactg tctacagcct actcttgcaa ctaccatctc attttctgac 7020
atcctgtcta catcttctgc catactctgc catctaccat accacctctt accatctacc 7080
acaccatctt ttatctccat ccctctcaga agcctccaag ctgaatcctg ctttatgtgt 7140
tcatctcagc ccctgcatgg aaagctgacc ccagaggcag aactattccc agagagcttg 7200
gccaagaaaa acaaaactac cagcctggcc aggctcagga gtagtaagct gcagtgtctg 7260
ttgtgttcta gcttcaacag ctgcaggagt tccactctca aatgctccac atttctcaca 7320

tcctcctgat tctggtcact acccatcttc aaagaacaga atatctcaca tcagcatact 7380
gtgaaggact agtcatgggt gcagctgctc agagctgcaa agtcattctg gatggtggag 7440
agcttacaaa catttcatga tgctcccccc gctctgatgg ctggagccca atccctacac 7500

agactcctgc tgtatgtgtt ttcctttcac tctgagccac agccagaggg caggcattca 7560
gtctcctctt caggctgggg ctggggcact gagaactcac ccaacacctt gctctcactc 7620
cttctgcaaa acaagaaaga gctttgtgct gcagtagcca tgaagaatga aaggaaggct 7680
ttaactaaaa aatgtcagag attattttca acccccttact gtggatcacc agcaaggagg 7740

aaacacaaca cagagacatt ttttcccctc aaattatcaa aagaatcact gcatttgtta 7800
aagagagcaa ctgaatcagg aagcagagtt ttgaacatat cagaagttag gaatctgcat 7860
cagagacaaa tgcagtcatg gttgtttgct gcataccagc cctaatcatt agaagcctca 7920
tggacttcaa acatcattcc ctctgacaag atgctctagc ctaactccat gagataaaat 7980
aaatctgcct ttcagagcca aagaagagtc caccagcttc ttctcagtgt gaacaagagc 8040
tccagtcagg ttagtcagtc cagtgcagta gaggagacca gtctgcatcc tctaattttc 8100
aaaggcaaga agatttgttt accctggaca ccaggcacaa gtgaggtcac agagctctta 8160

gatatgcagt cctcatgagt gaggagacta aagcgcatgc catcaagact tcagtgtaga 8220
gaaaacctcc aaaaaagcct cctcactact tctggaatag ctcagaggcc gaggcggcct 8280
cggcctctgc ataaataaaa aaaattagtc agccatgggg cggagaatgg gcggaactgg 8340
gcggagttag gggcgggatg ggcggagtta ggggcgggac tatggttgct gactaattga 8400
gatgcatgct ttgcatactt ctgcctgctg gggagcctgg ggactttcca cacctggttg 8460
ctgactaatt gagatgcatg ctttgcatac ttctgcctgc tggggagcct ggggactttc 8520
cacaccctaa ctgacacaca ttccacagct gcattaatga atcggccaac gcgcggggag 8580

aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt 8640
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga 8700
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 8760
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc ccccctgacg agcatcacaa 8820
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 8880
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 8940
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 9000

cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 9060
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 9120
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 9180
tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat 9240
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 9300
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 9360

aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 9420

aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct 9480
tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga 9540
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc 9600
catagttgcc tgactcctgc aaaccacgtt gtgtctcaaa atctctgatg ttacattgca 9660
caagataaaa atatatcatc atgaacaata aaactgtctg cttacataaa cagtaataca 9720
aggggtgtta tgagccatat tcaacgggaa acgtcttgct cgaggccgcg attaaattcc 9780
aacatggatg ctgatttata tgggtataaa tgggctcgcg ataatgtcgg gcaatcaggt 9840

gcgacaatct atcgattgta tgggaagccc gatgcgccag agttgtttct gaaacatggc 9900
aaaggtagcg ttgccaatga tgttacagat gagatggtca gactaaactg gctgacggaa 9960
tttatgcctc ttccgaccat caagcatttt atccgtactc ctgatgatgc atggttactc 10020
accactgcga tccccgggaa aacagcattc caggtattag aagaatatcc tgattcaggt 10080
gaaaatattg ttgatgcgct ggcagtgttc ctgcgccggt tgcattcgat tcctgtttgt 10140
aattgtcctt ttaacagcga tcgcgtattt cgtctcgctc aggcgcaatc acgaatgaat 10200
aacggtttgg ttgatgcgag tgattttgat gacgagcgta atggctggcc tgttgaacaa 10260

gtctggaaag aaatgcataa gcttttgcca ttctcaccgg attcagtcgt cactcatggt 10320
gatttctcac ttgataacct tatttttgac gaggggaaat taataggttg tattgatgtt 10380
ggacgagtcg gaatcgcaga ccgataccag gatcttgcca tcctatggaa ctgcctcggt 10440
gagttttctc cttcattaca gaaacggctt tttcaaaaat atggtattga taatcctgat 10500
atgaataaat tgcagtttca tttgatgctc gatgagtttt tctaagggcg gcctgccacc 10560
atacccacgc cgaaacaagc gctcatgagc ccgaagtggc gagcccgatc ttccccatcg 10620
gtgatgtcgg cgatataggc gccagcaacc gcacctgtgg cgccggtgat gagggcgcgc 10680

caagtcgacg tccggcagtc 10700

<210> 8
<211> 10700
<212> DNA
<213> Artificial sequence
<220><223> Synthetic polynucleotide
<400> 8

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactatt agatctgatg gccgcgctag ctctgggtat ttaagcccga gtgagcacgc 180
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240
tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300

ctagttatta atagtaatca attacggggt cattagttca tagcccatat atggagttcc 360
gcgttacata acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat 420
tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc 480
aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc 540
caagtacgcc ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt 600
acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta 660
ccatggtcga ggtgagcccc acgttctgct tcactctccc catctccccc ccctccccac 720

ccccaatttt gtatttattt atttttaat tattttgtgc agcgatgggg gcgggggggg 780
gggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga 840
gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa gtttcctttt atggcgaggc 900
ggcggcggcg gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgacgc 960
tgccttcgcc ccgtgccccg ctccgccgcc gcctcgcgcc gcccgccccg gctctgactg 1020
accgcgttac tcccacaggt gagcgggcgg gacggccctt ctcctccggg ctgtaattag 1080
cgcttggttt aatgacggct tgtttctttt ctgtggctgc gtgaaagcct tgaggggctc 1140

cgggagctag agcctctgct aaccatgttc atgccttctt ctttttccta cagctcctgg 1200
gcaacgtgct ggttattgtg ctgtctcatc attttggcaa agaattcctc gaagatccga 1260
agggaaagtc ttccacgact gtgggatccg ttcgaagata tcaccggttg agccaccatg 1320
gaattcagca gccccagcag agaggaatgc cccaagcctc tgagccgggt gtcaatcatg 1380
gccggatctc tgacaggact gctgctgctt caggccgtgt cttgggcttc tggcgctaga 1440
ccttgcatcc ccaagagctt cggctacagc agcgtcgtgt gcgtgtgcaa tgccacctac 1500
tgcgacagct tcgaccctcc tacctttcct gctctgggca ccttcagcag atacgagagc 1560

accagatccg gcagacggat ggaactgagc atgggaccca tccaggccaa tcacacaggc 1620
actggcctgc tgctgacact gcagcctgag cagaaattcc agaaagtgaa aggcttcggc 1680
ggagccatga cagatgccgc cgctctgaat atcctggctc tgtctccacc agctcagaac 1740
ctgctgctca agagctactt cagcgaggaa ggcatcggct acaacatcat cagagtgccc 1800
atggccagct gcgacttcag catcaggacc tacacctacg ccgacacacc cgacgatttc 1860
cagctgcaca acttcagcct gcctgaagag gacaccaagc tgaagatccc tctgatccac 1920

agagccctgc agctggcaca aagacccgtg tcactgctgg cctctccatg gacatctccc 1980

acctggctga aaacaaatgg cgccgtgaat ggcaagggca gcctgaaagg ccaacctggc 2040
gacatctacc accagacctg ggccagatac ttcgtgaagt tcctggacgc ctatgccgag 2100
cacaagctgc agttttgggc cgtgacagcc gagaacgaac cttctgctgg actgctgagc 2160
ggctacccct ttcagtgcct gggctttaca cccgagcacc agcgggactt tatcgcccgt 2220
gatctgggac ccacactggc caatagcacc caccataatg tgcggctgct gatgctggac 2280
gaccagagac tgcttctgcc ccactgggct aaagtggtgc tgacagatcc tgaggccgcc 2340
aaatacgtgc acggaatcgc cgtgcactgg tatctggact ttctggcccc tgccaaggcc 2400

acactgggag agacacacag actgttcccc aacaccatgc tgttcgccag cgaagcctgt 2460
gtgggcagca agttttggga acagagcgtg cggctcggca gctgggatag aggcatgcag 2520
tacagccaca gcatcatcac caacctgctg taccacgtcg tcggctggac cgactggaat 2580
ctggccctga atcctgaagg cggccctaac tgggtccgaa acttcgtgga cagccccatc 2640
atcgtggaca tcaccaagga caccttctac aagcagccca tgttctacca cctgggacac 2700
ttcagcaagt tcatccccga gggctctcag cgcgttggac tggtggcttc ccagaagaac 2760
gatctggacg ccgtggctct gatgcaccct gatggatctg ctgtggtggt ggtcctgaac 2820

cgcagcagca aagatgtgcc cctgaccatc aaggatcccg ccgtgggatt cctggaaaca 2880
atcagccctg gctactccat ccacacctac ctgtggcgta gacagtgaca attgttaatt 2940
aagtttaaac cctcgaggcc gcaagcttat cgataatcaa cctctggatt acaaaatttg 3000
tgaaagattg actggtattc ttaactatgt tgctcctttt acgctatgtg gatacgctgc 3060
tttaatgcct ttgtatcatg ctattgcttc ccgtatggct ttcattttct cctccttgta 3120
taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtcaggc aacgtggcgt 3180
ggtgtgcact gtgtttgctg acgcaacccc cactggttgg ggcattgcca ccacctgtca 3240

gctcctttcc gggactttcg ctttccccct ccctattgcc acggcggaac tcatcgccgc 3300
ctgccttgcc cgctgctgga caggggctcg gctgttgggc actgacaatt ccgtggtgtt 3360
gtcggggaaa tcatcgtcct ttccttggct gctcgcctgt gttgccacct ggattctgcg 3420
cgggacgtcc ttctgctacg tcccttcggc cctcaatcca gcggaccttc cttcccgcgg 3480
cctgctgccg gctctgcggc ctcttccgcg tcttcgcctt cgccctcaga cgagtcggat 3540
ctccctttgg gccgcctccc cgcatcgata ccgtcgacta gagctcgctg atcagcctcg 3600
actgtgcctt ctagttgcca gccatctgtt gtttgcccct cccccgtgcc ttccttgacc 3660

ctggaaggtg ccactcccac tgtcctttcc taataaaatg aggaaattgc atcgcattgt 3720
ctgagtaggt gtcattctat tctgggggt ggggtggggc aggacagcaa gggggaggat 3780
tgggaagaca atagcaggca tgctggggag agatccacga taacaaacag ctttttggg 3840
gtgaacatat tgactgaatt ccctgcaggt tggccactcc ctctctgcgc gctcgctcgc 3900
tcactgaggc cgcccgggca aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag 3960
tgagcgagcg agcgcgcaga gagggagtgg ccaactccat cactaggggt tcctgcggcc 4020
gctcgtacgg tctcgaggaa ttcctgcagg ataacttgcc aacctcattc taaaatgtat 4080

atagaagccc aaaagacaat aacaaaaata ttcttgtaga acaaaatggg aaagaatgtt 4140
ccactaaata tcaagattta gagcaaagca tgagatgtgt ggggatagac agtgaggctg 4200
ataaaataga gtagagctca gaaacagacc cattgatata tgtaagtgac ctatgaaaaa 4260
aatatggcat tttacaatgg gaaaatgatg gtctttttct tttttagaaa aacagggaaa 4320
tatatttata tgtaaaaaat aaaagggaac ccatatgtca taccatacac acaaaaaaat 4380
tccagtgaat tataagtcta aatggagaag gcaaaacttt aaatctttta gaaaataata 4440
tagaagcatg cagaccagcc tggccaacat gatgaaaccc tctctactaa taataaaatc 4500

agtagaacta ctcaggacta ctttgagtgg gaagtccttt tctatgaaga cttctttggc 4560
caaaattagg ctctaaatgc aaggagatag tgcatcatgc ctggctgcac ttactgataa 4620
atgatgttat caccatcttt aaccaaatgc acaggaacaa gttatggtac tgatgtgctg 4680
gattgagaag gagctctact tccttgacag gacacatttg tatcaactta aaaaagcaga 4740
tttttgccag cagaactatt cattcagagg taggaaactt agaatagatg atgtcactga 4800
ttagcatggc ttccccatct ccacagctgc ttcccaccca ggttgcccac agttgagttt 4860
gtccagtgct cagggctgcc cactctcagt aagaagcccc acaccagccc ctctccaaat 4920

atgttggctg ttccttccat taaagtgacc ccactttaga gcagcaagtg gatttctgtt 4980
tcttacagtt caggaaggag gagtcagctg tgagaacctg gagcctgaga tgcttctaag 5040
tcccactgct actggggtca gggaagccag actccagcat cagcagtcag gagcactaag 5100
cccttgccaa catcctgttt ctcagagaaa ctgcttccat tataatggtt gtccttttt 5160
aagctatcaa gccaaacaac cagtgtctac cattattctc atcacctgaa gccaagggtt 5220
ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct ccagcttctg tcttcagtca 5280
ctccactctt agcctgctct gaatcaactc tgaccacagt tccctggagc ccctgccacc 5340

tgctgcccct gccaccttct ccatctgcag tgctgtgcag ccttctgcac tcttgcagag 5400
ctaataggtg gagacttgaa ggaagaggag gaaagtttct cataatagcc ttgctgcaag 5460
ctcaaatggg aggtgggcac tgtgcccagg agccttggag caaaggctgt gcccaacctc 5520

tgactgcatc caggtttggt cttgacagag ataagaagcc ctggcttttg gagccaaaat 5580
ctaggtcaga cttaggcagg attctcaaag tttatcagca gaacatgagg cagaagaccc 5640
tttctgctcc agcttcttca ggctcaacct tcatcagaat agatagaaag agaggctgtg 5700
agggttctta aaacagaagc aaatctgact cagagaataa acaacctcct agtaaactac 5760

agcttagaca gagcatctgg tggtgagtgt gctcagtgtc ctactcaact gtctggtatc 5820
agccctcatg aggacttctc ttctttccct catagacctc catctctgtt ttccttagcc 5880
tgcagaaatc tggatggcta ttcacagaat gcctgtgctt tcagagttgc atttttctc 5940
tggtattctg gttcaagcat ttgaaggtag gaaaggttct ccaagtgcaa gaaagccagc 6000
cctgagcctc aactgcctgg ctagtgtggt cagtaggatg caaaggctgt tgaatgccac 6060
aaggccaaac tttaacctgt gtaccacaag cctagcagca gaggcagctc tgctcactgg 6120
aactctctgt cttctttctc ctgagccttt tcttttcctg agttttctag ctctcctcaa 6180

ccttacctct gccctaccca ggacaaaccc aagagccact gtttctgtga tgtcctctcc 6240
agccctaatt aggcatcatg acttcagcct gaccttccat gctcagaagc agtgctaatc 6300
cacttcagat gagctgctct atgcaacaca ggcagagcct acaaaccttt gcaccagagc 6360
cctccacata tcagtgtttg ttcatactca cttcaacagc aaatgtgact gctgagatta 6420
agattttaca caagatggtc tgtaatttca cagttagttt tatcccatta ggtatgaaag 6480
aattagcata attcccctta aacatgaatg aatcttagat tttttaataa atagttttgg 6540
aagtaaagac agagacatca ggagcacaag gaatagcctg agaggacaaa cagaacaaga 6600

aagagtctgg aaatacacag gatgttcttg gcctcctcaa agcaagtgca agcagatagt 6660
accagcagcc ccaggctatc agagcccagt gaagagaagt accatgaaag ccacagctct 6720
aaccaccctg ttccagagtg acagacagtc cccaagacaa gccagcctga gccagagaga 6780
gaactgcaag agaaagtttc taatttaggt tctgttagat tcagacaagt gcaggtcatc 6840
ctctctccac agctactcac ctctccagcc taacaaagcc tgcagtccac actccaaccc 6900
tggtgtctca cctcctagcc tctcccaaca tcctgctctc tgaccatctt ctgcatctct 6960
catctcacca tctcccactg tctacagcct actcttgcaa ctaccatctc attttctgac 7020

atcctgtcta catcttctgc catactctgc catctaccat accacctctt accatctacc 7080
acaccatctt ttatctccat ccctctcaga agcctccaag ctgaatcctg ctttatgtgt 7140
tcatctcagc ccctgcatgg aaagctgacc ccagaggcag aactattccc agagagcttg 7200
gccaagaaaa acaaaactac cagcctggcc aggctcagga gtagtaagct gcagtgtctg 7260
ttgtgttcta gcttcaacag ctgcaggagt tccactctca aatgctccac atttctcaca 7320
tcctcctgat tctggtcact acccatcttc aaagaacaga atatctcaca tcagcatact 7380

gtgaaggact agtcatgggt gcagctgctc agagctgcaa agtcattctg gatggtggag 7440

agcttacaaa catttcatga tgctcccccc gctctgatgg ctggagccca atccctacac 7500
agactcctgc tgtatgtgtt ttcctttcac tctgagccac agccagaggg caggcattca 7560
gtctcctctt caggctgggg ctggggcact gagaactcac ccaacacctt gctctcactc 7620
cttctgcaaa acaagaaaga gctttgtgct gcagtagcca tgaagaatga aaggaaggct 7680
ttaactaaaa aatgtcagag attattttca acccccttact gtggatcacc agcaaggagg 7740
aaacacaaca cagagacatt ttttcccctc aaattatcaa aagaatcact gcatttgtta 7800
aagagagcaa ctgaatcagg aagcagagtt ttgaacatat cagaagttag gaatctgcat 7860

cagagacaaa tgcagtcatg gttgtttgct gcataccagc cctaatcatt agaagcctca 7920
tggacttcaa acatcattcc ctctgacaag atgctctagc ctaactccat gagataaaat 7980
aaatctgcct ttcagagcca aagaagagtc caccagcttc ttctcagtgt gaacaagagc 8040
tccagtcagg ttagtcagtc cagtgcagta gaggagacca gtctgcatcc tctaattttc 8100
aaaggcaaga agatttgttt accctggaca ccaggcacaa gtgaggtcac agagctctta 8160
gatatgcagt cctcatgagt gaggagacta aagcgcatgc catcaagact tcagtgtaga 8220
gaaaacctcc aaaaaagcct cctcactact tctggaatag ctcagaggcc gaggcggcct 8280

cggcctctgc ataaataaaa aaaattagtc agccatgggg cggagaatgg gcggaactgg 8340
gcggagttag gggcgggatg ggcggagtta ggggcgggac tatggttgct gactaattga 8400
gatgcatgct ttgcatactt ctgcctgctg gggagcctgg ggactttcca cacctggttg 8460
ctgactaatt gagatgcatg ctttgcatac ttctgcctgc tggggagcct ggggactttc 8520
cacaccctaa ctgacacaca ttccacagct gcattaatga atcggccaac gcgcggggag 8580
aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt 8640
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga 8700

atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 8760
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc ccccctgacg agcatcacaa 8820
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 8880
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 8940
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 9000
cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 9060
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 9120

atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 9180

tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat 9240

ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 9300

acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 9360

aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 9420

aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct 9480

tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga 9540

cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc 9600

catagttgcc tgactcctgc aaaccacgtt gtgtctcaaa atctctgatg ttacattgca 9660

caagataaaa atatatcatc atgaacaata aaactgtctg cttacataaa cagtaataca 9720

aggggtgtta tgagccatat tcaacgggaa acgtcttgct cgaggccgcg attaaattcc 9780

aacatggatg ctgatttata tgggtataaa tgggctcgcg ataatgtcgg gcaatcaggt 9840

gcgacaatct atcgattgta tgggaagccc gatgcgccag agttgtttct gaaacatggc 9900

aaaggtagcg ttgccaatga tgttacagat gagatggtca gactaaactg gctgacggaa 9960

tttatgcctc ttccgaccat caagcatttt atccgtactc ctgatgatgc atggttactc 10020

accactgcga tccccgggaa aacagcattc caggtattag aagaatatcc tgattcaggt 10080

gaaaatattg ttgatgcgct ggcagtgttc ctgcgccggt tgcattcgat tcctgtttgt 10140

aattgtcctt ttaacagcga tcgcgtattt cgtctcgctc aggcgcaatc acgaatgaat 10200

aacggtttgg ttgatgcgag tgattttgat gacgagcgta atggctggcc tgttgaacaa 10260

gtctggaaag aaatgcataa gcttttgcca ttctcaccgg attcagtcgt cactcatggt 10320

gatttctcac ttgataacct tatttttgac gaggggaaat taataggttg tattgatgtt 10380

ggacgagtcg gaatcgcaga ccgataccag gatcttgcca tcctatggaa ctgcctcggt 10440

gagttttctc cttcattaca gaaacggctt tttcaaaaat atggtattga taatcctgat 10500

atgaataaat tgcagtttca tttgatgctc gatgagtttt tctaagggcg gcctgccacc 10560

atacccacgc cgaaacaagc gctcatgagc ccgaagtggc gagcccgatc ttccccatcg 10620

gtgatgtcgg cgatataggc gccagcaacc gcacctgtgg cgccggtgat gagggcgcgc 10680

caagtcgacg tccggcagtc 10700

<210> 9

<211> 10700

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 9

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300

ctagttatta atagtaatca attacggggt cattagttca tagcccatat atggagttcc 360

gcgttacata acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat 420

tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc 480

aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc 540

caagtacgcc ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt 600

acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta 660

ccatggtcga ggtgagcccc acgttctgct tcactctccc catctccccc ccctccccac 720

ccccaatttt gtatttattt atttttaat tattttgtgc agcgatgggg gcgggggggg 780

gggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga 840

gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa gtttcctttt atggcgaggc 900

ggcggcggcg gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgacgc 960

tgccttcgcc ccgtgccccg ctccgccgcc gcctcgcgcc gcccgccccg gctctgactg 1020

accgcgttac tcccacaggt gagcgggcgg gacggccctt ctcctccggg ctgtaattag 1080

cgcttggttt aatgacggct tgtttctttt ctgtggctgc gtgaaagcct tgaggggctc 1140

cgggagctag agcctctgct aaccatgttc atgccttctt ctttttccta cagctcctgg 1200

gcaacgtgct ggttattgtg ctgtctcatc attttggcaa agaattcctc gaagatccga 1260

agggaaagtc ttccacgact gtgggatccg ttcgaagata tcaccggttg agccaccatg 1320

gaattcagca gccccagcag agaggaatgc cccaagcctc tgagccgggt gtcaatcatg 1380

gccggatctc tgacaggact gctgctgctt caggccgtgt cttgggcttc tggcgctaga 1440

ccttgcatcc ccaagagctt cggctacagc agcgtcgtgt gcgtgtgcaa tgccacctac 1500

tgcgacagct tcgaccctcc tacctttcct gctctgggca ccttcagcag atacgagagc 1560

accagatccg gcagacggat ggaactgagc atgggaccca tccaggccaa tcacacaggc 1620

actggcctgc tgctgacact gcagcctgag cagaaattcc agaaagtgaa aggcttcggc 1680

ggagccatga cagatgccgc cgctctgaat atcctggctc tgtctccacc agctcagaac 1740

ctgctgctca agagctactt cagcgaggaa ggcatcggct acaacatcat cagagtgccc 1800
atggccagct gcgacttcag catcaggacc tacacctacg ccgacacacc cgacgatttc 1860
cagctgcaca acttcagcct gcctgaagag gacaccaagc tgaagatccc tctgatccac 1920
agagccctgc agctggcaca aagacccgtg tcactgctgg cctctccatg gacatctccc 1980
acctggctga aaacaaatgg cgccgtgaat ggcaagggca gcctgaaagg ccaacctggc 2040
gacatctacc accagacctg ggccagatac ttcgtgaagt tcctggacgc ctatgccgag 2100

cacaagctgc agttttgggc cgtgacagcc gagaacgaac cttctgctgg actgctgagc 2160
ggctacccct ttcagtgcct gggctttaca cccgagcacc agcgggactt tatcgcccgt 2220
gatctgggac ccacactggc caatagcacc caccataatg tgcggctgct gatgctggac 2280
gaccagagac tgcttctgcc ccactgggct aaagtggtgc tgacagatcc tgaggccgcc 2340
aaatacgtgc acggaatcgc cgtgcactgg tatctggact ttctggcccc tgccaaggcc 2400
acactgggag agacacacag actgttcccc aacaccatgc tgttcgccag cgaagcctgt 2460
gtgggcagca agttttggga acagagcgtg cggctcggca gctgggatag aggcatgcag 2520

tacagccaca gcatcatcac caacctgctg taccacgtcg tcggctggac cgactggaat 2580
ctggccctga atcctgaagg cggccctaac tgggtccgaa acttcgtgga cagccccatc 2640
atcgtggaca tcaccaagga caccttctac aagcagccca tgttctacca cctgggacac 2700
ttcagcaagt tcatccccga gggctctcag cgcgttggac tggtggcttc ccagaagaac 2760
gatctggacg ccgtggctct gatgcaccct gatggatctg ctgtggtggt ggtcctgaac 2820
cgcagcagca aagatgtgcc cctgaccatc aaggatcccg ccgtgggatt cctggaaaca 2880
atcagccctg gctactccat ccacacctac ctgtggcgta gacagtgaca attgttaatt 2940

aagtttaaac cctcgaggcc gcaagcttat cgataatcaa cctctggatt acaaaatttg 3000
tgaaagattg actggtattc ttaactatgt tgctcctttt acgctatgtg gatacgctgc 3060
tttaatgcct ttgtatcatg ctattgcttc ccgtatggct ttcattttct cctccttgta 3120
taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtcaggc aacgtggcgt 3180
ggtgtgcact gtgtttgctg acgcaacccc cactggttgg ggcattgcca ccacctgtca 3240
gctcctttcc gggactttcg ctttccccct ccctattgcc acggcggaac tcatcgccgc 3300
ctgccttgcc cgctgctgga caggggctcg gctgttgggc actgacaatt ccgtggtgtt 3360

gtcggggaaa tcatcgtcct ttccttggct gctcgcctgt gttgccacct ggattctgcg 3420
cgggacgtcc ttctgctacg tcccttcggc cctcaatcca gcggaccttc cttcccgcgg 3480
cctgctgccg gctctgcggc ctcttccgcg tcttcgcctt cgccctcaga cgagtcggat 3540
ctccctttgg gccgcctccc cgcatcgata ccgtcgacta gagctcgctg atcagcctcg 3600

actgtgcctt ctagttgcca gccatctgtt gtttgcccct cccccgtgcc ttccttgacc 3660
ctggaaggtg ccactcccac tgtcctttcc taataaaatg aggaaattgc atcgcattgt 3720
ctgagtaggt gtcattctat tctgggggt ggggtggggc aggacagcaa gggggaggat 3780

tgggaagaca atagcaggca tgctggggag agatccacga taacaaacag ctttttggg 3840
gtgaacatat tgactgaatt ccctgcaggt tggccactcc ctctctgcgc gctcgctcgc 3900
tcactgaggc cgcccgggca aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag 3960
tgagcgagcg agcgcgcaga gagggagtgg ccaactccat cactaggggt tcctgcggcc 4020
gctcgtacgg tctcgaggaa ttcctgcagg ataacttgcc aacctcattc taaaatgtat 4080
atagaagccc aaaagacaat aacaaaaata ttcttgtaga acaaaatggg aaagaatgtt 4140
ccactaaata tcaagattta gagcaaagca tgagatgtgt ggggatagac agtgaggctg 4200

ataaaataga gtagagctca gaaacagacc cattgatata tgtaagtgac ctatgaaaaa 4260
aatatggcat tttacaatgg gaaaatgatg gtctttttct tttttagaaa aacagggaaa 4320
tatatttata tgtaaaaaat aaaagggaac ccatatgtca taccatacac acaaaaaaat 4380
tccagtgaat tataagtcta aatggagaag gcaaaacttt aaatctttta gaaaataata 4440
tagaagcatg cagaccagcc tggccaacat gatgaaaccc tctctactaa taataaaatc 4500
agtagaacta ctcaggacta ctttgagtgg gaagtccttt tctatgaaga cttctttggc 4560
caaaattagg ctctaaatgc aaggagatag tgcatcatgc ctggctgcac ttactgataa 4620

atgatgttat caccatcttt aaccaaatgc acaggaacaa gttatggtac tgatgtgctg 4680
gattgagaag gagctctact tccttgacag gacacatttg tatcaactta aaaaagcaga 4740
tttttgccag cagaactatt cattcagagg taggaaactt agaatagatg atgtcactga 4800
ttagcatggc ttccccatct ccacagctgc ttcccaccca ggttgcccac agttgagttt 4860
gtccagtgct cagggctgcc cactctcagt aagaagcccc acaccagccc ctctccaaat 4920
atgttggctg ttccttccat taaagtgacc ccactttaga gcagcaagtg gatttctgtt 4980
tcttacagtt caggaaggag gagtcagctg tgagaacctg gagcctgaga tgcttctaag 5040

tcccactgct actggggtca gggaagccag actccagcat cagcagtcag gagcactaag 5100
cccttgccaa catcctgttt ctcagagaaa ctgcttccat tataatggtt gtcctttttt 5160
aagctatcaa gccaaacaac cagtgtctac cattattctc atcacctgaa gccaagggtt 5220
ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct ccagcttctg tcttcagtca 5280
ctccactctt agcctgctct gaatcaactc tgaccacagt tccctggagc ccctgccacc 5340
tgctgcccct gccaccttct ccatctgcag tgctgtgcag ccttctgcac tcttgcagag 5400

ctaataggtg gagacttgaa ggaagaggag gaaagtttct cataatagcc ttgctgcaag 5460

ctcaaatggg aggtgggcac tgtgcccagg agccttggag caaaggctgt gcccaacctc 5520
tgactgcatc caggtttggt cttgacagag ataagaagcc ctggcttttg gagccaaaat 5580
ctaggtcaga cttaggcagg attctcaaag tttatcagca gaacatgagg cagaagaccc 5640
tttctgctcc agcttcttca ggctcaacct tcatcagaat agatagaaag agaggctgtg 5700
agggttctta aaacagaagc aaatctgact cagagaataa acaacctcct agtaaactac 5760
agcttagaca gagcatctgg tggtgagtgt gctcagtgtc ctactcaact gtctggtatc 5820
agccctcatg aggacttctc ttctttccct catagacctc catctctgtt ttccttagcc 5880

tgcagaaatc tggatggcta ttcacagaat gcctgtgctt tcagagttgc atttttctc 5940
tggtattctg gttcaagcat ttgaaggtag gaaaggttct ccaagtgcaa gaaagccagc 6000
cctgagcctc aactgcctgg ctagtgtggt cagtaggatg caaaggctgt tgaatgccac 6060
aaggccaaac tttaacctgt gtaccacaag cctagcagca gaggcagctc tgctcactgg 6120
aactctctgt cttctttctc ctgagccttt tcttttcctg agttttctag ctctcctcaa 6180
ccttacctct gccctaccca ggacaaaccc aagagccact gtttctgtga tgtcctctcc 6240
agccctaatt aggcatcatg acttcagcct gaccttccat gctcagaagc agtgctaatc 6300

cacttcagat gagctgctct atgcaacaca ggcagagcct acaaaccttt gcaccagagc 6360
cctccacata tcagtgtttg ttcatactca cttcaacagc aaatgtgact gctgagatta 6420
agattttaca caagatggtc tgtaatttca cagttagttt tatcccatta ggtatgaaag 6480
aattagcata attcccctta aacatgaatg aatcttagat tttttaataa atagttttgg 6540
aagtaaagac agagacatca ggagcacaag gaatagcctg agaggacaaa cagaacaaga 6600
aagagtctgg aaatacacag gatgttcttg gcctcctcaa agcaagtgca agcagatagt 6660
accagcagcc ccaggctatc agagcccagt gaagagaagt accatgaaag ccacagctct 6720

aaccaccctg ttccagagtg acagacagtc cccaagacaa gccagcctga gccagagaga 6780
gaactgcaag agaaagtttc taatttaggt tctgttagat tcagacaagt gcaggtcatc 6840
ctctctccac agctactcac ctctccagcc taacaaagcc tgcagtccac actccaaccc 6900
tggtgtctca cctcctagcc tctcccaaca tcctgctctc tgaccatctt ctgcatctct 6960
catctcacca tctcccactg tctacagcct actcttgcaa ctaccatctc attttctgac 7020
atcctgtcta catcttctgc catactctgc catctaccat accacctctt accatctacc 7080
acaccatctt ttatctccat ccctctcaga agcctccaag ctgaatcctg ctttatgtgt 7140

tcatctcagc ccctgcatgg aaagctgacc ccagaggcag aactattccc agagagcttg 7200
gccaagaaaa acaaaactac cagcctggcc aggctcagga gtagtaagct gcagtgtctg 7260
ttgtgttcta gcttcaacag ctgcaggagt tccactctca aatgctccac atttctcaca 7320
tcctcctgat tctggtcact acccatcttc aaagaacaga atatctcaca tcagcatact 7380
gtgaaggact agtcatgggt gcagctgctc agagctgcaa agtcattctg gatggtggag 7440
agcttacaaa catttcatga tgctcccccc gctctgatgg ctggagccca atccctacac 7500
agactcctgc tgtatgtgtt ttcctttcac tctgagccac agccagaggg caggcattca 7560

gtctcctctt caggctgggg ctggggcact gagaactcac ccaacacctt gctctcactc 7620
cttctgcaaa acaagaaaga gctttgtgct gcagtagcca tgaagaatga aaggaaggct 7680
ttaactaaaa aatgtcagag attattttca acccccttact gtggatcacc agcaaggagg 7740
aaacacaaca cagagacatt ttttcccctc aaattatcaa aagaatcact gcatttgtta 7800
aagagagcaa ctgaatcagg aagcagagtt ttgaacatat cagaagttag gaatctgcat 7860
cagagacaaa tgcagtcatg gttgtttgct gcataccagc cctaatcatt agaagcctca 7920
tggacttcaa acatcattcc ctctgacaag atgctctagc ctaactccat gagataaaat 7980

aaatctgcct ttcagagcca aagaagagtc caccagcttc ttctcagtgt gaacaagagc 8040
tccagtcagg ttagtcagtc cagtgcagta gaggagacca gtctgcatcc tctaattttc 8100
aaaggcaaga agatttgttt accctggaca ccaggcacaa gtgaggtcac agagctctta 8160
gatatgcagt cctcatgagt gaggagacta aagcgcatgc catcaagact tcagtgtaga 8220
gaaaacctcc aaaaaagcct cctcactact tctggaatag ctcagaggcc gaggcggcct 8280
cggcctctgc ataaataaaa aaaattagtc agccatgggg cggagaatgg gcggaactgg 8340
gcggagttag gggcgggatg ggcggagtta ggggcgggac tatggttgct gactaattga 8400

gatgcatgct ttgcatactt ctgcctgctg gggagcctgg ggactttcca cacctggttg 8460
ctgactaatt gagatgcatg ctttgcatac ttctgcctgc tggggagcct ggggactttc 8520
cacaccctaa ctgacacaca ttccacagct gcattaatga atcggccaac gcgcggggag 8580
aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt 8640
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga 8700
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 8760
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc ccccctgacg agcatcacaa 8820

aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 8880
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 8940
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 9000

cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 9060

cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 9120

atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 9180

tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat 9240

ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 9300

acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 9360

aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 9420

aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct 9480

tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga 9540

cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc 9600

catagttgcc tgactcctgc aaaccacgtt gtgtctcaaa atctctgatg ttacattgca 9660

caagataaaa atatatcatc atgaacaata aaactgtctg cttacataaa cagtaataca 9720

aggggtgtta tgagccatat tcaacgggaa acgtcttgct cgaggccgcg attaaattcc 9780

aacatggatg ctgatttata tgggtataaa tgggctcgcg ataatgtcgg gcaatcaggt 9840

gcgacaatct atcgattgta tgggaagccc gatgcgccag agttgtttct gaaacatggc 9900

aaaggtagcg ttgccaatga tgttacagat gagatggtca gactaaactg gctgacggaa 9960

tttatgcctc ttccgaccat caagcatttt atccgtactc ctgatgatgc atggttactc 10020

accactgcga tccccgggaa aacagcattc caggtattag aagaatatcc tgattcaggt 10080

gaaaatattg ttgatgcgct ggcagtgttc ctgcgccggt tgcattcgat tcctgtttgt 10140

aattgtcctt ttaacagcga tcgcgtattt cgtctcgctc aggcgcaatc acgaatgaat 10200

aacggtttgg ttgatgcgag tgattttgat gacgagcgta atggctggcc tgttgaacaa 10260

gtctggaaag aaatgcataa gcttttgcca ttctcaccgg attcagtcgt cactcatggt 10320

gatttctcac ttgataacct tatttttgac gaggggaaat taataggttg tattgatgtt 10380

ggacgagtcg gaatcgcaga ccgataccag gatcttgcca tcctatggaa ctgcctcggt 10440

gagttttctc cttcattaca gaaacggctt tttcaaaaat atggtattga taatcctgat 10500

atgaataaat tgcagtttca tttgatgctc gatgagtttt tctaagggcg gcctgccacc 10560

atacccacgc cgaaacaagc gctcatgagc ccgaagtggc gagcccgatc ttccccatcg 10620

gtgatgtcgg cgatataggc gccagcaacc gcacctgtgg cgccggtgat gagggcgcgc 10680

caagtcgacg tccggcagtc 10700

<210> 10

<211> 10700

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 10

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgcccgggc aaagcccggg 60

cgtcgggcga cctttggtcg cccggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300

ctagttatta atagtaatca attacggggt cattagttca tagcccatat atggagttcc 360

gcgttacata acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat 420

tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc 480

aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc 540

caagtacgcc ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt 600

acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta 660

ccatggtcga ggtgagcccc acgttctgct tcactctccc catctccccc ccctccccac 720

ccccaatttt gtatttattt atttttaat tattttgtgc agcgatgggg gcgggggggg 780

gggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga 840

gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa gtttcctttt atggcgaggc 900

ggcggcggcg gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgacgc 960

tgccttcgcc ccgtgccccg ctccgccgcc gcctcgcgcc gcccgccccg gctctgactg 1020

accgcgttac tcccacaggt gagcgggcgg gacggccctt ctcctccggg ctgtaattag 1080

cgcttggttt aatgacggct tgtttctttt ctgtggctgc gtgaaagcct tgaggggctc 1140

cgggagctag agcctctgct aaccatgttc atgccttctt ctttttccta cagctcctgg 1200

gcaacgtgct ggttattgtg ctgtctcatc attttggcaa agaattcctc gaagatccga 1260

agggaaagtc ttccacgact gtgggatccg ttcgaagata tcaccggttg agccaccatg 1320

gaattcagca gccccagcag agaggaatgc cccaagcctc tgagccgggt gtcaatcatg 1380

gccggatctc tgacaggact gctgctgctt caggccgtgt cttgggcttc tggcgctaga 1440

ccttgcatcc ccaagagctt cggctacagc agcgtcgtgt gcgtgtgcaa tgccacctac 1500

tgcgacagct tcgaccctcc tacctttcct gctctgggca ccttcagcag atacgagagc 1560

accagatccg gcagacggat ggaactgagc atgggaccca tccaggccaa tcacacaggc 1620

actggcctgc tgctgacact gcagcctgag cagaaattcc agaaagtgaa aggcttcggc 1680
ggagccatga cagatgccgc cgctctgaat atcctggctc tgtctccacc agctcagaac 1740
ctgctgctca agagctactt cagcgaggaa ggcatcggct acaacatcat cagagtgccc 1800

atggccagct gcgacttcag catcaggacc tacacctacg ccgacacacc cgacgatttc 1860
cagctgcaca acttcagcct gcctgaagag gacaccaagc tgaagatccc tctgatccac 1920
agagccctgc agctggcaca aagacccgtg tcactgctgg cctctccatg gacatctccc 1980
acctggctga aaacaaatgg cgccgtgaat ggcaagggca gcctgaaagg ccaacctggc 2040
gacatctacc accagacctg ggccagatac ttcgtgaagt tcctggacgc ctatgccgag 2100
cacaagctgc agttttgggc cgtgacagcc gagaacgaac cttctgctgg actgctgagc 2160
ggctacccct ttcagtgcct gggctttaca cccgagcacc agcgggactt tatcgcccgt 2220

gatctgggac ccacactggc caatagcacc caccataatg tgcggctgct gatgctggac 2280
gaccagagac tgcttctgcc ccactgggct aaagtggtgc tgacagatcc tgaggccgcc 2340
aaatacgtgc acggaatcgc cgtgcactgg tatctggact ttctggcccc tgccaaggcc 2400
acactgggag agacacacag actgttcccc aacaccatgc tgttcgccag cgaagcctgt 2460
gtgggcagca agttttggga acagagcgtg cggctcggca gctgggatag aggcatgcag 2520
tacagccaca gcatcatcac caacctgctg taccacgtcg tcggctggac cgactggaat 2580
ctggccctga atcctgaagg cggccctaac tgggtccgaa acttcgtgga cagccccatc 2640

atcgtggaca tcaccaagga caccttctac aagcagccca tgttctacca cctgggacac 2700
ttcagcaagt tcatccccga gggctctcag cgcgttggac tggtggcttc ccagaagaac 2760
gatctggacg ccgtggctct gatgcaccct gatggatctg ctgtggtggt ggtcctgaac 2820
cgcagcagca aagatgtgcc cctgaccatc aaggatcccg ccgtgggatt cctggaaaca 2880
atcagccctg gctactccat ccacacctac ctgtggcgta gacagtgaca attgttaatt 2940
aagtttaaac cctcgaggcc gcaagcttat cgataatcaa cctctggatt acaaaatttg 3000
tgaaagattg actggtattc ttaactatgt tgctcctttt acgctatgtg gatacgctgc 3060

tttaatgcct ttgtatcatg ctattgcttc ccgtatggct ttcattttct cctccttgta 3120
taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtcaggc aacgtggcgt 3180
ggtgtgcact gtgtttgctg acgcaacccc cactggttgg ggcattgcca ccacctgtca 3240
gctcctttcc gggactttcg ctttccccct ccctattgcc acggcggaac tcatcgccgc 3300
ctgccttgcc cgctgctgga caggggctcg gctgttgggc actgacaatt ccgtggtgtt 3360
gtcggggaaa tcatcgtcct ttccttggct gctcgcctgt gttgccacct ggattctgcg 3420

cgggacgtcc ttctgctacg tcccttcggc cctcaatcca gcggaccttc cttcccgcgg 3480

cctgctgccg gctctgcggc ctcttccgcg tcttcgcctt cgccctcaga cgagtcggat 3540
ctccctttgg gccgcctccc cgcatcgata ccgtcgacta gagctcgctg atcagcctcg 3600
actgtgcctt ctagttgcca gccatctgtt gtttgcccct cccccgtgcc ttccttgacc 3660
ctggaaggtg ccactcccac tgtcctttcc taataaaatg aggaaattgc atcgcattgt 3720
ctgagtaggt gtcattctat tctggggggt ggggtggggc aggacagcaa gggggaggat 3780
tgggaagaca atagcaggca tgctggggag agatccacga taacaaacag ctttttttggg 3840
gtgaacatat tgactgaatt ccctgcagga ggaaccccta gtgatggagt tggccactcc 3900

ctctctgcgc gctcgctcgc tcactgaggc cgcccgggca aagcccgggc gtcgggcgac 3960
ctttggtcgc ccggcctcag tgagcgagcg agcgcgcaga gagggagtgg ccaagcggcc 4020
gctcgtacgg tctcgaggaa ttcctgcagg ataacttgcc aacctcattc taaaatgtat 4080
atagaagccc aaaagacaat aacaaaaata ttcttgtaga acaaaatggg aaagaatgtt 4140
ccactaaata tcaagattta gagcaaagca tgagatgtgt ggggatagac agtgaggctg 4200
ataaaataga gtagagctca gaaacagacc cattgatata tgtaagtgac ctatgaaaaa 4260
aatatggcat tttacaatgg gaaaatgatg gtctttttct tttttagaaa aacagggaaa 4320

tatatttata tgtaaaaaat aaaagggaac ccatatgtca taccatacac acaaaaaaat 4380
tccagtgaat tataagtcta aatggagaag gcaaaacttt aaatctttta gaaaataata 4440
tagaagcatg cagaccagcc tggccaacat gatgaaaccc tctctactaa taataaaatc 4500
agtagaacta ctcaggacta ctttgagtgg gaagtccttt tctatgaaga cttctttggc 4560
caaaattagg ctctaaatgc aaggagatag tgcatcatgc ctggctgcac ttactgataa 4620
atgatgttat caccatcttt aaccaaatgc acaggaacaa gttatggtac tgatgtgctg 4680
gattgagaag gagctctact tccttgacag gacacatttg tatcaactta aaaaagcaga 4740

tttttgccag cagaactatt cattcagagg taggaaactt agaatagatg atgtcactga 4800
ttagcatggc ttccccatct ccacagctgc ttcccaccca ggttgcccac agttgagttt 4860
gtccagtgct cagggctgcc cactctcagt aagaagcccc acaccagccc ctctccaaat 4920
atgttggctg ttccttccat taaagtgacc ccactttaga gcagcaagtg gatttctgtt 4980
tcttacagtt caggaaggag gagtcagctg tgagaacctg gagcctgaga tgcttctaag 5040
tcccactgct actggggtca gggaagccag actccagcat cagcagtcag gagcactaag 5100
cccttgccaa catcctgttt ctcagagaaa ctgcttccat tataatggtt gtccttttt 5160

aagctatcaa gccaaacaac cagtgtctac cattattctc atcacctgaa gccaagggtt 5220

ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct ccagcttctg tcttcagtca 5280

ctccactctt agcctgctct gaatcaactc tgaccacagt tccctggagc ccctgccacc 5340

tgctgcccct gccaccttct ccatctgcag tgctgtgcag ccttctgcac tcttgcagag 5400

ctaataggtg gagacttgaa ggaagaggag gaaagtttct cataatagcc ttgctgcaag 5460

ctcaaatggg aggtgggcac tgtgcccagg agccttggag caaaggctgt gcccaacctc 5520

tgactgcatc caggtttggt cttgacagag ataagaagcc ctggcttttg gagccaaaat 5580

ctaggtcaga cttaggcagg attctcaaag tttatcagca gaacatgagg cagaagaccc 5640

tttctgctcc agcttcttca ggctcaacct tcatcagaat agatagaaag agaggctgtg 5700

agggttctta aaacagaagc aaatctgact cagagaataa acaacctcct agtaaactac 5760

agcttagaca gagcatctgg tggtgagtgt gctcagtgtc ctactcaact gtctggtatc 5820

agccctcatg aggacttctc ttctttccct catagacctc catctctgtt ttccttagcc 5880

tgcagaaatc tggatggcta ttcacagaat gcctgtgctt tcagagttgc atttttttctc 5940

tggtattctg gttcaagcat ttgaaggtag gaaaggttct ccaagtgcaa gaaagccagc 6000

cctgagcctc aactgcctgg ctagtgtggt cagtaggatg caaaggctgt tgaatgccac 6060

aaggccaaac tttaacctgt gtaccacaag cctagcagca gaggcagctc tgctcactgg 6120

aactctctgt cttctttctc ctgagccttt tcttttcctg agttttctag ctctcctcaa 6180

ccttacctct gccctaccca ggacaaaccc aagagccact gtttctgtga tgtcctctcc 6240

agccctaatt aggcatcatg acttcagcct gaccttccat gctcagaagc agtgctaatc 6300

cacttcagat gagctgctct atgcaacaca ggcagagcct acaaaccttt gcaccagagc 6360

cctccacata tcagtgtttg ttcatactca cttcaacagc aaatgtgact gctgagatta 6420

agattttaca caagatggtc tgtaatttca cagttagttt tatcccatta ggtatgaaag 6480

aattagcata attcccctta aacatgaatg aatcttagat tttttaataa atagttttgg 6540

aagtaaagac agagacatca ggagcacaag gaatagcctg agaggacaaa cagaacaaga 6600

aagagtctgg aaatacacag gatgttcttg gcctcctcaa agcaagtgca agcagatagt 6660

accagcagcc ccaggctatc agagcccagt gaagagaagt accatgaaag ccacagctct 6720

aaccaccctg ttccagagtg acagacagtc cccaagacaa gccagcctga gccagagaga 6780

gaactgcaag agaaagtttc taatttaggt tctgttagat tcagacaagt gcaggtcatc 6840

ctctctccac agctactcac ctctccagcc taacaaagcc tgcagtccac actccaaccc 6900

tggtgtctca cctcctagcc tctcccaaca tcctgctctc tgaccatctt ctgcatctct 6960

catctcacca tctcccactg tctacagcct actcttgcaa ctaccatctc attttctgac 7020

atcctgtcta catcttctgc catactctgc catctaccat accacctctt accatctacc 7080
acaccatctt ttatctccat ccctctcaga agcctccaag ctgaatcctg ctttatgtgt 7140
tcatctcagc ccctgcatgg aaagctgacc ccagaggcag aactattccc agagagcttg 7200
gccaagaaaa acaaaactac cagcctggcc aggctcagga gtagtaagct gcagtgtctg 7260

ttgtgttcta gcttcaacag ctgcaggagt tccactctca aatgctccac atttctcaca 7320
tcctcctgat tctggtcact acccatcttc aaagaacaga atatctcaca tcagcatact 7380
gtgaaggact agtcatgggt gcagctgctc agagctgcaa agtcattctg gatggtggag 7440
agcttacaaa catttcatga tgctcccccc gctctgatgg ctggagccca atccctacac 7500
agactcctgc tgtatgtgtt ttcctttcac tctgagccac agccagaggg caggcattca 7560
gtctcctctt caggctgggg ctggggcact gagaactcac ccaacacctt gctctcactc 7620
cttctgcaaa acaagaaaga gctttgtgct gcagtagcca tgaagaatga aaggaaggct 7680

ttaactaaaa aatgtcagag attattttca acccccttact gtggatcacc agcaaggagg 7740
aaacacaaca cagagacatt ttttcccctc aaattatcaa aagaatcact gcatttgtta 7800
aagagagcaa ctgaatcagg aagcagagtt ttgaacatat cagaagttag gaatctgcat 7860
cagagacaaa tgcagtcatg gttgtttgct gcataccagc cctaatcatt agaagcctca 7920
tggacttcaa acatcattcc ctctgacaag atgctctagc ctaactccat gagataaaat 7980
aaatctgcct ttcagagcca aagaagagtc caccagcttc ttctcagtgt gaacaagagc 8040
tccagtcagg ttagtcagtc cagtgcagta gaggagacca gtctgcatcc tctaattttc 8100

aaaggcaaga agatttgttt accctggaca ccaggcacaa gtgaggtcac agagctctta 8160
gatatgcagt cctcatgagt gaggagacta aagcgcatgc catcaagact tcagtgtaga 8220
gaaaacctcc aaaaaagcct cctcactact tctggaatag ctcagaggcc gaggcggcct 8280
cggcctctgc ataaataaaa aaaattagtc agccatgggg cggagaatgg gcggaactgg 8340
gcggagttag gggcgggatg ggcggagtta ggggcgggac tatggttgct gactaattga 8400
gatgcatgct ttgcatactt ctgcctgctg gggagcctgg ggactttcca cacctggttg 8460
ctgactaatt gagatgcatg ctttgcatac ttctgcctgc tggggagcct ggggactttc 8520

cacaccctaa ctgacacaca ttccacagct gcattaatga atcggccaac gcgcggggag 8580
aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt 8640
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga 8700
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 8760
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc ccccctgacg agcatcacaa 8820
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 8880

tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 8940

gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 9000
cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 9060
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 9120
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 9180
tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat 9240
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 9300
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 9360

aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 9420
aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct 9480
tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga 9540
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc 9600
catagttgcc tgactcctgc aaaccacgtt gtgtctcaaa atctctgatg ttacattgca 9660
caagataaaa atatatcatc atgaacaata aaactgtctg cttacataaa cagtaataca 9720
aggggtgtta tgagccatat tcaacgggaa acgtcttgct cgaggccgcg attaaattcc 9780

aacatggatg ctgatttata tgggtataaa tgggctcgcg ataatgtcgg gcaatcaggt 9840
gcgacaatct atcgattgta tgggaagccc gatgcgccag agttgtttct gaaacatggc 9900
aaaggtagcg ttgccaatga tgttacagat gagatggtca gactaaactg gctgacggaa 9960
tttatgcctc ttccgaccat caagcatttt atccgtactc ctgatgatgc atggttactc 10020
accactgcga tccccgggaa aacagcattc caggtattag aagaatatcc tgattcaggt 10080
gaaaatattg ttgatgcgct ggcagtgttc ctgcgccggt tgcattcgat tcctgtttgt 10140
aattgtcctt ttaacagcga tcgcgtattt cgtctcgctc aggcgcaatc acgaatgaat 10200

aacggtttgg ttgatgcgag tgattttgat gacgagcgta atggctggcc tgttgaacaa 10260
gtctggaaag aaatgcataa gcttttgcca ttctcaccgg attcagtcgt cactcatggt 10320
gatttctcac ttgataacct tatttttgac gaggggaaat taataggttg tattgatgtt 10380
ggacgagtcg gaatcgcaga ccgataccag gatcttgcca tcctatggaa ctgcctcggt 10440
gagttttctc cttcattaca gaaacggctt tttcaaaaat atggtattga taatcctgat 10500
atgaataaat tgcagtttca tttgatgctc gatgagtttt tctaagggcg gcctgccacc 10560
atacccacgc cgaaacaagc gctcatgagc ccgaagtggc gagcccgatc ttccccatcg 10620

gtgatgtcgg cgatataggc gccagcaacc gcacctgtgg cgccggtgat gagggcgcgc 10680

caagtcgacg tccggcagtc 10700

<210> 11

<211> 11188

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 11

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactatt agatctgatg gccgcgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

gtggtgactg agatgttttc taggaaacac aaaagataca aaaaagaaca cgtggaagga 300

tagccaaaaa ggggggctgc ccccatttcc tgcaccccgc tgcgatggct ggcaccattt 360

ggaagacttc gagatacact gttgagcgca gtaagacaac agtgtatctc gaagtcttcc 420

agatggggcc agccggtcca ctctgtatcc aggccagttc tgcaaggcgt tcgaggacca 480

ccccccctccc ctcgccacca gggtggtctc atacagaact tataagattc ccaaatccaa 540

agacatttca cgtttatggt gatttcccag aacacatagc gacatgcaaa tattgcaggg 600

cgccactccc ctgtccctca cagccatctt cctgccaggg cgcacgcgcg ctgggtgttc 660

ccgcctagtg acactgggcc cgcgattcct tggagcgggt tgatgacgtc agcgtttccc 720

atggtgaatc cctaggttct agaaccggtg acgtctccca tggtgaagct tggatctgaa 780

ttcggtacct agttattaat agtaatcaat tacggggtca ttagttcata gcccatatat 840

ggagttccgc gttacataac ttacggtaaa tggcccgcct ggctgaccgc ccaacgaccc 900

ccgcccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca 960

ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagtac atcaagtgta 1020

tcatatgcca agtacgcccc ctattgacgt caatgacggt aaatggcccg cctggcatta 1080

tgcccagtac atgaccttat gggactttcc tacttggcag tacatctacg tattagtcat 1140

cgctattacc atggtcgagg tgagccccac gttctgcttc actctcccca tctccccccc 1200

ctccccaccc ccaattttgt atttatttat tttttaatta ttttgtgcag cgatgggggc 1260

gggggggggg gggggcgcg cgccaggcgg ggcggggcgg ggcgaggggc ggggcggggc 1320

gaggcggaga ggtgcggcgg cagccaatca gagcggcgcg ctccgaaagt ttccttttat 1380

ggcgaggcgg cggcggcggc ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc 1440

tgcgacgctg ccttcgcccc gtgccccgct ccgccgccgc ctcgcgccgc ccgccccggc 1500

tctgactgac cgcgttactc ccacaggtga gcgggcggga cggcccttct cctccgggct 1560
gtaattagcg cttggtttaa tgacggcttg tttcttttct gtggctgcgt gaaagccttg 1620
aggggctccg ggagctagag cctctgctaa ccatgttcat gccttcttct ttttcctaca 1680
gctcctgggc aacgtgctgg ttattgtgct gtctcatcat tttggcaaag aattcctcga 1740
agatccgaag ggaaagtctt ccacgactgt gggatccgtt cgaagatatc accggttgag 1800
ccaccatgga attcagcagc cccagcagag aggaatgccc caagcctctg agccgggtgt 1860
caatcatggc cggatctctg acaggactgc tgctgcttca ggccgtgtct tgggcttctg 1920

gcgctagacc ttgcatcccc aagagcttcg gctacagcag cgtcgtgtgc gtgtgcaatg 1980
ccacctactg cgacagcttc gaccctccta cctttcctgc tctgggcacc ttcagcagat 2040
acgagagcac cagatccggc agacggatgg aactgagcat gggacccatc caggccaatc 2100
acacaggcac tggcctgctg ctgacactgc agcctgagca gaaattccag aaagtgaaag 2160
gcttcggcgg agccatgaca gatgccgccg ctctgaatat cctggctctg tctccaccag 2220
ctcagaacct gctgctcaag agctacttca gcgaggaagg catcggctac aacatcatca 2280
gagtgcccat ggccagctgc gacttcagca tcaggaccta cacctacgcc gacacacccg 2340

acgatttcca gctgcacaac ttcagcctgc ctgaagagga caccaagctg aagatccctc 2400
tgatccacag agccctgcag ctggcacaaa gacccgtgtc actgctggcc tctccatgga 2460
catctcccac ctggctgaaa acaaatggcg ccgtgaatgg caagggcagc ctgaaaggcc 2520
aacctggcga catctaccac cagacctggg ccagatactt cgtgaagttc ctggacgcct 2580
atgccgagca caagctgcag ttttgggccg tgacagccga gaacgaacct tctgctggac 2640
tgctgagcgg ctaccccttt cagtgcctgg gctttacacc cgagcaccag cgggacttta 2700
tcgcccgtga tctgggaccc acactggcca atagcaccca ccataatgtg cggctgctga 2760

tgctggacga ccagagactg cttctgcccc actgggctaa agtggtgctg acagatcctg 2820
aggccgccaa atacgtgcac ggaatcgccg tgcactggta tctggacttt ctggcccctg 2880
ccaaggccac actgggagag acacacagac tgttccccaa caccatgctg ttcgccagcg 2940
aagcctgtgt gggcagcaag ttttgggaac agagcgtgcg gctcggcagc tgggatagag 3000
gcatgcagta cagccacagc atcatcacca acctgctgta ccacgtcgtc ggctggaccg 3060
actggaatct ggccctgaat cctgaaggcg gccctaactg ggtccgaaac ttcgtggaca 3120
gccccatcat cgtggacatc accaaggaca ccttctacaa gcagcccatg ttctaccacc 3180

tgggacactt cagcaagttc atccccgagg gctctcagcg cgttggactg gtggcttccc 3240
agaagaacga tctggacgcc gtggctctga tgcaccctga tggatctgct gtggtggtgg 3300
tcctgaaccg cagcagcaaa gatgtgcccc tgaccatcaa ggatcccgcc gtgggattcc 3360
tggaaacaat cagccctggc tactccatcc acacctacct gtggcgtaga cagtgacaat 3420
tgttaattaa gtttaaaccc tcgaggccgc aagcttatcg ataatcaacc tctggattac 3480
aaaatttgtg aaagattgac tggtattctt aactatgttg ctccttttac gctatgtgga 3540
tacgctgctt taatgccttt gtatcatgct attgcttccc gtatggcttt cattttctcc 3600

tccttgtata aatcctggtt gctgtctctt tatgaggagt tgtggcccgt tgtcaggcaa 3660
cgtggcgtgg tgtgcactgt gtttgctgac gcaacccccа ctggttgggg cattgccacc 3720
acctgtcagc tcctttccgg gactttcgct ttccccctcc ctattgccac ggcggaactc 3780
atcgccgcct gccttgcccg ctgctggaca ggggctcggc tgttgggcac tgacaattcc 3840
gtggtgttgt cggggaaatc atcgtccttt ccttggctgc tcgcctgtgt tgccacctgg 3900
attctgcgcg ggacgtcctt ctgctacgtc ccttcggccc tcaatccagc ggaccttcct 3960
tcccgcggcc tgctgccggc tctgcggcct cttccgcgtc ttcgccttcg ccctcagacg 4020

agtcggatct ccctttgggc cgcctccccg catcgatacc gtcgactaga gctcgctgat 4080
cagcctcgac tgtgccttct agttgccagc catctgttgt ttgcccctcc cccgtgcctt 4140
ccttgaccct ggaaggtgcc actcccactg tcctttccta ataaaatgag gaaattgcat 4200
cgcattgtct gagtaggtgt cattctattc tggggggtgg ggtggggcag gacagcaagg 4260
gggaggattg ggaagacaat agcaggcatg ctggggagag atccacgata acaaacagct 4320
tttttggggt gaacatattg actgaattcc ctgcaggttg gccactccct ctctgcgcgc 4380
tcgctcgctc actgaggccg cccgggcaaa gcccgggcgt cgggcgacct ttggtcgccc 4440

ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aactccatca ctaggggttc 4500
ctgcggccgc tcgtacggtc tcgaggaatt cctgcaggat aacttgccaa cctcattcta 4560
aaatgtatat agaagcccaa aagacaataa caaaaatatt cttgtagaac aaaatgggaa 4620
agaatgttcc actaaatatc aagatttaga gcaaagcatg agatgtgtgg ggatagacag 4680
tgaggctgat aaaatagagt agagctcaga aacagaccca ttgatatatg taagtgacct 4740
atgaaaaaaa tatggcattt tacaatggga aaatgatggt ctttttcttt tttagaaaaa 4800
cagggaaata tatttatatg taaaaaataa aagggaaccc atatgtcata ccatacacac 4860

aaaaaaattc cagtgaatta taagtctaaa tggagaaggc aaaactttaa atcttttaga 4920
aaataatata gaagcatgca gaccagcctg gccaacatga tgaaaccctc tctactaata 4980
ataaaatcag tagaactact caggactact ttgagtggga agtccttttc tatgaagact 5040

tctttggcca aaattaggct ctaaatgcaa ggagatagtg catcatgcct ggctgcactt 5100
actgataaat gatgttatca ccatctttaa ccaaatgcac aggaacaagt tatggtactg 5160
atgtgctgga ttgagaagga gctctacttc cttgacagga cacatttgta tcaacttaaa 5220
aaagcagatt tttgccagca gaactattca ttcagaggta ggaaacttag aatagatgat 5280

gtcactgatt agcatggctt ccccatctcc acagctgctt cccacccagg ttgcccacag 5340
ttgagtttgt ccagtgctca gggctgccca ctctcagtaa gaagccccac accagcccct 5400
ctccaaatat gttggctgtt ccttccatta aagtgacccc actttagagc agcaagtgga 5460
tttctgtttc ttacagttca ggaaggagga gtcagctgtg agaacctgga gcctgagatg 5520
cttctaagtc ccactgctac tggggtcagg gaagccagac tccagcatca gcagtcagga 5580
gcactaagcc cttgccaaca tcctgtttct cagagaaact gcttccatta taatggttgt 5640
cctttttaa gctatcaagc caaacaacca gtgtctacca ttattctcat cacctgaagc 5700

caagggttct agcaaaagtc aagctgtctt gtaatggttg atgtgcctcc agcttctgtc 5760
ttcagtcact ccactcttag cctgctctga atcaactctg accacagttc cctggagccc 5820
ctgccacctg ctgcccctgc caccttctcc atctgcagtg ctgtgcagcc ttctgcactc 5880
ttgcagagct aataggtgga gacttgaagg aagaggagga aagtttctca taatagcctt 5940
gctgcaagct caaatgggag gtgggcactg tgcccaggag ccttggagca aaggctgtgc 6000
ccaacctctg actgcatcca ggtttggtct tgacagagat aagaagccct ggcttttgga 6060
gccaaaatct aggtcagact taggcaggat tctcaaagtt tatcagcaga acatgaggca 6120

gaagaccctt tctgctccag cttcttcagg ctcaaccttc atcagaatag atagaaagag 6180
aggctgtgag ggttcttaaa acagaagcaa atctgactca gagaataaac aacctcctag 6240
taaactacag cttagacaga gcatctggtg gtgagtgtgc tcagtgtcct actcaactgt 6300
ctggtatcag ccctcatgag gacttctctt ctttccctca tagacctcca tctctgtttt 6360
ccttagcctg cagaaatctg gatggctatt cacagaatgc ctgtgctttc agagttgcat 6420
tttttctctg gtattctggt tcaagcattt gaaggtagga aaggttctcc aagtgcaaga 6480
aagccagccc tgagcctcaa ctgcctggct agtgtggtca gtaggatgca aaggctgttg 6540

aatgccacaa ggccaaactt taacctgtgt accacaagcc tagcagcaga ggcagctctg 6600
ctcactggaa ctctctgtct tctttctcct gagccttttc ttttcctgag ttttctagct 6660
ctcctcaacc ttacctctgc cctacccagg acaaacccaa gagccactgt ttctgtgatg 6720
tcctctccag ccctaattag gcatcatgac ttcagcctga ccttccatgc tcagaagcag 6780
tgctaatcca cttcagatga gctgctctat gcaacacagg cagagcctac aaacctttgc 6840
accagagccc tccacatatc agtgtttgtt catactcact tcaacagcaa atgtgactgc 6900

tgagattaag attttacaca agatggtctg taatttcaca gttagtttta tcccattagg 6960

tatgaaagaa ttagcataat tccccttaaa catgaatgaa tcttagattt tttaataaat 7020
agttttggaa gtaaagacag agacatcagg agcacaagga atagcctgag aggacaaaca 7080
gaacaagaaa gagtctggaa atacacagga tgttcttggc ctcctcaaag caagtgcaag 7140
cagatagtac cagcagcccc aggctatcag agcccagtga agagaagtac catgaaagcc 7200
acagctctaa ccaccctgtt ccagagtgac agacagtccc caagacaagc cagcctgagc 7260
cagagagaga actgcaagag aaagtttcta atttaggttc tgttagattc agacaagtgc 7320
aggtcatcct ctctccacag ctactcacct ctccagccta acaaagcctg cagtccacac 7380

tccaaccctg gtgtctcacc tcctagcctc tcccaacatc ctgctctctg accatcttct 7440
gcatctctca tctcaccatc tcccactgtc tacagcctac tcttgcaact accatctcat 7500
tttctgacat cctgtctaca tcttctgcca tactctgcca tctaccatac cacctcttac 7560
catctaccac accatctttt atctccatcc ctctcagaag cctccaagct gaatcctgct 7620
ttatgtgttc atctcagccc ctgcatggaa agctgacccc agaggcagaa ctattcccag 7680
agagcttggc caagaaaaac aaaactacca gcctggccag gctcaggagt agtaagctgc 7740
agtgtctgtt gtgttctagc ttcaacagct gcaggagttc cactctcaaa tgctccacat 7800

ttctcacatc ctcctgattc tggtcactac ccatcttcaa agaacagaat atctcacatc 7860
agcatactgt gaaggactag tcatgggtgc agctgctcag agctgcaaag tcattctgga 7920
tggtggagag cttacaaaca tttcatgatg ctccccccgc tctgatggct ggagcccaat 7980
ccctacacag actcctgctg tatgtgtttt cctttcactc tgagccacag ccagagggca 8040
ggcattcagt ctcctcttca ggctggggct ggggcactga gaactcaccc aacaccttgc 8100
tctcactcct tctgcaaaac aagaaagagc tttgtgctgc agtagccatg aagaatgaaa 8160
ggaaggcttt aactaaaaaa tgtcagagat tattttcaac cccttactgt ggatcaccag 8220

caaggaggaa acacaacaca gagacatttt ttcccctcaa attatcaaaa gaatcactgc 8280
atttgttaaa gagagcaact gaatcaggaa gcagagtttt gaacatatca gaagttagga 8340
atctgcatca gagacaaatg cagtcatggt tgtttgctgc ataccagccc taatcattag 8400
aagcctcatg gacttcaaac atcattccct ctgacaagat gctctagcct aactccatga 8460
gataaaataa atctgccttt cagagccaaa gaagagtcca ccagcttctt ctcagtgtga 8520
acaagagctc cagtcaggtt agtcagtcca gtgcagtaga ggagaccagt ctgcatcctc 8580
taattttcaa aggcaagaag atttgtttac cctggacacc aggcacaagt gaggtcacag 8640

agctcttaga tatgcagtcc tcatgagtga ggagactaaa gcgcatgcca tcaagacttc 8700
agtgtagaga aaacctccaa aaaagcctcc tcactacttc tggaatagct cagaggccga 8760
ggcggcctcg gcctctgcat aaataaaaaa aattagtcag ccatggggcg gagaatgggc 8820
ggaactgggc ggagttaggg gcgggatggg cggagttagg ggcgggacta tggttgctga 8880
ctaattgaga tgcatgcttt gcatacttct gcctgctggg gagcctgggg actttccaca 8940
cctggttgct gactaattga gatgcatgct ttgcatactt ctgcctgctg gggagcctgg 9000
ggactttcca caccctaact gacacacatt ccacagctgc attaatgaat cggccaacgc 9060

gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg 9120
cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta 9180
tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc 9240
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag 9300
catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 9360
caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc 9420
ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt 9480

aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc 9540
gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga 9600
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 9660
ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag aagaacagta 9720
tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 9780
tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg 9840
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 9900

tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc 9960
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 10020
tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt 10080
cgttcatcca tagttgcctg actcctgcaa accacgttgt gtctcaaaat ctctgatgtt 10140
acattgcaca agataaaaat atatcatcat gaacaataaa actgtctgct tacataaaca 10200
gtaatacaag gggtgttatg agccatattc aacgggaaac gtcttgctcg aggccgcgat 10260
taaattccaa catggatgct gatttatatg ggtataaatg ggctcgcgat aatgtcgggc 10320

aatcaggtgc gacaatctat cgattgtatg ggaagcccga tgcgccagag ttgtttctga 10380
aacatggcaa aggtagcgtt gccaatgatg ttacagatga gatggtcaga ctaaactggc 10440
tgacggaatt tatgcctctt ccgaccatca agcattttat ccgtactcct gatgatgcat 10500

ggttactcac cactgcgatc cccgggaaaa cagcattcca ggtattagaa gaatatcctg 10560

attcaggtga aaatattgtt gatgcgctgg cagtgttcct gcgccggttg cattcgattc 10620

ctgtttgtaa ttgtcctttt aacagcgatc gcgtatttcg tctcgctcag gcgcaatcac 10680

gaatgaataa cggtttggtt gatgcgagtg attttgatga cgagcgtaat ggctggcctg 10740

ttgaacaagt ctggaaagaa atgcataagc ttttgccatt ctcaccggat tcagtcgtca 10800

ctcatggtga tttctcactt gataacctta tttttgacga ggggaaatta ataggttgta 10860

ttgatgttgg acgagtcgga atcgcagacc gataccagga tcttgccatc ctatggaact 10920

gcctcggtga gttttctcct tcattacaga aacggctttt tcaaaaatat ggtattgata 10980

atcctgatat gaataaattg cagtttcatt tgatgctcga tgagtttttc taagggcggc 11040

ctgccaccat acccacgccg aaacaagcgc tcatgagccc gaagtggcga gcccgatctt 11100

ccccatcggt gatgtcggcg atataggcgc cagcaaccgc acctgtggcg ccggtgatga 11160

gggcgcgcca agtcgacgtc cggcagtc 11188

<210> 12

<211> 11187

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 12

ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac ctagttataa 60

tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg cgttacataa 120

cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata 180

atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggag 240

tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc aagtacgccc 300

cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta 360

tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac catggtcgag 420

gtgagcccca cgttctgctt cactctcccc atctcccccc cctccccacc cccaattttg 480

tatttattta ttttttaatt attttgtgca gcgatggggg cggggggggg gggggggcgc 540

gcgccaggcg gggcggggcg gggcgagggg cggggcgggg cgaggcggag aggtgcggcg 600

gcagccaatc agagcggcgc gctccgaaag tttcctttta tggcgaggcg gcggcggcgg 660

cggccctata aaaagcgaag cgcgcggcgg gcgggagtcg ctgcgacgct gccttcgccc 720

cgtgccccgc tccgccgccg cctcgcgccg cccgccccgg ctctgactga ccgcgttact 780

cccacaggtg agcgggcggg acggcccttc tcctccgggc tgtaattagc gcttggttta 840

atgacggctt gtttcttttc tgtggctgcg tgaaagcctt gaggggctcc gggagctaga 900

gcctctgcta accatgttca tgccttcttc tttttcctac agctcctggg caacgtgctg 960

gttattgtgc tgtctcatca ttttggcaaa gaattcctcg aagatccgaa gggaaagtct 1020

tccacgactg tgggatccgt tcgaagatat caccggttga gccaccatgg aattcagcag 1080

ccccagcaga gaggaatgcc ccaagcctct gagccgggtg tcaatcatgg ccggatctct 1140

gacaggactg ctgctgcttc aggccgtgtc ttgggcttct ggcgctagac cttgcatccc 1200

caagagcttc ggctacagca gcgtcgtgtg cgtgtgcaat gccacctact gcgacagctt 1260

cgaccctcct acctttcctg ctctgggcac cttcagcaga tacgagagca ccagatccgg 1320

cagacggatg gaactgagca tgggacccat ccaggccaat cacacaggca ctggcctgct 1380

gctgacactg cagcctgagc agaaattcca gaaagtgaaa ggcttcggcg gagccatgac 1440

agatgccgcc gctctgaata tcctggctct gtctccacca gctcagaacc tgctgctcaa 1500

gagctacttc agcgaggaag gcatcggcta caacatcatc agagtgccca tggccagctg 1560

cgacttcagc atcaggacct acacctacgc cgacacaccc gacgatttcc agctgcacaa 1620

cttcagcctg cctgaagagg acaccaagct gaagatccct ctgatccaca gagccctgca 1680

gctggcacaa agacccgtgt cactgctggc ctctccatgg acatctccca cctggctgaa 1740

aacaaatggc gccgtgaatg gcaagggcag cctgaaaggc caacctggcg acatctacca 1800

ccagacctgg gccagatact tcgtgaagtt cctggacgcc tatgccgagc acaagctgca 1860

gttttgggcc gtgacagccg agaacgaacc ttctgctgga ctgctgagcg gctacccctt 1920

tcagtgcctg ggctttacac ccgagcacca gcgggacttt atcgcccgtg atctgggacc 1980

cacactggcc aatagcaccc accataatgt gcggctgctg atgctggacg accagagact 2040

gcttctgccc cactgggcta aagtggtgct gacagatcct gaggccgcca aatacgtgca 2100

cggaatcgcc gtgcactggt atctggactt tctggcccct gccaaggcca cactgggaga 2160

gacacacaga ctgttcccca acaccatgct gttcgccagc gaagcctgtg tgggcagcaa 2220

gttttgggaa cagagcgtgc ggctcggcag ctgggataga ggcatgcagt acagccacag 2280

catcatcacc aacctgctgt accacgtcgt cggctggacc gactggaatc tggccctgaa 2340

tcctgaaggc ggccctaact gggtccgaaa cttcgtggac agccccatca tcgtggacat 2400

caccaaggac accttctaca agcagcccat gttctaccac ctgggacact tcagcaagtt 2460

catccccgag ggctctcagc gcgttggact ggtggcttcc cagaagaacg atctggacgc 2520

cgtggctctg atgcaccctg atggatctgc tgtggtggtg gtcctgaacc gcagcagcaa 2580

agatgtgccc ctgaccatca aggatcccgc cgtgggattc ctggaaacaa tcagccctgg 2640

ctactccatc cacacctacc tgtggcgtag acagtgacaa ttgttaatta agtttaaacc 2700
ctcgaggccg caagcttatc gataatcaac ctctggatta caaaatttgt gaaagattga 2760
ctggtattct taactatgtt gctcctttta cgctatgtgg atacgctgct ttaatgcctt 2820

tgtatcatgc tattgcttcc cgtatggctt tcattttctc ctccttgtat aaatcctggt 2880
tgctgtctct ttatgaggag ttgtggcccg ttgtcaggca acgtggcgtg gtgtgcactg 2940
tgtttgctga cgcaaccccc actggttggg gcattgccac cacctgtcag ctcctttccg 3000
ggactttcgc tttccccctc cctattgcca cggcggaact catcgccgcc tgccttgccc 3060
gctgctggac aggggctcgg ctgttgggca ctgacaattc cgtggtgttg tcggggaaat 3120
catcgtcctt tccttggctg ctcgcctgtg ttgccacctg gattctgcgc gggacgtcct 3180
tctgctacgt cccttcggcc ctcaatccag cggaccttcc ttcccgcggc ctgctgccgg 3240

ctctgcggcc tcttccgcgt cttcgccttc gccctcagac gagtcggatc tccctttggg 3300
ccgcctcccc gcatcgatac cgtcgactag agctcgctga tcagcctcga ctgtgccttc 3360
tagttgccag ccatctgttg tttgcccctc ccccgtgcct tccttgaccc tggaaggtgc 3420
cactcccact gtcctttcct aataaaatga ggaaattgca tcgcattgtc tgagtaggtg 3480
tcattctatt ctgggggtg gggtggggca ggacagcaag ggggaggatt gggaagacaa 3540
tagcaggcat gctggggaga gatccacgat aacaaacagc ttttttgggg tgaacatatt 3600
gactgaattc cctgcaggtt ggccactccc tctctgcgcg ctcgctcgct cactgaggcc 3660

gcccgggcaa agcccgggcg tcgggcgacc tttggtcgcc cggcctcagt gagcgagcga 3720
gcgcgcagag agggagtggc caactccatc actaggggtt cctgcggccg ctcgtacggt 3780
ctcgaggaat tcctgcagga taacttgcca acctcattct aaaatgtata tagaagccca 3840
aaagacaata acaaaaatat tcttgtagaa caaaatggga aagaatgttc cactaaatat 3900
caagatttag agcaaagcat gagatgtgtg gggatagaca gtgaggctga taaaatagag 3960
tagagctcag aaacagaccc attgatatat gtaagtgacc tatgaaaaaa atatggcatt 4020
ttacaatggg aaaatgatgg tctttttctt ttttagaaaa acagggaaat atatttatat 4080

gtaaaaaata aaagggaacc catatgtcat accatacaca caaaaaaatt ccagtgaatt 4140
ataagtctaa atggagaagg caaaacttta aatcttttag aaaataatat agaagcatgc 4200
agaccagcct ggccaacatg atgaaaccct ctctactaat aataaaatca gtagaactac 4260
tcaggactac tttgagtggg aagtcctttt ctatgaagac ttctttggcc aaaattaggc 4320
tctaaatgca aggagatagt gcatcatgcc tggctgcact tactgataaa tgatgttatc 4380
accatcttta accaaatgca caggaacaag ttatggtact gatgtgctgg attgagaagg 4440

agctctactt ccttgacagg acacatttgt atcaacttaa aaaagcagat ttttgccagc 4500

agaactattc attcagaggt aggaaactta gaatagatga tgtcactgat tagcatggct 4560
tccccatctc cacagctgct tcccacccag gttgcccaca gttgagtttg tccagtgctc 4620
agggctgccc actctcagta agaagcccca caccagcccc tctccaaata tgttggctgt 4680
tccttccatt aaagtgaccc cactttagag cagcaagtgg atttctgttt cttacagttc 4740
aggaaggagg agtcagctgt gagaacctgg agcctgagat gcttctaagt cccactgcta 4800
ctggggtcag ggaagccaga ctccagcatc agcagtcagg agcactaagc ccttgccaac 4860
atcctgtttc tcagagaaac tgcttccatt ataatggttg tccttttta agctatcaag 4920

ccaaacaacc agtgtctacc attattctca tcacctgaag ccaagggttc tagcaaaagt 4980
caagctgtct tgtaatggtt gatgtgcctc cagcttctgt cttcagtcac tccactctta 5040
gcctgctctg aatcaactct gaccacagtt ccctggagcc cctgccacct gctgcccctg 5100
ccaccttctc catctgcagt gctgtgcagc cttctgcact cttgcagagc taataggtgg 5160
agacttgaag gaagaggagg aaagtttctc ataatagcct tgctgcaagc tcaaatggga 5220
ggtgggcact gtgcccagga gccttggagc aaaggctgtg cccaacctct gactgcatcc 5280
aggtttggtc ttgacagaga taagaagccc tggcttttgg agccaaaatc taggtcagac 5340

ttaggcagga ttctcaaagt ttatcagcag aacatgaggc agaagaccct ttctgctcca 5400
gcttcttcag gctcaacctt catcagaata gatagaaaga gaggctgtga gggttcttaa 5460
aacagaagca aatctgactc agagaataaa caacctccta gtaaactaca gcttagacag 5520
agcatctggt ggtgagtgtg ctcagtgtcc tactcaactg tctggtatca gccctcatga 5580
ggacttctct tctttccctc atagacctcc atctctgttt tccttagcct gcagaaatct 5640
ggatggctat tcacagaatg cctgtgcttt cagagttgca ttttttctct ggtattctgg 5700
ttcaagcatt tgaaggtagg aaaggttctc caagtgcaag aaagccagcc ctgagcctca 5760

actgcctggc tagtgtggtc agtaggatgc aaaggctgtt gaatgccaca aggccaaact 5820
ttaacctgtg taccacaagc ctagcagcag aggcagctct gctcactgga actctctgtc 5880
ttctttctcc tgagcctttt cttttcctga gttttctagc tctcctcaac cttacctctg 5940
ccctacccag gacaaaccca agagccactg tttctgtgat gtcctctcca gccctaatta 6000
ggcatcatga cttcagcctg accttccatg ctcagaagca gtgctaatcc acttcagatg 6060
agctgctcta tgcaacacag gcagagccta caaacctttg caccagagcc ctccacatat 6120
cagtgtttgt tcatactcac ttcaacagca aatgtgactg ctgagattaa gattttacac 6180

aagatggtct gtaatttcac agttagtttt atcccattag gtatgaaaga attagcataa 6240
ttccccttaa acatgaatga atcttagatt ttttaataaa tagttttgga agtaaagaca 6300
gagacatcag gagcacaagg aatagcctga gaggacaaac agaacaagaa agagtctgga 6360
aatacacagg atgttcttgg cctcctcaaa gcaagtgcaa gcagatagta ccagcagccc 6420
caggctatca gagcccagtg aagagaagta ccatgaaagc cacagctcta accaccctgt 6480
tccagagtga cagacagtcc ccaagacaag ccagcctgag ccagagagag aactgcaaga 6540
gaaagtttct aatttaggtt ctgttagatt cagacaagtg caggtcatcc tctctccaca 6600

gctactcacc tctccagcct aacaaagcct gcagtccaca ctccaaccct ggtgtctcac 6660
ctcctagcct ctcccaacat cctgctctct gaccatcttc tgcatctctc atctcaccat 6720
ctcccactgt ctacagccta ctcttgcaac taccatctca ttttctgaca tcctgtctac 6780
atcttctgcc atactctgcc atctaccata ccacctctta ccatctacca caccatcttt 6840
tatctccatc cctctcagaa gcctccaagc tgaatcctgc tttatgtgtt catctcagcc 6900
cctgcatgga aagctgaccc cagaggcaga actattccca gagagcttgg ccaagaaaaa 6960
caaaactacc agcctggcca ggctcaggag tagtaagctg cagtgtctgt tgtgttctag 7020

cttcaacagc tgcaggagtt ccactctcaa atgctccaca tttctcacat cctcctgatt 7080
ctggtcacta cccatcttca aagaacagaa tatctcacat cagcatactg tgaaggacta 7140
gtcatgggtg cagctgctca gagctgcaaa gtcattctgg atggtggaga gcttacaaac 7200
atttcatgat gctccccccg ctctgatggc tggagcccaa tccctacaca gactcctgct 7260
gtatgtgttt tcctttcact ctgagccaca gccagagggc aggcattcag tctcctcttc 7320
aggctggggc tggggcactg agaactcacc caacaccttg ctctcactcc ttctgcaaaa 7380
caagaaagag ctttgtgctg cagtagccat gaagaatgaa aggaaggctt taactaaaaa 7440

atgtcagaga ttattttcaa ccccttactg tggatcacca gcaaggagga aacacaacac 7500
agagacattt tttcccctca aattatcaaa agaatcactg catttgttaa agagagcaac 7560
tgaatcagga agcagagttt tgaacatatc agaagttagg aatctgcatc agagacaaat 7620
gcagtcatgg ttgtttgctg cataccagcc ctaatcatta gaagcctcat ggacttcaaa 7680
catcattccc tctgacaaga tgctctagcc taactccatg agataaaata aatctgcctt 7740
tcagagccaa agaagagtcc accagcttct tctcagtgtg aacaagagct ccagtcaggt 7800
tagtcagtcc agtgcagtag aggagaccag tctgcatcct ctaattttca aaggcaagaa 7860

gatttgttta ccctggacac caggcacaag tgaggtcaca gagctcttag atatgcagtc 7920
ctcatgagtg aggagactaa agcgcatgcc atcaagactt cagtgtagag aaaacctcca 7980
aaaaagcctc ctcactactt ctggaatagc tcagaggccg aggcggcctc ggcctctgca 8040

taaataaaaa aaattagtca gccatggggc ggagaatggg cggaactggg cggagttagg 8100
ggcgggatgg gcggagttag gggcgggact atggttgctg actaattgag atgcatgctt 8160
tgcatacttc tgcctgctgg ggagcctggg gactttccac acctggttgc tgactaattg 8220
agatgcatgc tttgcatact tctgcctgct ggggagcctg gggactttcc acaccctaac 8280

tgacacacat tccacagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc 8340
gtattgggcg ctcttccgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc 8400
ggcgagcggt atcagctcac tcaaaggcgg taatacggtt atccacagaa tcaggggata 8460
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 8520
cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct 8580
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccctggaa 8640
gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc 8700

tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt 8760
aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 8820
ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta tcgccactgg 8880
cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct 8940
tgaagtggtg gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc 9000
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 9060
ctggtagcgg tggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 9120

aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 9180
aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa 9240
aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 9300
gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct 9360
gactcctgca aaccacgttg tgtctcaaaa tctctgatgt tacattgcac aagataaaaa 9420
tatatcatca tgaacaataa aactgtctgc ttacataaac agtaatacaa ggggtgttat 9480
gagccatatt caacgggaaa cgtcttgctc gaggccgcga ttaaattcca acatggatgc 9540

tgatttatat gggtataaat gggctcgcga taatgtcggg caatcaggtg cgacaatcta 9600
tcgattgtat gggaagcccg atgcgccaga gttgtttctg aaacatggca aaggtagcgt 9660
tgccaatgat gttacagatg agatggtcag actaaactgg ctgacggaat ttatgcctct 9720
tccgaccatc aagcatttta tccgtactcc tgatgatgca tggttactca ccactgcgat 9780
ccccgggaaa acagcattcc aggtattaga agaatatcct gattcaggtg aaaatattgt 9840
tgatgcgctg gcagtgttcc tgcgccggtt gcattcgatt cctgtttgta attgtccttt 9900

taacagcgat cgcgtatttc gtctcgctca ggcgcaatca cgaatgaata acggtttggt 9960

tgatgcgagt gattttgatg acgagcgtaa tggctggcct gttgaacaag tctggaaaga 10020
aatgcataag cttttgccat tctcaccgga ttcagtcgtc actcatggtg atttctcact 10080
tgataacctt atttttgacg aggggaaatt aataggttgt attgatgttg gacgagtcgg 10140
aatcgcagac cgataccagg atcttgccat cctatggaac tgcctcggtg agttttctcc 10200
ttcattacag aaacggcttt ttcaaaaata tggtattgat aatcctgata tgaataaatt 10260
gcagtttcat ttgatgctcg atgagttttt ctaagggcgg cctgccacca tacccacgcc 10320
gaaacaagcg ctcatgagcc cgaagtggcg agcccgatct tccccatcgg tgatgtcggc 10380

gatataggcg ccagcaaccg cacctgtggc gccggtgatg agggcgcgcc aagtcgacgt 10440
ccggcagtct tggccactcc ctctctgcgc gctcgctcgc tcactgaggc cgggcgacca 10500
aaggtcgccc gacgcccggg ctttgcccgg gcggcctcag tgagcgagcg agcgcgcaga 10560
gagggagtgg ccaactccat cactaggggt tcctgctagc tctgggtatt taagcccgag 10620
tgagcacgca gggtctccat tttgaagcgg gaggttacgc gttcgtcgac tactagtggg 10680
taccagagcg tggtgactga gatgttttct aggaaacaca aaagatacaa aaaagaacac 10740
gtggaaggat agccaaaaag gggggctgcc cccatttcct gcaccccgct gcgatggctg 10800

gcaccatttg gaagacttcg agatacactg ttgagcgcag taagacaaca gtgtatctcg 10860
aagtcttcca gatggggcca gccggtccac tctgtatcca ggccagttct gcaaggcgtt 10920
cgaggaccac ccccctcccc tcgccaccag ggtggtctca tacagaactt ataagattcc 10980
caaatccaaa gacatttcac gtttatggtg atttcccaga acacatagcg acatgcaaat 11040
attgcagggc gccactcccc tgtccctcac agccatcttc ctgccagggc gcacgcgcgc 11100
tgggtgttcc cgcctagtga cactgggccc gcgattcctt ggagcgggtt gatgacgtca 11160
gcgtttccca tggtgaatcc ctaggtt 11187

<210> 13
<211> 10960
<212> DNA
<213> Artificial sequence
<220><223> Synthetic polynucleotide
<400> 13
ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120
gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240
tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300
cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc 360

cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga cccccgccca 420
ttgacgtcaa taatgacgta tgttcccata gtaacgccaa tagggacttt ccattgacgt 480
caatgggtgg actatttacg gtaaactgcc cacttggcag tacatcaagt gtatcatatg 540
ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgcccag 600
tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt 660
accatggtcg aggtgagccc cacgttctgc ttcactctcc ccatctcccc cccctcccca 720
cccccaattt tgtatttatt tatttttaa ttattttgtg cagcgatggg ggcgggggg 780

gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg 840
agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg 900
cggcggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg 960
ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cgcccgcccc ggctctgact 1020
gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctcctccgg gctgtaatta 1080
gcgcttggtt taatgacggc ttgtcctggt ggcgagggga ggggggtggt cctcgaacgc 1140
cttgcagaac tggcctggat acagagtgga ccggctggcc ccatctggaa gacttcgaga 1200

tacactgttg tcttactgcg ctcaacagtg tatctcgaag tcttccaaat ggtgccagcc 1260
atcgcagcgg ggtgcaggaa atgggggcag cccccctttt tggctatcct tccacgtgtt 1320
ctttttgta tcttttgtgt ttcctagaaa acatctcagt caccaccttt ctgtggctgc 1380
gtgaaagcct tgaggggctc cgggagctag agcctctgct aaccatgttc atgccttctt 1440
ctttttccta cagctcctgg gcaacgtgct ggttattgtg ctgtctcatc attttggcaa 1500
agaattcctc gaagatccga agggaaagtc ttccacgact gtgggatccg ttcgaagata 1560
tcaccggttg agccaccatg gaattcagca gccccagcag agaggaatgc cccaagcctc 1620

tgagccgggt gtcaatcatg gccggatctc tgacaggact gctgctgctt caggccgtgt 1680
cttgggcttc tggcgctaga ccttgcatcc ccaagagctt cggctacagc agcgtcgtgt 1740
gcgtgtgcaa tgccacctac tgcgacagct tcgaccctcc tacctttcct gctctgggca 1800
ccttcagcag atacgagagc accagatccg gcagacggat ggaactgagc atgggaccca 1860
tccaggccaa tcacacaggc actggcctgc tgctgacact gcagcctgag cagaaattcc 1920
agaaagtgaa aggcttcggc ggagccatga cagatgccgc cgctctgaat atcctggctc 1980

tgtctccacc agctcagaac ctgctgctca agagctactt cagcgaggaa ggcatcggct 2040

acaacatcat cagagtgccc atggccagct gcgacttcag catcaggacc tacacctacg 2100
ccgacacacc cgacgatttc cagctgcaca acttcagcct gcctgaagag gacaccaagc 2160
tgaagatccc tctgatccac agagccctgc agctggcaca aagacccgtg tcactgctgg 2220
cctctccatg gacatctccc acctggctga aaacaaatgg cgccgtgaat ggcaagggca 2280
gcctgaaagg ccaacctggc gacatctacc accagacctg ggccagatac ttcgtgaagt 2340
tcctggacgc ctatgccgag cacaagctgc agttttgggc cgtgacagcc gagaacgaac 2400
cttctgctgg actgctgagc ggctacccct ttcagtgcct gggctttaca cccgagcacc 2460

agcgggactt tatcgcccgt gatctgggac ccacactggc caatagcacc caccataatg 2520
tgcggctgct gatgctggac gaccagagac tgcttctgcc ccactgggct aaagtggtgc 2580
tgacagatcc tgaggccgcc aaatacgtgc acggaatcgc cgtgcactgg tatctggact 2640
ttctggcccc tgccaaggcc acactgggag agacacacag actgttcccc aacaccatgc 2700
tgttcgccag cgaagcctgt gtgggcagca agttttggga acagagcgtg cggctcggca 2760
gctgggatag aggcatgcag tacagccaca gcatcatcac caacctgctg taccacgtcg 2820
tcggctggac cgactggaat ctggccctga atcctgaagg cggccctaac tgggtccgaa 2880

acttcgtgga cagccccatc atcgtggaca tcaccaagga caccttctac aagcagccca 2940
tgttctacca cctgggacac ttcagcaagt tcatccccga gggctctcag cgcgttggac 3000
tggtggcttc ccagaagaac gatctggacg ccgtggctct gatgcaccct gatggatctg 3060
ctgtggtggt ggtcctgaac cgcagcagca aagatgtgcc cctgaccatc aaggatcccg 3120
ccgtgggatt cctggaaaca atcagccctg gctactccat ccacacctac ctgtggcgta 3180
gacagtgaca attgttaatt aagtttaaac cctcgaggcc gcaagcttat cgataatcaa 3240
cctctggatt acaaaatttg tgaaagattg actggtattc ttaactatgt tgctcctttt 3300

acgctatgtg gatacgctgc tttaatgcct ttgtatcatg ctattgcttc ccgtatggct 3360
ttcattttct cctccttgta taaatcctgg ttgctgtctc tttatgagga gttgtggccc 3420
gttgtcaggc aacgtggcgt ggtgtgcact gtgtttgctg acgcaacccc cactggttgg 3480
ggcattgcca ccacctgtca gctcctttcc gggactttcg ctttccccct ccctattgcc 3540
acggcggaac tcatcgccgc ctgccttgcc cgctgctgga caggggctcg gctgttgggc 3600
actgacaatt ccgtggtgtt gtcggggaaa tcatcgtcct ttccttggct gctcgcctgt 3660
gttgccacct ggattctgcg cgggacgtcc ttctgctacg tcccttcggc cctcaatcca 3720

gcggaccttc cttcccgcgg cctgctgccg gctctgcggc ctcttccgcg tcttcgcctt 3780
cgccctcaga cgagtcggat ctccctttgg gccgcctccc cgcatcgata ccgtcgacta 3840
gagctcgctg atcagcctcg actgtgcctt ctagttgcca gccatctgtt gtttgcccct 3900
cccccgtgcc ttccttgacc ctggaaggtg ccactcccac tgtcctttcc taataaaatg 3960
aggaaattgc atcgcattgt ctgagtaggt gtcattctat tctggggggt ggggtggggc 4020
aggacagcaa gggggaggat tgggaagaca atagcaggca tgctggggag agatccacga 4080
taacaaacag ctttttggg gtgaacatat tgactgaatt ccctgcaggt tggccactcc 4140

ctctctgcgc gctcgctcgc tcactgaggc cgcccgggca aagcccgggc gtcgggcgac 4200
ctttggtcgc ccggcctcag tgagcgagcg agcgcgcaga gagggagtgg ccaactccat 4260
cactaggggt tcctgcggcc gctcgtacgg tctcgaggaa ttcctgcagg ataacttgcc 4320
aacctcattc taaaatgtat atagaagccc aaaagacaat aacaaaaata ttcttgtaga 4380
acaaaatggg aaagaatgtt ccactaaata tcaagattta gagcaaagca tgagatgtgt 4440
ggggatagac agtgaggctg ataaaataga gtagagctca gaaacagacc cattgatata 4500
tgtaagtgac ctatgaaaaa aatatggcat tttacaatgg gaaaatgatg gtctttttct 4560

tttttagaaa aacagggaaa tatatttata tgtaaaaaat aaaagggaac ccatatgtca 4620
taccatacac acaaaaaaat tccagtgaat tataagtcta aatggagaag gcaaaacttt 4680
aaatctttta gaaaataata tagaagcatg cagaccagcc tggccaacat gatgaaaccc 4740
tctctactaa taataaaatc agtagaacta ctcaggacta ctttgagtgg gaagtccttt 4800
tctatgaaga cttctttggc caaaattagg ctctaaatgc aaggagatag tgcatcatgc 4860
ctggctgcac ttactgataa atgatgttat caccatcttt aaccaaatgc acaggaacaa 4920
gttatggtac tgatgtgctg gattgagaag gagctctact tccttgacag gacacatttg 4980

tatcaactta aaaaagcaga tttttgccag cagaactatt cattcagagg taggaaactt 5040
agaatagatg atgtcactga ttagcatggc ttccccatct ccacagctgc ttcccaccca 5100
ggttgcccac agttgagttt gtccagtgct cagggctgcc cactctcagt aagaagcccc 5160
acaccagccc ctctccaaat atgttggctg ttccttccat taaagtgacc ccactttaga 5220
gcagcaagtg gatttctgtt tcttacagtt caggaaggag gagtcagctg tgagaacctg 5280
gagcctgaga tgcttctaag tcccactgct actggggtca gggaagccag actccagcat 5340
cagcagtcag gagcactaag cccttgccaa catcctgttt ctcagagaaa ctgcttccat 5400

tataatggtt gtcctttttt aagctatcaa gccaaacaac cagtgtctac cattattctc 5460
atcacctgaa gccaagggtt ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct 5520
ccagcttctg tcttcagtca ctccactctt agcctgctct gaatcaactc tgaccacagt 5580

tccctggagc ccctgccacc tgctgcccct gccaccttct ccatctgcag tgctgtgcag 5640
ccttctgcac tcttgcagag ctaataggtg gagacttgaa ggaagaggag gaaagtttct 5700
cataatagcc ttgctgcaag ctcaaatggg aggtgggcac tgtgcccagg agccttggag 5760
caaaggctgt gcccaacctc tgactgcatc caggtttggt cttgacagag ataagaagcc 5820

ctggcttttg gagccaaaat ctaggtcaga cttaggcagg attctcaaag tttatcagca 5880
gaacatgagg cagaagaccc tttctgctcc agcttcttca ggctcaacct tcatcagaat 5940
agatagaaag agaggctgtg agggttctta aaacagaagc aaatctgact cagagaataa 6000
acaacctcct agtaaactac agcttagaca gagcatctgg tggtgagtgt gctcagtgtc 6060
ctactcaact gtctggtatc agccctcatg aggacttctc ttctttccct catagacctc 6120
catctctgtt ttccttagcc tgcagaaatc tggatggcta ttcacagaat gcctgtgctt 6180
tcagagttgc atttttctc tggtattctg gttcaagcat ttgaaggtag gaaaggttct 6240

ccaagtgcaa gaaagccagc cctgagcctc aactgcctgg ctagtgtggt cagtaggatg 6300
caaaggctgt tgaatgccac aaggccaaac tttaacctgt gtaccacaag cctagcagca 6360
gaggcagctc tgctcactgg aactctctgt cttctttctc ctgagccttt tcttttcctg 6420
agttttctag ctctcctcaa ccttacctct gccctaccca ggacaaaccc aagagccact 6480
gtttctgtga tgtcctctcc agccctaatt aggcatcatg acttcagcct gaccttccat 6540
gctcagaagc agtgctaatc cacttcagat gagctgctct atgcaacaca ggcagagcct 6600
acaaaccttt gcaccagagc cctccacata tcagtgtttg ttcatactca cttcaacagc 6660

aaatgtgact gctgagatta agattttaca caagatggtc tgtaatttca cagttagttt 6720
tatcccatta ggtatgaaag aattagcata attcccctta aacatgaatg aatcttagat 6780
tttttaataa atagttttgg aagtaaagac agagacatca ggagcacaag gaatagcctg 6840
agaggacaaa cagaacaaga aagagtctgg aaatacacag gatgttcttg gcctcctcaa 6900
agcaagtgca agcagatagt accagcagcc ccaggctatc agagcccagt gaagagaagt 6960
accatgaaag ccacagctct aaccaccctg ttccagagtg acagacagtc cccaagacaa 7020
gccagcctga gccagagaga gaactgcaag agaaagtttc taatttaggt tctgttagat 7080

tcagacaagt gcaggtcatc ctctctccac agctactcac ctctccagcc taacaaagcc 7140
tgcagtccac actccaaccc tggtgtctca cctcctagcc tctcccaaca tcctgctctc 7200
tgaccatctt ctgcatctct catctcacca tctcccactg tctacagcct actcttgcaa 7260
ctaccatctc attttctgac atcctgtcta catcttctgc catactctgc catctaccat 7320
accacctctt accatctacc acaccatctt ttatctccat ccctctcaga agcctccaag 7380
ctgaatcctg ctttatgtgt tcatctcagc ccctgcatgg aaagctgacc ccagaggcag 7440

aactattccc agagagcttg gccaagaaaa acaaaactac cagcctggcc aggctcagga 7500

gtagtaagct gcagtgtctg ttgtgttcta gcttcaacag ctgcaggagt tccactctca 7560
aatgctccac atttctcaca tcctcctgat tctggtcact acccatcttc aaagaacaga 7620
atatctcaca tcagcatact gtgaaggact agtcatgggt gcagctgctc agagctgcaa 7680
agtcattctg gatggtggag agcttacaaa catttcatga tgctcccccc gctctgatgg 7740
ctggagccca atccctacac agactcctgc tgtatgtgtt ttcctttcac tctgagccac 7800
agccagaggg caggcattca gtctcctctt caggctgggg ctggggcact gagaactcac 7860
ccaacacctt gctctcactc cttctgcaaa acaagaaaga gctttgtgct gcagtagcca 7920

tgaagaatga aaggaaggct ttaactaaaa aatgtcagag attattttca acccccttact 7980
gtggatcacc agcaaggagg aaacacaaca cagagacatt ttttcccctc aaattatcaa 8040
aagaatcact gcatttgtta aagagagcaa ctgaatcagg aagcagagtt ttgaacatat 8100
cagaagttag gaatctgcat cagagacaaa tgcagtcatg gttgtttgct gcataccagc 8160
cctaatcatt agaagcctca tggacttcaa acatcattcc ctctgacaag atgctctagc 8220
ctaactccat gagataaaat aaatctgcct ttcagagcca aagaagagtc caccagcttc 8280
ttctcagtgt gaacaagagc tccagtcagg ttagtcagtc cagtgcagta gaggagacca 8340

gtctgcatcc tctaattttc aaaggcaaga agatttgttt accctggaca ccaggcacaa 8400
gtgaggtcac agagctctta gatatgcagt cctcatgagt gaggagacta aagcgcatgc 8460
catcaagact tcagtgtaga gaaaacctcc aaaaaagcct cctcactact tctggaatag 8520
ctcagaggcc gaggcggcct cggcctctgc ataaataaaa aaaattagtc agccatgggg 8580
cggagaatgg gcggaactgg gcggagttag gggcgggatg ggcggagtta ggggcgggac 8640
tatggttgct gactaattga gatgcatgct ttgcatactt ctgcctgctg gggagcctgg 8700
ggactttcca cacctggttg ctgactaatt gagatgcatg ctttgcatac ttctgcctgc 8760

tggggagcct ggggactttc cacaccctaa ctgacacaca ttccacagct gcattaatga 8820
atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc 8880
actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg 8940
gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc 9000
cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc 9060
ccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga 9120
ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttccgacc 9180

ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat 9240

agctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg 9300

cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc 9360

aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga 9420

gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact 9480

agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt 9540

ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag 9600

cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg 9660

tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa 9720

aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata 9780

tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg 9840

atctgtctat ttcgttcatc catagttgcc tgactcctgc aaaccacgtt gtgtctcaaa 9900

atctctgatg ttacattgca caagataaaa atatatcatc atgaacaata aaactgtctg 9960

cttacataaa cagtaataca aggggtgtta tgagccatat tcaacgggaa acgtcttgct 10020

cgaggccgcg attaaattcc aacatggatg ctgatttata tgggtataaa tgggctcgcg 10080

ataatgtcgg gcaatcaggt gcgacaatct atcgattgta tgggaagccc gatgcgccag 10140

agttgtttct gaaacatggc aaaggtagcg ttgccaatga tgttacagat gagatggtca 10200

gactaaactg gctgacggaa tttatgcctc ttccgaccat caagcatttt atccgtactc 10260

ctgatgatgc atggttactc accactgcga tccccgggaa aacagcattc caggtattag 10320

aagaatatcc tgattcaggt gaaaatattg ttgatgcgct ggcagtgttc ctgcgccggt 10380

tgcattcgat tcctgtttgt aattgtcctt ttaacagcga tcgcgtattt cgtctcgctc 10440

aggcgcaatc acgaatgaat aacggtttgg ttgatgcgag tgattttgat gacgagcgta 10500

atggctggcc tgttgaacaa gtctggaaag aaatgcataa gcttttgcca ttctcaccgg 10560

attcagtcgt cactcatggt gatttctcac ttgataacct tatttttgac gaggggaaat 10620

taataggttg tattgatgtt ggacgagtcg gaatcgcaga ccgataccag gatcttgcca 10680

tcctatggaa ctgcctcggt gagttttctc cttcattaca gaaacggctt tttcaaaaat 10740

atggtattga taatcctgat atgaataaat tgcagtttca tttgatgctc gatgagtttt 10800

tctaagggcg gcctgccacc atacccacgc cgaaacaagc gctcatgagc ccgaagtggc 10860

gagcccgatc ttccccatcg gtgatgtcgg cgatataggc gccagcaacc gcacctgtgg 10920

cgccggtgat gagggcgcgc caagtcgacg tccggcagtc 10960

<210> 14

<211> 536

<212> PRT

<213> Homo sapiens

<400> 14

Met Glu Phe Ser Ser Pro Ser Arg Glu Glu Cys Pro Lys Pro Leu Ser

1 5 10 15

Arg Val Ser Ile Met Ala Gly Ser Leu Thr Gly Leu Leu Leu Leu Gln

20 25 30

Ala Val Ser Trp Ala Ser Gly Ala Arg Pro Cys Ile Pro Lys Ser Phe

35 40 45

Gly Tyr Ser Ser Val Val Cys Val Cys Asn Ala Thr Tyr Cys Asp Ser

50 55 60

Phe Asp Pro Pro Thr Phe Pro Ala Leu Gly Thr Phe Ser Arg Tyr Glu

65 70 75 80

Ser Thr Arg Ser Gly Arg Arg Met Glu Leu Ser Met Gly Pro Ile Gln

85 90 95

Ala Asn His Thr Gly Thr Gly Leu Leu Leu Thr Leu Gln Pro Glu Gln

100 105 110

Lys Phe Gln Lys Val Lys Gly Phe Gly Gly Ala Met Thr Asp Ala Ala

115 120 125

Ala Leu Asn Ile Leu Ala Leu Ser Pro Pro Ala Gln Asn Leu Leu Leu

130 135 140

Lys Ser Tyr Phe Ser Glu Glu Gly Ile Gly Tyr Asn Ile Ile Arg Val

145 150 155 160

Pro Met Ala Ser Cys Asp Phe Ser Ile Arg Thr Tyr Thr Tyr Ala Asp

165 170 175

Thr Pro Asp Asp Phe Gln Leu His Asn Phe Ser Leu Pro Glu Glu Asp

180 185 190

Thr Lys Leu Lys Ile Pro Leu Ile His Arg Ala Leu Gln Leu Ala Gln

195 200 205

Arg Pro Val Ser Leu Leu Ala Ser Pro Trp Thr Ser Pro Thr Trp Leu

210 215 220

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

225 230 235 240

Gly Asp Ile Tyr His Gln Thr Trp Ala Arg Tyr Phe Val Lys Phe Leu

245 250 255

Asp Ala Tyr Ala Glu His Lys Leu Gln Phe Trp Ala Val Thr Ala Glu

260 265 270

Asn Glu Pro Ser Ala Gly Leu Leu Ser Gly Tyr Pro Phe Gln Cys Leu

275 280 285

Gly Phe Thr Pro Glu His Gln Arg Asp Phe Ile Ala Arg Asp Leu Gly

290 295 300

Pro Thr Leu Ala Asn Ser Thr His His Asn Val Arg Leu Leu Met Leu

305 310 315 320

Asp Asp Gln Arg Leu Leu Leu Pro His Trp Ala Lys Val Val Leu Thr

325 330 335

Asp Pro Glu Ala Ala Lys Tyr Val His Gly Ile Ala Val His Trp Tyr

340 345 350

Leu Asp Phe Leu Ala Pro Ala Lys Ala Thr Leu Gly Glu Thr His Arg

355 360 365

Leu Phe Pro Asn Thr Met Leu Phe Ala Ser Glu Ala Cys Val Gly Ser

370 375 380

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

385 390 395 400

Gln Tyr Ser His Ser Ile Ile Thr Asn Leu Leu Tyr His Val Val Gly

405 410 415

Trp Thr Asp Trp Asn Leu Ala Leu Asn Pro Glu Gly Gly Pro Asn Trp

420 425 430

Val Arg Asn Phe Val Asp Ser Pro Ile Ile Val Asp Ile Thr Lys Asp

435 440 445

Thr Phe Tyr Lys Gln Pro Met Phe Tyr His Leu Gly His Phe Ser Lys

450 455 460

Phe Ile Pro Glu Gly Ser Gln Arg Val Gly Leu Val Ala Ser Gln Lys

465 470 475 480

Asn Asp Leu Asp Ala Val Ala Leu Met His Pro Asp Gly Ser Ala Val

485 490 495

Val Val Val Leu Asn Arg Ser Ser Lys Asp Val Pro Leu Thr Ile Lys

500 505 510

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

515 520 525

His Thr Tyr Leu Trp Arg Arg Gln

530 535

<210> 15

<211> 1608

<212> DNA

<213> Homo sapiens

<400> 15

atggaattca gcagccccag cagagaggaa tgccccaagc ctctgagccg ggtgtcaatc 60

atggccggat ctctgacagg actgctgctg cttcaggccg tgtcttgggc ttctggcgct 120

agaccttgca tccccaagag cttcggctac agcagcgtcg tgtgcgtgtg caatgccacc 180

tactgcgaca gcttcgaccc tcctaccttt cctgctctgg gcaccttcag cagatacgag 240

agcaccagat ccggcagacg gatggaactg agcatgggac ccatccaggc caatcacaca 300

ggcactggcc tgctgctgac actgcagcct gagcagaaat tccagaaagt gaaaggcttc 360

ggcggagcca tgacagatgc cgccgctctg aatatcctgg ctctgtctcc accagctcag 420

aacctgctgc tcaagagcta cttcagcgag gaaggcatcg gctacaacat catcagagtg 480

cccatggcca gctgcgactt cagcatcagg acctacacct acgccgacac acccgacgat 540

ttccagctgc acaacttcag cctgcctgaa gaggacacca agctgaagat ccctctgatc 600

cacagagccc tgcagctggc acaaagaccc gtgtcactgc tggcctctcc atggacatct 660

cccacctggc tgaaaacaaa tggcgccgtg aatggcaagg gcagcctgaa aggccaacct 720

ggcgacatct accaccagac ctgggccaga tacttcgtga agttcctgga cgcctatgcc 780

gagcacaagc tgcagttttg ggccgtgaca gccgagaacg aaccttctgc tggactgctg 840

agcggctacc cctttcagtg cctgggcttt acacccgagc accagcggga ctttatcgcc 900

cgtgatctgg gacccacact ggccaatagc acccaccata atgtgcggct gctgatgctg 960

gacgaccaga gactgcttct gccccactgg gctaaagtgg tgctgacaga tcctgaggcc 1020

gccaaatacg tgcacggaat cgccgtgcac tggtatctgg actttctggc ccctgccaag 1080

gccacactgg gagagacaca cagactgttc cccaacacca tgctgttcgc cagcgaagcc 1140

tgtgtgggca gcaagttttg ggaacagagc gtgcggctcg gcagctggga tagaggcatg 1200

cagtacagcc acagcatcat caccaacctg ctgtaccacg tcgtcggctg gaccgactgg 1260

aatctggccc tgaatcctga aggcggccct aactgggtcc gaaacttcgt ggacagcccc 1320

atcatcgtgg acatcaccaa ggacaccttc tacaagcagc ccatgttcta ccacctggga 1380

cacttcagca agttcatccc cgagggctct cagcgcgttg gactggtggc ttcccagaag 1440

aacgatctgg acgccgtggc tctgatgcac cctgatggat ctgctgtggt ggtggtcctg 1500

aaccgcagca gcaaagatgt gcccctgacc atcaaggatc ccgccgtggg attcctggaa 1560

acaatcagcc ctggctactc catccacacc tacctgtggc gtagacag 1608

<210> 16

<211> 524

<212> PRT

<213> Homo sapiens

<400> 16

Met Tyr Ala Leu Phe Leu Leu Ala Ser Leu Leu Gly Ala Ala Leu Ala

1 5 10 15

Gly Pro Val Leu Gly Leu Lys Glu Cys Thr Arg Gly Ser Ala Val Trp

20 25 30

Cys Gln Asn Val Lys Thr Ala Ser Asp Cys Gly Ala Val Lys His Cys

35 40 45

Leu Gln Thr Val Trp Asn Lys Pro Thr Val Lys Ser Leu Pro Cys Asp

50 55 60

Ile Cys Lys Asp Val Val Thr Ala Ala Gly Asp Met Leu Lys Asp Asn

65 70 75 80

Ala Thr Glu Glu Glu Ile Leu Val Tyr Leu Glu Lys Thr Cys Asp Trp

85 90 95

Leu Pro Lys Pro Asn Met Ser Ala Ser Cys Lys Glu Ile Val Asp Ser

100 105 110

Tyr Leu Pro Val Ile Leu Asp Ile Ile Lys Gly Glu Met Ser Arg Pro

115 120 125

Gly Glu Val Cys Ser Ala Leu Asn Leu Cys Glu Ser Leu Gln Lys His

130 135 140

Leu Ala Glu Leu Asn His Gln Lys Gln Leu Glu Ser Asn Lys Ile Pro

145 150 155 160

Glu Leu Asp Met Thr Glu Val Val Ala Pro Phe Met Ala Asn Ile Pro

165 170 175

Leu Leu Leu Tyr Pro Gln Asp Gly Pro Arg Ser Lys Pro Gln Pro Lys

180 185 190

Asp Asn Gly Asp Val Cys Gln Asp Cys Ile Gln Met Val Thr Asp Ile

195 200 205

Gln Thr Ala Val Arg Thr Asn Ser Thr Phe Val Gln Ala Leu Val Glu

210 215 220

His Val Lys Glu Glu Cys Asp Arg Leu Gly Pro Gly Met Ala Asp Ile

225 230 235 240

Cys Lys Asn Tyr Ile Ser Gln Tyr Ser Glu Ile Ala Ile Gln Met Met

245 250 255

Met His Met Gln Pro Lys Glu Ile Cys Ala Leu Val Gly Phe Cys Asp

260 265 270

Glu Val Lys Glu Met Pro Met Gln Thr Leu Val Pro Ala Lys Val Ala

275 280 285

Ser Lys Asn Val Ile Pro Ala Leu Glu Leu Val Glu Pro Ile Lys Lys

290 295 300

His Glu Val Pro Ala Lys Ser Asp Val Tyr Cys Glu Val Cys Glu Phe

305 310 315 320

Leu Val Lys Glu Val Thr Lys Leu Ile Asp Asn Asn Lys Thr Glu Lys

325 330 335

Glu Ile Leu Asp Ala Phe Asp Lys Met Cys Ser Lys Leu Pro Lys Ser

340 345 350

Leu Ser Glu Glu Cys Gln Glu Val Val Asp Thr Tyr Gly Ser Ser Ile

355 360 365

Leu Ser Ile Leu Leu Glu Glu Val Ser Pro Glu Leu Val Cys Ser Met

370 375 380

Leu His Leu Cys Ser Gly Thr Arg Leu Pro Ala Leu Thr Val His Val

385 390 395 400

Thr Gln Pro Lys Asp Gly Gly Phe Cys Glu Val Cys Lys Lys Leu Val

405 410 415

Gly Tyr Leu Asp Arg Asn Leu Glu Lys Asn Ser Thr Lys Gln Glu Ile

420 425 430

Leu Ala Ala Leu Glu Lys Gly Cys Ser Phe Leu Pro Asp Pro Tyr Gln

435 440 445

Lys Gln Cys Asp Gln Phe Val Ala Glu Tyr Glu Pro Val Leu Ile Glu

450 455 460

Ile Leu Val Glu Val Met Asp Pro Ser Phe Val Cys Leu Lys Ile Gly

465 470 475 480

Ala Cys Pro Ser Ala His Lys Pro Leu Leu Gly Thr Glu Lys Cys Ile

485 490 495

Trp Gly Pro Ser Tyr Trp Cys Gln Asn Thr Glu Thr Ala Ala Gln Cys

500 505 510

Asn Ala Val Glu His Cys Lys Arg His Val Trp Asn

515 520

<210> 17

<211> 1572

<212> DNA

<213> Homo sapiens

<400> 17

atgtacgccc tgttcctgct ggccagcctg ctgggcgccg ccctggccgg ccccgtgctg 60

ggcctgaagg agtgcacccg cggcagcgcc gtgtggtgcc agaacgtgaa gaccgccagc 120

gactgcggcg ccgtgaagca ctgcctgcag accgtgtgga acaagcccac cgtgaagagc 180

ctgccctgcg acatctgcaa ggacgtggtg accgccgccg gcgacatgct gaaggacaac 240

gccaccgagg aggagatcct ggtgtacctg gagaagacct gcgactggct gcccaagccc 300

aacatgagcg ccagctgcaa ggagatcgtg gacagctacc tgcccgtgat cctggacatc 360

atcaagggcg agatgagccg ccccggcgag gtgtgcagcg ccctgaacct gtgcgagagc 420

ctgcagaagc acctggccga gctgaaccac cagaagcagc tggagagcaa caagatcccc 480

gagctggaca tgaccgaggt ggtggccccc ttcatggcca acatcccccct gctgctgtac 540

ccccaggacg gcccccgcag caagccccag cccaaggaca acggcgacgt gtgccaggac 600

tgcatccaga tggtgaccga catccagacc gccgtgcgca ccaacagcac cttcgtgcag 660

gccctggtgg agcacgtgaa ggaggagtgc gaccgcctgg gccccggcat ggccgacatc 720

tgcaagaact acatcagcca gtacagcgag atcgccatcc agatgatgat gcacatgcag 780

cccaaggaga tctgcgccct ggtgggcttc tgcgacgagg tgaaggagat gcccatgcag 840

accctggtgc ccgccaaggt ggccagcaag aacgtgatcc ccgccctgga gctggtggag 900

cccatcaaga agcacgaggt gcccgccaag agcgacgtgt actgcgaggt gtgcgagttc 960

ctggtgaagg aggtgaccaa gctgatcgac aacaacaaga ccgagaagga gatcctggac 1020

gccttcgaca agatgtgcag caagctgccc aagagcctga gcgaggagtg ccaggaggtg 1080

gtggacacct acggcagcag catcctgagc atcctgctgg aggaggtgag ccccgagctg 1140

gtgtgcagca tgctgcacct gtgcagcggc acccgcctgc ccgccctgac cgtgcacgtg 1200

acccagccca aggacggcgg cttctgcgag gtgtgcaaga agctggtggg ctacctggac 1260

cgcaacctgg agaagaacag caccaagcag gagatcctgg ccgccctgga gaagggctgc 1320

agcttcctgc ccgaccccta ccagaagcag tgcgaccagt tcgtggccga gtacgagccc 1380

gtgctgatcg agatcctggt ggaggtgatg gaccccagct tcgtgtgcct gaagatcggc 1440

gcctgcccca gcgcccacaa gcccctgctg ggcaccgaga agtgcatctg gggccccagc 1500

tactggtgcc agaacaccga gaccgccgcc cagtgcaacg ccgtggagca ctgcaagcgc 1560

cacgtgtgga ac 1572

<210> 18

<211> 478

<212> PRT

<213> Homo sapiens

<400> 18

Met Gly Arg Cys Cys Phe Tyr Thr Ala Gly Thr Leu Ser Leu Leu Leu

1 5 10 15

Leu Val Thr Ser Val Thr Leu Leu Val Ala Arg Val Phe Gln Lys Ala

20 25 30

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

35 40 45

Ala Phe Asp Ser Trp Glu Lys Pro Pro Leu Pro Val Tyr Thr Gln Phe

50 55 60

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

65 70 75 80

Pro Arg Val Glu Glu Val Gly Pro Tyr Thr Tyr Arg Glu Leu Arg Asn

85 90 95

Lys Ala Asn Ile Gln Phe Gly Asp Asn Gly Thr Thr Ile Ser Ala Val

100 105 110

Ser Asn Lys Ala Tyr Val Phe Glu Arg Asp Gln Ser Val Gly Asp Pro

115 120 125

Lys Ile Asp Leu Ile Arg Thr Leu Asn Ile Pro Val Leu Thr Val Ile

130 135 140

Glu Trp Ser Gln Val His Phe Leu Arg Glu Ile Ile Glu Ala Met Leu

145 150 155 160

Lys Ala Tyr Gln Gln Lys Leu Phe Val Thr His Thr Val Asp Glu Leu

165 170 175

Leu Trp Gly Tyr Lys Asp Glu Ile Leu Ser Leu Ile His Val Phe Arg

180 185 190

Pro Asp Ile Ser Pro Tyr Phe Gly Leu Phe Tyr Glu Lys Asn Gly Thr

195 200 205

Asn Asp Gly Asp Tyr Val Phe Leu Thr Gly Glu Asp Ser Tyr Leu Asn

210 215 220

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

225 230 235 240

Ile Thr Asp Lys Cys Asn Met Ile Asn Gly Thr Asp Gly Asp Ser Phe

245 250 255

His Pro Leu Ile Thr Lys Asp Glu Val Leu Tyr Val Phe Pro Ser Asp

260 265 270

Phe Cys Arg Ser Val Tyr Ile Thr Phe Ser Asp Tyr Glu Ser Val Gln

275 280 285

Gly Leu Pro Ala Phe Arg Tyr Lys Val Pro Ala Glu Ile Leu Ala Asn

290 295 300

Thr Ser Asp Asn Ala Gly Phe Cys Ile Pro Glu Gly Asn Cys Leu Gly

305 310 315 320

Ser Gly Val Leu Asn Val Ser Ile Cys Lys Asn Gly Ala Pro Ile Ile

325 330 335

Met Ser Phe Pro His Phe Tyr Gln Ala Asp Glu Arg Phe Val Ser Ala

340 345 350

Ile Glu Gly Met His Pro Asn Gln Glu Asp His Glu Thr Phe Val Asp

355 360 365

Ile Asn Pro Leu Thr Gly Ile Ile Leu Lys Ala Ala Lys Arg Phe Gln

370 375 380

Ile Asn Ile Tyr Val Lys Lys Leu Asp Asp Phe Val Glu Thr Gly Asp

385 390 395 400

Ile Arg Thr Met Val Phe Pro Val Met Tyr Leu Asn Glu Ser Val His

405 410 415

Ile Asp Lys Glu Thr Ala Ser Arg Leu Lys Ser Met Ile Asn Thr Thr

420 425 430

Leu Ile Ile Thr Asn Ile Pro Tyr Ile Ile Met Ala Leu Gly Val Phe

435 440 445

Phe Gly Leu Val Phe Thr Trp Leu Ala Cys Lys Gly Gln Gly Ser Met

450 455 460

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

465 470 475

<210> 19

<211> 1434

<212> DNA

<213> Homo sapiens

<400> 19

atgggccgct gctgcttcta caccgccggc accctgagcc tgctgctgct ggtgaccagc 60

gtgaccctgc tggtggcccg cgtgttccag aaggccgtgg accagagcat cgagaagaag 120

atcgtgctgc gcaacggcac cgaggccttc gacagctggg agaagccccc cctgcccgtg 180

tacacccagt tctacttctt caacgtgacc aaccccgagg agatcctgcg cggcgagacc 240

ccccgcgtgg aggaggtggg cccctacacc taccgcgagc tgcgcaacaa ggccaacatc 300

cagttcggcg acaacggcac caccatcagc gccgtgagca acaaggccta cgtgttcgag 360

cgcgaccaga gcgtgggcga ccccaagatc gacctgatcc gcaccctgaa catccccgtg 420

ctgaccgtga tcgagtggag ccaggtgcac ttcctgcgcg agatcatcga ggccatgctg 480

aaggcctacc agcagaagct gttcgtgacc cacaccgtgg acgagctgct gtggggctac 540

aaggacgaga tcctgagcct gatccacgtg ttccgccccg acatcagccc ctacttcggc 600

ctgttctacg agaagaacgg caccaacgac ggcgactacg tgttcctgac cggcgaggac 660

agctacctga acttcaccaa gatcgtggag tggaacggca agaccagcct ggactggtgg 720

atcaccgaca agtgcaacat gatcaacggc accgacggcg acagcttcca ccccctgatc 780

accaaggacg aggtgctgta cgtgttcccc agcgacttct gccgcagcgt gtacatcacc 840

ttcagcgact acgagagcgt gcagggcctg cccgccttcc gctacaaggt gcccgccgag 900

atcctggcca acaccagcga caacgccggc ttctgcatcc ccgagggcaa ctgcctgggc 960

agcggcgtgc tgaacgtgag catctgcaag aacggcgccc ccatcatcat gagcttcccc 1020

cacttctacc aggccgacga gcgcttcgtg agcgccatcg agggcatgca ccccaaccag 1080

gaggaccacg agaccttcgt ggacatcaac cccctgaccg gcatcatcct gaaggccgcc 1140

aagcgcttcc agatcaacat ctacgtgaag aagctggacg acttcgtgga gaccggcgac 1200

atccgcacca tggtgttccc cgtgatgtac ctgaacgaga gcgtgcacat cgacaaggag 1260

accgccagcc gcctgaagag catgatcaac accaccctga tcatcaccaa catcccctac 1320

atcatcatgg ccctgggcgt gttcttcggc ctggtgttca cctggctggc ctgcaagggc 1380

cagggcagca tggacgaggg caccgccgac gagcgcgccc ccctgatccg cacc 1434

<210> 20

<211> 23

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 20

tggaagactt cgagatacac tgt 23

<210> 21

<211> 23

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 21

acagtgtatc tcgaagtctt cca 23

<210> 22

<211> 21

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 22

tttagaaata agtggtagtc a 21

<210> 23

<211> 21

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 23

tgactaccac ttatttctaa a 21

<210> 24

<211> 19

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 24

agggtatcaa gactacgaa 19

<210> 25

<211> 19

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 25

ttcgtagtct tgataccct 19

<210> 26

<211> 19

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 26

tattagatct gatggccgc 19

<210> 27

<211> 20

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 27

ctccatcact aggggttcct 20

<210> 28

<211> 60

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 28

agctctgggt atttaagccc gagtgagcac gcagggtctc cattttgaag cgggaggtta 60

<210> 29

<211> 145

<212> DNA

<213> Unknown

<220><223> AAV2 ITR

<400> 29

aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc 120

gagcgcgcag agagggagtg gccaa 145

<210> 30

<211> 927

<212> PRT

<213> Homo sapiens

<400> 30

Met Gly Thr Gln Asp Pro Gly Asn Met Gly Thr Gly Val Pro Ala Ser

1 5 10 15

Glu Gln Ile Ser Cys Ala Lys Glu Asp Pro Gln Val Tyr Cys Pro Glu

20 25 30

Glu Thr Gly Gly Thr Lys Asp Val Gln Val Thr Asp Cys Lys Ser Pro

35 40 45

Glu Asp Ser Arg Pro Pro Lys Glu Thr Asp Cys Cys Asn Pro Glu Asp

50 55 60

Ser Gly Gln Leu Met Val Ser Tyr Glu Gly Lys Ala Met Gly Tyr Gln

65 70 75 80

Val Pro Pro Phe Gly Trp Arg Ile Cys Leu Ala His Glu Phe Thr Glu

85 90 95

Lys Arg Lys Pro Phe Gln Ala Asn Asn Val Ser Leu Ser Asn Met Ile

100 105 110

Lys His Ile Gly Met Gly Leu Arg Tyr Leu Gln Trp Trp Tyr Arg Lys

115 120 125

Thr His Val Glu Lys Lys Thr Pro Phe Ile Asp Met Ile Asn Ser Val

130 135 140

Pro Leu Arg Gln Ile Tyr Gly Cys Pro Leu Gly Gly Ile Gly Gly Gly

145 150 155 160

Thr Ile Thr Arg Gly Trp Arg Gly Gln Phe Cys Arg Trp Gln Leu Asn

165 170 175

Pro Gly Met Tyr Gln His Arg Thr Val Ile Ala Asp Gln Phe Thr Val

180 185 190

Cys Leu Arg Arg Glu Gly Gln Thr Val Tyr Gln Gln Val Leu Ser Leu

195 200 205

Glu Arg Pro Ser Val Leu Arg Ser Trp Asn Trp Gly Leu Cys Gly Tyr

210 215 220

Phe Ala Phe Tyr His Ala Leu Tyr Pro Arg Ala Trp Thr Val Tyr Gln

225 230 235 240

Leu Pro Gly Gln Asn Val Thr Leu Thr Cys Arg Gln Ile Thr Pro Ile

245 250 255

Leu Pro His Asp Tyr Gln Asp Ser Ser Leu Pro Val Gly Val Phe Val

260 265 270

Trp Asp Val Glu Asn Glu Gly Asp Glu Ala Leu Asp Val Ser Ile Met

275 280 285

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

290 295 300

Leu Trp Asn Glu Pro Phe Cys Leu Glu Arg Ser Gly Glu Thr Val Arg

305 310 315 320

Gly Leu Leu Leu His His Pro Thr Leu Pro Asn Pro Tyr Thr Met Ala

325 330 335

Val Ala Ala Arg Val Thr Ala Ala Thr Thr Val Thr His Ile Thr Ala

340 345 350

Phe Asp Pro Asp Ser Thr Gly Gln Gln Val Trp Gln Asp Leu Leu Gln

355 360 365

Asp Gly Gln Leu Asp Ser Pro Thr Gly Gln Ser Thr Pro Thr Gln Lys

370 375 380

Gly Val Gly Ile Ala Gly Ala Val Cys Val Ser Ser Lys Leu Arg Pro

385 390 395 400

Arg Gly Gln Cys Arg Leu Glu Phe Ser Leu Ala Trp Asp Met Pro Arg

405 410 415

Ile Met Phe Gly Ala Lys Gly Gln Val His Tyr Arg Arg Tyr Thr Arg

420 425 430

Phe Phe Gly Gln Asp Gly Asp Ala Ala Pro Ala Leu Ser His Tyr Ala

435 440 445

Leu Cys Arg Tyr Ala Glu Trp Glu Glu Arg Ile Ser Ala Trp Gln Ser

450 455 460

Pro Val Leu Asp Asp Arg Ser Leu Pro Ala Trp Tyr Lys Ser Ala Leu

465 470 475 480

Phe Asn Glu Leu Tyr Phe Leu Ala Asp Gly Gly Thr Val Trp Leu Glu

485 490 495

Val Leu Glu Asp Ser Leu Pro Glu Glu Leu Gly Arg Asn Met Cys His

500 505 510

Leu Arg Pro Thr Leu Arg Asp Tyr Gly Arg Phe Gly Tyr Leu Glu Gly

515 520 525

Gln Glu Tyr Arg Met Tyr Asn Thr Tyr Asp Val His Phe Tyr Ala Ser

530 535 540

Phe Ala Leu Ile Met Leu Trp Pro Lys Leu Glu Leu Ser Leu Gln Tyr

545 550 555 560

Asp Met Ala Leu Ala Thr Leu Arg Glu Asp Leu Thr Arg Arg Arg Tyr

565 570 575

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

580 585 590

His Asp Ile Gly Asp Pro Asp Asp Glu Pro Trp Leu Arg Val Asn Ala

595 600 605

Tyr Leu Ile His Asp Thr Ala Asp Trp Lys Asp Leu Asn Leu Lys Phe

610 615 620

Val Leu Gln Val Tyr Arg Asp Tyr Tyr Leu Thr Gly Asp Gln Asn Phe

625 630 635 640

Leu Lys Asp Met Trp Pro Val Cys Leu Ala Val Met Glu Ser Glu Met

645 650 655

Lys Phe Asp Lys Asp His Asp Gly Leu Ile Glu Asn Gly Gly Tyr Ala

660 665 670

Asp Gln Thr Tyr Asp Gly Trp Val Thr Thr Gly Pro Ser Ala Tyr Cys

675 680 685

Gly Gly Leu Trp Leu Ala Ala Val Ala Val Met Val Gln Met Ala Ala

690 695 700

Leu Cys Gly Ala Gln Asp Ile Gln Asp Lys Phe Ser Ser Ile Leu Ser

705 710 715 720

Arg Gly Gln Glu Ala Tyr Glu Arg Leu Leu Trp Asn Gly Arg Tyr Tyr

725 730 735

Asn Tyr Asp Ser Ser Ser Arg Pro Gln Ser Arg Ser Val Met Ser Asp

740 745 750

Gln Cys Ala Gly Gln Trp Phe Leu Lys Ala Cys Gly Leu Gly Glu Gly

755 760 765

Asp Thr Glu Val Phe Pro Thr Gln His Val Val Arg Ala Leu Gln Thr

770 775 780

Ile Phe Glu Leu Asn Val Gln Ala Phe Ala Gly Gly Ala Met Gly Ala

785 790 795 800

Val Asn Gly Met Gln Pro His Gly Val Pro Asp Lys Ser Ser Val Gln

805 810 815

Ser Asp Glu Val Trp Val Gly Val Val Tyr Gly Leu Ala Ala Thr Met

820 825 830

Ile Gln Glu Gly Leu Thr Trp Glu Gly Phe Gln Thr Ala Glu Gly Cys

835 840 845

Tyr Arg Thr Val Trp Glu Arg Leu Gly Leu Ala Phe Gln Thr Pro Glu

850 855 860

Ala Tyr Cys Gln Gln Arg Val Phe Arg Ser Leu Ala Tyr Met Arg Pro

865 870 875 880

Leu Ser Ile Trp Ala Met Gln Leu Ala Leu Gln Gln Gln Gln His Lys

885 890 895

Lys Ala Ser Trp Pro Lys Val Lys Gln Gly Thr Gly Leu Arg Thr Gly

900 905 910

Pro Met Phe Gly Pro Lys Glu Ala Met Ala Asn Leu Ser Pro Glu

915 920 925

<210> 31

<211> 2781

<212> DNA

<213> Homo sapiens

<400> 31

atgggcaccc aggaccccgg caacatgggc accggcgtgc ccgccagcga gcagatcagc 60

tgcgccaagg aggacccccca ggtgtactgc cccgaggaga ccggcggcac caaggacgtg 120

caggtgaccg actgcaagag ccccgaggac agccgccccc ccaaggagac cgactgctgc 180

aaccccgagg acagcggcca gctgatggtg agctacgagg gcaaggccat gggctaccag 240

gtgccccccct tcggctggcg catctgcctg gcccacgagt tcaccgagaa gcgcaagccc 300

ttccaggcca acaacgtgag cctgagcaac atgatcaagc acatcggcat gggcctgcgc 360

tacctgcagt ggtggtaccg caagacccac gtggagaaga agaccccctt catcgacatg 420

atcaacagcg tgcccctgcg ccagatctac ggctgccccc tgggcggcat cggcggcggc 480

accatcaccc gcggctggcg cggccagttc tgccgctggc agctgaaccc cggcatgtac 540

cagcaccgca ccgtgatcgc cgaccagttc accgtgtgcc tgcgccgcga gggccagacc 600

gtgtaccagc aggtgctgag cctggagcgc cccagcgtgc tgcgcagctg gaactggggc 660

ctgtgcggct acttcgcctt ctaccacgcc ctgtaccccc gcgcctggac cgtgtaccag 720

ctgcccggcc agaacgtgac cctgacctgc cgccagatca cccccatcct gccccacgac 780

taccaggaca gcagcctgcc cgtgggcgtg ttcgtgtggg acgtggagaa cgagggcgac 840

gaggccctgg acgtgagcat catgttcagc atgcgcaacg gcctgggcgg cggcgacgac 900

gcccccggcg gcctgtggaa cgagcccttc tgcctggagc gcagcggcga gaccgtgcgc 960

ggcctgctgc tgcaccaccc caccctgccc aacccctaca ccatggccgt ggccgcccgc 1020

gtgaccgccg ccaccaccgt gacccacatc accgccttcg accccgacag caccggccag 1080

caggtgtggc aggacctgct gcaggacggc cagctggaca gccccaccgg ccagagcacc 1140

cccacccaga agggcgtggg catcgccggc gccgtgtgcg tgagcagcaa gctgcgcccc 1200

cgcggccagt gccgcctgga gttcagcctg gcctgggaca tgccccgcat catgttcggc 1260

gccaagggcc aggtgcacta ccgccgctac acccgcttct tcggccagga cggcgacgcc 1320

gcccccgccc tgagccacta cgccctgtgc cgctacgccg agtgggagga gcgcatcagc 1380

gcctggcaga gccccgtgct ggacgaccgc agcctgcccg cctggtacaa gagcgccctg 1440

ttcaacgagc tgtacttcct ggccgacggc ggcaccgtgt ggctggaggt gctggaggac 1500

agcctgcccg aggagctggg ccgcaacatg tgccacctgc gccccaccct gcgcgactac 1560

ggccgcttcg gctacctgga gggccaggag taccgcatgt acaacaccta cgacgtgcac 1620

ttctacgcca gcttcgccct gatcatgctg tggcccaagc tggagctgag cctgcagtac 1680

gacatggccc tggccaccct gcgcgaggac ctgacccgcc gccgctacct gatgagcggc 1740

gtgatggccc ccgtgaagcg ccgcaacgtg atcccccacg acatcggcga ccccgacgac 1800

gagccctggc tgcgcgtgaa cgcctacctg atccacgaca ccgccgactg gaaggacctg 1860

aacctgaagt tcgtgctgca ggtgtaccgc gactactacc tgaccggcga ccagaacttc 1920

ctgaaggaca tgtggcccgt gtgcctggcc gtgatggaga gcgagatgaa gttcgacaag 1980

gaccacgacg gcctgatcga gaacggcggc tacgccgacc agacctacga cggctgggtg 2040

accaccggcc ccagcgccta ctgcggcggc ctgtggctgg ccgccgtggc cgtgatggtg 2100

cagatggccg ccctgtgcgg cgcccaggac atccaggaca agttcagcag catcctgagc 2160

cgcggccagg aggcctacga gcgcctgctg tggaacggcc gctactacaa ctacgacagc 2220

agcagccgcc cccagagccg cagcgtgatg agcgaccagt gcgccggcca gtggttcctg 2280

aaggcctgcg gcctgggcga gggcgacacc gaggtgttcc ccacccagca cgtggtgcgc 2340

gccctgcaga ccatcttcga gctgaacgtg caggccttcg ccggcggcgc catgggcgcc 2400

gtgaacggca tgcagcccca cggcgtgccc gacaagagca gcgtgcagag cgacgaggtg 2460

tgggtgggcg tggtgtacgg cctggccgcc accatgatcc aggagggcct gacctgggag 2520

ggcttccaga ccgccgaggg ctgctaccgc accgtgtggg agcgcctggg cctggccttc 2580

cagacccccg aggcctactg ccagcagcgc gtgttccgca gcctggccta catgcgcccc 2640

ctgagcatct gggccatgca gctggccctg cagcagcagc agcacaagaa ggccagctgg 2700

cccaaggtga agcagggcac cggcctgcgc accggcccca tgttcggccc caaggaggcc 2760

atggccaacc tgagccccga g 2781

<210> 32

<211> 11264

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 32

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agtaagtcac 300

tgactgtcta tgcctgggaa agggtgggca ggagatgggg cagtgcagga aaagtggcac 360

tatgaaccct cctggtggcg aggggagggg ggtggtcctc gaacgccttg cagaactggc 420

ctggatacag agtggaccgg ctggccccat ctggaagact tcgagataca ctgttgtctt 480

actgcgctca acagtgtatc tcgaagtctt ccaaatggtg ccagccatcg cagcggggtg 540

caggaaatgg gggcagcccc cctttttggc tatccttcca cgtgttcttt tttgtatctt 600

ttgtgtttcc tagaaaacat ctcagtcacc accgcagccc taggaatgca tctagacaat 660

tgtactaacc ttcttctctt tcctctcctg acagtccgga aagccaccat gggcacccag 720

gaccccggca acatgggcac cggcgtgccc gccagcgagc agatcagctg cgccaaggag 780

gacccccagg tgtactgccc cgaggagacc ggcggcacca aggacgtgca ggtgaccgac 840

tgcaagagcc ccgaggacag ccgcccccccc aaggagaccg actgctgcaa ccccgaggac 900

agcggccagc tgatggtgag ctacgagggc aaggccatgg gctaccaggt gcccccccttc 960

ggctggcgca tctgcctggc ccacgagttc accgagaagc gcaagccctt ccaggccaac 1020

aacgtgagcc tgagcaacat gatcaagcac atcggcatgg gcctgcgcta cctgcagtgg 1080

tggtaccgca agacccacgt ggagaagaag accccccttca tcgacatgat caacagcgtg 1140

cccctgcgcc agatctacgg ctgcccccctg ggcggcatcg gcggcggcac catcacccgc 1200

ggctggcgcg gccagttctg ccgctggcag ctgaaccccg gcatgtacca gcaccgcacc 1260

gtgatcgccg accagttcac cgtgtgcctg cgccgcgagg gccagaccgt gtaccagcag 1320

gtgctgagcc tggagcgccc cagcgtgctg cgcagctgga actggggcct gtgcggctac 1380

ttcgccttct accacgccct gtacccccgc gcctggaccg tgtaccagct gcccggccag 1440

aacgtgaccc tgacctgccg ccagatcacc cccatcctgc cccacgacta ccaggacagc 1500

agcctgcccg tgggcgtgtt cgtgtgggac gtggagaacg agggcgacga ggccctggac 1560

gtgagcatca tgttcagcat gcgcaacggc ctgggcggcg gcgacgacgc ccccggcggc 1620

ctgtggaacg agcccttctg cctggagcgc agcggcgaga ccgtgcgcgg cctgctgctg 1680

caccacccca ccctgcccaa cccctacacc atggccgtgg ccgcccgcgt gaccgccgcc 1740

accaccgtga cccacatcac cgccttcgac cccgacagca ccggccagca ggtgtggcag 1800

gacctgctgc aggacggcca gctggacagc cccaccggcc agagcacccc cacccagaag 1860

ggcgtgggca tcgccggcgc cgtgtgcgtg agcagcaagc tgcgcccccg cggccagtgc 1920

cgcctggagt tcagcctggc ctgggacatg ccccgcatca tgttcggcgc caagggccag 1980

gtgcactacc gccgctacac ccgcttcttc ggccaggacg gcgacgccgc ccccgccctg 2040

agccactacg ccctgtgccg ctacgccgag tgggaggagc gcatcagcgc ctggcagagc 2100

cccgtgctgg acgaccgcag cctgcccgcc tggtacaaga gcgccctgtt caacgagctg 2160

tacttcctgg ccgacggcgg caccgtgtgg ctggaggtgc tggaggacag cctgcccgag 2220

gagctgggcc gcaacatgtg ccacctgcgc cccaccctgc gcgactacgg ccgcttcggc 2280

tacctggagg gccaggagta ccgcatgtac aacacctacg acgtgcactt ctacgccagc 2340

ttcgccctga tcatgctgtg gcccaagctg gagctgagcc tgcagtacga catggccctg 2400

gccaccctgc gcgaggacct gacccgccgc cgctacctga tgagcggcgt gatggccccc 2460

gtgaagcgcc gcaacgtgat cccccacgac atcggcgacc ccgacgacga gccctggctg 2520

cgcgtgaacg cctacctgat ccacgacacc gccgactgga aggacctgaa cctgaagttc 2580

gtgctgcagg tgtaccgcga ctactacctg accggcgacc agaacttcct gaaggacatg 2640
tggcccgtgt gcctggccgt gatggagagc gagatgaagt tcgacaagga ccacgacggc 2700

ctgatcgaga acggcggcta cgccgaccag acctacgacg gctgggtgac caccggcccc 2760
agcgcctact gcggcggcct gtggctggcc gccgtggccg tgatggtgca gatggccgcc 2820
ctgtgcggcg cccaggacat ccaggacaag ttcagcagca tcctgagccg cggccaggag 2880
gcctacgagc gcctgctgtg gaacggccgc tactacaact acgacagcag cagccgcccc 2940
cagagccgca gcgtgatgag cgaccagtgc gccggccagt ggttcctgaa ggcctgcggc 3000
ctgggcgagg gcgacaccga ggtgttcccc acccagcacg tggtgcgcgc cctgcagacc 3060
atcttcgagc tgaacgtgca ggccttcgcc ggcggcgcca tgggcgccgt gaacggcatg 3120

cagccccacg gcgtgcccga caagagcagc gtgcagagcg acgaggtgtg ggtgggcgtg 3180
gtgtacggcc tggccgccac catgatccag gagggcctga cctgggaggg cttccagacc 3240
gccgagggct gctaccgcac cgtgtgggag cgcctgggcc tggccttcca gacccccgag 3300
gcctactgcc agcagcgcgt gttccgcagc ctggcctaca tgcgccccct gagcatctgg 3360
gccatgcagc tggccctgca gcagcagcag cacaagaagg ccagctggcc caaggtgaag 3420
cagggcaccg gcctgcgcac cggccccatg ttcggcccca aggaggccat ggccaacctg 3480
agccccgagt gacaattgtt aattaagttt aaaccctcga ggccgcaagc ttatcgataa 3540

tcaacctctg gattacaaaa tttgtgaaag attgactggt attcttaact atgttgctcc 3600
ttttacgcta tgtggatacg ctgctttaat gcctttgtat catgctattg cttcccgtat 3660
ggctttcatt ttctcctcct tgtataaatc ctggttgctg tctctttatg aggagttgtg 3720
gcccgttgtc aggcaacgtg gcgtggtgtg cactgtgttt gctgacgcaa cccccactgg 3780
ttggggcatt gccaccacct gtcagctcct ttccgggact ttcgctttcc ccctccctat 3840
tgccacggcg gaactcatcg ccgcctgcct tgcccgctgc tggacagggg ctcggctgtt 3900
gggcactgac aattccgtgg tgttgtcggg gaaatcatcg tcctttcctt ggctgctcgc 3960

ctgtgttgcc acctggattc tgcgcgggac gtccttctgc tacgtccctt cggccctcaa 4020
tccagcggac cttccttccc gcggcctgct gccggctctg cggcctcttc cgcgtcttcg 4080
ccttcgccct cagacgagtc ggatctccct ttgggccgcc tccccgcatc gataccgtcg 4140
actagagctc gctgatcagc ctcgactgtg ccttctagtt gccagccatc tgttgtttgc 4200
ccctcccccg tgccttcctt gaccctggaa ggtgccactc ccactgtcct ttcctaataa 4260
aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg gggtggggtg 4320

gggcaggaca gcaaggggga ggattgggaa gacaatagca ggcatgctgg ggagagatcc 4380

acgataacaa acagcttttt tggggtgaac atattgactg aattccctgc aggttggcca 4440
ctccctctct gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg 4500
cgacctttgg tcgcccggcc tcagtgagcg agcgagcgcg cagagaggga gtggccaact 4560
ccatcactag gggttcctgc ggccgctcgt acggtctcga ggaattcctg caggataact 4620
tgccaacctc attctaaaat gtatatagaa gcccaaaaga caataacaaa aatattcttg 4680
tagaacaaaa tgggaaagaa tgttccacta aatatcaaga tttagagcaa agcatgagat 4740
gtgtggggat agacagtgag gctgataaaa tagagtagag ctcagaaaca gacccattga 4800

tatatgtaag tgacctatga aaaaaatatg gcattttaca atgggaaaat gatggtcttt 4860
ttctttttta gaaaaacagg gaaatatatt tatatgtaaa aaataaaagg gaacccatat 4920
gtcataccat acacacaaaa aaattccagt gaattataag tctaaatgga gaaggcaaaa 4980
ctttaaatct tttagaaaat aatatagaag catgcagacc agcctggcca acatgatgaa 5040
accctctcta ctaataataa aatcagtaga actactcagg actactttga gtgggaagtc 5100
cttttctatg aagacttctt tggccaaaat taggctctaa atgcaaggag atagtgcatc 5160
atgcctggct gcacttactg ataaatgatg ttatcaccat ctttaaccaa atgcacagga 5220

acaagttatg gtactgatgt gctggattga gaaggagctc tacttccttg acaggacaca 5280
tttgtatcaa cttaaaaaag cagatttttg ccagcagaac tattcattca gaggtaggaa 5340
acttagaata gatgatgtca ctgattagca tggcttcccc atctccacag ctgcttccca 5400
cccaggttgc ccacagttga gtttgtccag tgctcagggc tgcccactct cagtaagaag 5460
ccccacacca gcccctctcc aaatatgttg gctgttcctt ccattaaagt gaccccactt 5520
tagagcagca agtggatttc tgtttcttac agttcaggaa ggaggagtca gctgtgagaa 5580
cctggagcct gagatgcttc taagtcccac tgctactggg gtcagggaag ccagactcca 5640

gcatcagcag tcaggagcac taagcccttg ccaacatcct gtttctcaga gaaactgctt 5700
ccattataat ggttgtcctt ttttaagcta tcaagccaaa caaccagtgt ctaccattat 5760
tctcatcacc tgaagccaag ggttctagca aaagtcaagc tgtcttgtaa tggttgatgt 5820
gcctccagct tctgtcttca gtcactccac tcttagcctg ctctgaatca actctgacca 5880
cagttccctg gagcccctgc cacctgctgc ccctgccacc ttctccatct gcagtgctgt 5940
gcagccttct gcactcttgc agagctaata ggtggagact tgaaggaaga ggaggaaagt 6000
ttctcataat agccttgctg caagctcaaa tgggaggtgg gcactgtgcc caggagcctt 6060

ggagcaaagg ctgtgcccaa cctctgactg catccaggtt tggtcttgac agagataaga 6120
agccctggct tttggagcca aaatctaggt cagacttagg caggattctc aaagtttatc 6180
agcagaacat gaggcagaag accctttctg ctccagcttc ttcaggctca accttcatca 6240
gaatagatag aaagagaggc tgtgagggtt cttaaaacag aagcaaatct gactcagaga 6300
ataaacaacc tcctagtaaa ctacagctta gacagagcat ctggtggtga gtgtgctcag 6360
tgtcctactc aactgtctgg tatcagccct catgaggact tctcttcttt ccctcataga 6420
cctccatctc tgttttcctt agcctgcaga aatctggatg gctattcaca gaatgcctgt 6480

gctttcagag ttgcattttt tctctggtat tctggttcaa gcatttgaag gtaggaaagg 6540
ttctccaagt gcaagaaagc cagccctgag cctcaactgc ctggctagtg tggtcagtag 6600
gatgcaaagg ctgttgaatg ccacaaggcc aaactttaac ctgtgtacca caagcctagc 6660
agcagaggca gctctgctca ctggaactct ctgtcttctt tctcctgagc cttttctttt 6720
cctgagtttt ctagctctcc tcaaccttac ctctgcccta cccaggacaa acccaagagc 6780
cactgtttct gtgatgtcct ctccagccct aattaggcat catgacttca gcctgacctt 6840
ccatgctcag aagcagtgct aatccacttc agatgagctg ctctatgcaa cacaggcaga 6900

gcctacaaac ctttgcacca gagccctcca catatcagtg tttgttcata ctcacttcaa 6960
cagcaaatgt gactgctgag attaagattt tacacaagat ggtctgtaat ttcacagtta 7020
gttttatccc attaggtatg aaagaattag cataattccc cttaaacatg aatgaatctt 7080
agatttttta ataaatagtt ttggaagtaa agacagagac atcaggagca caaggaatag 7140
cctgagagga caaacagaac aagaaagagt ctggaaatac acaggatgtt cttggcctcc 7200
tcaaagcaag tgcaagcaga tagtaccagc agccccaggc tatcagagcc cagtgaagag 7260
aagtaccatg aaagccacag ctctaaccac cctgttccag agtgacagac agtccccaag 7320

acaagccagc ctgagccaga gagagaactg caagagaaag tttctaattt aggttctgtt 7380
agattcagac aagtgcaggt catcctctct ccacagctac tcacctctcc agcctaacaa 7440
agcctgcagt ccacactcca accctggtgt ctcacctcct agcctctccc aacatcctgc 7500
tctctgacca tcttctgcat ctctcatctc accatctccc actgtctaca gcctactctt 7560
gcaactacca tctcattttc tgacatcctg tctacatctt ctgccatact ctgccatcta 7620
ccataccacc tcttaccatc taccacacca tcttttatct ccatccctct cagaagcctc 7680
caagctgaat cctgctttat gtgttcatct cagcccctgc atggaaagct gaccccagag 7740

gcagaactat tcccagagag cttggccaag aaaaacaaaa ctaccagcct ggccaggctc 7800
aggagtagta agctgcagtg tctgttgtgt tctagcttca acagctgcag gagttccact 7860
ctcaaatgct ccacatttct cacatcctcc tgattctggt cactacccat cttcaaagaa 7920

cagaatatct cacatcagca tactgtgaag gactagtcat gggtgcagct gctcagagct 7980
gcaaagtcat tctggatggt ggagagctta caaacatttc atgatgctcc ccccgctctg 8040
atggctggag cccaatccct acacagactc ctgctgtatg tgttttcctt tcactctgag 8100
ccacagccag agggcaggca ttcagtctcc tcttcaggct ggggctgggg cactgagaac 8160

tcacccaaca ccttgctctc actccttctg caaaacaaga aagagctttg tgctgcagta 8220
gccatgaaga atgaaaggaa ggctttaact aaaaaatgtc agagattatt ttcaacccct 8280
tactgtggat caccagcaag gaggaaacac aacacagaga catttttcc cctcaaatta 8340
tcaaaagaat cactgcattt gttaaagaga gcaactgaat caggaagcag agttttgaac 8400
atatcagaag ttaggaatct gcatcagaga caaatgcagt catggttgtt tgctgcatac 8460
cagccctaat cattagaagc ctcatggact tcaaacatca ttccctctga caagatgctc 8520
tagcctaact ccatgagata aaataaatct gcctttcaga gccaaagaag agtccaccag 8580

cttcttctca gtgtgaacaa gagctccagt caggttagtc agtccagtgc agtagaggag 8640
accagtctgc atcctctaat tttcaaaggc aagaagattt gtttaccctg gacaccaggc 8700
acaagtgagg tcacagagct cttagatatg cagtcctcat gagtgaggag actaaagcgc 8760
atgccatcaa gacttcagtg tagagaaaac ctccaaaaaa gcctcctcac tacttctgga 8820
atagctcaga ggccgaggcg gcctcggcct ctgcataaat aaaaaaaatt agtcagccat 8880
ggggcggaga atgggcggaa ctgggcggag ttaggggcgg gatgggcgga gttaggggcg 8940
ggactatggt tgctgactaa ttgagatgca tgctttgcat acttctgcct gctggggagc 9000

ctggggactt tccacacctg gttgctgact aattgagatg catgctttgc atacttctgc 9060
ctgctgggga gcctggggac tttccacacc ctaactgaca cacattccac agctgcatta 9120
atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc 9180
gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa 9240
ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 9300
aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct 9360
ccgcccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 9420

aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc 9480
gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc 9540
tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg 9600
tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga 9660
gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta acaggattag 9720
cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtggccta actacggcta 9780

cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag 9840

agttggtagc tcttgatccg gcaaacaaac caccgctggt agcggtggtt tttttgtttg 9900
caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac 9960
ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc 10020
aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag 10080
tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg cacctatctc 10140
agcgatctgt ctatttcgtt catccatagt tgcctgactc ctgcaaacca cgttgtgtct 10200
caaaatctct gatgttacat tgcacaagat aaaaatatat catcatgaac aataaaactg 10260

tctgcttaca taaacagtaa tacaaggggt gttatgagcc atattcaacg ggaaacgtct 10320
tgctcgaggc cgcgattaaa ttccaacatg gatgctgatt tatatgggta taaatgggct 10380
cgcgataatg tcgggcaatc aggtgcgaca atctatcgat tgtatgggaa gcccgatgcg 10440
ccagagttgt ttctgaaaca tggcaaaggt agcgttgcca atgatgttac agatgagatg 10500
gtcagactaa actggctgac ggaatttatg cctcttccga ccatcaagca ttttatccgt 10560
actcctgatg atgcatggtt actcaccact gcgatccccg ggaaaacagc attccaggta 10620
ttagaagaat atcctgattc aggtgaaaat attgttgatg cgctggcagt gttcctgcgc 10680

cggttgcatt cgattcctgt ttgtaattgt ccttttaaca gcgatcgcgt atttcgtctc 10740
gctcaggcgc aatcacgaat gaataacggt ttggttgatg cgagtgattt tgatgacgag 10800
cgtaatggct ggcctgttga acaagtctgg aaagaaatgc ataagctttt gccattctca 10860
ccggattcag tcgtcactca tggtgatttc tcacttgata accttatttt tgacgagggg 10920
aaattaatag gttgtattga tgttggacga gtcggaatcg cagaccgata ccaggatctt 10980
gccatcctat ggaactgcct cggtgagttt tctccttcat tacagaaacg gctttttcaa 11040
aaatatggta ttgataatcc tgatatgaat aaattgcagt ttcatttgat gctcgatgag 11100

tttttctaag ggcggcctgc caccataccc acgccgaaac aagcgctcat gagcccgaag 11160
tggcgagccc gatcttcccc atcggtgatg tcggcgatat aggcgccagc aaccgcacct 11220
gtggcgccgg tgatgagggc gcgccaagtc gacgtccggc agtc 11264

<210> 33
<211> 685
<212> PRT
<213> Artificial sequence
<220><223> Synthetic polypeptide
<400> 33

Met Ala Glu Trp Leu Leu Ser Ala Ser Trp Gln Arg Arg Ala Lys Ala
1 5 10 15
Met Thr Ala Ala Ala Gly Ser Ala Gly Arg Ala Ala Val Pro Leu Leu

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

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

145 150 155 160
Phe Asp Trp Pro Tyr Val Asn Leu Gln Leu Thr Ala Tyr Tyr Val Val
165 170 175
Thr Trp Ile Val Gly Ala Lys Arg Tyr His Asp Leu Asp Ile Asp Tyr
180 185 190
Ile Gly Ile Trp Asn Glu Arg Ser Tyr Asn Ala Asn Tyr Ile Lys Ile
195 200 205
Leu Arg Lys Met Leu Asn Tyr Gln Gly Leu Gln Arg Val Lys Ile Ile

210 215 220
Ala Ser Asp Asn Leu Trp Glu Ser Ile Ser Ala Ser Met Leu Leu Asp
225 230 235 240
Ala Glu Leu Phe Lys Val Val Asp Val Ile Gly Ala His Tyr Pro Gly

245 250 255

Thr His Ser Ala Lys Asp Ala Lys Leu Thr Gly Lys Lys Leu Trp Ser

260 265 270

Ser Glu Asp Phe Ser Thr Leu Asn Ser Asp Met Gly Ala Gly Cys Trp

275 280 285

Gly Arg Ile Leu Asn Gln Asn Tyr Ile Asn Gly Tyr Met Thr Ser Thr

290 295 300

Ile Ala Trp Asn Leu Val Ala Ser Tyr Tyr Glu Gln Leu Pro Tyr Gly

305 310 315 320

Arg Cys Gly Leu Met Thr Ala Gln Glu Pro Trp Ser Gly His Tyr Val

325 330 335

Val Glu Ser Pro Val Trp Val Ser Ala His Thr Thr Gln Phe Thr Gln

340 345 350

Pro Gly Trp Tyr Tyr Leu Lys Thr Val Gly His Leu Glu Lys Gly Gly

355 360 365

Ser Tyr Val Ala Leu Thr Asp Gly Leu Gly Asn Leu Thr Ile Ile Ile

370 375 380

Glu Thr Met Ser His Lys His Ser Lys Cys Ile Arg Pro Phe Leu Pro

385 390 395 400

Tyr Phe Asn Val Ser Gln Gln Phe Ala Thr Phe Val Leu Lys Gly Ser

405 410 415

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

420 425 430

Thr Ser Glu Arg Phe Leu Phe Lys Gln Leu Asp Ser Leu Trp Leu Leu

435 440 445

Asp Ser Asp Gly Ser Phe Thr Leu Ser Leu His Glu Asp Glu Leu Phe

450 455 460

Thr Leu Thr Thr Leu Thr Thr Gly Arg Lys Gly Ser Tyr Pro Leu Pro

465 470 475 480

Pro Lys Ser Gln Pro Phe Pro Ser Thr Tyr Lys Asp Asp Phe Asn Val

485 490 495

Asp Tyr Pro Phe Phe Ser Glu Ala Pro Asn Phe Ala Asp Gln Thr Gly

500 505 510

Val Phe Glu Tyr Phe Thr Asn Ile Glu Asp Pro Gly Glu His His Phe

515 520 525

Thr Leu Arg Gln Val Leu Asn Gln Arg Pro Ile Thr Trp Ala Ala Asp

530 535 540

Ala Ser Asn Thr Ile Ser Ile Ile Gly Asp Tyr Asn Trp Thr Asn Leu

545 550 555 560

Thr Ile Lys Cys Asp Val Tyr Ile Glu Thr Pro Asp Thr Gly Gly Val

565 570 575

Phe Ile Ala Gly Arg Val Asn Lys Gly Gly Ile Leu Ile Arg Ser Ala

580 585 590

Arg Gly Ile Phe Phe Trp Ile Phe Ala Asn Gly Ser Tyr Arg Val Thr

595 600 605

Gly Asp Leu Ala Gly Trp Ile Ile Tyr Ala Leu Gly Arg Val Glu Val

610 615 620

Thr Ala Lys Lys Trp Tyr Thr Leu Thr Leu Thr Ile Lys Gly His Phe

625 630 635 640

Thr Ser Gly Met Leu Asn Asp Lys Ser Leu Trp Thr Asp Ile Pro Val

645 650 655

Asn Phe Pro Lys Asn Gly Trp Ala Ala Ile Gly Thr His Ser Phe Glu

660 665 670

Phe Ala Gln Phe Asp Asn Phe Leu Val Glu Ala Thr Arg

675 680 685

<210> 34

<211> 2055

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 34

atggccgagt ggctgctgag cgccagctgg cagcgccgcg ccaaggccat gaccgccgcc 60

gccggcagcg ccggccgcgc cgccgtgccc ctgctgctgt gcgccctgct ggcccccggc 120

ggcgcctacg tgctggacga cagcgacggc ctgggccgcg agttcgacgg catcggcgcc 180

gtgagcggcg gcggcgccac cagccgcctg ctggtgaact accccgagcc ctaccgcagc 240
cagatcctgg actacctgtt caagcccaac ttcggcgcca gcctgcacat cctgaaggtg 300
gagatcggcg gcgacggcca gaccaccgac ggcaccgagc ccagccacat gcactacgcc 360
ctggacgaga actacttccg cggctacgag tggtggctga tgaaggaggc caagaagcgc 420
aaccccaaca tcaccctgat cggcctgccc tggagcttcc ccggctggct gggcaagggc 480
ttcgactggc cctacgtgaa cctgcagctg accgcctact acgtggtgac ctggatcgtg 540
ggcgccaagc gctaccacga cctggacatc gactacatcg gcatctggaa cgagcgcagc 600

tacaacgcca actacatcaa gatcctgcgc aagatgctga actaccaggg cctgcagcgc 660
gtgaagatca tcgccagcga caacctgtgg gagagcatca gcgccagcat gctgctggac 720
gccgagctgt tcaaggtggt ggacgtgatc ggcgcccact accccggcac ccacagcgcc 780
aaggacgcca agctgaccgg caagaagctg tggagcagcg aggacttcag caccctgaac 840
agcgacatgg gcgccggctg ctggggccgc atcctgaacc agaactacat caacggctac 900
atgaccagca ccatcgcctg gaacctggtg gccagctact acgagcagct gccctacggc 960
cgctgcggcc tgatgaccgc ccaggagccc tggagcggcc actacgtggt ggagagcccc 1020

gtgtgggtga gcgcccacac cacccagttc acccagcccg gctggtacta cctgaagacc 1080
gtgggccacc tggagaaggg cggcagctac gtggccctga ccgacggcct gggcaacctg 1140
accatcatca tcgagaccat gagccacaag cacagcaagt gcatccgccc cttcctgccc 1200
tacttcaacg tgagccagca gttcgccacc ttcgtgctga agggcagctt cagcgagatc 1260
cccgagctgc aggtgtggta caccaagctg ggcaagacca gcgagcgctt cctgttcaag 1320
cagctggaca gcctgtggct gctggacagc gacggcagct tcaccctgag cctgcacgag 1380
gacgagctgt tcaccctgac caccctgacc accggccgca agggcagcta ccccctgccc 1440

cccaagagcc agcccttccc cagcacctac aaggacgact tcaacgtgga ctacccttc 1500
ttcagcgagg cccccaactt cgccgaccag accggcgtgt tcgagtactt caccaacatc 1560
gaggaccccg gcgagcacca cttcaccctg cgccaggtgc tgaaccagcg ccccatcacc 1620
tgggccgccg acgccagcaa caccatcagc atcatcggcg actacaactg gaccaacctg 1680
accatcaagt gcgacgtgta catcgagacc cccgacaccg gcggcgtgtt catcgccggc 1740
cgcgtgaaca agggcggcat cctgatccgc agcgcccgcg gcatcttctt ctggatcttc 1800
gccaacggca gctaccgcgt gaccggcgac ctggccggct ggatcatcta cgccctgggc 1860

cgcgtggagg tgaccgccaa gaagtggtac accctgaccc tgaccatcaa gggccacttc 1920

accagcggca tgctgaacga caagagcctg tggaccgaca tccccgtgaa cttccccaag 1980

aacggctggg ccgccatcgg cacccacagc ttcgagttcg cccagttcga caacttcctg 2040

gtggaggcca cccgc 2055

<210> 35

<211> 339

<212> PRT

<213> Artificial sequence

<220><223> Synthetic polypeptide

<400> 35

Met Trp Gln Leu Trp Ala Ser Leu Cys Cys Leu Leu Val Leu Ala Asn
1 5 10 15

Ala Arg Ser Arg Pro Ser Phe His Pro Leu Ser Asp Glu Leu Val Asn
20 25 30

Tyr Val Asn Lys Arg Asn Thr Thr Trp Gln Ala Gly His Asn Phe Tyr
35 40 45

Asn Val Asp Met Ser Tyr Leu Lys Arg Leu Cys Gly Thr Phe Leu Gly
50 55 60

Gly Pro Lys Pro Pro Gln Arg Val Met Phe Thr Glu Asp Leu Lys Leu
65 70 75 80

Pro Ala Ser Phe Asp Ala Arg Glu Gln Trp Pro Gln Cys Pro Thr Ile
85 90 95

Lys Glu Ile Arg Asp Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Gly
100 105 110

Ala Val Glu Ala Ile Ser Asp Arg Ile Cys Ile His Thr Asn Ala His
115 120 125

Val Ser Val Glu Val Ser Ala Glu Asp Leu Leu Thr Cys Cys Gly Ser
130 135 140

Met Cys Gly Asp Gly Cys Asn Gly Gly Tyr Pro Ala Glu Ala Trp Asn
145 150 155 160

Phe Trp Thr Arg Lys Gly Leu Val Ser Gly Gly Leu Tyr Glu Ser His
165 170 175

Val Gly Cys Arg Pro Tyr Ser Ile Pro Pro Cys Glu His His Val Asn
180 185 190
Gly Ser Arg Pro Pro Cys Thr Gly Glu Gly Asp Thr Pro Lys Cys Ser
195 200 205

Lys Ile Cys Glu Pro Gly Tyr Ser Pro Thr Tyr Lys Gln Asp Lys His
210 215 220
Tyr Gly Tyr Asn Ser Tyr Ser Val Ser Asn Ser Glu Lys Asp Ile Met
225 230 235 240
Ala Glu Ile Tyr Lys Asn Gly Pro Val Glu Gly Ala Phe Ser Val Tyr
245 250 255
Ser Asp Phe Leu Leu Tyr Lys Ser Gly Val Tyr Gln His Val Thr Gly
260 265 270

Glu Met Met Gly Gly His Ala Ile Arg Ile Leu Gly Trp Gly Val Glu
275 280 285
Asn Gly Thr Pro Tyr Trp Leu Val Ala Asn Ser Trp Asn Thr Asp Trp
290 295 300
Gly Asp Asn Gly Phe Phe Lys Ile Leu Arg Gly Gln Asp His Cys Gly
305 310 315 320
Ile Glu Ser Glu Val Val Ala Gly Ile Pro Arg Thr Asp Gln Tyr Trp
325 330 335

Glu Lys Ile

<210> 36
<211> 1017
<212> DNA
<213> Artificial sequence
<220><223> Synthetic polynucleotide
<400> 36
atgtggcagc tgtgggccag cctgtgctgc ctgctggtgc tggccaacgc ccgcagccgc 60
cccagcttcc acccccctgag cgacgagctg gtgaactacg tgaacaagcg caacaccacc 120
tggcaggccg gccacaactt ctacaacgtg gacatgagct acctgaagcg cctgtgcggc 180
accttcctgg gcggccccaa gcccccccag cgcgtgatgt tcaccgagga cctgaagctg 240

cccgccagct tcgacgcccg cgagcagtgg ccccagtgcc ccaccatcaa ggagatccgc 300

gaccagggca gctgcggcag ctgctgggcc ttcggcgccg tggaggccat cagcgaccgc 360

atctgcatcc acaccaacgc ccacgtgagc gtggaggtga gcgccgagga cctgctgacc 420

tgctgcggca gcatgtgcgg cgacggctgc aacggcggct accccgccga ggcctggaac 480

ttctggaccc gcaagggcct ggtgagcggc ggcctgtacg agagccacgt gggctgccgc 540

ccctacagca tcccccccctg cgagcaccac gtgaacggca gccgcccccc ctgcaccggc 600

gagggcgaca cccccaagtg cagcaagatc tgcgagcccg gctacagccc cacctacaag 660

caggacaagc actacggcta caacagctac agcgtgagca acagcgagaa ggacatcatg 720

gccgagatct acaagaacgg ccccgtggag ggcgccttca gcgtgtacag cgacttcctg 780

ctgtacaaga gcggcgtgta ccagcacgtg accggcgaga tgatgggcgg ccacgccatc 840

cgcatcctgg gctggggcgt ggagaacggc acccccctact ggctggtggc caacagctgg 900

aacaccgact ggggcgacaa cggcttcttc aagatcctgc gcggccagga ccactgcggc 960

atcgagagcg aggtggtggc cggcatcccc cgcaccgacc agtactggga gaagatc 1017

<210> 37

<211> 631

<212> PRT

<213> Artificial sequence

<220><223> Synthetic polypeptide

<400> 37

Met Pro Arg Tyr Gly Ala Ser Leu Arg Gln Ser Cys Pro Arg Ser Gly

1 5 10 15

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

20 25 30

Met Gly Leu Val Leu Ala Leu Ala Leu Ala Leu Ala Leu Ala Leu Ala

35 40 45

Leu Ser Asp Ser Arg Val Leu Trp Ala Pro Ala Glu Ala His Pro Leu

50 55 60

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

65 70 75 80

Arg Asp Val Phe Gly Trp Gly Asn Leu Thr Cys Pro Ile Cys Lys Gly

85 90 95

Leu Phe Thr Ala Ile Asn Leu Gly Leu Lys Lys Glu Pro Asn Val Ala

100 105 110

Arg Val Gly Ser Val Ala Ile Lys Leu Cys Asn Leu Leu Lys Ile Ala

115 120 125

Pro Pro Ala Val Cys Gln Ser Ile Val His Leu Phe Glu Asp Asp Met

130 135 140

Val Glu Val Trp Arg Arg Ser Val Leu Ser Pro Ser Glu Ala Cys Gly

145 150 155 160

Leu Leu Leu Gly Ser Thr Cys Gly His Trp Asp Ile Phe Ser Ser Trp

165 170 175

Asn Ile Ser Leu Pro Thr Val Pro Lys Pro Pro Pro Lys Pro Pro Ser

180 185 190

Pro Pro Ala Pro Gly Ala Pro Val Ser Arg Ile Leu Phe Leu Thr Asp

195 200 205

Leu His Trp Asp His Asp Tyr Leu Glu Gly Thr Asp Pro Asp Cys Ala

210 215 220

Asp Pro Leu Cys Cys Arg Arg Gly Ser Gly Leu Pro Pro Ala Ser Arg

225 230 235 240

Pro Gly Ala Gly Tyr Trp Gly Glu Tyr Ser Lys Cys Asp Leu Pro Leu

245 250 255

Arg Thr Leu Glu Ser Leu Leu Ser Gly Leu Gly Pro Ala Gly Pro Phe

260 265 270

Asp Met Val Tyr Trp Thr Gly Asp Ile Pro Ala His Asp Val Trp His

275 280 285

Gln Thr Arg Gln Asp Gln Leu Arg Ala Leu Thr Thr Val Thr Ala Leu

290 295 300

Val Arg Lys Phe Leu Gly Pro Val Pro Val Tyr Pro Ala Val Gly Asn

305 310 315 320

His Glu Ser Thr Pro Val Asn Ser Phe Pro Pro Pro Phe Ile Glu Gly

325 330 335

Asn His Ser Ser Arg Trp Leu Tyr Glu Ala Met Ala Lys Ala Trp Glu

340 345 350

Pro Trp Leu Pro Ala Glu Ala Leu Arg Thr Leu Arg Ile Gly Gly Phe

355 360 365

Tyr Ala Leu Ser Pro Tyr Pro Gly Leu Arg Leu Ile Ser Leu Asn Met

370 375 380

Asn Phe Cys Ser Arg Glu Asn Phe Trp Leu Leu Ile Asn Ser Thr Asp

385 390 395 400

Pro Ala Gly Gln Leu Gln Trp Leu Val Gly Glu Leu Gln Ala Ala Glu

405 410 415

Asp Arg Gly Asp Lys Val His Ile Ile Gly His Ile Pro Pro Gly His

420 425 430

Cys Leu Lys Ser Trp Ser Trp Asn Tyr Tyr Arg Ile Val Ala Arg Tyr

435 440 445

Glu Asn Thr Leu Ala Ala Gln Phe Phe Gly His Thr His Val Asp Glu

450 455 460

Phe Glu Val Phe Tyr Asp Glu Glu Thr Leu Ser Arg Pro Leu Ala Val

465 470 475 480

Ala Phe Leu Ala Pro Ser Ala Thr Thr Tyr Ile Gly Leu Asn Pro Gly

485 490 495

Tyr Arg Val Tyr Gln Ile Asp Gly Asn Tyr Ser Gly Ser Ser His Val

500 505 510

Val Leu Asp His Glu Thr Tyr Ile Leu Asn Leu Thr Gln Ala Asn Ile

515 520 525

Pro Gly Ala Ile Pro His Trp Gln Leu Leu Tyr Arg Ala Arg Glu Thr

530 535 540

Tyr Gly Leu Pro Asn Thr Leu Pro Thr Ala Trp His Asn Leu Val Tyr

545 550 555 560

Arg Met Arg Gly Asp Met Gln Leu Phe Gln Thr Phe Trp Phe Leu Tyr

565 570 575

His Lys Gly His Pro Pro Ser Glu Pro Cys Gly Thr Pro Cys Arg Leu

580 585 590

Ala Thr Leu Cys Ala Gln Leu Ser Ala Arg Ala Asp Ser Pro Ala Leu

595 600 605

Cys Arg His Leu Met Pro Asp Gly Ser Leu Pro Glu Ala Gln Ser Leu

610 615 620

Trp Pro Arg Pro Leu Phe Cys

625 630

<210> 38

<211> 1896

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 38

atgccccgct acggcgccag cctgcgccag agctgccccc gcagcggccg cgagcagggc 60

caggacggca ccgccggcgc ccccggcctg ctgtggatgg gcctggtgct ggccctggcc 120

ctggccctgg ccctggccct ggccctgagc gacagccgcg tgctgtgggc ccccgccgag 180

gcccaccccc tgagccccca gggccacccc gcccgcctgc accgcatcgt gccccgcctg 240

cgcgacgtgt tcggctgggg caacctgacc tgccccatct gcaagggcct gttcaccgcc 300

atcaacctgg gcctgaagaa ggagcccaac gtggcccgcg tgggcagcgt ggccatcaag 360

ctgtgcaacc tgctgaagat cgcccccccc gccgtgtgcc agagcatcgt gcacctgttc 420

gaggacgaca tggtggaggt gtggcgccgc agcgtgctga gccccagcga ggcctgcggc 480

ctgctgctgg gcagcacctg cggccactgg gacatcttca gcagctggaa catcagcctg 540

cccaccgtgc ccaagccccc ccccaagccc cccagccccc ccgcccccgg cgcccccgtg 600

agccgcatcc tgttcctgac cgacctgcac tgggaccacg actacctgga gggcaccgac 660

cccgactgcg ccgacccccct gtgctgccgc cgcggcagcg gcctgccccc cgccagccgc 720

cccggcgccg gctactgggg cgagtacagc aagtgcgacc tgcccctgcg caccctggag 780

agcctgctga gcggcctggg ccccgccggc cccttcgaca tggtgtactg gaccggcgac 840

atccccgccc acgacgtgtg gcaccagacc cgccaggacc agctgcgcgc cctgaccacc 900

gtgaccgccc tggtgcgcaa gttcctgggc cccgtgcccg tgtaccccgc cgtgggcaac 960

cacgagagca cccccgtgaa cagcttcccc cccccccttca tcgagggcaa ccacagcagc 1020

cgctggctgt acgaggccat ggccaaggcc tgggagccct ggctgcccgc cgaggccctg 1080

cgcaccctgc gcatcggcgg cttctacgcc ctgagcccct accccggcct gcgcctgatc 1140

agcctgaaca tgaacttctg cagccgcgag aacttctggc tgctgatcaa cagcaccgac 1200

cccgccggcc agctgcagtg gctggtgggc gagctgcagg ccgccgagga ccgcggcgac 1260

aaggtgcaca tcatcggcca catccccccc ggccactgcc tgaagagctg gagctggaac 1320
tactaccgca tcgtggcccg ctacgagaac accctggccg cccagttctt cggccacacc 1380
cacgtggacg agttcgaggt gttctacgac gaggagaccc tgagccgccc cctggccgtg 1440
gccttcctgg cccccagcgc caccacctac atcggcctga accccggcta ccgcgtgtac 1500
cagatcgacg gcaactacag cggcagcagc cacgtggtgc tggaccacga gacctacatc 1560
ctgaacctga cccaggccaa catccccggc gccatccccc actggcagct gctgtaccgc 1620
gcccgcgaga cctacggcct gcccaacacc ctgcccaccg cctggcacaa cctggtgtac 1680

cgcatgcgcg gcgacatgca gctgttccag accttctggt tcctgtacca caagggccac 1740
ccccccagcg agccctgcgg cacccccctgc cgcctggcca ccctgtgcgc ccagctgagc 1800
gcccgcgccg acagccccgc cctgtgccgc cacctgatgc ccgacggcag cctgcccgag 1860
gcccagagcc tgtggccccg ccccctgttc tgctaa 1896

<210> 39
<211> 11329
<212> DNA
<213> Artificial sequence
<220><223> Synthetic polynucleotide
<400> 39

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240
tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300
tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360
atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420
tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480
agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctgcag ccctaggaat gcatctagac aattgtacta accttcttct 600
ctttcctctc ctgacagtcc ggaaagccac catggaattc agcagcccca gcagagagga 660
atgccccaag cctctgagcc gggtgtcaat catggccgga tctctgacag gactgctgct 720
gcttcaggcc gtgtcttggg cttctggcgc tagaccttgc atccccaaga gcttcggcta 780

cagcagcgtc gtgtgcgtgt gcaatgccac ctactgcgac agcttcgacc ctcctacctt 840

tcctgctctg ggcaccttca gcagatacga gagcaccaga tccggcagac ggatggaact 900

gagcatggga cccatccagg ccaatcacac aggcactggc ctgctgctga cactgcagcc 960

tgagcagaaa ttccagaaag tgaaaggctt cggcggagcc atgacagatg ccgccgctct 1020

gaatatcctg gctctgtctc caccagctca gaacctgctg ctcaagagct acttcagcga 1080

ggaaggcatc ggctacaaca tcatcagagt gcccatggcc agctgcgact tcagcatcag 1140

gacctacacc tacgccgaca cacccgacga tttccagctg cacaacttca gcctgcctga 1200

agaggacacc aagctgaaga tccctctgat ccacagagcc ctgcagctgg cacaaagacc 1260

cgtgtcactg ctggcctctc catggacatc tcccacctgg ctgaaaacaa atggcgccgt 1320

gaatggcaag ggcagcctga aaggccaacc tggcgacatc taccaccaga cctgggccag 1380

atacttcgtg aagttcctgg acgcctatgc cgagcacaag ctgcagtttt gggccgtgac 1440

agccgagaac gaaccttctg ctggactgct gagcggctac ccctttcagt gcctgggctt 1500

tacacccgag caccagcggg actttatcgc ccgtgatctg ggacccacac tggccaatag 1560

cacccaccat aatgtgcggc tgctgatgct ggacgaccag agactgcttc tgccccactg 1620

ggctaaagtg gtgctgacag atcctgaggc cgccaaatac gtgcacggaa tcgccgtgca 1680

ctggtatctg gactttctgg cccctgccaa ggccacactg ggagagacac acagactgtt 1740

ccccaacacc atgctgttcg ccagcgaagc ctgtgtgggc agcaagtttt gggaacagag 1800

cgtgcggctc ggcagctggg atagaggcat gcagtacagc cacagcatca tcaccaacct 1860

gctgtaccac gtcgtcggct ggaccgactg gaatctggcc ctgaatcctg aaggcggccc 1920

taactgggtc cgaaacttcg tggacagccc catcatcgtg gacatcacca aggacacctt 1980

ctacaagcag cccatgttct accacctggg acacttcagc aagttcatcc ccgagggctc 2040

tcagcgcgtt ggactggtgg cttcccagaa gaacgatctg gacgccgtgg ctctgatgca 2100

ccctgatgga tctgctgtgg tggtggtcct gaaccgcagc agcaaagatg tgcccctgac 2160

catcaaggat cccgccgtgg gattcctgga aacaatcagc cctggctact ccatccacac 2220

ctacctgtgg cgtagacagg agggcagagg aagtcttctg acatgcggag acgtggaaga 2280

gaatcccggc cctatggccg agtggctgct gagcgccagc tggcagcgcc gcgccaaggc 2340

catgaccgcc gccgccggca gcgccggccg cgccgccgtg cccctgctgc tgtgcgccct 2400

gctggccccc ggcggcgcct acgtgctgga cgacagcgac ggcctgggcc gcgagttcga 2460

cggcatcggc gccgtgagcg gcggcggcgc caccagccgc ctgctggtga actaccccga 2520

gccctaccgc agccagatcc tggactacct gttcaagccc aacttcggcg ccagcctgca 2580

catcctgaag gtggagatcg gcggcgacgg ccagaccacc gacggcaccg agcccagcca 2640

catgcactac gccctggacg agaactactt ccgcggctac gagtggtggc tgatgaagga 2700
ggccaagaag cgcaacccca acatcaccct gatcggcctg ccctggagct tccccggctg 2760
gctgggcaag ggcttcgact ggccctacgt gaacctgcag ctgaccgcct actacgtggt 2820
gacctggatc gtgggcgcca agcgctacca cgacctggac atcgactaca tcggcatctg 2880
gaacgagcgc agctacaacg ccaactacat caagatcctg cgcaagatgc tgaactacca 2940
gggcctgcag cgcgtgaaga tcatcgccag cgacaacctg tgggagagca tcagcgccag 3000
catgctgctg gacgccgagc tgttcaaggt ggtggacgtg atcggcgccc actaccccgg 3060

cacccacagc gccaaggacg ccaagctgac cggcaagaag ctgtggagca gcgaggactt 3120
cagcaccctg aacagcgaca tgggcgccgg ctgctggggc cgcatcctga accagaacta 3180
catcaacggc tacatgacca gcaccatcgc ctggaacctg gtggccagct actacgagca 3240
gctgccctac ggccgctgcg gcctgatgac cgcccaggag ccctggagcg gccactacgt 3300
ggtggagagc cccgtgtggg tgagcgccca caccacccag ttcacccagc ccggctggta 3360
ctacctgaag accgtgggcc acctggagaa gggcggcagc tacgtggccc tgaccgacgg 3420
cctgggcaac ctgaccatca tcatcgagac catgagccac aagcacagca agtgcatccg 3480

ccccttcctg ccctacttca acgtgagcca gcagttcgcc accttcgtgc tgaagggcag 3540
cttcagcgag atccccgagc tgcaggtgtg gtacaccaag ctgggcaaga ccagcgagcg 3600
cttcctgttc aagcagctgg acagcctgtg gctgctggac agcgacggca gcttcaccct 3660
gagcctgcac gaggacgagc tgttcaccct gaccaccctg accaccggcc gcaagggcag 3720
ctaccccctg ccccccaaga gccagccctt ccccagcacc tacaaggacg acttcaacgt 3780
ggactacccc ttcttcagcg aggcccccaa cttcgccgac cagaccggcg tgttcgagta 3840
cttcaccaac atcgaggacc ccggcgagca ccacttcacc ctgcgccagg tgctgaacca 3900

gcgccccatc acctgggccg ccgacgccag caacaccatc agcatcatcg gcgactacaa 3960
ctggaccaac ctgaccatca agtgcgacgt gtacatcgag acccccgaca ccggcggcgt 4020
gttcatcgcc ggccgcgtga acaagggcgg catcctgatc cgcagcgccc gcggcatctt 4080
cttctggatc ttcgccaacg gcagctaccg cgtgaccggc gacctggccg gctggatcat 4140
ctacgccctg ggccgcgtgg aggtgaccgc caagaagtgg tacaccctga ccctgaccat 4200
caagggccac ttcaccagcg gcatgctgaa cgacaagagc ctgtggaccg acatccccgt 4260
gaacttcccc aagaacggct gggccgccat cggcacccac agcttcgagt tcgcccagtt 4320

cgacaacttc ctggtggagg ccacccgctg acaattgtta attaagttta aaccctcgag 4380
gccgcaagca ataaaatatc tttattttca ttacatctgt gtgttggttt tttgtgtgga 4440
gatccacgat aacaaacagc ttttttgggg tgaacatatt gactgaattc cctgcaggtt 4500
ggccactccc tctctgcgcg ctcgctcgct cactgaggcc gcccgggcaa agcccgggcg 4560
tcgggcgacc tttggtcgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc 4620
caactccatc actaggggtt cctgcggccg ctcgtacggt ctcgaggaat tcctgcagga 4680
taacttgcca acctcattct aaaatgtata tagaagccca aaagacaata acaaaaatat 4740

tcttgtagaa caaaatggga aagaatgttc cactaaatat caagatttag agcaaagcat 4800
gagatgtgtg gggatagaca gtgaggctga taaaatagag tagagctcag aaacagaccc 4860
attgatatat gtaagtgacc tatgaaaaaa atatggcatt ttacaatggg aaaatgatgg 4920
tctttttctt ttttagaaaa acagggaaat atatttatat gtaaaaaata aaagggaacc 4980
catatgtcat accatacaca caaaaaaatt ccagtgaatt ataagtctaa atggagaagg 5040
caaaacttta aatcttttag aaaataatat agaagcatgc agaccagcct ggccaacatg 5100
atgaaaccct ctctactaat aataaaatca gtagaactac tcaggactac tttgagtggg 5160

aagtcctttt ctatgaagac ttctttggcc aaaattaggc tctaaatgca aggagatagt 5220
gcatcatgcc tggctgcact tactgataaa tgatgttatc accatcttta accaaatgca 5280
caggaacaag ttatggtact gatgtgctgg attgagaagg agctctactt ccttgacagg 5340
acacatttgt atcaacttaa aaaagcagat ttttgccagc agaactattc attcagaggt 5400
aggaaactta gaatagatga tgtcactgat tagcatggct tccccatctc cacagctgct 5460
tcccacccag gttgcccaca gttgagtttg tccagtgctc agggctgccc actctcagta 5520
agaagcccca caccagcccc tctccaaata tgttggctgt tccttccatt aaagtgaccc 5580

cactttagag cagcaagtgg atttctgttt cttacagttc aggaaggagg agtcagctgt 5640
gagaacctgg agcctgagat gcttctaagt cccactgcta ctggggtcag ggaagccaga 5700
ctccagcatc agcagtcagg agcactaagc ccttgccaac atcctgtttc tcagagaaac 5760
tgcttccatt ataatggttg tccttttta agctatcaag ccaaacaacc agtgtctacc 5820
attattctca tcacctgaag ccaagggttc tagcaaaagt caagctgtct tgtaatggtt 5880
gatgtgcctc cagcttctgt cttcagtcac tccactctta gcctgctctg aatcaactct 5940
gaccacagtt ccctggagcc cctgccacct gctgcccctg ccaccttctc catctgcagt 6000

gctgtgcagc cttctgcact cttgcagagc taataggtgg agacttgaag gaagaggagg 6060
aaagtttctc ataatagcct tgctgcaagc tcaaatggga ggtgggcact gtgcccagga 6120
gccttggagc aaaggctgtg cccaacctct gactgcatcc aggtttggtc ttgacagaga 6180

taagaagccc tggcttttgg agccaaaatc taggtcagac ttaggcagga ttctcaaagt 6240
ttatcagcag aacatgaggc agaagaccct ttctgctcca gcttcttcag gctcaacctt 6300
catcagaata gatagaaaga gaggctgtga gggttcttaa aacagaagca aatctgactc 6360
agagaataaa caacctccta gtaaactaca gcttagacag agcatctggt ggtgagtgtg 6420

ctcagtgtcc tactcaactg tctggtatca gccctcatga ggacttctct tctttccctc 6480
atagacctcc atctctgttt tccttagcct gcagaaatct ggatggctat tcacagaatg 6540
cctgtgcttt cagagttgca ttttttctct ggtattctgg ttcaagcatt tgaaggtagg 6600
aaaggttctc caagtgcaag aaagccagcc ctgagcctca actgcctggc tagtgtggtc 6660
agtaggatgc aaaggctgtt gaatgccaca aggccaaact ttaacctgtg taccacaagc 6720
ctagcagcag aggcagctct gctcactgga actctctgtc ttctttctcc tgagcctttt 6780
cttttcctga gttttctagc tctcctcaac cttacctctg ccctacccag gacaaaccca 6840

agagccactg tttctgtgat gtcctctcca gccctaatta ggcatcatga cttcagcctg 6900
accttccatg ctcagaagca gtgctaatcc acttcagatg agctgctcta tgcaacacag 6960
gcagagccta caaacctttg caccagagcc ctccacatat cagtgtttgt tcatactcac 7020
ttcaacagca aatgtgactg ctgagattaa gattttacac aagatggtct gtaatttcac 7080
agttagtttt atcccattag gtatgaaaga attagcataa ttccccttaa acatgaatga 7140
atcttagatt ttttaataaa tagttttgga agtaaagaca gagacatcag gagcacaagg 7200
aatagcctga gaggacaaac agaacaagaa agagtctgga aatacacagg atgttcttgg 7260

cctcctcaaa gcaagtgcaa gcagatagta ccagcagccc caggctatca gagcccagtg 7320
aagagaagta ccatgaaagc cacagctcta accaccctgt tccagagtga cagacagtcc 7380
ccaagacaag ccagcctgag ccagagagag aactgcaaga gaaagtttct aatttaggtt 7440
ctgttagatt cagacaagtg caggtcatcc tctctccaca gctactcacc tctccagcct 7500
aacaaagcct gcagtccaca ctccaaccct ggtgtctcac ctcctagcct ctcccaacat 7560
cctgctctct gaccatcttc tgcatctctc atctcaccat ctcccactgt ctacagccta 7620
ctcttgcaac taccatctca ttttctgaca tcctgtctac atcttctgcc atactctgcc 7680

atctaccata ccacctctta ccatctacca caccatcttt tatctccatc cctctcagaa 7740
gcctccaagc tgaatcctgc tttatgtgtt catctcagcc cctgcatgga aagctgaccc 7800
cagaggcaga actattccca gagagcttgg ccaagaaaaa caaaactacc agcctggcca 7860
ggctcaggag tagtaagctg cagtgtctgt tgtgttctag cttcaacagc tgcaggagtt 7920
ccactctcaa atgctccaca tttctcacat cctcctgatt ctggtcacta cccatcttca 7980
aagaacagaa tatctcacat cagcatactg tgaaggacta gtcatgggtg cagctgctca 8040

gagctgcaaa gtcattctgg atggtggaga gcttacaaac atttcatgat gctccccccg 8100

ctctgatggc tggagcccaa tccctacaca gactcctgct gtatgtgttt tcctttcact 8160
ctgagccaca gccagagggc aggcattcag tctcctcttc aggctggggc tggggcactg 8220
agaactcacc caacaccttg ctctcactcc ttctgcaaaa caagaaagag ctttgtgctg 8280
cagtagccat gaagaatgaa aggaaggctt taactaaaaa atgtcagaga ttattttcaa 8340
ccccttactg tggatcacca gcaaggagga aacacaacac agagacattt tttcccctca 8400
aattatcaaa agaatcactg catttgttaa agagagcaac tgaatcagga agcagagttt 8460
tgaacatatc agaagttagg aatctgcatc agagacaaat gcagtcatgg ttgtttgctg 8520

cataccagcc ctaatcatta gaagcctcat ggacttcaaa catcattccc tctgacaaga 8580
tgctctagcc taactccatg agataaaata aatctgcctt tcagagccaa agaagagtcc 8640
accagcttct tctcagtgtg aacaagagct ccagtcaggt tagtcagtcc agtgcagtag 8700
aggagaccag tctgcatcct ctaattttca aaggcaagaa gatttgttta ccctggacac 8760
caggcacaag tgaggtcaca gagctcttag atatgcagtc ctcatgagtg aggagactaa 8820
agcgcatgcc atcaagactt cagtgtagag aaaacctcca aaaaagcctc ctcactactt 8880
ctggaatagc tcagaggccg aggcggcctc ggcctctgca taaataaaaa aaattagtca 8940

gccatggggc ggagaatggg cggaactggg cggagttagg ggcgggatgg gcggagttag 9000
gggcgggact atggttgctg actaattgag atgcatgctt tgcatacttc tgcctgctgg 9060
ggagcctggg gactttccac acctggttgc tgactaattg agatgcatgc tttgcatact 9120
tctgcctgct ggggagcctg gggactttcc acaccctaac tgacacacat tccacagctg 9180
cattaatgaa tcggccaacg cgcggggaga ggcggtttgc gtattgggcg ctcttccgct 9240
tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac 9300
tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga 9360

gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat 9420
aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac 9480
ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct 9540
gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg 9600
ctttctcata gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg 9660
ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt 9720
cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg 9780

attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac 9840

ggctacacta gaagaacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga 9900

aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggtttttt 9960

gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt 10020

tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga 10080

ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc 10140

taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct 10200

atctcagcga tctgtctatt tcgttcatcc atagttgcct gactcctgca aaccacgttg 10260

tgtctcaaaa tctctgatgt tacattgcac aagataaaaa tatatcatca tgaacaataa 10320

aactgtctgc ttacataaac agtaatacaa ggggtgttat gagccatatt caacgggaaa 10380

cgtcttgctc gaggccgcga ttaaattcca acatggatgc tgatttatat gggtataaat 10440

gggctcgcga taatgtcggg caatcaggtg cgacaatcta tcgattgtat gggaagcccg 10500

atgcgccaga gttgtttctg aaacatggca aaggtagcgt tgccaatgat gttacagatg 10560

agatggtcag actaaactgg ctgacggaat ttatgcctct tccgaccatc aagcatttta 10620

tccgtactcc tgatgatgca tggttactca ccactgcgat ccccgggaaa acagcattcc 10680

aggtattaga agaatatcct gattcaggtg aaaatattgt tgatgcgctg gcagtgttcc 10740

tgcgccggtt gcattcgatt cctgtttgta attgtccttt taacagcgat cgcgtatttc 10800

gtctcgctca ggcgcaatca cgaatgaata acggtttggt tgatgcgagt gattttgatg 10860

acgagcgtaa tggctggcct gttgaacaag tctggaaaga aatgcataag cttttgccat 10920

tctcaccgga ttcagtcgtc actcatggtg atttctcact tgataacctt atttttgacg 10980

aggggaaatt aataggttgt attgatgttg gacgagtcgg aatcgcagac cgataccagg 11040

atcttgccat cctatggaac tgcctcggtg agttttctcc ttcattacag aaacggcttt 11100

ttcaaaaata tggtattgat aatcctgata tgaataaatt gcagtttcat ttgatgctcg 11160

atgagttttt ctaagggcgg cctgccacca tacccacgcc gaaacaagcg ctcatgagcc 11220

cgaagtggcg agcccgatct tccccatcgg tgatgtcggc gatataggcg ccagcaaccg 11280

cacctgtggc gccggtgatg agggcgcgcc aagtcgacgt ccggcagtc 11329

<210> 40

<211> 11776

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 40

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120
gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240
tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300
tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360
atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420
tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540
tggcactatg aaccctgcag ccctaggaat gcatctagac aattgtacta accttcttct 600
ctttcctctc ctgacagtcc ggaaagccac catggccgag tggctgctga gcgccagctg 660
gcagcgccgc gccaaggcca tgaccgccgc cgccggcagc gccggccgcg ccgccgtgcc 720
cctgctgctg tgcgccctgc tggcccccgg cggcgcctac gtgctggacg acagcgacgg 780
cctgggccgc gagttcgacg gcatcggcgc cgtgagcggc ggcggcgcca ccagccgcct 840
gctggtgaac taccccgagc cctaccgcag ccagatcctg gactacctgt tcaagcccaa 900

cttcggcgcc agcctgcaca tcctgaaggt ggagatcggc ggcgacggcc agaccaccga 960
cggcaccgag cccagccaca tgcactacgc cctggacgag aactacttcc gcggctacga 1020
gtggtggctg atgaaggagg ccaagaagcg caaccccaac atcaccctga tcggcctgcc 1080
ctggagcttc cccggctggc tgggcaaggg cttcgactgg ccctacgtga acctgcagct 1140
gaccgcctac tacgtggtga cctggatcgt gggcgccaag cgctaccacg acctggacat 1200
cgactacatc ggcatctgga acgagcgcag ctacaacgcc aactacatca agatcctgcg 1260
caagatgctg aactaccagg gcctgcagcg cgtgaagatc atcgccagcg acaacctgtg 1320

ggagagcatc agcgccagca tgctgctgga cgccgagctg ttcaaggtgg tggacgtgat 1380
cggcgcccac taccccggca cccacagcgc caaggacgcc aagctgaccg gcaagaagct 1440
gtggagcagc gaggacttca gcaccctgaa cagcgacatg ggcgccggct gctggggccg 1500
catcctgaac cagaactaca tcaacggcta catgaccagc accatcgcct ggaacctggt 1560
ggccagctac tacgagcagc tgccctacgg ccgctgcggc ctgatgaccg cccaggagcc 1620
ctggagcggc cactacgtgg tggagagccc cgtgtgggtg agcgcccaca ccacccagtt 1680
cacccagccc ggctggtact acctgaagac cgtgggccac ctggagaagg gcggcagcta 1740

cgtggccctg accgacggcc tgggcaacct gaccatcatc atcgagacca tgagccacaa 1800
gcacagcaag tgcatccgcc ccttcctgcc ctacttcaac gtgagccagc agttcgccac 1860
cttcgtgctg aagggcagct tcagcgagat ccccgagctg caggtgtggt acaccaagct 1920
gggcaagacc agcgagcgct tcctgttcaa gcagctggac agcctgtggc tgctggacag 1980
cgacggcagc ttcaccctga gcctgcacga ggacgagctg ttcaccctga ccaccctgac 2040
caccggccgc aagggcagct acccccctgcc ccccaagagc cagcccttcc ccagcaccta 2100
caaggacgac ttcaacgtgg actacccctt cttcagcgag gcccccaact tcgccgacca 2160

gaccggcgtg ttcgagtact tcaccaacat cgaggacccc ggcgagcacc acttcaccct 2220
gcgccaggtg ctgaaccagc gccccatcac ctgggccgcc gacgccagca acaccatcag 2280
catcatcggc gactacaact ggaccaacct gaccatcaag tgcgacgtgt acatcgagac 2340
ccccgacacc ggcggcgtgt tcatcgccgg ccgcgtgaac aagggcggca tcctgatccg 2400
cagcgcccgc ggcatcttct tctggatctt cgccaacggc agctaccgcg tgaccggcga 2460
cctggccggc tggatcatct acgccctggg ccgcgtggag gtgaccgcca agaagtggta 2520
caccctgacc ctgaccatca agggccactt caccagcggc atgctgaacg acaagagcct 2580

gtggaccgac atccccgtga acttccccaa gaacggctgg gccgccatcg gcacccacag 2640
cttcgagttc gcccagttcg acaacttcct ggtggaggcc acccgctgat tgtggccgaa 2700
ccgccgaact cagaggccgg ccccagaaaa cccgagcgag taggggggcgg cgcgcaggag 2760
ggaggagaac tgggggcgcg ggaggctggt gggtgtgggg ggtggagatg tagaagatgt 2820
gacgccgcgg cccggcgggt gccagattag cggacgcggt gcccgcggtt gcaacgggat 2880
cccgggcgct gcagcttggg aggcggctct ccccaggcgg cgtccgcgga gacacccatc 2940
cgtgaacccc aggtcccggg ccgccggctc gccgcgcacc aggggccggc ggacagaaga 3000

gcggccgagc ggctcgaggc tgggggaccg cgggcgcggc cgcgcgctgc cgggcgggag 3060
gctgggggggc cggggccggg gccgtgcccc ggagcgggtc ggaggccggg gccggggccg 3120
ggggacggcg gctccccgcg cggctccagc ggctcgggga tcccggccgg gccccgcagg 3180
gaccatgatg gaattcagca gccccagcag agaggaatgc cccaagcctc tgagccgggt 3240
gtcaatcatg gccggatctc tgacaggact gctgctgctt caggccgtgt cttgggcttc 3300
tggcgctaga ccttgcatcc ccaagagctt cggctacagc agcgtcgtgt gcgtgtgcaa 3360
tgccacctac tgcgacagct tcgaccctcc tacctttcct gctctgggca ccttcagcag 3420

atacgagagc accagatccg gcagacggat ggaactgagc atgggaccca tccaggccaa 3480
tcacacaggc actggcctgc tgctgacact gcagcctgag cagaaattcc agaaagtgaa 3540
aggcttcggc ggagccatga cagatgccgc cgctctgaat atcctggctc tgtctccacc 3600

agctcagaac ctgctgctca agagctactt cagcgaggaa ggcatcggct acaacatcat 3660
cagagtgccc atggccagct gcgacttcag catcaggacc tacacctacg ccgacacacc 3720
cgacgatttc cagctgcaca acttcagcct gcctgaagag gacaccaagc tgaagatccc 3780
tctgatccac agagccctgc agctggcaca aagacccgtg tcactgctgg cctctccatg 3840

gacatctccc acctggctga aaacaaatgg cgccgtgaat ggcaagggca gcctgaaagg 3900
ccaacctggc gacatctacc accagacctg ggccagatac ttcgtgaagt tcctggacgc 3960
ctatgccgag cacaagctgc agttttgggc cgtgacagcc gagaacgaac cttctgctgg 4020
actgctgagc ggctacccct ttcagtgcct gggctttaca cccgagcacc agcgggactt 4080
tatcgcccgt gatctgggac ccacactggc caatagcacc caccataatg tgcggctgct 4140
gatgctggac gaccagagac tgcttctgcc ccactgggct aaagtggtgc tgacagatcc 4200
tgaggccgcc aaatacgtgc acggaatcgc cgtgcactgg tatctggact ttctggcccc 4260

tgccaaggcc acactgggag agacacacag actgttcccc aacaccatgc tgttcgccag 4320
cgaagcctgt gtgggcagca agttttggga acagagcgtg cggctcggca gctgggatag 4380
aggcatgcag tacagccaca gcatcatcac caacctgctg taccacgtcg tcggctggac 4440
cgactggaat ctggccctga atcctgaagg cggccctaac tgggtccgaa acttcgtgga 4500
cagccccatc atcgtggaca tcaccaagga caccttctac aagcagccca tgttctacca 4560
cctgggacac ttcagcaagt tcatccccga gggctctcag cgcgttggac tggtggcttc 4620
ccagaagaac gatctggacg ccgtggctct gatgcaccct gatggatctg ctgtggtggt 4680

ggtcctgaac cgcagcagca aagatgtgcc cctgaccatc aaggatcccg ccgtgggatt 4740
cctggaaaca atcagccctg gctactccat ccacacctac ctgtggcgta gacagtgaca 4800
attgttaatt aagtttaaac cctcgaggcc gcaagcaata aaatatcttt attttcatta 4860
catctgtgtg ttggtttttt gtgtggagat ccacgataac aaacagcttt tttggggtga 4920
acatattgac tgaattccct gcaggttggc cactccctct ctgcgcgctc gctcgctcac 4980
tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt ggtcgcccgg cctcagtgag 5040
cgagcgagcg cgcagagagg gagtggccaa ctccatcact aggggttcct gcggccgctc 5100

gtacggtctc gaggaattcc tgcaggataa cttgccaacc tcattctaaa atgtatatag 5160
aagcccaaaa gacaataaca aaaatattct tgtagaacaa aatgggaaag aatgttccac 5220
taaatatcaa gatttagagc aaagcatgag atgtgtgggg atagacagtg aggctgataa 5280
aatagagtag agctcagaaa cagacccatt gatatatgta agtgacctat gaaaaaaata 5340
tggcatttta caatgggaaa atgatggtct ttttcttttt tagaaaaaca gggaaatata 5400
tttatatgta aaaaataaaa gggaacccat atgtcatacc atacacacaa aaaaattcca 5460

gtgaattata agtctaaatg gagaaggcaa aactttaaat cttttagaaa ataatataga 5520

agcatgcaga ccagcctggc caacatgatg aaaccctctc tactaataat aaaatcagta 5580
gaactactca ggactacttt gagtgggaag tccttttcta tgaagacttc tttggccaaa 5640
attaggctct aaatgcaagg agatagtgca tcatgcctgg ctgcacttac tgataaatga 5700
tgttatcacc atctttaacc aaatgcacag gaacaagtta tggtactgat gtgctggatt 5760
gagaaggagc tctacttcct tgacaggaca catttgtatc aacttaaaaa agcagatttt 5820
tgccagcaga actattcatt cagaggtagg aaacttagaa tagatgatgt cactgattag 5880
catggcttcc ccatctccac agctgcttcc cacccaggtt gcccacagtt gagtttgtcc 5940

agtgctcagg gctgcccact ctcagtaaga agccccacac cagcccctct ccaaatatgt 6000
tggctgttcc ttccattaaa gtgaccccac tttagagcag caagtggatt tctgtttctt 6060
acagttcagg aaggaggagt cagctgtgag aacctggagc ctgagatgct tctaagtccc 6120
actgctactg gggtcaggga agccagactc cagcatcagc agtcaggagc actaagccct 6180
tgccaacatc ctgtttctca gagaaactgc ttccattata atggttgtcc ttttttaagc 6240
tatcaagcca aacaaccagt gtctaccatt attctcatca cctgaagcca agggttctag 6300
caaaagtcaa gctgtcttgt aatggttgat gtgcctccag cttctgtctt cagtcactcc 6360

actcttagcc tgctctgaat caactctgac cacagttccc tggagcccct gccacctgct 6420
gcccctgcca ccttctccat ctgcagtgct gtgcagcctt ctgcactctt gcagagctaa 6480
taggtggaga cttgaaggaa gaggaggaaa gtttctcata atagccttgc tgcaagctca 6540
aatgggaggt gggcactgtg cccaggagcc ttggagcaaa ggctgtgccc aacctctgac 6600
tgcatccagg tttggtcttg acagagataa gaagccctgg cttttggagc caaaatctag 6660
gtcagactta ggcaggattc tcaaagttta tcagcagaac atgaggcaga agaccctttc 6720
tgctccagct tcttcaggct caaccttcat cagaatagat agaaagagag gctgtgaggg 6780

ttcttaaaac agaagcaaat ctgactcaga gaataaacaa cctcctagta aactacagct 6840
tagacagagc atctggtggt gagtgtgctc agtgtcctac tcaactgtct ggtatcagcc 6900
ctcatgagga cttctcttct ttccctcata gacctccatc tctgttttcc ttagcctgca 6960
gaaatctgga tggctattca cagaatgcct gtgctttcag agttgcattt tttctctggt 7020
attctggttc aagcatttga aggtaggaaa ggttctccaa gtgcaagaaa gccagccctg 7080
agcctcaact gcctggctag tgtggtcagt aggatgcaaa ggctgttgaa tgccacaagg 7140
ccaaacttta acctgtgtac cacaagccta gcagcagagg cagctctgct cactggaact 7200

ctctgtcttc tttctcctga gccttttctt ttcctgagtt ttctagctct cctcaacctt 7260

acctctgccc tacccaggac aaacccaaga gccactgttt ctgtgatgtc ctctccagcc 7320

ctaattaggc atcatgactt cagcctgacc ttccatgctc agaagcagtg ctaatccact 7380

tcagatgagc tgctctatgc aacacaggca gagcctacaa acctttgcac cagagccctc 7440

cacatatcag tgtttgttca tactcacttc aacagcaaat gtgactgctg agattaagat 7500

tttacacaag atggtctgta atttcacagt tagttttatc ccattaggta tgaaagaatt 7560

agcataattc cccttaaaca tgaatgaatc ttagattttt taataaatag ttttggaagt 7620

aaagacagag acatcaggag cacaaggaat agcctgagag gacaaacaga acaagaaaga 7680

gtctggaaat acacaggatg ttcttggcct cctcaaagca agtgcaagca gatagtacca 7740

gcagccccag gctatcagag cccagtgaag agaagtacca tgaaagccac agctctaacc 7800

accctgttcc agagtgacag acagtcccca agacaagcca gcctgagcca gagagagaac 7860

tgcaagagaa agtttctaat ttaggttctg ttagattcag acaagtgcag gtcatcctct 7920

ctccacagct actcacctct ccagcctaac aaagcctgca gtccacactc caaccctggt 7980

gtctcacctc ctagcctctc ccaacatcct gctctctgac catcttctgc atctctcatc 8040

tcaccatctc ccactgtcta cagcctactc ttgcaactac catctcattt tctgacatcc 8100

tgtctacatc ttctgccata ctctgccatc taccatacca cctcttacca tctaccacac 8160

catcttttat ctccatccct ctcagaagcc tccaagctga atcctgcttt atgtgttcat 8220

ctcagcccct gcatggaaag ctgaccccag aggcagaact attcccagag agcttggcca 8280

agaaaaacaa aactaccagc ctggccaggc tcaggagtag taagctgcag tgtctgttgt 8340

gttctagctt caacagctgc aggagttcca ctctcaaatg ctccacattt ctcacatcct 8400

cctgattctg gtcactaccc atcttcaaag aacagaatat ctcacatcag catactgtga 8460

aggactagtc atgggtgcag ctgctcagag ctgcaaagtc attctggatg gtggagagct 8520

tacaaacatt tcatgatgct ccccccgctc tgatggctgg agcccaatcc ctacacagac 8580

tcctgctgta tgtgttttcc tttcactctg agccacagcc agagggcagg cattcagtct 8640

cctcttcagg ctggggctgg ggcactgaga actcacccaa caccttgctc tcactccttc 8700

tgcaaaacaa gaaagagctt tgtgctgcag tagccatgaa gaatgaaagg aaggctttaa 8760

ctaaaaaatg tcagagatta ttttcaaccc cttactgtgg atcaccagca aggaggaaac 8820

acaacacaga gacatttttt cccctcaaat tatcaaaaga atcactgcat ttgttaaaga 8880

gagcaactga atcaggaagc agagttttga acatatcaga agttaggaat ctgcatcaga 8940

gacaaatgca gtcatggttg tttgctgcat accagcccta atcattagaa gcctcatgga 9000

cttcaaacat cattccctct gacaagatgc tctagcctaa ctccatgaga taaaataaat 9060

ctgcctttca gagccaaaga agagtccacc agcttcttct cagtgtgaac aagagctcca 9120
gtcaggttag tcagtccagt gcagtagagg agaccagtct gcatcctcta attttcaaag 9180
gcaagaagat ttgtttaccc tggacaccag gcacaagtga ggtcacagag ctcttagata 9240
tgcagtcctc atgagtgagg agactaaagc gcatgccatc aagacttcag tgtagagaaa 9300

acctccaaaa aagcctcctc actacttctg gaatagctca gaggccgagg cggcctcggc 9360
ctctgcataa ataaaaaaaa ttagtcagcc atggggcgga gaatgggcgg aactgggcgg 9420
agttaggggc gggatgggcg gagttagggg cgggactatg gttgctgact aattgagatg 9480
catgctttgc atacttctgc ctgctgggga gcctggggac tttccacacc tggttgctga 9540
ctaattgaga tgcatgcttt gcatacttct gcctgctggg gagcctgggg actttccaca 9600
ccctaactga cacacattcc acagctgcat taatgaatcg gccaacgcgc ggggagaggc 9660
ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt 9720

cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca 9780
ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa 9840
aaggccgcgt tgctggcgtt tttccatagg ctccgccccc ctgacgagca tcacaaaaat 9900
cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc 9960
cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc 10020
gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt 10080
tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aaccccccgt tcagcccgac 10140

cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg 10200
ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca 10260
gagttcttga agtggtggcc taactacggc tacactagaa gaacagtatt tggtatctgc 10320
gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacaa 10380
accaccgctg gtagcggtgg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa 10440
ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac 10500
tcacgttaag ggatttggt catgagatta tcaaaaagga tcttcaccta gatcctttta 10560

aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt 10620
taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata 10680
gttgcctgac tcctgcaaac cacgttgtgt ctcaaaatct ctgatgttac attgcacaag 10740
ataaaaatat atcatcatga acaataaaac tgtctgctta cataaacagt aatacaaggg 10800
gtgttatgag ccatattcaa cgggaaacgt cttgctcgag gccgcgatta aattccaaca 10860
tggatgctga tttatatggg tataaatggg ctcgcgataa tgtcgggcaa tcaggtgcga 10920

caatctatcg attgtatggg aagcccgatg cgccagagtt gtttctgaaa catggcaaag 10980

gtagcgttgc caatgatgtt acagatgaga tggtcagact aaactggctg acggaattta 11040
tgcctcttcc gaccatcaag cattttatcc gtactcctga tgatgcatgg ttactcacca 11100
ctgcgatccc cgggaaaaca gcattccagg tattagaaga atatcctgat tcaggtgaaa 11160
atattgttga tgcgctggca gtgttcctgc gccggttgca ttcgattcct gtttgtaatt 11220
gtccttttaa cagcgatcgc gtatttcgtc tcgctcaggc gcaatcacga atgaataacg 11280
gtttggttga tgcgagtgat tttgatgacg agcgtaatgg ctggcctgtt gaacaagtct 11340
ggaaagaaat gcataagctt ttgccattct caccggattc agtcgtcact catggtgatt 11400

tctcacttga taaccttatt tttgacgagg ggaaattaat aggttgtatt gatgttggac 11460
gagtcggaat cgcagaccga taccaggatc ttgccatcct atggaactgc ctcggtgagt 11520
tttctccttc attacagaaa cggctttttc aaaaatatgg tattgataat cctgatatga 11580
ataaattgca gtttcatttg atgctcgatg agtttttcta agggcggcct gccaccatac 11640
ccacgccgaa acaagcgctc atgagcccga agtggcgagc ccgatcttcc ccatcggtga 11700
tgtcggcgat ataggcgcca gcaaccgcac ctgtggcgcc ggtgatgagg gcgcgccaag 11760
tcgacgtccg gcagtc 11776

<210> 41
<211> 11348
<212> DNA
<213> Artificial sequence
<220><223> Synthetic polynucleotide
<400> 41
ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120
gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240
tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300
tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420
tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480
agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540
tggcactatg aaccctcctg gtggcgaggg gagggggtg gtcctcgaac gccttgcaga 600

actggcctgg atacagagtg gaccggctgg ccccatctgg aagacttcga gatacactgt 660
tgtcttactg cgctcaacag tgtatctcga agtcttccaa atggtgccag ccatcgcagc 720
ggggtgcagg aaatgggggc agcccccctt tttggctatc cttccacgtg ttctttttg 780

tatcttttgt gtttcctaga aaacatctca gtcaccaccg cagccctagg aatgcatcta 840
gacaattgta ctaaccttct tctctttcct ctcctgacag tccggaaagc caccatgtgg 900
cagctgtggg ccagcctgtg ctgcctgctg gtgctggcca acgcccgcag ccgccccagc 960
ttccaccccc tgagcgacga gctggtgaac tacgtgaaca agcgcaacac cacctggcag 1020
gccggccaca acttctacaa cgtggacatg agctacctga agcgcctgtg cggcaccttc 1080
ctgggcggcc ccaagccccc ccagcgcgtg atgttcaccg aggacctgaa gctgcccgcc 1140
agcttcgacg cccgcgagca gtggccccag tgccccacca tcaaggagat ccgcgaccag 1200

ggcagctgcg gcagctgctg ggccttcggc gccgtggagg ccatcagcga ccgcatctgc 1260
atccacacca acgcccacgt gagcgtggag gtgagcgccg aggacctgct gacctgctgc 1320
ggcagcatgt gcggcgacgg ctgcaacggc ggctaccccg ccgaggcctg gaacttctgg 1380
acccgcaagg gcctggtgag cggcggcctg tacgagagcc acgtgggctg ccgcccctac 1440
agcatccccc cctgcgagca ccacgtgaac ggcagccgcc ccccctgcac cggcgagggc 1500
gacaccccca agtgcagcaa gatctgcgag cccggctaca gccccaccta caagcaggac 1560
aagcactacg gctacaacag ctacagcgtg agcaacagcg agaaggacat catggccgag 1620

atctacaaga acggccccgt ggagggcgcc ttcagcgtgt acagcgactt cctgctgtac 1680
aagagcggcg tgtaccagca cgtgaccggc gagatgatgg gcggccacgc catccgcatc 1740
ctgggctggg gcgtggagaa cggcaccccc tactggctgg tggccaacag ctggaacacc 1800
gactggggcg acaacggctt cttcaagatc ctgcgcggcc aggaccactg cggcatcgag 1860
agcgaggtgg tggccggcat cccccgcacc gaccagtact gggagaagat cgagggcaga 1920
ggaagtcttc tgacatgcgg agacgtggaa gagaatcccg gccctatgga attcagcagc 1980
cccagcagag aggaatgccc caagcctctg agccgggtgt caatcatggc cggatctctg 2040

acaggactgc tgctgcttca ggccgtgtct tgggcttctg gcgctagacc ttgcatcccc 2100
aagagcttcg gctacagcag cgtcgtgtgc gtgtgcaatg ccacctactg cgacagcttc 2160
gaccctccta cctttcctgc tctgggcacc ttcagcagat acgagagcac cagatccggc 2220
agacggatgg aactgagcat gggacccatc caggccaatc acacaggcac tggcctgctg 2280
ctgacactgc agcctgagca gaaattccag aaagtgaaag gcttcggcgg agccatgaca 2340
gatgccgccg ctctgaatat cctggctctg tctccaccag ctcagaacct gctgctcaag 2400

agctacttca gcgaggaagg catcggctac aacatcatca gagtgcccat ggccagctgc 2460

gacttcagca tcaggaccta cacctacgcc gacacacccg acgatttcca gctgcacaac 2520
ttcagcctgc ctgaagagga caccaagctg aagatccctc tgatccacag agccctgcag 2580
ctggcacaaa gacccgtgtc actgctggcc tctccatgga catctcccac ctggctgaaa 2640
acaaatggcg ccgtgaatgg caagggcagc ctgaaaggcc aacctggcga catctaccac 2700
cagacctggg ccagatactt cgtgaagttc ctggacgcct atgccgagca caagctgcag 2760
ttttgggccg tgacagccga gaacgaacct tctgctggac tgctgagcgg ctaccccttt 2820
cagtgcctgg gctttacacc cgagcaccag cgggacttta tcgcccgtga tctgggaccc 2880

acactggcca atagcaccca ccataatgtg cggctgctga tgctggacga ccagagactg 2940
cttctgcccc actgggctaa agtggtgctg acagatcctg aggccgccaa atacgtgcac 3000
ggaatcgccg tgcactggta tctggacttt ctggcccctg ccaaggccac actgggagag 3060
acacacagac tgttccccaa caccatgctg ttcgccagcg aagcctgtgt gggcagcaag 3120
ttttgggaac agagcgtgcg gctcggcagc tgggatagag gcatgcagta cagccacagc 3180
atcatcacca acctgctgta ccacgtcgtc ggctggaccg actggaatct ggccctgaat 3240
cctgaaggcg gccctaactg ggtccgaaac ttcgtggaca gccccatcat cgtggacatc 3300

accaaggaca ccttctacaa gcagcccatg ttctaccacc tgggacactt cagcaagttc 3360
atccccgagg gctctcagcg cgttggactg gtggcttccc agaagaacga tctggacgcc 3420
gtggctctga tgcaccctga tggatctgct gtggtggtgg tcctgaaccg cagcagcaaa 3480
gatgtgcccc tgaccatcaa ggatcccgcc gtgggattcc tggaaacaat cagccctggc 3540
tactccatcc acacctacct gtggcgtaga cagtgacaat tgttaattaa gtttaaaccc 3600
tcgaggccgc aagcttatcg ataatcaacc tctggattac aaaatttgtg aaagattgac 3660
tggtattctt aactatgttg ctccttttac gctatgtgga tacgctgctt taatgccttt 3720

gtatcatgct attgcttccc gtatggcttt cattttctcc tccttgtata aatcctggtt 3780
gctgtctctt tatgaggagt tgtggcccgt tgtcaggcaa cgtggcgtgg tgtgcactgt 3840
gtttgctgac gcaaccccca ctggttgggg cattgccacc acctgtcagc tcctttccgg 3900
gactttcgct ttccccctcc ctattgccac ggcggaactc atcgccgcct gccttgcccg 3960
ctgctggaca ggggctcggc tgttgggcac tgacaattcc gtggtgttgt cggggaaatc 4020
atcgtccttt ccttggctgc tcgcctgtgt tgccacctgg attctgcgcg ggacgtcctt 4080
ctgctacgtc ccttcggccc tcaatccagc ggaccttcct tcccgcggcc tgctgccggc 4140

tctgcggcct cttccgcgtc ttcgccttcg ccctcagacg agtcggatct ccctttgggc 4200
cgcctccccg catcgatacc gtcgactaga gctcgctgat cagcctcgac tgtgccttct 4260
agttgccagc catctgttgt ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc 4320
actcccactg tcctttccta ataaaatgag gaaattgcat cgcattgtct gagtaggtgt 4380
cattctattc tgggggtgg ggtggggcag gacagcaagg gggaggattg ggaagacaat 4440
agcaggcatg ctggggagag atccacgata acaaacagct tttttggggt gaacatattg 4500
actgaattcc ctgcaggttg gccactccct ctctgcgcgc tcgctcgctc actgaggccg 4560

cccgggcaaa gcccgggcgt cgggcgacct ttggtcgccc ggcctcagtg agcgagcgag 4620
cgcgcagaga gggagtggcc aactccatca ctaggggttc ctgcggccgc tcgtacggtc 4680
tcgaggaatt cctgcaggat aacttgccaa cctcattcta aaatgtatat agaagcccaa 4740
aagacaataa caaaaatatt cttgtagaac aaaatgggaa agaatgttcc actaaatatc 4800
aagatttaga gcaaagcatg agatgtgtgg ggatagacag tgaggctgat aaaatagagt 4860
agagctcaga aacagaccca ttgatatatg taagtgacct atgaaaaaaa tatggcattt 4920
tacaatggga aaatgatggt ctttttcttt tttagaaaaa cagggaaata tatttatatg 4980

taaaaaataa aagggaaccc atatgtcata ccatacacac aaaaaaattc cagtgaatta 5040
taagtctaaa tggagaaggc aaaactttaa atcttttaga aaataatata gaagcatgca 5100
gaccagcctg gccaacatga tgaaaccctc tctactaata ataaaatcag tagaactact 5160
caggactact ttgagtggga agtccttttc tatgaagact tctttggcca aaattaggct 5220
ctaaatgcaa ggagatagtg catcatgcct ggctgcactt actgataaat gatgttatca 5280
ccatctttaa ccaaatgcac aggaacaagt tatggtactg atgtgctgga ttgagaagga 5340
gctctacttc cttgacagga cacatttgta tcaacttaaa aaagcagatt tttgccagca 5400

gaactattca ttcagaggta ggaaacttag aatagatgat gtcactgatt agcatggctt 5460
ccccatctcc acagctgctt cccacccagg ttgcccacag ttgagtttgt ccagtgctca 5520
gggctgccca ctctcagtaa gaagccccac accagcccct ctccaaatat gttggctgtt 5580
ccttccatta aagtgacccc actttagagc agcaagtgga tttctgtttc ttacagttca 5640
ggaaggagga gtcagctgtg agaacctgga gcctgagatg cttctaagtc ccactgctac 5700
tggggtcagg gaagccagac tccagcatca gcagtcagga gcactaagcc cttgccaaca 5760
tcctgtttct cagagaaact gcttccatta taatggttgt cctttttaa gctatcaagc 5820

caaacaacca gtgtctacca ttattctcat cacctgaagc caagggttct agcaaaagtc 5880
aagctgtctt gtaatggttg atgtgcctcc agcttctgtc ttcagtcact ccactcttag 5940
cctgctctga atcaactctg accacagttc cctggagccc ctgccacctg ctgcccctgc 6000

caccttctcc atctgcagtg ctgtgcagcc ttctgcactc ttgcagagct aataggtgga 6060
gacttgaagg aagaggagga aagtttctca taatagcctt gctgcaagct caaatgggag 6120
gtgggcactg tgcccaggag ccttggagca aaggctgtgc ccaacctctg actgcatcca 6180
ggtttggtct tgacagagat aagaagccct ggcttttgga gccaaaatct aggtcagact 6240

taggcaggat tctcaaagtt tatcagcaga acatgaggca gaagaccctt tctgctccag 6300
cttcttcagg ctcaaccttc atcagaatag atagaaagag aggctgtgag ggttcttaaa 6360
acagaagcaa atctgactca gagaataaac aacctcctag taaactacag cttagacaga 6420
gcatctggtg gtgagtgtgc tcagtgtcct actcaactgt ctggtatcag ccctcatgag 6480
gacttctctt ctttccctca tagacctcca tctctgtttt ccttagcctg cagaaatctg 6540
gatggctatt cacagaatgc ctgtgctttc agagttgcat tttttctctg gtattctggt 6600
tcaagcattt gaaggtagga aaggttctcc aagtgcaaga aagccagccc tgagcctcaa 6660

ctgcctggct agtgtggtca gtaggatgca aaggctgttg aatgccacaa ggccaaactt 6720
taacctgtgt accacaagcc tagcagcaga ggcagctctg ctcactggaa ctctctgtct 6780
tctttctcct gagccttttc ttttcctgag ttttctagct ctcctcaacc ttacctctgc 6840
cctacccagg acaaacccaa gagccactgt ttctgtgatg tcctctccag ccctaattag 6900
gcatcatgac ttcagcctga ccttccatgc tcagaagcag tgctaatcca cttcagatga 6960
gctgctctat gcaacacagg cagagcctac aaacctttgc accagagccc tccacatatc 7020
agtgtttgtt catactcact tcaacagcaa atgtgactgc tgagattaag attttacaca 7080

agatggtctg taatttcaca gttagtttta tcccattagg tatgaaagaa ttagcataat 7140
tccccttaaa catgaatgaa tcttagattt tttaataaat agttttggaa gtaaagacag 7200
agacatcagg agcacaagga atagcctgag aggacaaaca gaacaagaaa gagtctggaa 7260
atacacagga tgttcttggc ctcctcaaag caagtgcaag cagatagtac cagcagcccc 7320
aggctatcag agcccagtga agagaagtac catgaaagcc acagctctaa ccaccctgtt 7380
ccagagtgac agacagtccc caagacaagc cagcctgagc cagagagaga actgcaagag 7440
aaagtttcta atttaggttc tgttagattc agacaagtgc aggtcatcct ctctccacag 7500

ctactcacct ctccagccta acaaagcctg cagtccacac tccaaccctg gtgtctcacc 7560
tcctagcctc tcccaacatc ctgctctctg accatcttct gcatctctca tctcaccatc 7620
tcccactgtc tacagcctac tcttgcaact accatctcat tttctgacat cctgtctaca 7680
tcttctgcca tactctgcca tctaccatac cacctcttac catctaccac accatctttt 7740
atctccatcc ctctcagaag cctccaagct gaatcctgct ttatgtgttc atctcagccc 7800
ctgcatggaa agctgacccc agaggcagaa ctattcccag agagcttggc caagaaaaac 7860

aaaactacca gcctggccag gctcaggagt agtaagctgc agtgtctgtt gtgttctagc 7920

ttcaacagct gcaggagttc cactctcaaa tgctccacat ttctcacatc ctcctgattc 7980
tggtcactac ccatcttcaa agaacagaat atctcacatc agcatactgt gaaggactag 8040
tcatgggtgc agctgctcag agctgcaaag tcattctgga tggtggagag cttacaaaca 8100
tttcatgatg ctccccccgc tctgatggct ggagcccaat ccctacacag actcctgctg 8160
tatgtgtttt cctttcactc tgagccacag ccagagggca ggcattcagt ctcctcttca 8220
ggctggggct ggggcactga gaactcaccc aacaccttgc tctcactcct tctgcaaaac 8280
aagaaagagc tttgtgctgc agtagccatg aagaatgaaa ggaaggcttt aactaaaaaa 8340

tgtcagagat tattttcaac cccttactgt ggatcaccag caaggaggaa acacaacaca 8400
gagacatttt ttcccctcaa attatcaaaa gaatcactgc atttgttaaa gagagcaact 8460
gaatcaggaa gcagagtttt gaacatatca gaagttagga atctgcatca gagacaaatg 8520
cagtcatggt tgtttgctgc ataccagccc taatcattag aagcctcatg gacttcaaac 8580
atcattccct ctgacaagat gctctagcct aactccatga gataaaataa atctgccttt 8640
cagagccaaa gaagagtcca ccagcttctt ctcagtgtga acaagagctc cagtcaggtt 8700
agtcagtcca gtgcagtaga ggagaccagt ctgcatcctc taattttcaa aggcaagaag 8760

atttgtttac cctggacacc aggcacaagt gaggtcacag agctcttaga tatgcagtcc 8820
tcatgagtga ggagactaaa gcgcatgcca tcaagacttc agtgtagaga aaacctccaa 8880
aaaagcctcc tcactacttc tggaatagct cagaggccga ggcggcctcg gcctctgcat 8940
aaataaaaaa aattagtcag ccatggggcg gagaatgggc ggaactgggc ggagttaggg 9000
gcgggatggg cggagttagg ggcgggacta tggttgctga ctaattgaga tgcatgcttt 9060
gcatacttct gcctgctggg gagcctgggg actttccaca cctggttgct gactaattga 9120
gatgcatgct ttgcatactt ctgcctgctg gggagcctgg ggactttcca caccctaact 9180

gacacacatt ccacagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg 9240
tattgggcgc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg 9300
gcgagcggta tcagctcact caaaggcggt aatacggtta tccacagaat caggggataa 9360
cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc 9420
gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc 9480
aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag 9540
ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct 9600

cccttcggga agcgtggcgc tttctcatag ctcacgctgt aggtatctca gttcggtgta 9660

ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc gttcagcccg accgctgcgc 9720

cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc 9780

agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt 9840

gaagtggtgg cctaactacg gctacactag aagaacagta tttggtatct gcgctctgct 9900

gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc 9960

tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca 10020

agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta 10080

agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa 10140

atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca gttaccaatg 10200

cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgcctg 10260

actcctgcaa accacgttgt gtctcaaaat ctctgatgtt acattgcaca agataaaaat 10320

atatcatcat gaacaataaa actgtctgct tacataaaca gtaatacaag gggtgttatg 10380

agccatattc aacgggaaac gtcttgctcg aggccgcgat taaattccaa catggatgct 10440

gatttatatg ggtataaatg ggctcgcgat aatgtcgggc aatcaggtgc gacaatctat 10500

cgattgtatg ggaagcccga tgcgccagag ttgtttctga aacatggcaa aggtagcgtt 10560

gccaatgatg ttacagatga gatggtcaga ctaaactggc tgacggaatt tatgcctctt 10620

ccgaccatca agcattttat ccgtactcct gatgatgcat ggttactcac cactgcgatc 10680

cccgggaaaa cagcattcca ggtattagaa gaatatcctg attcaggtga aaatattgtt 10740

gatgcgctgg cagtgttcct gcgccggttg cattcgattc ctgtttgtaa ttgtcctttt 10800

aacagcgatc gcgtatttcg tctcgctcag gcgcaatcac gaatgaataa cggtttggtt 10860

gatgcgagtg attttgatga cgagcgtaat ggctggcctg ttgaacaagt ctggaaagaa 10920

atgcataagc ttttgccatt ctcaccggat tcagtcgtca ctcatggtga tttctcactt 10980

gataacctta tttttgacga ggggaaatta ataggttgta ttgatgttgg acgagtcgga 11040

atcgcagacc gataccagga tcttgccatc ctatggaact gcctcggtga gttttctcct 11100

tcattacaga aacggctttt tcaaaaatat ggtattgata atcctgatat gaataaattg 11160

cagtttcatt tgatgctcga tgagtttttc taagggcggc ctgccaccat acccacgccg 11220

aaacaagcgc tcatgagccc gaagtggcga gcccgatctt ccccatcggt gatgtcggcg 11280

atataggcgc cagcaaccgc acctgtggcg ccggtgatga gggcgcgcca agtcgacgtc 11340

cggcagtc 11348

<210> 42

<211> 11433

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 42

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420

tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctcctg gtggcgaggg gagggggtg gtcctcgaac gccttgcaga 600

actggcctgg atacagagtg gaccggctgg ccccatctgg aagacttcga gatacactgt 660

tgtcttactg cgctcaacag tgtatctcga agtcttccaa atggtgccag ccatcgcagc 720

ggggtgcagg aaatgggggc agcccccctt tttggctatc cttccacgtg ttctttttg 780

tatcttttgt gtttcctaga aaacatctca gtcaccaccg cagccctagg aatgcatcta 840

gacaattgta ctaaccttct tctctttcct ctcctgacag tccggaaagc caccatggaa 900

ttcagcagcc ccagcagaga ggaatgcccc aagcctctga gccgggtgtc aatcatggcc 960

ggatctctga caggactgct gctgcttcag gccgtgtctt gggcttctgg cgctagacct 1020

tgcatcccca agagcttcgg ctacagcagc gtcgtgtgcg tgtgcaatgc cacctactgc 1080

gacagcttcg accctcctac ctttcctgct ctgggcacct tcagcagata cgagagcacc 1140

agatccggca gacggatgga actgagcatg ggacccatcc aggccaatca cacaggcact 1200

ggcctgctgc tgacactgca gcctgagcag aaattccaga aagtgaaagg cttcggcgga 1260

gccatgacag atgccgccgc tctgaatatc ctggctctgt ctccaccagc tcagaacctg 1320

ctgctcaaga gctacttcag cgaggaaggc atcggctaca acatcatcag agtgcccatg 1380

gccagctgcg acttcagcat caggacctac acctacgccg acacacccga cgatttccag 1440

ctgcacaact tcagcctgcc tgaagaggac accaagctga agatccctct gatccacaga 1500

gccctgcagc tggcacaaag acccgtgtca ctgctggcct ctccatggac atctcccacc 1560

tggctgaaaa caaatggcgc cgtgaatggc aagggcagcc tgaaaggcca acctggcgac 1620

atctaccacc agacctgggc cagatacttc gtgaagttcc tggacgccta tgccgagcac 1680

aagctgcagt tttgggccgt gacagccgag aacgaacctt ctgctggact gctgagcggc 1740

taccccttc agtgcctggg ctttacaccc gagcaccagc gggactttat cgcccgtgat 1800

ctgggaccca cactggccaa tagcacccac cataatgtgc ggctgctgat gctggacgac 1860

cagagactgc ttctgcccca ctgggctaaa gtggtgctga cagatcctga ggccgccaaa 1920

tacgtgcacg gaatcgccgt gcactggtat ctggactttc tggcccctgc caaggccaca 1980

ctgggagaga cacacagact gttccccaac accatgctgt tcgccagcga agcctgtgtg 2040

ggcagcaagt tttgggaaca gagcgtgcgg ctcggcagct gggatagagg catgcagtac 2100

agccacagca tcatcaccaa cctgctgtac cacgtcgtcg gctggaccga ctggaatctg 2160

gccctgaatc ctgaaggcgg ccctaactgg gtccgaaact tcgtggacag ccccatcatc 2220

gtggacatca ccaaggacac cttctacaag cagcccatgt tctaccacct gggacacttc 2280

agcaagttca tccccgaggg ctctcagcgc gttggactgg tggcttccca gaagaacgat 2340

ctggacgccg tggctctgat gcaccctgat ggatctgctg tggtggtggt cctgaaccgc 2400

agcagcaaag atgtgcccct gaccatcaag gatcccgccg tgggattcct ggaaacaatc 2460

agccctggct actccatcca cacctacctg tggcgtagac aggagggcag aggaagtctt 2520

ctgacatgcg gagacgtgga agagaatccc ggccctatgc cccgctacgg cgccagcctg 2580

cgccagagct gcccccgcag cggccgcgag cagggccagg acggcaccgc cggcgccccc 2640

ggcctgctgt ggatgggcct ggtgctggcc ctggccctgg ccctggccct ggccctggcc 2700

ctgagcgaca gccgcgtgct gtgggccccc gccgaggccc acccccctgag cccccagggc 2760

caccccgccc gcctgcaccg catcgtgccc cgcctgcgcg acgtgttcgg ctggggcaac 2820

ctgacctgcc ccatctgcaa gggcctgttc accgccatca acctgggcct gaagaaggag 2880

cccaacgtgg cccgcgtggg cagcgtggcc atcaagctgt gcaacctgct gaagatcgcc 2940

ccccccgccg tgtgccagag catcgtgcac ctgttcgagg acgacatggt ggaggtgtgg 3000

cgccgcagcg tgctgagccc cagcgaggcc tgcggcctgc tgctgggcag cacctgcggc 3060

cactgggaca tcttcagcag ctggaacatc agcctgccca ccgtgcccaa gcccccccc 3120

aagcccccca gcccccccgc ccccggcgcc cccgtgagcc gcatcctgtt cctgaccgac 3180

ctgcactggg accacgacta cctggagggc accgaccccg actgcgccga ccccctgtgc 3240

tgccgccgcg gcagcggcct gcccccccgcc agccgccccg gcgccggcta ctggggcgag 3300

tacagcaagt gcgacctgcc cctgcgcacc ctggagagcc tgctgagcgg cctgggcccc 3360

gccggcccct tcgacatggt gtactggacc ggcgacatcc ccgcccacga cgtgtggcac 3420

cagacccgcc aggaccagct gcgcgccctg accaccgtga ccgccctggt gcgcaagttc 3480
ctgggccccg tgcccgtgta ccccgccgtg ggcaaccacg agagcacccc cgtgaacagc 3540
ttcccccccc ccttcatcga gggcaaccac agcagccgct ggctgtacga ggccatggcc 3600

aaggcctggg agccctggct gcccgccgag gccctgcgca ccctgcgcat cggcggcttc 3660
tacgccctga gcccctaccc cggcctgcgc ctgatcagcc tgaacatgaa cttctgcagc 3720
cgcgagaact tctggctgct gatcaacagc accgaccccg ccggccagct gcagtggctg 3780
gtgggcgagc tgcaggccgc cgaggaccgc ggcgacaagg tgcacatcat cggccacatc 3840
ccccccggcc actgcctgaa gagctggagc tggaactact accgcatcgt ggcccgctac 3900
gagaacaccc tggccgccca gttcttcggc cacacccacg tggacgagtt cgaggtgttc 3960
tacgacgagg agaccctgag ccgccccctg gccgtggcct tcctggcccc cagcgccacc 4020

acctacatcg gcctgaaccc cggctaccgc gtgtaccaga tcgacggcaa ctacagcggc 4080
agcagccacg tggtgctgga ccacgagacc tacatcctga acctgaccca ggccaacatc 4140
cccggcgcca tcccccactg gcagctgctg taccgcgccc gcgagaccta cggcctgccc 4200
aacaccctgc ccaccgcctg gcacaacctg gtgtaccgca tgcgcggcga catgcagctg 4260
ttccagacct tctggttcct gtaccacaag ggccaccccc ccagcgagcc ctgcggcacc 4320
ccctgccgcc tggccaccct gtgcgcccag ctgagcgccc gcgccgacag ccccgccctg 4380
tgccgccacc tgatgcccga cggcagcctg cccgaggccc agagcctgtg gccccgcccc 4440

ctgttctgct aatgacaatt gttaattaag tttaaaccct cgaggccgca agcaataaaa 4500
tatctttatt ttcattacat ctgtgtgttg gtttttgtg tggagatcca cgataacaaa 4560
cagctttttt ggggtgaaca tattgactga attccctgca ggttggccac tccctctctg 4620
cgcgctcgct cgctcactga ggccgcccgg gcaaagcccg ggcgtcgggc gacctttggt 4680
cgcccggcct cagtgagcga gcgagcgcgc agagagggag tggccaactc catcactagg 4740
ggttcctgcg gccgctcgta cggtctcgag gaattcctgc aggataactt gccaacctca 4800
ttctaaaatg tatatagaag cccaaaagac aataacaaaa atattcttgt agaacaaaat 4860

gggaaagaat gttccactaa atatcaagat ttagagcaaa gcatgagatg tgtggggata 4920
gacagtgagg ctgataaaat agagtagagc tcagaaacag acccattgat atatgtaagt 4980
gacctatgaa aaaaatatgg cattttacaa tgggaaaatg atggtctttt tctttttag 5040
aaaaacaggg aaatatattt atatgtaaaa aataaaaggg aacccatatg tcataccata 5100
cacacaaaaa aattccagtg aattataagt ctaaatggag aaggcaaaac tttaaatctt 5160
ttagaaaata atatagaagc atgcagacca gcctggccaa catgatgaaa ccctctctac 5220

taataataaa atcagtagaa ctactcagga ctactttgag tgggaagtcc ttttctatga 5280

agacttcttt ggccaaaatt aggctctaaa tgcaaggaga tagtgcatca tgcctggctg 5340
cacttactga taaatgatgt tatcaccatc tttaaccaaa tgcacaggaa caagttatgg 5400
tactgatgtg ctggattgag aaggagctct acttccttga caggacacat ttgtatcaac 5460
ttaaaaaagc agatttttgc cagcagaact attcattcag aggtaggaaa cttagaatag 5520
atgatgtcac tgattagcat ggcttcccca tctccacagc tgcttcccac ccaggttgcc 5580
cacagttgag tttgtccagt gctcagggct gcccactctc agtaagaagc cccacaccag 5640
cccctctcca aatatgttgg ctgttccttc cattaaagtg accccacttt agagcagcaa 5700

gtggatttct gtttcttaca gttcaggaag gaggagtcag ctgtgagaac ctggagcctg 5760
agatgcttct aagtcccact gctactgggg tcagggaagc cagactccag catcagcagt 5820
caggagcact aagcccttgc caacatcctg tttctcagag aaactgcttc cattataatg 5880
gttgtccttt tttaagctat caagccaaac aaccagtgtc taccattatt ctcatcacct 5940
gaagccaagg gttctagcaa aagtcaagct gtcttgtaat ggttgatgtg cctccagctt 6000
ctgtcttcag tcactccact cttagcctgc tctgaatcaa ctctgaccac agttccctgg 6060
agcccctgcc acctgctgcc cctgccacct tctccatctg cagtgctgtg cagccttctg 6120

cactcttgca gagctaatag gtggagactt gaaggaagag gaggaaagtt tctcataata 6180
gccttgctgc aagctcaaat gggaggtggg cactgtgccc aggagccttg gagcaaaggc 6240
tgtgcccaac ctctgactgc atccaggttt ggtcttgaca gagataagaa gccctggctt 6300
ttggagccaa aatctaggtc agacttaggc aggattctca aagtttatca gcagaacatg 6360
aggcagaaga ccctttctgc tccagcttct tcaggctcaa ccttcatcag aatagataga 6420
aagagaggct gtgagggttc ttaaaacaga agcaaatctg actcagagaa taaacaacct 6480
cctagtaaac tacagcttag acagagcatc tggtggtgag tgtgctcagt gtcctactca 6540

actgtctggt atcagccctc atgaggactt ctcttctttc cctcatagac ctccatctct 6600
gttttcctta gcctgcagaa atctggatgg ctattcacag aatgcctgtg ctttcagagt 6660
tgcatttttt ctctggtatt ctggttcaag catttgaagg taggaaaggt tctccaagtg 6720
caagaaagcc agccctgagc ctcaactgcc tggctagtgt ggtcagtagg atgcaaaggc 6780
tgttgaatgc cacaaggcca aactttaacc tgtgtaccac aagcctagca gcagaggcag 6840
ctctgctcac tggaactctc tgtcttcttt ctcctgagcc ttttcttttc ctgagttttc 6900
tagctctcct caaccttacc tctgccctac ccaggacaaa cccaagagcc actgtttctg 6960

tgatgtcctc tccagcccta attaggcatc atgacttcag cctgaccttc catgctcaga 7020
agcagtgcta atccacttca gatgagctgc tctatgcaac acaggcagag cctacaaacc 7080
tttgcaccag agccctccac atatcagtgt ttgttcatac tcacttcaac agcaaatgtg 7140
actgctgaga ttaagatttt acacaagatg gtctgtaatt tcacagttag ttttatccca 7200
ttaggtatga aagaattagc ataattcccc ttaaacatga atgaatctta gatttttaa 7260
taaatagttt tggaagtaaa gacagagaca tcaggagcac aaggaatagc ctgagaggac 7320
aaacagaaca agaaagagtc tggaaataca caggatgttc ttggcctcct caaagcaagt 7380

gcaagcagat agtaccagca gccccaggct atcagagccc agtgaagaga agtaccatga 7440
aagccacagc tctaaccacc ctgttccaga gtgacagaca gtccccaaga caagccagcc 7500
tgagccagag agagaactgc aagagaaagt ttctaattta ggttctgtta gattcagaca 7560
agtgcaggtc atcctctctc cacagctact cacctctcca gcctaacaaa gcctgcagtc 7620
cacactccaa ccctggtgtc tcacctccta gcctctccca acatcctgct ctctgaccat 7680
cttctgcatc tctcatctca ccatctccca ctgtctacag cctactcttg caactaccat 7740
ctcattttct gacatcctgt ctacatcttc tgccatactc tgccatctac cataccacct 7800

cttaccatct accacaccat cttttatctc catccctctc agaagcctcc aagctgaatc 7860
ctgctttatg tgttcatctc agcccctgca tggaaagctg accccagagg cagaactatt 7920
cccagagagc ttggccaaga aaaacaaaac taccagcctg gccaggctca ggagtagtaa 7980
gctgcagtgt ctgttgtgtt ctagcttcaa cagctgcagg agttccactc tcaaatgctc 8040
cacatttctc acatcctcct gattctggtc actacccatc ttcaaagaac agaatatctc 8100
acatcagcat actgtgaagg actagtcatg ggtgcagctg ctcagagctg caaagtcatt 8160
ctggatggtg gagagcttac aaacatttca tgatgctccc cccgctctga tggctggagc 8220

ccaatcccta cacagactcc tgctgtatgt gttttccttt cactctgagc cacagccaga 8280
gggcaggcat tcagtctcct cttcaggctg gggctggggc actgagaact cacccaacac 8340
cttgctctca ctccttctgc aaaacaagaa agagctttgt gctgcagtag ccatgaagaa 8400
tgaaaggaag gctttaacta aaaaatgtca gagattattt tcaacccctt actgtggatc 8460
accagcaagg aggaaacaca acacagagac atttttttccc ctcaaattat caaaagaatc 8520
actgcatttg ttaaagagag caactgaatc aggaagcaga gttttgaaca tatcagaagt 8580
taggaatctg catcagagac aaatgcagtc atggttgttt gctgcatacc agccctaatc 8640

attagaagcc tcatggactt caaacatcat tccctctgac aagatgctct agcctaactc 8700
catgagataa aataaatctg cctttcagag ccaaagaaga gtccaccagc ttcttctcag 8760
tgtgaacaag agctccagtc aggttagtca gtccagtgca gtagaggaga ccagtctgca 8820

tcctctaatt ttcaaaggca agaagatttg tttaccctgg acaccaggca caagtgaggt 8880
cacagagctc ttagatatgc agtcctcatg agtgaggaga ctaaagcgca tgccatcaag 8940
acttcagtgt agagaaaacc tccaaaaaag cctcctcact acttctggaa tagctcagag 9000
gccgaggcgg cctcggcctc tgcataaata aaaaaaatta gtcagccatg gggcggagaa 9060

tgggcggaac tgggcggagt taggggcggg atgggcggag ttaggggcgg gactatggtt 9120
gctgactaat tgagatgcat gctttgcata cttctgcctg ctggggagcc tggggacttt 9180
ccacacctgg ttgctgacta attgagatgc atgctttgca tacttctgcc tgctggggag 9240
cctggggact ttccacaccc taactgacac acattccaca gctgcattaa tgaatcggcc 9300
aacgcgcggg gagaggcggt ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact 9360
cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac 9420
ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa 9480

aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgcccccctg 9540
acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa 9600
gataccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc 9660
ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac 9720
gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac 9780
cccccgttca gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg 9840
taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt 9900

atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa 9960
cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct 10020
cttgatccgg caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga 10080
ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg 10140
ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct 10200
tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt 10260
aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc 10320

tatttcgttc atccatagtt gcctgactcc tgcaaaccac gttgtgtctc aaaatctctg 10380
atgttacatt gcacaagata aaaatatatc atcatgaaca ataaaactgt ctgcttacat 10440
aaacagtaat acaaggggtg ttatgagcca tattcaacgg gaaacgtctt gctcgaggcc 10500
gcgattaaat tccaacatgg atgctgattt atatgggtat aaatgggctc gcgataatgt 10560
cgggcaatca ggtgcgacaa tctatcgatt gtatgggaag cccgatgcgc cagagttgtt 10620
tctgaaacat ggcaaaggta gcgttgccaa tgatgttaca gatgagatgg tcagactaaa 10680

ctggctgacg gaatttatgc ctcttccgac catcaagcat tttatccgta ctcctgatga 10740

tgcatggtta ctcaccactg cgatccccgg gaaaacagca ttccaggtat tagaagaata 10800
tcctgattca ggtgaaaata ttgttgatgc gctggcagtg ttcctgcgcc ggttgcattc 10860
gattcctgtt tgtaattgtc cttttaacag cgatcgcgta tttcgtctcg ctcaggcgca 10920
atcacgaatg aataacggtt tggttgatgc gagtgatttt gatgacgagc gtaatggctg 10980
gcctgttgaa caagtctgga aagaaatgca taagcttttg ccattctcac cggattcagt 11040
cgtcactcat ggtgatttct cacttgataa ccttattttt gacgagggga aattaatagg 11100
ttgtattgat gttggacgag tcggaatcgc agaccgatac caggatcttg ccatcctatg 11160

gaactgcctc ggtgagtttt ctccttcatt acagaaacgg ctttttcaaa aatatggtat 11220
tgataatcct gatatgaata aattgcagtt tcatttgatg ctcgatgagt ttttctaagg 11280
gcggcctgcc accataccca cgccgaaaca agcgctcatg agcccgaagt ggcgagcccg 11340
atcttcccca tcggtgatgt cggcgatata ggcgccagca accgcacctg tggcgccggt 11400
gatgagggcg cgccaagtcg acgtccggca gtc 11433

<210> 43
<211> 11776
<212> DNA
<213> Artificial sequence
<220><223> Synthetic polynucleotide
<400> 43

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120
gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240
tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300
tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360
atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420
tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540
tggcactatg aaccctgcag ccctaggaat gcatctagac aattgtacta accttcttct 600
ctttcctctc ctgacagtcc ggaaagccac catggccgag tggctgctga gcgccagctg 660
gcagcgccgc gccaaggcca tgaccgccgc cgccggcagc gccggccgcg ccgccgtgcc 720

cctgctgctg tgcgccctgc tggcccccgg cggcgcctac gtgctggacg acagcgacgg 780

cctgggccgc gagttcgacg gcatcggcgc cgtgagcggc ggcggcgcca ccagccgcct 840

gctggtgaac taccccgagc cctaccgcag ccagatcctg gactacctgt tcaagcccaa 900

cttcggcgcc agcctgcaca tcctgaaggt ggagatcggc ggcgacggcc agaccaccga 960

cggcaccgag cccagccaca tgcactacgc cctggacgag aactacttcc gcggctacga 1020

gtggtggctg atgaaggagg ccaagaagcg caaccccaac atcaccctga tcggcctgcc 1080

ctggagcttc cccggctggc tgggcaaggg cttcgactgg ccctacgtga acctgcagct 1140

gaccgcctac tacgtggtga cctggatcgt gggcgccaag cgctaccacg acctggacat 1200

cgactacatc ggcatctgga acgagcgcag ctacaacgcc aactacatca agatcctgcg 1260

caagatgctg aactaccagg gcctgcagcg cgtgaagatc atcgccagcg acaacctgtg 1320

ggagagcatc agcgccagca tgctgctgga cgccgagctg ttcaaggtgg tggacgtgat 1380

cggcgcccac taccccggca cccacagcgc caaggacgcc aagctgaccg gcaagaagct 1440

gtggagcagc gaggacttca gcaccctgaa cagcgacatg ggcgccggct gctggggccg 1500

catcctgaac cagaactaca tcaacggcta catgaccagc accatcgcct ggaacctggt 1560

ggccagctac tacgagcagc tgccctacgg ccgctgcggc ctgatgaccg cccaggagcc 1620

ctggagcggc cactacgtgg tggagagccc cgtgtgggtg agcgcccaca ccacccagtt 1680

cacccagccc ggctggtact acctgaagac cgtgggccac ctggagaagg gcggcagcta 1740

cgtggccctg accgacggcc tgggcaacct gaccatcatc atcgagacca tgagccacaa 1800

gcacagcaag tgcatccgcc ccttcctgcc ctacttcaac gtgagccagc agttcgccac 1860

cttcgtgctg aagggcagct tcagcgagat ccccgagctg caggtgtggt acaccaagct 1920

gggcaagacc agcgagcgct tcctgttcaa gcagctggac agcctgtggc tgctggacag 1980

cgacggcagc ttcaccctga gcctgcacga ggacgagctg ttcaccctga ccaccctgac 2040

caccggccgc aagggcagct acccccctgcc ccccaagagc cagcccttcc ccagcaccta 2100

caaggacgac ttcaacgtgg actacccctt cttcagcgag gcccccaact tcgccgacca 2160

gaccggcgtg ttcgagtact tcaccaacat cgaggacccc ggcgagcacc acttcaccct 2220

gcgccaggtg ctgaaccagc gccccatcac ctgggccgcc gacgccagca acaccatcag 2280

catcatcggc gactacaact ggaccaacct gaccatcaag tgcgacgtgt acatcgagac 2340

ccccgacacc ggcggcgtgt tcatcgccgg ccgcgtgaac aagggcggca tcctgatccg 2400

cagcgcccgc ggcatcttct tctggatctt cgccaacggc agctaccgcg tgaccggcga 2460

cctggccggc tggatcatct acgccctggg ccgcgtggag gtgaccgcca agaagtggta 2520

caccctgacc ctgaccatca agggccactt caccagcggc atgctgaacg acaagagcct 2580

gtggaccgac atccccgtga acttccccaa gaacggctgg gccgccatcg gcacccacag 2640
cttcgagttc gcccagttcg acaacttcct ggtggaggcc acccgctgat tgtggccgaa 2700
ccgccgaact cagaggccgg ccccagaaaa cccgagcgag taggggggcgg cgcgcaggag 2760
ggaggagaac tgggggcgcg ggaggctggt gggtgtgggg ggtggagatg tagaagatgt 2820
gacgccgcgg cccggcgggt gccagattag cggacgcggt gcccgcggtt gcaacgggat 2880
cccgggcgct gcagcttggg aggcggctct ccccaggcgg cgtccgcgga gacacccatc 2940
cgtgaacccc aggtcccggg ccgccggctc gccgcgcacc aggggccggc ggacagaaga 3000

gcggccgagc ggctcgaggc tgggggaccg cgggcgcggc cgcgcgctgc cgggcgggag 3060
gctgggggggc cggggccggg gccgtgcccc ggagcgggtc ggaggccggg gccggggccg 3120
ggggacggcg gctccccgcg cggctccagc ggctcgggga tcccggccgg gccccgcagg 3180
gaccatgatg gaattcagca gccccagcag agaggaatgc cccaagcctc tgagccgggt 3240
gtcaatcatg gccggatctc tgacaggact gctgctgctt caggccgtgt cttgggcttc 3300
tggcgctaga ccttgcatcc ccaagagctt cggctacagc agcgtcgtgt gcgtgtgcaa 3360
tgccacctac tgcgacagct tcgaccctcc tacctttcct gctctgggca ccttcagcag 3420

atacgagagc accagatccg gcagacggat ggaactgagc atgggaccca tccaggccaa 3480
tcacacaggc actggcctgc tgctgacact gcagcctgag cagaaattcc agaaagtgaa 3540
aggcttcggc ggagccatga cagatgccgc cgctctgaat atcctggctc tgtctccacc 3600
agctcagaac ctgctgctca agagctactt cagcgaggaa ggcatcggct acaacatcat 3660
cagagtgccc atggccagct gcgacttcag catcaggacc tacacctacg ccgacacacc 3720
cgacgatttc cagctgcaca acttcagcct gcctgaagag gacaccaagc tgaagatccc 3780
tctgatccac agagccctgc agctggcaca aagacccgtg tcactgctgg cctctccatg 3840

gacatctccc acctggctga aaacaaatgg cgccgtgaat ggcaagggca gcctgaaagg 3900
ccaacctggc gacatctacc accagacctg ggccagatac ttcgtgaagt tcctggacgc 3960
ctatgccgag cacaagctgc agttttgggc cgtgacagcc gagaacgaac cttctgctgg 4020
actgctgagc ggctacccct ttcagtgcct gggctttaca cccgagcacc agcgggactt 4080
tatcgcccgt gatctgggac ccacactggc caatagcacc caccataatg tgcggctgct 4140
gatgctggac gaccagagac tgcttctgcc ccactgggct aaagtggtgc tgacagatcc 4200
tgaggccgcc aaatacgtgc acggaatcgc cgtgcactgg tatctggact ttctggcccc 4260

tgccaaggcc acactgggag agacacacag actgttcccc aacaccatgc tgttcgccag 4320
cgaagcctgt gtgggcagca agttttggga acagagcgtg cggctcggca gctgggatag 4380
aggcatgcag tacagccaca gcatcatcac caacctgctg taccacgtcg tcggctggac 4440
cgactggaat ctggccctga atcctgaagg cggccctaac tgggtccgaa acttcgtgga 4500
cagccccatc atcgtggaca tcaccaagga caccttctac aagcagccca tgttctacca 4560
cctgggacac ttcagcaagt tcatccccga gggctctcag cgcgttggac tggtggcttc 4620
ccagaagaac gatctggacg ccgtggctct gatgcaccct gatggatctg ctgtggtggt 4680

ggtcctgaac cgcagcagca aagatgtgcc cctgaccatc aaggatcccg ccgtgggatt 4740
cctggaaaca atcagccctg gctactccat ccacacctac ctgtggcgta gacagtgaca 4800
attgttaatt aagtttaaac cctcgaggcc gcaagcaata aaatatcttt attttcatta 4860
catctgtgtg ttggtttttt gtgtggagat ccacgataac aaacagcttt tttggggtga 4920
acatattgac tgaattccct gcaggttggc cactccctct ctgcgcgctc gctcgctcac 4980
tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt ggtcgcccgg cctcagtgag 5040
cgagcgagcg cgcagagagg gagtggccaa ctccatcact aggggttcct gcggccgctc 5100

gtacggtctc gaggaattcc tgcaggataa cttgccaacc tcattctaaa atgtatatag 5160
aagcccaaaa gacaataaca aaaatattct tgtagaacaa aatgggaaag aatgttccac 5220
taaatatcaa gatttagagc aaagcatgag atgtgtgggg atagacagtg aggctgataa 5280
aatagagtag agctcagaaa cagacccatt gatatatgta agtgacctat gaaaaaaata 5340
tggcatttta caatgggaaa atgatggtct ttttcttttt tagaaaaaca gggaaatata 5400
tttatatgta aaaaataaaa gggaacccat atgtcatacc atacacacaa aaaaattcca 5460
gtgaattata agtctaaatg gagaaggcaa aactttaaat cttttagaaa ataatataga 5520

agcatgcaga ccagcctggc caacatgatg aaaccctctc tactaataat aaaatcagta 5580
gaactactca ggactacttt gagtgggaag tccttttcta tgaagacttc tttggccaaa 5640
attaggctct aaatgcaagg agatagtgca tcatgcctgg ctgcacttac tgataaatga 5700
tgttatcacc atctttaacc aaatgcacag gaacaagtta tggtactgat gtgctggatt 5760
gagaaggagc tctacttcct tgacaggaca catttgtatc aacttaaaaa agcagatttt 5820
tgccagcaga actattcatt cagaggtagg aaacttagaa tagatgatgt cactgattag 5880
catggcttcc ccatctccac agctgcttcc cacccaggtt gcccacagtt gagtttgtcc 5940

agtgctcagg gctgcccact ctcagtaaga agccccacac cagcccctct ccaaatatgt 6000
tggctgttcc ttccattaaa gtgaccccac tttagagcag caagtggatt tctgtttctt 6060
acagttcagg aaggaggagt cagctgtgag aacctggagc ctgagatgct tctaagtccc 6120

actgctactg gggtcaggga agccagactc cagcatcagc agtcaggagc actaagccct 6180
tgccaacatc ctgtttctca gagaaactgc ttccattata atggttgtcc ttttttaagc 6240
tatcaagcca aacaaccagt gtctaccatt attctcatca cctgaagcca agggttctag 6300
caaaagtcaa gctgtcttgt aatggttgat gtgcctccag cttctgtctt cagtcactcc 6360

actcttagcc tgctctgaat caactctgac cacagttccc tggagcccct gccacctgct 6420
gcccctgcca ccttctccat ctgcagtgct gtgcagcctt ctgcactctt gcagagctaa 6480
taggtggaga cttgaaggaa gaggaggaaa gtttctcata atagccttgc tgcaagctca 6540
aatgggaggt gggcactgtg cccaggagcc ttggagcaaa ggctgtgccc aacctctgac 6600
tgcatccagg tttggtcttg acagagataa gaagccctgg cttttggagc caaaatctag 6660
gtcagactta ggcaggattc tcaaagttta tcagcagaac atgaggcaga agaccctttc 6720
tgctccagct tcttcaggct caaccttcat cagaatagat agaaagagag gctgtgaggg 6780

ttcttaaaac agaagcaaat ctgactcaga gaataaacaa cctcctagta aactacagct 6840
tagacagagc atctggtggt gagtgtgctc agtgtcctac tcaactgtct ggtatcagcc 6900
ctcatgagga cttctcttct ttccctcata gacctccatc tctgttttcc ttagcctgca 6960
gaaatctgga tggctattca cagaatgcct gtgctttcag agttgcattt tttctctggt 7020
attctggttc aagcatttga aggtaggaaa ggttctccaa gtgcaagaaa gccagccctg 7080
agcctcaact gcctggctag tgtggtcagt aggatgcaaa ggctgttgaa tgccacaagg 7140
ccaaacttta acctgtgtac cacaagccta gcagcagagg cagctctgct cactggaact 7200

ctctgtcttc tttctcctga gccttttctt ttcctgagtt ttctagctct cctcaacctt 7260
acctctgccc tacccaggac aaacccaaga gccactgttt ctgtgatgtc ctctccagcc 7320
ctaattaggc atcatgactt cagcctgacc ttccatgctc agaagcagtg ctaatccact 7380
tcagatgagc tgctctatgc aacacaggca gagcctacaa acctttgcac cagagccctc 7440
cacatatcag tgtttgttca tactcacttc aacagcaaat gtgactgctg agattaagat 7500
tttacacaag atggtctgta atttcacagt tagttttatc ccattaggta tgaaagaatt 7560
agcataattc cccttaaaca tgaatgaatc ttagattttt taataaatag ttttggaagt 7620

aaagacagag acatcaggag cacaaggaat agcctgagag gacaaacaga acaagaaaga 7680
gtctggaaat acacaggatg ttcttggcct cctcaaagca agtgcaagca gatagtacca 7740
gcagccccag gctatcagag cccagtgaag agaagtacca tgaaagccac agctctaacc 7800
accctgttcc agagtgacag acagtcccca agacaagcca gcctgagcca gagagagaac 7860
tgcaagagaa agtttctaat ttaggttctg ttagattcag acaagtgcag gtcatcctct 7920
ctccacagct actcacctct ccagcctaac aaagcctgca gtccacactc caaccctggt 7980

gtctcacctc ctagcctctc ccaacatcct gctctctgac catcttctgc atctctcatc 8040

tcaccatctc ccactgtcta cagcctactc ttgcaactac catctcattt tctgacatcc 8100
tgtctacatc ttctgccata ctctgccatc taccatacca cctcttacca tctaccacac 8160
catcttttat ctccatccct ctcagaagcc tccaagctga atcctgcttt atgtgttcat 8220
ctcagcccct gcatggaaag ctgaccccag aggcagaact attcccagag agcttggcca 8280
agaaaaacaa aactaccagc ctggccaggc tcaggagtag taagctgcag tgtctgttgt 8340
gttctagctt caacagctgc aggagttcca ctctcaaatg ctccacattt ctcacatcct 8400
cctgattctg gtcactaccc atcttcaaag aacagaatat ctcacatcag catactgtga 8460

aggactagtc atgggtgcag ctgctcagag ctgcaaagtc attctggatg gtggagagct 8520
tacaaacatt tcatgatgct ccccccgctc tgatggctgg agcccaatcc ctacacagac 8580
tcctgctgta tgtgttttcc tttcactctg agccacagcc agagggcagg cattcagtct 8640
cctcttcagg ctggggctgg ggcactgaga actcacccaa caccttgctc tcactccttc 8700
tgcaaaacaa gaaagagctt tgtgctgcag tagccatgaa gaatgaaagg aaggctttaa 8760
ctaaaaaatg tcagagatta ttttcaaccc cttactgtgg atcaccagca aggaggaaac 8820
acaacacaga gacatttttt cccctcaaat tatcaaaaga atcactgcat ttgttaaaga 8880

gagcaactga atcaggaagc agagttttga acatatcaga agttaggaat ctgcatcaga 8940
gacaaatgca gtcatggttg tttgctgcat accagcccta atcattagaa gcctcatgga 9000
cttcaaacat cattccctct gacaagatgc tctagcctaa ctccatgaga taaaataaat 9060
ctgcctttca gagccaaaga agagtccacc agcttcttct cagtgtgaac aagagctcca 9120
gtcaggttag tcagtccagt gcagtagagg agaccagtct gcatcctcta attttcaaag 9180
gcaagaagat ttgtttaccc tggacaccag gcacaagtga ggtcacagag ctcttagata 9240
tgcagtcctc atgagtgagg agactaaagc gcatgccatc aagacttcag tgtagagaaa 9300

acctccaaaa aagcctcctc actacttctg gaatagctca gaggccgagg cggcctcggc 9360
ctctgcataa ataaaaaaaa ttagtcagcc atggggcgga gaatgggcgg aactgggcgg 9420
agttaggggc gggatgggcg gagttagggg cgggactatg gttgctgact aattgagatg 9480
catgctttgc atacttctgc ctgctgggga gcctggggac tttccacacc tggttgctga 9540
ctaattgaga tgcatgcttt gcatacttct gcctgctggg gagcctgggg actttccaca 9600
ccctaactga cacacattcc acagctgcat taatgaatcg gccaacgcgc ggggagaggc 9660
ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt 9720

cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca 9780
ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa 9840
aaggccgcgt tgctggcgtt tttccatagg ctccgccccc ctgacgagca tcacaaaaat 9900
cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc 9960
cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc 10020
gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt 10080
tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aaccccccgt tcagcccgac 10140

cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg 10200
ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca 10260
gagttcttga agtggtggcc taactacggc tacactagaa gaacagtatt tggtatctgc 10320
gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacaa 10380
accaccgctg gtagcggtgg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa 10440
ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac 10500
tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatcctttta 10560

aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt 10620
taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata 10680
gttgcctgac tcctgcaaac cacgttgtgt ctcaaaatct ctgatgttac attgcacaag 10740
ataaaaatat atcatcatga acaataaaac tgtctgctta cataaacagt aatacaaggg 10800
gtgttatgag ccatattcaa cgggaaacgt cttgctcgag gccgcgatta aattccaaca 10860
tggatgctga tttatatggg tataaatggg ctcgcgataa tgtcgggcaa tcaggtgcga 10920
caatctatcg attgtatggg aagcccgatg cgccagagtt gtttctgaaa catggcaaag 10980

gtagcgttgc caatgatgtt acagatgaga tggtcagact aaactggctg acggaattta 11040
tgcctcttcc gaccatcaag cattttatcc gtactcctga tgatgcatgg ttactcacca 11100
ctgcgatccc cgggaaaaca gcattccagg tattagaaga atatcctgat tcaggtgaaa 11160
atattgttga tgcgctggca gtgttcctgc gccggttgca ttcgattcct gtttgtaatt 11220
gtccttttaa cagcgatcgc gtatttcgtc tcgctcaggc gcaatcacga atgaataacg 11280
gtttggttga tgcgagtgat tttgatgacg agcgtaatgg ctggcctgtt gaacaagtct 11340
ggaaagaaat gcataagctt ttgccattct caccggattc agtcgtcact catggtgatt 11400

tctcacttga taaccttatt tttgacgagg ggaaattaat aggttgtatt gatgttggac 11460
gagtcggaat cgcagaccga taccaggatc ttgccatcct atggaactgc ctcggtgagt 11520
tttctccttc attacagaaa cggctttttc aaaaatatgg tattgataat cctgatatga 11580

ataaattgca gtttcatttg atgctcgatg agtttttcta agggcggcct gccaccatac 11640

ccacgccgaa acaagcgctc atgagcccga agtggcgagc ccgatcttcc ccatcggtga 11700

tgtcggcgat ataggcgcca gcaaccgcac ctgtggcgcc ggtgatgagg gcgcgccaag 11760

tcgacgtccg gcagtc 11776

<210> 44

<211> 11064

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 44

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420

tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctgcag ccctaggaat gcatctagac aattgtacta accttcttct 600

ctttcctctc ctgacagtcc ggaaagccac catggaattc agcagcccca gcagagagga 660

atgccccaag cctctgagcc gggtgtcaat catggccgga tctctgacag gactgctgct 720

gcttcaggcc gtgtcttggg cttctggcgc tagaccttgc atccccaaga gcttcggcta 780

cagcagcgtc gtgtgcgtgt gcaatgccac ctactgcgac agcttcgacc ctcctacctt 840

tcctgctctg ggcaccttca gcagatacga gagcaccaga tccggcagac ggatggaact 900

gagcatggga cccatccagg ccaatcacac aggcactggc ctgctgctga cactgcagcc 960

tgagcagaaa ttccagaaag tgaaaggctt cggcggagcc atgacagatg ccgccgctct 1020

gaatatcctg gctctgtctc caccagctca gaacctgctg ctcaagagct acttcagcga 1080

ggaaggcatc ggctacaaca tcatcagagt gcccatggcc agctgcgact tcagcatcag 1140

gacctacacc tacgccgaca cacccgacga tttccagctg cacaacttca gcctgcctga 1200

agaggacacc aagctgaaga tccctctgat ccacagagcc ctgcagctgg cacaaagacc 1260

cgtgtcactg ctggcctctc catggacatc tcccacctgg ctgaaaacaa atggcgccgt 1320
gaatggcaag ggcagcctga aaggccaacc tggcgacatc taccaccaga cctgggccag 1380
atacttcgtg aagttcctgg acgcctatgc cgagcacaag ctgcagtttt gggccgtgac 1440
agccgagaac gaaccttctg ctggactgct gagcggctac ccctttcagt gcctgggctt 1500
tacacccgag caccagcggg actttatcgc ccgtgatctg ggacccacac tggccaatag 1560
cacccaccat aatgtgcggc tgctgatgct ggacgaccag agactgcttc tgccccactg 1620

ggctaaagtg gtgctgacag atcctgaggc cgccaaatac gtgcacggaa tcgccgtgca 1680
ctggtatctg gactttctgg cccctgccaa ggccacactg ggagagacac acagactgtt 1740
ccccaacacc atgctgttcg ccagcgaagc ctgtgtgggc agcaagtttt gggaacagag 1800
cgtgcggctc ggcagctggg atagaggcat gcagtacagc cacagcatca tcaccaacct 1860
gctgtaccac gtcgtcggct ggaccgactg gaatctggcc ctgaatcctg aaggcggccc 1920
taactgggtc cgaaacttcg tggacagccc catcatcgtg gacatcacca aggacacctt 1980
ctacaagcag cccatgttct accacctggg acacttcagc aagttcatcc ccgagggctc 2040

tcagcgcgtt ggactggtgg cttcccagaa gaacgatctg gacgccgtgg ctctgatgca 2100
ccctgatgga tctgctgtgg tggtggtcct gaaccgcagc agcaaagatg tgcccctgac 2160
catcaaggat cccgccgtgg gattcctgga aacaatcagc cctggctact ccatccacac 2220
ctacctgtgg cgtagacagt gacaattgtt aattaagttt aaaccctcga ggccgcaagc 2280
cgcatcgata ccgtcgacta gagctcgctg atcagcctcg actgtgcctt ctagttgcca 2340
gccatctgtt gtttgcccct cccccgtgcc ttccttgacc ctggaaggtg ccactcccac 2400
tgtcctttcc taataaaatg aggaaattgc atcgcattgt ctgagtaggt gtcattctat 2460

tctggggggt ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca 2520
tgctggggag agatccacga taacaaacag ctttttggg ggggcggagt tagggcggag 2580
ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga atgggcggtg 2640
aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg tcgcagccgg 2700
gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta agtcactgac 2760
tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag tggcactatg 2820
aaccctgcag ccctaggaat gcatctagac aattgtacta accttcttct ctttcctctc 2880

ctgacagtcc ggaaagccac catgtggcag ctgtgggcca gcctgtgctg cctgctggtg 2940
ctggccaacg cccgcagccg ccccagcttc cacccccctga gcgacgagct ggtgaactac 3000
gtgaacaagc gcaacaccac ctggcaggcc ggccacaact tctacaacgt ggacatgagc 3060
tacctgaagc gcctgtgcgg caccttcctg ggcggcccca agcccccccca gcgcgtgatg 3120

ttcaccgagg acctgaagct gcccgccagc ttcgacgccc gcgagcagtg gccccagtgc 3180
cccaccatca aggagatccg cgaccagggc agctgcggca gctgctgggc cttcggcgcc 3240
gtggaggcca tcagcgaccg catctgcatc cacaccaacg cccacgtgag cgtggaggtg 3300

agcgccgagg acctgctgac ctgctgcggc agcatgtgcg gcgacggctg caacggcggc 3360
taccccgccg aggcctggaa cttctggacc cgcaagggcc tggtgagcgg cggcctgtac 3420
gagagccacg tgggctgccg cccctacagc atcccccccct gcgagcacca cgtgaacggc 3480
agccgccccc cctgcaccgg cgagggcgac acccccaagt gcagcaagat ctgcgagccc 3540
ggctacagcc ccacctacaa gcaggacaag cactacggct acaacagcta cagcgtgagc 3600
aacagcgaga aggacatcat ggccgagatc tacaagaacg gccccgtgga gggcgccttc 3660
agcgtgtaca gcgacttcct gctgtacaag agcggcgtgt accagcacgt gaccggcgag 3720

atgatgggcg gccacgccat ccgcatcctg ggctggggcg tggagaacgg cacccccctac 3780
tggctggtgg ccaacagctg gaacaccgac tggggcgaca acggcttctt caagatcctg 3840
cgcggccagg accactgcgg catcgagagc gaggtggtgg ccggcatccc ccgcaccgac 3900
cagtactggg agaagatctg acccagggga ctcagcggcc gctcgagtct agagggcccg 3960
tttaaacccg ctgatcagcc tcgaagacat gataagatac attgatgagt ttggacaaac 4020
cacaacaaga atgcagtgaa aaaaatgctt tatttgtgaa atttgtgatg ctattgcttt 4080
atttgtaacc attataagct gcaataaaca agttaacaac aacaattgca ttcattttat 4140

gtttcaggtt caggggggaga tgtgggaggt tttttaaagc aagtaaaacc tctacaaatg 4200
tggtatgaac atattgactg aattccctgc aggttggcca ctccctctct gcgcgctcgc 4260
tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg cgacctttgg tcgcccggcc 4320
tcagtgagcg agcgagcgcg cagagaggga gtggccaact ccatcactag gggttcctgc 4380
ggccgctcgt acggtctcga ggaattcctg caggataact tgccaacctc attctaaaat 4440
gtatatagaa gcccaaaaga caataacaaa aatattcttg tagaacaaaa tgggaaagaa 4500
tgttccacta aatatcaaga tttagagcaa agcatgagat gtgtggggat agacagtgag 4560

gctgataaaa tagagtagag ctcagaaaca gacccattga tatatgtaag tgacctatga 4620
aaaaaatatg gcattttaca atgggaaaat gatggtcttt ttctttttta gaaaaacagg 4680
gaaatatatt tatatgtaaa aaataaaagg gaacccatat gtcataccat acacacaaaa 4740
aaattccagt gaattataag tctaaatgga gaaggcaaaa ctttaaatct tttagaaaat 4800
aatatagaag catgcagacc agcctggcca acatgatgaa accctctcta ctaataataa 4860
aatcagtaga actactcagg actactttga gtgggaagtc cttttctatg aagacttctt 4920

tggccaaaat taggctctaa atgcaaggag atagtgcatc atgcctggct gcacttactg 4980

ataaatgatg ttatcaccat ctttaaccaa atgcacagga acaagttatg gtactgatgt 5040
gctggattga gaaggagctc tacttccttg acaggacaca tttgtatcaa cttaaaaaag 5100
cagatttttg ccagcagaac tattcattca gaggtaggaa acttagaata gatgatgtca 5160
ctgattagca tggcttcccc atctccacag ctgcttccca cccaggttgc ccacagttga 5220
gtttgtccag tgctcagggc tgcccactct cagtaagaag ccccacacca gcccctctcc 5280
aaatatgttg gctgttcctt ccattaaagt gaccccactt tagagcagca agtggatttc 5340
tgtttcttac agttcaggaa ggaggagtca gctgtgagaa cctggagcct gagatgcttc 5400

taagtcccac tgctactggg gtcagggaag ccagactcca gcatcagcag tcaggagcac 5460
taagcccttg ccaacatcct gtttctcaga gaaactgctt ccattataat ggttgtcctt 5520
ttttaagcta tcaagccaaa caaccagtgt ctaccattat tctcatcacc tgaagccaag 5580
ggttctagca aaagtcaagc tgtcttgtaa tggttgatgt gcctccagct tctgtcttca 5640
gtcactccac tcttagcctg ctctgaatca actctgacca cagttccctg gagcccctgc 5700
cacctgctgc ccctgccacc ttctccatct gcagtgctgt gcagccttct gcactcttgc 5760
agagctaata ggtggagact tgaaggaaga ggaggaaagt ttctcataat agccttgctg 5820

caagctcaaa tgggaggtgg gcactgtgcc caggagcctt ggagcaaagg ctgtgcccaa 5880
cctctgactg catccaggtt tggtcttgac agagataaga agccctggct tttggagcca 5940
aaatctaggt cagacttagg caggattctc aaagtttatc agcagaacat gaggcagaag 6000
acccttttctg ctccagcttc ttcaggctca accttcatca gaatagatag aaagagaggc 6060
tgtgagggtt cttaaaacag aagcaaatct gactcagaga ataaacaacc tcctagtaaa 6120
ctacagctta gacagagcat ctggtggtga gtgtgctcag tgtcctactc aactgtctgg 6180
tatcagccct catgaggact tctcttcttt ccctcataga cctccatctc tgttttcctt 6240

agcctgcaga aatctggatg gctattcaca gaatgcctgt gctttcagag ttgcattttt 6300
tctctggtat tctggttcaa gcatttgaag gtaggaaagg ttctccaagt gcaagaaagc 6360
cagccctgag cctcaactgc ctggctagtg tggtcagtag gatgcaaagg ctgttgaatg 6420
ccacaaggcc aaactttaac ctgtgtacca caagcctagc agcagaggca gctctgctca 6480
ctggaactct ctgtcttctt tctcctgagc cttttctttt cctgagtttt ctagctctcc 6540
tcaaccttac ctctgcccta cccaggacaa acccaagagc cactgtttct gtgatgtcct 6600
ctccagccct aattaggcat catgacttca gcctgacctt ccatgctcag aagcagtgct 6660

aatccacttc agatgagctg ctctatgcaa cacaggcaga gcctacaaac ctttgcacca 6720
gagccctcca catatcagtg tttgttcata ctcacttcaa cagcaaatgt gactgctgag 6780
attaagattt tacacaagat ggtctgtaat ttcacagtta gttttatccc attaggtatg 6840
aaagaattag cataattccc cttaaacatg aatgaatctt agatttttta ataaatagtt 6900
ttggaagtaa agacagagac atcaggagca caaggaatag cctgagagga caaacagaac 6960
aagaaagagt ctggaaatac acaggatgtt cttggcctcc tcaaagcaag tgcaagcaga 7020
tagtaccagc agccccaggc tatcagagcc cagtgaagag aagtaccatg aaagccacag 7080

ctctaaccac cctgttccag agtgacagac agtccccaag acaagccagc ctgagccaga 7140
gagagaactg caagagaaag tttctaattt aggttctgtt agattcagac aagtgcaggt 7200
catcctctct ccacagctac tcacctctcc agcctaacaa agcctgcagt ccacactcca 7260
accctggtgt ctcacctcct agcctctccc aacatcctgc tctctgacca tcttctgcat 7320
ctctcatctc accatctccc actgtctaca gcctactctt gcaactacca tctcattttc 7380
tgacatcctg tctacatctt ctgccatact ctgccatcta ccataccacc tcttaccatc 7440
taccacacca tcttttatct ccatccctct cagaagcctc caagctgaat cctgctttat 7500

gtgttcatct cagcccctgc atggaaagct gaccccagag gcagaactat tcccagagag 7560
cttggccaag aaaaacaaaa ctaccagcct ggccaggctc aggagtagta agctgcagtg 7620
tctgttgtgt tctagcttca acagctgcag gagttccact ctcaaatgct ccacatttct 7680
cacatcctcc tgattctggt cactacccat cttcaaagaa cagaatatct cacatcagca 7740
tactgtgaag gactagtcat gggtgcagct gctcagagct gcaaagtcat tctggatggt 7800
ggagagctta caaacatttc atgatgctcc ccccgctctg atggctggag cccaatccct 7860
acacagactc ctgctgtatg tgttttcctt tcactctgag ccacagccag agggcaggca 7920

ttcagtctcc tcttcaggct ggggctgggg cactgagaac tcacccaaca ccttgctctc 7980
actccttctg caaaacaaga aagagctttg tgctgcagta gccatgaaga atgaaaggaa 8040
ggctttaact aaaaaatgtc agagattatt ttcaacccct tactgtggat caccagcaag 8100
gaggaaacac aacacagaga catttttcc cctcaaatta tcaaaagaat cactgcattt 8160
gttaaagaga gcaactgaat caggaagcag agttttgaac atatcagaag ttaggaatct 8220
gcatcagaga caaatgcagt catggttgtt tgctgcatac cagccctaat cattagaagc 8280
ctcatggact tcaaacatca ttccctctga caagatgctc tagcctaact ccatgagata 8340

aaataaatct gcctttcaga gccaaagaag agtccaccag cttcttctca gtgtgaacaa 8400
gagctccagt caggttagtc agtccagtgc agtagaggag accagtctgc atcctctaat 8460
tttcaaaggc aagaagattt gtttaccctg gacaccaggc acaagtgagg tcacagagct 8520

cttagatatg cagtcctcat gagtgaggag actaaagcgc atgccatcaa gacttcagtg 8580
tagagaaaac ctccaaaaaa gcctcctcac tacttctgga atagctcaga ggccgaggcg 8640
gcctcggcct ctgcataaat aaaaaaaatt agtcagccat ggggcggaga atgggcggaa 8700
ctgggcggag ttaggggcgg gatgggcgga gttaggggcg ggactatggt tgctgactaa 8760

ttgagatgca tgctttgcat acttctgcct gctggggagc ctggggactt tccacacctg 8820
gttgctgact aattgagatg catgctttgc atacttctgc ctgctgggga gcctggggac 8880
tttccacacc ctaactgaca cacattccac agctgcatta atgaatcggc caacgcgcgg 8940
ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc gctcactgac tcgctgcgct 9000
cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca 9060
cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga 9120
accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgcccccct gacgagcatc 9180

acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg 9240
cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat 9300
acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt 9360
atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc 9420
agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg 9480
acttatcgcc actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg 9540
gtgctacaga gttcttgaag tggtggccta actacggcta cactagaaga acagtatttg 9600

gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg 9660
gcaaacaaac caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca 9720
gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga 9780
acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga 9840
tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt 9900
ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt 9960
catccatagt tgcctgactc ctgcaaacca cgttgtgtct caaaatctct gatgttacat 10020

tgcacaagat aaaaatatat catcatgaac aataaaactg tctgcttaca taaacagtaa 10080
tacaaggggt gttatgagcc atattcaacg ggaaacgtct tgctcgaggc cgcgattaaa 10140
ttccaacatg gatgctgatt tatatgggta taaatgggct cgcgataatg tcgggcaatc 10200
aggtgcgaca atctatcgat tgtatgggaa gcccgatgcg ccagagttgt ttctgaaaca 10260
tggcaaaggt agcgttgcca atgatgttac agatgagatg gtcagactaa actggctgac 10320
ggaatttatg cctcttccga ccatcaagca ttttatccgt actcctgatg atgcatggtt 10380

actcaccact gcgatccccg ggaaaacagc attccaggta ttagaagaat atcctgattc 10440

aggtgaaaat attgttgatg cgctggcagt gttcctgcgc cggttgcatt cgattcctgt 10500
ttgtaattgt ccttttaaca gcgatcgcgt atttcgtctc gctcaggcgc aatcacgaat 10560
gaataacggt ttggttgatg cgagtgattt tgatgacgag cgtaatggct ggcctgttga 10620
acaagtctgg aaagaaatgc ataagctttt gccattctca ccggattcag tcgtcactca 10680
tggtgatttc tcacttgata accttatttt tgacgagggg aaattaatag gttgtattga 10740
tgttggacga gtcggaatcg cagaccgata ccaggatctt gccatcctat ggaactgcct 10800
cggtgagttt tctccttcat tacagaaacg gctttttcaa aaatatggta ttgataatcc 10860

tgatatgaat aaattgcagt ttcatttgat gctcgatgag tttttctaag ggcggcctgc 10920
caccataccc acgccgaaac aagcgctcat gagcccgaag tggcgagccc gatcttcccc 10980
atcggtgatg tcggcgatat aggcgccagc aaccgcacct gtggcgccgg tgatgagggc 11040
gcgccaagtc gacgtccggc agtc 11064

<210> 45
<211> 250
<212> PRT
<213> Artificial sequence
<220><223> Synthetic polypeptide
<400> 45

Met Glu Lys Gly Pro Val Arg Ala Pro Ala Glu Lys Pro Arg Gly Ala
1 5 10 15

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

Ser Leu Gly Glu Asn Pro Gln Arg Gln Gly Leu Leu Lys Thr Pro Trp
85 90 95
Arg Ala Ala Ser Ala Met Gln Phe Phe Thr Lys Gly Tyr Gln Glu Thr

100 105 110

Ile Ser Asp Val Leu Asn Asp Ala Ile Phe Asp Glu Asp His Asp Glu

115 120 125

Met Val Ile Val Lys Asp Ile Asp Met Phe Ser Met Cys Glu His His

130 135 140

Leu Val Pro Phe Val Gly Lys Val His Ile Gly Tyr Leu Pro Asn Lys

145 150 155 160

Gln Val Leu Gly Leu Ser Lys Leu Ala Arg Ile Val Glu Ile Tyr Ser

165 170 175

Arg Arg Leu Gln Val Gln Glu Arg Leu Thr Lys Gln Ile Ala Val Ala

180 185 190

Ile Thr Glu Ala Leu Arg Pro Ala Gly Val Gly Val Val Val Glu Ala

195 200 205

Thr His Met Cys Met Val Met Arg Gly Val Gln Lys Met Asn Ser Lys

210 215 220

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

225 230 235 240

Arg Glu Glu Phe Leu Thr Leu Ile Arg Ser

245 250

<210> 46

<211> 750

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 46

atggagaagg gccccgtgcg cgcccccgcc gagaagcccc gcggcgcccg ctgcagcaac 60

ggcttccccg agcgcgaccc cccccgcccc ggccccagcc gccccgccga gaagcccccc 120

cgccccgagg ccaagagcgc ccagcccgcc gacggctgga agggcgagcg cccccgcagc 180

gaggaggaca acgagctgaa cctgcccaac ctggccgccg cctacagcag catcctgagc 240

agcctgggcg agaacccccа gcgccagggc ctgctgaaga ccccctggcg cgccgccagc 300

gccatgcagt tcttcaccaa gggctaccag gagaccatca gcgacgtgct gaacgacgcc 360

atcttcgacg aggaccacga cgagatggtg atcgtgaagg acatcgacat gttcagcatg 420

tgcgagcacc acctggtgcc cttcgtgggc aaggtgcaca tcggctacct gcccaacaag 480

caggtgctgg gcctgagcaa gctggcccgc atcgtggaga tctacagccg ccgcctgcag 540

gtgcaggagc gcctgaccaa gcagatcgcc gtggccatca ccgaggccct gcgccccgcc 600

ggcgtgggcg tggtggtgga ggccacccac atgtgcatgg tgatgcgcgg cgtgcagaag 660

atgaacagca agaccgtgac cagcaccatg ctgggcgtgt tccgcgagga ccccaagacc 720

cgcgaggagt tcctgaccct gatccgcagc 750

<210> 47

<211> 203

<212> PRT

<213> Artificial sequence

<220><223> Synthetic polypeptide

<400> 47

Met Gly Ser Arg Asp His Leu Phe Lys Val Leu Val Val Gly Asp Ala

1 5 10 15

Ala Val Gly Lys Thr Ser Leu Val Gln Arg Tyr Ser Gln Asp Ser Phe

20 25 30

Ser Lys His Tyr Lys Ser Thr Val Gly Val Asp Phe Ala Leu Lys Val

35 40 45

Leu Gln Trp Ser Asp Tyr Glu Ile Val Arg Leu Gln Leu Trp Asp Ile

50 55 60

Ala Gly Gln Glu Arg Phe Thr Ser Met Thr Arg Leu Tyr Tyr Arg Asp

65 70 75 80

Ala Ser Ala Cys Val Ile Met Phe Asp Val Thr Asn Ala Thr Thr Phe

85 90 95

Ser Asn Ser Gln Arg Trp Lys Gln Asp Leu Asp Ser Lys Leu Thr Leu

100 105 110

Pro Asn Gly Glu Pro Val Pro Cys Leu Leu Leu Ala Asn Lys Cys Asp

115 120 125

Leu Ser Pro Trp Ala Val Ser Arg Asp Gln Ile Asp Arg Phe Ser Lys

130 135 140

Glu Asn Gly Phe Thr Gly Trp Thr Glu Thr Ser Val Lys Glu Asn Lys

145 150 155 160

Asn Ile Asn Glu Ala Met Arg Val Leu Ile Glu Lys Met Met Arg Asn

165 170 175

Ser Thr Glu Asp Ile Met Ser Leu Ser Thr Gln Gly Asp Tyr Ile Asn

180 185 190

Leu Gln Thr Lys Ser Ser Ser Trp Ser Cys Cys

195 200

<210> 48

<211> 609

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 48

atgggcagcc gcgaccacct gttcaaggtg ctggtggtgg gcgacgccgc cgtgggcaag 60

accagcctgg tgcagcgcta cagccaggac agcttcagca agcactacaa gagcaccgtg 120

ggcgtggact tcgccctgaa ggtgctgcag tggagcgact acgagatcgt gcgcctgcag 180

ctgtgggaca tcgccggcca ggagcgcttc accagcatga cccgcctgta ctaccgcgac 240

gccagcgcct gcgtgatcat gttcgacgtg accaacgcca ccaccttcag caacagccag 300

cgctggaagc aggacctgga cagcaagctg accctgccca acggcgagcc cgtgccctgc 360

ctgctgctgg ccaacaagtg cgacctgagc ccctgggccg tgagccgcga ccagatcgac 420

cgcttcagca aggagaacgg cttcaccggc tggaccgaga ccagcgtgaa ggagaacaag 480

aacatcaacg aggccatgcg cgtgctgatc gagaagatga tgcgcaacag caccgaggac 540

atcatgagcc tgagcaccca gggcgactac atcaacctgc agaccaagag cagcagctgg 600

agctgctgc 609

<210> 49

<211> 796

<212> PRT

<213> Artificial sequence

<220><223> Synthetic polypeptide

<400> 49

Met Pro Thr Thr Gln Gln Ser Pro Gln Asp Glu Gln Glu Lys Leu Leu

1 5 10 15

Asp Glu Ala Ile Gln Ala Val Lys Val Gln Ser Phe Gln Met Lys Arg
20 25 30
Cys Leu Asp Lys Asn Lys Leu Met Asp Ala Leu Lys His Ala Ser Asn
35 40 45
Met Leu Gly Glu Leu Arg Thr Ser Met Leu Ser Pro Lys Ser Tyr Tyr
50 55 60
Glu Leu Tyr Met Ala Ile Ser Asp Glu Leu His Tyr Leu Glu Val Tyr

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

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

195 200 205
Arg Glu Arg Glu Arg Gln Glu Leu Arg Ile Leu Val Gly Thr Asn Leu
210 215 220
Val Arg Leu Ser Gln Leu Glu Gly Val Asn Val Glu Arg Tyr Lys Gln
225 230 235 240
Ile Val Leu Thr Gly Ile Leu Glu Gln Val Val Asn Cys Arg Asp Ala
245 250 255
Leu Ala Gln Glu Tyr Leu Met Glu Cys Ile Ile Gln Val Phe Pro Asp

260 265 270

Glu Phe His Leu Gln Thr Leu Asn Pro Phe Leu Arg Ala Cys Ala Glu

275 280 285

Leu His Gln Asn Val Asn Val Lys Asn Ile Ile Ile Ala Leu Ile Asp

290 295 300

Arg Leu Ala Leu Phe Ala His Arg Glu Asp Gly Pro Gly Ile Pro Ala

305 310 315 320

Asp Ile Lys Leu Phe Asp Ile Phe Ser Gln Gln Val Ala Thr Val Ile

325 330 335

Gln Ser Arg Gln Asp Met Pro Ser Glu Asp Val Val Ser Leu Gln Val

340 345 350

Ser Leu Ile Asn Leu Ala Met Lys Cys Tyr Pro Asp Arg Val Asp Tyr

355 360 365

Val Asp Lys Val Leu Glu Thr Thr Val Glu Ile Phe Asn Lys Leu Asn

370 375 380

Leu Glu His Ile Ala Thr Ser Ser Ala Val Ser Lys Glu Leu Thr Arg

385 390 395 400

Leu Leu Lys Ile Pro Val Asp Thr Tyr Asn Asn Ile Leu Thr Val Leu

405 410 415

Lys Leu Lys His Phe His Pro Leu Phe Glu Tyr Phe Asp Tyr Glu Ser

420 425 430

Arg Lys Ser Met Ser Cys Tyr Val Leu Ser Asn Val Leu Asp Tyr Asn

435 440 445

Thr Glu Ile Val Ser Gln Asp Gln Val Asp Ser Ile Met Asn Leu Val

450 455 460

Ser Thr Leu Ile Gln Asp Gln Pro Asp Gln Pro Val Glu Asp Pro Asp

465 470 475 480

Pro Glu Asp Phe Ala Asp Glu Gln Ser Leu Val Gly Arg Phe Ile His

485 490 495

Leu Leu Arg Ser Glu Asp Pro Asp Gln Gln Tyr Leu Ile Leu Asn Thr

500 505 510

Ala Arg Lys His Phe Gly Ala Gly Gly Asn Gln Arg Ile Arg Phe Thr

515 520 525

Leu Pro Pro Leu Val Phe Ala Ala Tyr Gln Leu Ala Phe Arg Tyr Lys

530 535 540

Glu Asn Ser Lys Val Asp Asp Lys Trp Glu Lys Lys Cys Gln Lys Ile

545 550 555 560

Phe Ser Phe Ala His Gln Thr Ile Ser Ala Leu Ile Lys Ala Glu Leu

565 570 575

Ala Glu Leu Pro Leu Arg Leu Phe Leu Gln Gly Ala Leu Ala Ala Gly

580 585 590

Glu Ile Gly Phe Glu Asn His Glu Thr Val Ala Tyr Glu Phe Met Ser

595 600 605

Gln Ala Phe Ser Leu Tyr Glu Asp Glu Ile Ser Asp Ser Lys Ala Gln

610 615 620

Leu Ala Ala Ile Thr Leu Ile Ile Gly Thr Phe Glu Arg Met Lys Cys

625 630 635 640

Phe Ser Glu Glu Asn His Glu Pro Leu Arg Thr Gln Cys Ala Leu Ala

645 650 655

Ala Ser Lys Leu Leu Lys Lys Pro Asp Gln Gly Arg Ala Val Ser Thr

660 665 670

Cys Ala His Leu Phe Trp Ser Gly Arg Asn Thr Asp Lys Asn Gly Glu

675 680 685

Glu Leu His Gly Gly Lys Arg Val Met Glu Cys Leu Lys Lys Ala Leu

690 695 700

Lys Ile Ala Asn Gln Cys Met Asp Pro Ser Leu Gln Val Gln Leu Phe

705 710 715 720

Ile Glu Ile Leu Asn Arg Tyr Ile Tyr Phe Tyr Glu Lys Glu Asn Asp

725 730 735

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

740 745 750

Asp Leu Pro Asn Leu Glu Ser Ser Glu Glu Thr Glu Gln Ile Asn Lys

755 760 765

His Phe His Asn Thr Leu Glu His Leu Arg Leu Arg Arg Glu Ser Pro

770 775 780

Glu Ser Glu Gly Pro Ile Tyr Glu Gly Leu Ile Leu

785 790 795

<210> 50

<211> 2388

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 50

atgcccacca cccagcagag cccccaggac gagcaggaga agctgctgga cgaggccatc 60

caggccgtga aggtgcagag cttccagatg aagcgctgcc tggacaagaa caagctgatg 120

gacgccctga agcacgccag caacatgctg ggcgagctgc gcaccagcat gctgagcccc 180

aagagctact acgagctgta catggccatc agcgacgagc tgcactacct ggaggtgtac 240

ctgaccgacg agttcgccaa gggccgcaag gtggccgacc tgtacgagct ggtgcagtac 300

gccggcaaca tcatcccccg cctgtacctg ctgatcaccg tgggcgtggt gtacgtgaag 360

agcttccccc agagccgcaa ggacatcctg aaggacctgg tggagatgtg ccgcggcgtg 420

cagcaccccc tgcgcggcct gttcctgcgc aactacctgc tgcagtgcac ccgcaacatc 480

ctgcccgacg agggcgagcc caccgacgag gagaccaccg gcgacatcag cgacagcatg 540

gacttcgtgc tgctgaactt cgccgagatg aacaagctgt gggtgcgcat gcagcaccag 600

ggccacagcc gcgaccgcga gaagcgcgag cgcgagcgcc aggagctgcg catcctggtg 660

ggcaccaacc tggtgcgcct gagccagctg gagggcgtga acgtggagcg ctacaagcag 720

atcgtgctga ccggcatcct ggagcaggtg gtgaactgcc gcgacgccct ggcccaggag 780

tacctgatgg agtgcatcat ccaggtgttc cccgacgagt tccacctgca gaccctgaac 840

cccttcctgc gcgcctgcgc cgagctgcac cagaacgtga acgtgaagaa catcatcatc 900

gccctgatcg accgcctggc cctgttcgcc caccgcgagg acggccccgg catccccgcc 960

gacatcaagc tgttcgacat cttcagccag caggtggcca ccgtgatcca gagccgccag 1020

gacatgccca gcgaggacgt ggtgagcctg caggtgagcc tgatcaacct ggccatgaag 1080

tgctaccccg accgcgtgga ctacgtggac aaggtgctgg agaccaccgt ggagatcttc 1140

aacaagctga acctggagca catcgccacc agcagcgccg tgagcaagga gctgacccgc 1200

ctgctgaaga tccccgtgga cacctacaac aacatcctga ccgtgctgaa gctgaagcac 1260

ttccaccccc tgttcgagta cttcgactac gagagccgca agagcatgag ctgctacgtg 1320

ctgagcaacg tgctggacta caacaccgag atcgtgagcc aggaccaggt ggacagcatc 1380

atgaacctgg tgagcaccct gatccaggac cagcccgacc agcccgtgga ggaccccgac 1440

cccgaggact tcgccgacga gcagagcctg gtgggccgct tcatccacct gctgcgcagc 1500

gaggaccccg accagcagta cctgatcctg aacaccgccc gcaagcactt cggcgccggc 1560

ggcaaccagc gcatccgctt caccctgccc cccctggtgt tcgccgccta ccagctggcc 1620

ttccgctaca aggagaacag caaggtggac gacaagtggg agaagaagtg ccagaagatc 1680

ttcagcttcg cccaccagac catcagcgcc ctgatcaagg ccgagctggc cgagctgccc 1740

ctgcgcctgt tcctgcaggg cgccctggcc gccggcgaga tcggcttcga gaaccacgag 1800

accgtggcct acgagttcat gagccaggcc ttcagcctgt acgaggacga gatcagcgac 1860

agcaaggccc agctggccgc catcaccctg atcatcggca ccttcgagcg catgaagtgc 1920

ttcagcgagg agaaccacga gcccctgcgc acccagtgcg ccctggccgc cagcaagctg 1980

ctgaagaagc ccgaccaggg ccgcgccgtg agcacctgcg cccacctgtt ctggagcggc 2040

cgcaacaccg acaagaacgg cgaggagctg cacggcggca agcgcgtgat ggagtgcctg 2100

aagaaggccc tgaagatcgc caaccagtgc atggacccca gcctgcaggt gcagctgttc 2160

atcgagatcc tgaaccgcta catctacttc tacgagaagg agaacgacgc cgtgaccatc 2220

caggtgctga accagctgat ccagaagatc cgcgaggacc tgcccaacct ggagagcagc 2280

gaggagaccg agcagatcaa caagcacttc cacaacaccc tggagcacct gcgcctgcgc 2340

cgcgagagcc ccgagagcga gggccccatc tacgagggcc tgatcctg 2388

<210> 51

<211> 11081

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 51

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360
atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420
tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480
agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540
tggcactatg aaccctcctg gtggcgaggg gagggggtg gtcctcgaac gccttgcaga 600
actggcctgg atacagagtg gaccggctgg ccccatctgg aagacttcga gatacactgt 660

tgtcttactg cgctcaacag tgtatctcga agtcttccaa atggtgccag ccatcgcagc 720
ggggtgcagg aaatgggggc agcccccctt tttggctatc cttccacgtg ttctttttg 780
tatcttttgt gtttcctaga aaacatctca gtcaccaccg cagccctagg aatgcatcta 840
gacaattgta ctaaccttct tctctttcct ctcctgacag tccggaaagc caccatggaa 900
ttcagcagcc ccagcagaga ggaatgcccc aagcctctga gccgggtgtc aatcatggcc 960
ggatctctga caggactgct gctgcttcag gccgtgtctt gggcttctgg cgctagacct 1020
tgcatcccca agagcttcgg ctacagcagc gtcgtgtgcg tgtgcaatgc cacctactgc 1080

gacagcttcg accctcctac ctttcctgct ctgggcacct tcagcagata cgagagcacc 1140
agatccggca gacggatgga actgagcatg ggacccatcc aggccaatca cacaggcact 1200
ggcctgctgc tgacactgca gcctgagcag aaattccaga aagtgaaagg cttcggcgga 1260
gccatgacag atgccgccgc tctgaatatc ctggctctgt ctccaccagc tcagaacctg 1320
ctgctcaaga gctacttcag cgaggaaggc atcggctaca acatcatcag agtgcccatg 1380
gccagctgcg acttcagcat caggacctac acctacgccg acacacccga cgatttccag 1440
ctgcacaact tcagcctgcc tgaagaggac accaagctga agatccctct gatccacaga 1500

gccctgcagc tggcacaaag acccgtgtca ctgctggcct ctccatggac atctcccacc 1560
tggctgaaaa caaatggcgc cgtgaatggc aagggcagcc tgaaaggcca acctggcgac 1620
atctaccacc agacctgggc cagatacttc gtgaagttcc tggacgccta tgccgagcac 1680
aagctgcagt tttgggccgt gacagccgag aacgaacctt ctgctggact gctgagcggc 1740
taccccttc agtgcctggg ctttacaccc gagcaccagc gggactttat cgcccgtgat 1800
ctgggaccca cactggccaa tagcacccac cataatgtgc ggctgctgat gctggacgac 1860
cagagactgc ttctgcccca ctgggctaaa gtggtgctga cagatcctga ggccgccaaa 1920

tacgtgcacg gaatcgccgt gcactggtat ctggactttc tggcccctgc caaggccaca 1980
ctgggagaga cacacagact gttccccaac accatgctgt tcgccagcga agcctgtgtg 2040
ggcagcaagt tttgggaaca gagcgtgcgg ctcggcagct gggatagagg catgcagtac 2100
agccacagca tcatcaccaa cctgctgtac cacgtcgtcg gctggaccga ctggaatctg 2160

gccctgaatc ctgaaggcgg ccctaactgg gtccgaaact tcgtggacag ccccatcatc 2220
gtggacatca ccaaggacac cttctacaag cagcccatgt tctaccacct gggacacttc 2280
agcaagttca tccccgaggg ctctcagcgc gttggactgg tggcttccca gaagaacgat 2340

ctggacgccg tggctctgat gcaccctgat ggatctgctg tggtggtggt cctgaaccgc 2400
agcagcaaag atgtgcccct gaccatcaag gatcccgccg tgggattcct ggaaacaatc 2460
agccctggct actccatcca cacctacctg tggcgtagac aggagggcag aggaagtctt 2520
ctgacatgcg gagacgtgga agagaatccc ggccctatgg agaagggccc cgtgcgcgcc 2580
cccgccgaga agccccgcgg cgcccgctgc agcaacggct tccccgagcg cgacccccccc 2640
cgccccggcc ccagccgccc cgccgagaag cccccccgcc ccgaggccaa gagcgcccag 2700
cccgccgacg gctggaaggg cgagcgcccc cgcagcgagg aggacaacga gctgaacctg 2760

cccaacctgg ccgccgccta cagcagcatc ctgagcagcc tgggcgagaa cccccagcgc 2820
cagggcctgc tgaagacccc ctggcgcgcc gccagcgcca tgcagttctt caccaagggc 2880
taccaggaga ccatcagcga cgtgctgaac gacgccatct tcgacgagga ccacgacgag 2940
atggtgatcg tgaaggacat cgacatgttc agcatgtgcg agcaccacct ggtgcccttc 3000
gtgggcaagg tgcacatcgg ctacctgccc aacaagcagg tgctgggcct gagcaagctg 3060
gcccgcatcg tggagatcta cagccgccgc ctgcaggtgc aggagcgcct gaccaagcag 3120
atcgccgtgg ccatcaccga ggccctgcgc cccgccggcg tgggcgtggt ggtggaggcc 3180

acccacatgt gcatggtgat gcgcggcgtg cagaagatga acagcaagac cgtgaccagc 3240
accatgctgg gcgtgttccg cgaggacccc aagacccgcg aggagttcct gaccctgatc 3300
cgcagctgac aattgttaat taagtttaaa ccctcgaggc cgcaagctta tcgataatca 3360
acctctggat tacaaaattt gtgaaagatt gactggtatt cttaactatg ttgctccttt 3420
tacgctatgt ggatacgctg ctttaatgcc tttgtatcat gctattgctt cccgtatggc 3480
tttcattttc tcctccttgt ataaatcctg gttgctgtct ctttatgagg agttgtggcc 3540
cgttgtcagg caacgtggcg tggtgtgcac tgtgtttgct gacgcaaccc ccactggttg 3600

gggcattgcc accacctgtc agctcctttc cgggactttc gctttccccc tccctattgc 3660
cacggcggaa ctcatcgccg cctgccttgc ccgctgctgg acaggggctc ggctgttggg 3720
cactgacaat tccgtggtgt tgtcggggaa atcatcgtcc tttccttggc tgctcgcctg 3780
tgttgccacc tggattctgc gcgggacgtc cttctgctac gtcccttcgg ccctcaatcc 3840
agcggacctt ccttcccgcg gcctgctgcc ggctctgcgg cctcttccgc gtcttcgcct 3900
tcgccctcag acgagtcgga tctccctttg ggccgcctcc ccgcatcgat accgtcgact 3960

agagctcgct gatcagcctc gactgtgcct tctagttgcc agccatctgt tgtttgcccc 4020

tcccccgtgc cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat 4080
gaggaaattg catcgcattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg 4140
caggacagca agggggagga ttgggaagac aatagcaggc atgctgggga gagatccacg 4200
ataacaaaca gctttttttgg ggtgaacata ttgactgaat tccctgcagg ttggccactc 4260
cctctctgcg cgctcgctcg ctcactgagg ccgcccgggc aaagcccggg cgtcgggcga 4320
cctttggtcg cccggcctca gtgagcgagc gagcgcgcag agagggagtg gccaactcca 4380
tcactagggg ttcctgcggc cgctcgtacg gtctcgagga attcctgcag gataacttgc 4440

caacctcatt ctaaaatgta tatagaagcc caaaagacaa taacaaaaat attcttgtag 4500
aacaaaatgg gaaagaatgt tccactaaat atcaagattt agagcaaagc atgagatgtg 4560
tggggataga cagtgaggct gataaaatag agtagagctc agaaacagac ccattgatat 4620
atgtaagtga cctatgaaaa aaatatggca ttttacaatg ggaaaatgat ggtctttttc 4680
ttttttagaa aaacagggaa atatatttat atgtaaaaaa taaaagggaa cccatatgtc 4740
ataccataca cacaaaaaaa ttccagtgaa ttataagtct aaatggagaa ggcaaaactt 4800
taaatctttt agaaaataat atagaagcat gcagaccagc ctggccaaca tgatgaaacc 4860

ctctctacta ataataaaat cagtagaact actcaggact actttgagtg ggaagtcctt 4920
ttctatgaag acttctttgg ccaaaattag gctctaaatg caaggagata gtgcatcatg 4980
cctggctgca cttactgata aatgatgtta tcaccatctt taaccaaatg cacaggaaca 5040
agttatggta ctgatgtgct ggattgagaa ggagctctac ttccttgaca ggacacattt 5100
gtatcaactt aaaaaagcag atttttgcca gcagaactat tcattcagag gtaggaaact 5160
tagaatagat gatgtcactg attagcatgg cttccccatc tccacagctg cttcccaccc 5220
aggttgccca cagttgagtt tgtccagtgc tcagggctgc ccactctcag taagaagccc 5280

cacaccagcc cctctccaaa tatgttggct gttccttcca ttaaagtgac cccactttag 5340
agcagcaagt ggatttctgt ttcttacagt tcaggaagga ggagtcagct gtgagaacct 5400
ggagcctgag atgcttctaa gtcccactgc tactggggtc agggaagcca gactccagca 5460
tcagcagtca ggagcactaa gcccttgcca acatcctgtt tctcagagaa actgcttcca 5520
ttataatggt tgtccttttt taagctatca agccaaacaa ccagtgtcta ccattattct 5580
catcacctga agccaagggt tctagcaaaa gtcaagctgt cttgtaatgg ttgatgtgcc 5640
tccagcttct gtcttcagtc actccactct tagcctgctc tgaatcaact ctgaccacag 5700

ttccctggag cccctgccac ctgctgcccc tgccaccttc tccatctgca gtgctgtgca 5760
gccttctgca ctcttgcaga gctaataggt ggagacttga aggaagagga ggaaagtttc 5820
tcataatagc cttgctgcaa gctcaaatgg gaggtgggca ctgtgcccag gagccttgga 5880
gcaaaggctg tgcccaacct ctgactgcat ccaggtttgg tcttgacaga gataagaagc 5940
cctggctttt ggagccaaaa tctaggtcag acttaggcag gattctcaaa gtttatcagc 6000
agaacatgag gcagaagacc ctttctgctc cagcttcttc aggctcaacc ttcatcagaa 6060
tagatagaaa gagaggctgt gagggttctt aaaacagaag caaatctgac tcagagaata 6120

aacaacctcc tagtaaacta cagcttagac agagcatctg gtggtgagtg tgctcagtgt 6180
cctactcaac tgtctggtat cagccctcat gaggacttct cttctttccc tcatagacct 6240
ccatctctgt tttccttagc ctgcagaaat ctggatggct attcacagaa tgcctgtgct 6300
ttcagagttg catttttct ctggtattct ggttcaagca tttgaaggta ggaaaggttc 6360
tccaagtgca agaaagccag ccctgagcct caactgcctg gctagtgtgg tcagtaggat 6420
gcaaaggctg ttgaatgcca caaggccaaa ctttaacctg tgtaccacaa gcctagcagc 6480
agaggcagct ctgctcactg gaactctctg tcttctttct cctgagcctt ttcttttcct 6540

gagttttcta gctctcctca accttacctc tgccctaccc aggacaaacc caagagccac 6600
tgtttctgtg atgtcctctc cagccctaat taggcatcat gacttcagcc tgaccttcca 6660
tgctcagaag cagtgctaat ccacttcaga tgagctgctc tatgcaacac aggcagagcc 6720
tacaaacctt tgcaccagag ccctccacat atcagtgttt gttcatactc acttcaacag 6780
caaatgtgac tgctgagatt aagattttac acaagatggt ctgtaatttc acagttagtt 6840
ttatcccatt aggtatgaaa gaattagcat aattcccctt aaacatgaat gaatcttaga 6900
ttttttaata aatagttttg gaagtaaaga cagagacatc aggagcacaa ggaatagcct 6960

gagaggacaa acagaacaag aaagagtctg gaaatacaca ggatgttctt ggcctcctca 7020
aagcaagtgc aagcagatag taccagcagc cccaggctat cagagcccag tgaagagaag 7080
taccatgaaa gccacagctc taaccaccct gttccagagt gacagacagt ccccaagaca 7140
agccagcctg agccagagag agaactgcaa gagaaagttt ctaatttagg ttctgttaga 7200
ttcagacaag tgcaggtcat cctctctcca cagctactca cctctccagc ctaacaaagc 7260
ctgcagtcca cactccaacc ctggtgtctc acctcctagc ctctcccaac atcctgctct 7320
ctgaccatct tctgcatctc tcatctcacc atctcccact gtctacagcc tactcttgca 7380

actaccatct cattttctga catcctgtct acatcttctg ccatactctg ccatctacca 7440
taccacctct taccatctac cacaccatct tttatctcca tccctctcag aagcctccaa 7500
gctgaatcct gctttatgtg ttcatctcag cccctgcatg gaaagctgac cccagaggca 7560

gaactattcc cagagagctt ggccaagaaa aacaaaacta ccagcctggc caggctcagg 7620
agtagtaagc tgcagtgtct gttgtgttct agcttcaaca gctgcaggag ttccactctc 7680
aaatgctcca catttctcac atcctcctga ttctggtcac tacccatctt caaagaacag 7740
aatatctcac atcagcatac tgtgaaggac tagtcatggg tgcagctgct cagagctgca 7800

aagtcattct ggatggtgga gagcttacaa acatttcatg atgctccccc cgctctgatg 7860
gctggagccc aatccctaca cagactcctg ctgtatgtgt tttcctttca ctctgagcca 7920
cagccagagg gcaggcattc agtctcctct tcaggctggg gctggggcac tgagaactca 7980
cccaacacct tgctctcact ccttctgcaa aacaagaaag agctttgtgc tgcagtagcc 8040
atgaagaatg aaaggaaggc tttaactaaa aaatgtcaga gattattttc aacccccttac 8100
tgtggatcac cagcaaggag gaaacacaac acagagacat tttttcccct caaattatca 8160
aaagaatcac tgcatttgtt aaagagagca actgaatcag gaagcagagt tttgaacata 8220

tcagaagtta ggaatctgca tcagagacaa atgcagtcat ggttgtttgc tgcataccag 8280
ccctaatcat tagaagcctc atggacttca aacatcattc cctctgacaa gatgctctag 8340
cctaactcca tgagataaaa taaatctgcc tttcagagcc aaagaagagt ccaccagctt 8400
cttctcagtg tgaacaagag ctccagtcag gttagtcagt ccagtgcagt agaggagacc 8460
agtctgcatc ctctaatttt caaaggcaag aagatttgtt taccctggac accaggcaca 8520
agtgaggtca cagagctctt agatatgcag tcctcatgag tgaggagact aaagcgcatg 8580
ccatcaagac ttcagtgtag agaaaacctc caaaaaagcc tcctcactac ttctggaata 8640

gctcagaggc cgaggcggcc tcggcctctg cataaataaa aaaaattagt cagccatggg 8700
gcggagaatg ggcggaactg ggcggagtta ggggcgggat gggcggagtt aggggcggga 8760
ctatggttgc tgactaattg agatgcatgc tttgcatact tctgcctgct ggggagcctg 8820
gggactttcc acacctggtt gctgactaat tgagatgcat gctttgcata cttctgcctg 8880
ctggggagcc tggggacttt ccacacccta actgacacac attccacagc tgcattaatg 8940
aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg cttcctcgct 9000
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc 9060

ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 9120
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 9180
ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 9240
actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc ctgttccgac 9300
cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 9360
tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt 9420

gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 9480

caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 9540
agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 9600
tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 9660
tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa 9720
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 9780
gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa 9840
aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat 9900

atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc 9960
gatctgtcta tttcgttcat ccatagttgc ctgactcctg caaaccacgt tgtgtctcaa 10020
aatctctgat gttacattgc acaagataaa aatatatcat catgaacaat aaaactgtct 10080
gcttacataa acagtaatac aaggggtgtt atgagccata ttcaacggga aacgtcttgc 10140
tcgaggccgc gattaaattc caacatggat gctgatttat atgggtataa atgggctcgc 10200
gataatgtcg ggcaatcagg tgcgacaatc tatcgattgt atgggaagcc cgatgcgcca 10260
gagttgtttc tgaaacatgg caaaggtagc gttgccaatg atgttacaga tgagatggtc 10320

agactaaact ggctgacgga atttatgcct cttccgacca tcaagcattt tatccgtact 10380
cctgatgatg catggttact caccactgcg atccccggga aaacagcatt ccaggtatta 10440
gaagaatatc ctgattcagg tgaaaatatt gttgatgcgc tggcagtgtt cctgcgccgg 10500
ttgcattcga ttcctgtttg taattgtcct tttaacagcg atcgcgtatt tcgtctcgct 10560
caggcgcaat cacgaatgaa taacggtttg gttgatgcga gtgattttga tgacgagcgt 10620
aatggctggc ctgttgaaca agtctggaaa gaaatgcata agcttttgcc attctcaccg 10680
gattcagtcg tcactcatgg tgatttctca cttgataacc ttatttttga cgaggggaaa 10740

ttaataggtt gtattgatgt tggacgagtc ggaatcgcag accgatacca ggatcttgcc 10800
atcctatgga actgcctcgg tgagttttct ccttcattac agaaacggct ttttcaaaaa 10860
tatggtattg ataatcctga tatgaataaa ttgcagtttc atttgatgct cgatgagttt 10920
ttctaagggc ggcctgccac catacccacg ccgaaacaag cgctcatgag cccgaagtgg 10980
cgagcccgat cttccccatc ggtgatgtcg gcgatatagg cgccagcaac cgcacctgtg 11040
gcgccggtga tgagggcgcg ccaagtcgac gtccggcagt c 11081

<210> 52

<211> 10940

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 52

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420

tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctcctg gtggcgaggg gagggggggtg gtcctcgaac gccttgcaga 600

actggcctgg atacagagtg gaccggctgg ccccatctgg aagacttcga gatacactgt 660

tgtcttactg cgctcaacag tgtatctcga agtcttccaa atggtgccag ccatcgcagc 720

ggggtgcagg aaatgggggc agcccccctt tttggctatc cttccacgtg ttctttttg 780

tatcttttgt gtttcctaga aaacatctca gtcaccaccg cagccctagg aatgcatcta 840

gacaattgta ctaaccttct tctctttcct ctcctgacag tccggaaagc caccatgggc 900

agccgcgacc acctgttcaa ggtgctggtg gtgggcgacg ccgccgtggg caagaccagc 960

ctggtgcagc gctacagcca ggacagcttc agcaagcact acaagagcac cgtgggcgtg 1020

gacttcgccc tgaaggtgct gcagtggagc gactacgaga tcgtgcgcct gcagctgtgg 1080

gacatcgccg gccaggagcg cttcaccagc atgacccgcc tgtactaccg cgacgccagc 1140

gcctgcgtga tcatgttcga cgtgaccaac gccaccacct tcagcaacag ccagcgctgg 1200

aagcaggacc tggacagcaa gctgaccctg cccaacggcg agcccgtgcc ctgcctgctg 1260

ctggccaaca agtgcgacct gagcccctgg gccgtgagcc gcgaccagat cgaccgcttc 1320

agcaaggaga acggcttcac cggctggacc gagaccagcg tgaaggagaa caagaacatc 1380

aacgaggcca tgcgcgtgct gatcgagaag atgatgcgca acagcaccga ggacatcatg 1440

agcctgagca cccagggcga ctacatcaac ctgcagacca agagcagcag ctggagctgc 1500

tgcgagggca gaggaagtct tctgacatgc ggagacgtgg aagagaatcc cggccctatg 1560

gaattcagca gccccagcag agaggaatgc cccaagcctc tgagccgggt gtcaatcatg 1620

gccggatctc tgacaggact gctgctgctt caggccgtgt cttgggcttc tggcgctaga 1680
ccttgcatcc ccaagagctt cggctacagc agcgtcgtgt gcgtgtgcaa tgccacctac 1740
tgcgacagct tcgaccctcc tacctttcct gctctgggca ccttcagcag atacgagagc 1800
accagatccg gcagacggat ggaactgagc atgggaccca tccaggccaa tcacacaggc 1860
actggcctgc tgctgacact gcagcctgag cagaaattcc agaaagtgaa aggcttcggc 1920
ggagccatga cagatgccgc cgctctgaat atcctggctc tgtctccacc agctcagaac 1980
ctgctgctca agagctactt cagcgaggaa ggcatcggct acaacatcat cagagtgccc 2040

atggccagct gcgacttcag catcaggacc tacacctacg ccgacacacc cgacgatttc 2100
cagctgcaca acttcagcct gcctgaagag gacaccaagc tgaagatccc tctgatccac 2160
agagccctgc agctggcaca aagacccgtg tcactgctgg cctctccatg gacatctccc 2220
acctggctga aaacaaatgg cgccgtgaat ggcaagggca gcctgaaagg ccaacctggc 2280
gacatctacc accagacctg ggccagatac ttcgtgaagt tcctggacgc ctatgccgag 2340
cacaagctgc agttttgggc cgtgacagcc gagaacgaac cttctgctgg actgctgagc 2400
ggctacccct ttcagtgcct gggctttaca cccgagcacc agcgggactt tatcgcccgt 2460

gatctgggac ccacactggc caatagcacc caccataatg tgcggctgct gatgctggac 2520
gaccagagac tgcttctgcc ccactgggct aaagtggtgc tgacagatcc tgaggccgcc 2580
aaatacgtgc acggaatcgc cgtgcactgg tatctggact ttctggcccc tgccaaggcc 2640
acactgggag agacacacag actgttcccc aacaccatgc tgttcgccag cgaagcctgt 2700
gtgggcagca agttttggga acagagcgtg cggctcggca gctgggatag aggcatgcag 2760
tacagccaca gcatcatcac caacctgctg taccacgtcg tcggctggac cgactggaat 2820
ctggccctga atcctgaagg cggccctaac tgggtccgaa acttcgtgga cagccccatc 2880

atcgtggaca tcaccaagga caccttctac aagcagccca tgttctacca cctgggacac 2940
ttcagcaagt tcatccccga gggctctcag cgcgttggac tggtggcttc ccagaagaac 3000
gatctggacg ccgtggctct gatgcaccct gatggatctg ctgtggtggt ggtcctgaac 3060
cgcagcagca aagatgtgcc cctgaccatc aaggatcccg ccgtgggatt cctggaaaca 3120
atcagccctg gctactccat ccacacctac ctgtggcgta gacagtgaca attgttaatt 3180
aagtttaaac cctcgaggcc gcaagcttat cgataatcaa cctctggatt acaaaatttg 3240
tgaaagattg actggtattc ttaactatgt tgctcctttt acgctatgtg gatacgctgc 3300

tttaatgcct ttgtatcatg ctattgcttc ccgtatggct ttcattttct cctccttgta 3360
taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtcaggc aacgtggcgt 3420
ggtgtgcact gtgtttgctg acgcaacccc cactggttgg ggcattgcca ccacctgtca 3480
gctcctttcc gggactttcg ctttccccct ccctattgcc acggcggaac tcatcgccgc 3540
ctgccttgcc cgctgctgga caggggctcg gctgttgggc actgacaatt ccgtggtgtt 3600
gtcggggaaa tcatcgtcct ttccttggct gctcgcctgt gttgccacct ggattctgcg 3660
cgggacgtcc ttctgctacg tcccttcggc cctcaatcca gcggaccttc cttcccgcgg 3720

cctgctgccg gctctgcggc ctcttccgcg tcttcgcctt cgccctcaga cgagtcggat 3780
ctccctttgg gccgcctccc cgcatcgata ccgtcgacta gagctcgctg atcagcctcg 3840
actgtgcctt ctagttgcca gccatctgtt gtttgcccct cccccgtgcc ttccttgacc 3900
ctggaaggtg ccactcccac tgtcctttcc taataaaatg aggaaattgc atcgcattgt 3960
ctgagtaggt gtcattctat tctggggggt ggggtggggc aggacagcaa gggggaggat 4020
tgggaagaca atagcaggca tgctggggag agatccacga taacaaacag ctttttttggg 4080
gtgaacatat tgactgaatt ccctgcaggt tggccactcc ctctctgcgc gctcgctcgc 4140

tcactgaggc cgcccgggca aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag 4200
tgagcgagcg agcgcgcaga gagggagtgg ccaactccat cactaggggt tcctgcggcc 4260
gctcgtacgg tctcgaggaa ttcctgcagg ataacttgcc aacctcattc taaaatgtat 4320
atagaagccc aaaagacaat aacaaaaata ttcttgtaga acaaaatggg aaagaatgtt 4380
ccactaaata tcaagattta gagcaaagca tgagatgtgt ggggatagac agtgaggctg 4440
ataaaataga gtagagctca gaaacagacc cattgatata tgtaagtgac ctatgaaaaa 4500
aatatggcat tttacaatgg gaaaatgatg gtctttttct tttttagaaa aacagggaaa 4560

tatatttata tgtaaaaaat aaaagggaac ccatatgtca taccatacac acaaaaaaat 4620
tccagtgaat tataagtcta aatggagaag gcaaaacttt aaatctttta gaaaataata 4680
tagaagcatg cagaccagcc tggccaacat gatgaaaccc tctctactaa taataaaatc 4740
agtagaacta ctcaggacta ctttgagtgg gaagtccttt tctatgaaga cttctttggc 4800
caaaattagg ctctaaatgc aaggagatag tgcatcatgc ctggctgcac ttactgataa 4860
atgatgttat caccatcttt aaccaaatgc acaggaacaa gttatggtac tgatgtgctg 4920
gattgagaag gagctctact tccttgacag gacacatttg tatcaactta aaaaagcaga 4980

tttttgccag cagaactatt cattcagagg taggaaactt agaatagatg atgtcactga 5040
ttagcatggc ttccccatct ccacagctgc ttcccaccca ggttgcccac agttgagttt 5100
gtccagtgct cagggctgcc cactctcagt aagaagcccc acaccagccc ctctccaaat 5160

atgttggctg ttccttccat taaagtgacc ccactttaga gcagcaagtg gatttctgtt 5220
tcttacagtt caggaaggag gagtcagctg tgagaacctg gagcctgaga tgcttctaag 5280
tcccactgct actggggtca gggaagccag actccagcat cagcagtcag gagcactaag 5340
cccttgccaa catcctgttt ctcagagaaa ctgcttccat tataatggtt gtccttttt 5400

aagctatcaa gccaaacaac cagtgtctac cattattctc atcacctgaa gccaagggtt 5460
ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct ccagcttctg tcttcagtca 5520
ctccactctt agcctgctct gaatcaactc tgaccacagt tccctggagc ccctgccacc 5580
tgctgcccct gccaccttct ccatctgcag tgctgtgcag ccttctgcac tcttgcagag 5640
ctaataggtg gagacttgaa ggaagaggag gaaagtttct cataatagcc ttgctgcaag 5700
ctcaaatggg aggtgggcac tgtgcccagg agccttggag caaaggctgt gcccaacctc 5760
tgactgcatc caggtttggt cttgacagag ataagaagcc ctggcttttg gagccaaaat 5820

ctaggtcaga cttaggcagg attctcaaag tttatcagca gaacatgagg cagaagaccc 5880
tttctgctcc agcttcttca ggctcaacct tcatcagaat agatagaaag agaggctgtg 5940
agggttctta aaacagaagc aaatctgact cagagaataa acaacctcct agtaaactac 6000
agcttagaca gagcatctgg tggtgagtgt gctcagtgtc ctactcaact gtctggtatc 6060
agccctcatg aggacttctc ttctttccct catagacctc catctctgtt ttccttagcc 6120
tgcagaaatc tggatggcta ttcacagaat gcctgtgctt tcagagttgc atttttctc 6180
tggtattctg gttcaagcat ttgaaggtag gaaaggttct ccaagtgcaa gaaagccagc 6240

cctgagcctc aactgcctgg ctagtgtggt cagtaggatg caaaggctgt tgaatgccac 6300
aaggccaaac tttaacctgt gtaccacaag cctagcagca gaggcagctc tgctcactgg 6360
aactctctgt cttctttctc ctgagccttt tcttttcctg agttttctag ctctcctcaa 6420
ccttacctct gccctaccca ggacaaaccc aagagccact gtttctgtga tgtcctctcc 6480
agccctaatt aggcatcatg acttcagcct gaccttccat gctcagaagc agtgctaatc 6540
cacttcagat gagctgctct atgcaacaca ggcagagcct acaaaccttt gcaccagagc 6600
cctccacata tcagtgtttg ttcatactca cttcaacagc aaatgtgact gctgagatta 6660

agattttaca caagatggtc tgtaatttca cagttagttt tatcccatta ggtatgaaag 6720
aattagcata attcccctta aacatgaatg aatcttagat tttttaataa atagttttgg 6780
aagtaaagac agagacatca ggagcacaag gaatagcctg agaggacaaa cagaacaaga 6840
aagagtctgg aaatacacag gatgttcttg gcctcctcaa agcaagtgca agcagatagt 6900
accagcagcc ccaggctatc agagcccagt gaagagaagt accatgaaag ccacagctct 6960
aaccaccctg ttccagagtg acagacagtc cccaagacaa gccagcctga gccagagaga 7020

gaactgcaag agaaagtttc taatttaggt tctgttagat tcagacaagt gcaggtcatc 7080

ctctctccac agctactcac ctctccagcc taacaaagcc tgcagtccac actccaaccc 7140
tggtgtctca cctcctagcc tctcccaaca tcctgctctc tgaccatctt ctgcatctct 7200
catctcacca tctcccactg tctacagcct actcttgcaa ctaccatctc attttctgac 7260
atcctgtcta catcttctgc catactctgc catctaccat accacctctt accatctacc 7320
acaccatctt ttatctccat ccctctcaga agcctccaag ctgaatcctg ctttatgtgt 7380
tcatctcagc ccctgcatgg aaagctgacc ccagaggcag aactattccc agagagcttg 7440
gccaagaaaa acaaaactac cagcctggcc aggctcagga gtagtaagct gcagtgtctg 7500

ttgtgttcta gcttcaacag ctgcaggagt tccactctca aatgctccac atttctcaca 7560
tcctcctgat tctggtcact acccatcttc aaagaacaga atatctcaca tcagcatact 7620
gtgaaggact agtcatgggt gcagctgctc agagctgcaa agtcattctg gatggtggag 7680
agcttacaaa catttcatga tgctcccccc gctctgatgg ctggagccca atccctacac 7740
agactcctgc tgtatgtgtt ttcctttcac tctgagccac agccagaggg caggcattca 7800
gtctcctctt caggctgggg ctggggcact gagaactcac ccaacacctt gctctcactc 7860
cttctgcaaa acaagaaaga gctttgtgct gcagtagcca tgaagaatga aaggaaggct 7920

ttaactaaaa aatgtcagag attattttca acccccttact gtggatcacc agcaaggagg 7980
aaacacaaca cagagacatt ttttcccctc aaattatcaa aagaatcact gcatttgtta 8040
aagagagcaa ctgaatcagg aagcagagtt ttgaacatat cagaagttag gaatctgcat 8100
cagagacaaa tgcagtcatg gttgtttgct gcataccagc cctaatcatt agaagcctca 8160
tggacttcaa acatcattcc ctctgacaag atgctctagc ctaactccat gagataaaat 8220
aaatctgcct ttcagagcca aagaagagtc caccagcttc ttctcagtgt gaacaagagc 8280
tccagtcagg ttagtcagtc cagtgcagta gaggagacca gtctgcatcc tctaattttc 8340

aaaggcaaga agatttgttt accctggaca ccaggcacaa gtgaggtcac agagctctta 8400
gatatgcagt cctcatgagt gaggagacta aagcgcatgc catcaagact tcagtgtaga 8460
gaaaacctcc aaaaaagcct cctcactact tctggaatag ctcagaggcc gaggcggcct 8520
cggcctctgc ataaataaaa aaaattagtc agccatgggg cggagaatgg gcggaactgg 8580
gcggagttag gggcgggatg ggcggagtta ggggcgggac tatggttgct gactaattga 8640
gatgcatgct ttgcatactt ctgcctgctg gggagcctgg ggactttcca cacctggttg 8700
ctgactaatt gagatgcatg ctttgcatac ttctgcctgc tggggagcct ggggactttc 8760

cacaccctaa ctgacacaca ttccacagct gcattaatga atcggccaac gcgcggggag 8820
aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt 8880
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga 8940
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 9000
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc ccccctgacg agcatcacaa 9060
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 9120
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 9180

gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 9240
cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 9300
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 9360
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 9420
tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat 9480
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 9540
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 9600

aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 9660
aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct 9720
tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga 9780
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc 9840
catagttgcc tgactcctgc aaaccacgtt gtgtctcaaa atctctgatg ttacattgca 9900
caagataaaa atatatcatc atgaacaata aaactgtctg cttacataaa cagtaataca 9960
aggggtgtta tgagccatat tcaacgggaa acgtcttgct cgaggccgcg attaaattcc 10020

aacatggatg ctgatttata tgggtataaa tgggctcgcg ataatgtcgg gcaatcaggt 10080
gcgacaatct atcgattgta tgggaagccc gatgcgccag agttgtttct gaaacatggc 10140
aaaggtagcg ttgccaatga tgttacagat gagatggtca gactaaactg gctgacggaa 10200
tttatgcctc ttccgaccat caagcatttt atccgtactc ctgatgatgc atggttactc 10260
accactgcga tccccgggaa aacagcattc caggtattag aagaatatcc tgattcaggt 10320
gaaaatattg ttgatgcgct ggcagtgttc ctgcgccggt tgcattcgat tcctgtttgt 10380
aattgtcctt ttaacagcga tcgcgtattt cgtctcgctc aggcgcaatc acgaatgaat 10440

aacggtttgg ttgatgcgag tgattttgat gacgagcgta atggctggcc tgttgaacaa 10500
gtctggaaag aaatgcataa gcttttgcca ttctcaccgg attcagtcgt cactcatggt 10560
gatttctcac ttgataacct tatttttgac gaggggaaat taataggttg tattgatgtt 10620

ggacgagtcg gaatcgcaga ccgataccag gatcttgcca tcctatggaa ctgcctcggt 10680

gagttttctc cttcattaca gaaacggctt tttcaaaaat atggtattga taatcctgat 10740

atgaataaat tgcagtttca tttgatgctc gatgagtttt tctaagggcg gcctgccacc 10800

atacccacgc cgaaacaagc gctcatgagc ccgaagtggc gagcccgatc ttccccatcg 10860

gtgatgtcgg cgatataggc gccagcaacc gcacctgtgg cgccggtgat gagggcgcgc 10920

caagtcgacg tccggcagtc 10940

<210> 53

<211> 10934

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 53

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420

tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctcctg gtggcgaggg gagggggtg gtcctcgaac gccttgcaga 600

actggcctgg atacagagtg gaccggctgg ccccatctgg aagacttcga gatacactgt 660

tgtcttactg cgctcaacag tgtatctcga agtcttccaa atggtgccag ccatcgcagc 720

ggggtgcagg aaatgggggc agcccccctt tttggctatc cttccacgtg ttctttttg 780

tatcttttgt gtttcctaga aaacatctca gtcaccaccg cagccctagg aatgcatcta 840

gacaattgta ctaaccttct tctctttcct ctcctgacag tccggaaagc caccatggaa 900

ttcagcagcc ccagcagaga ggaatgcccc aagcctctga gccgggtgtc aatcatggcc 960

ggatctctga caggactgct gctgcttcag gccgtgtctt gggcttctgg cgctagacct 1020

tgcatcccca agagcttcgg ctacagcagc gtcgtgtgcg tgtgcaatgc cacctactgc 1080

gacagcttcg accctcctac ctttcctgct ctgggcacct tcagcagata cgagagcacc 1140

agatccggca gacggatgga actgagcatg ggacccatcc aggccaatca cacaggcact 1200
ggcctgctgc tgacactgca gcctgagcag aaattccaga aagtgaaagg cttcggcgga 1260
gccatgacag atgccgccgc tctgaatatc ctggctctgt ctccaccagc tcagaacctg 1320
ctgctcaaga gctacttcag cgaggaaggc atcggctaca acatcatcag agtgcccatg 1380
gccagctgcg acttcagcat caggacctac acctacgccg acacacccga cgatttccag 1440
ctgcacaact tcagcctgcc tgaagaggac accaagctga agatccctct gatccacaga 1500

gccctgcagc tggcacaaag acccgtgtca ctgctggcct ctccatggac atctcccacc 1560
tggctgaaaa caaatggcgc cgtgaatggc aagggcagcc tgaaaggcca acctggcgac 1620
atctaccacc agacctgggc cagatacttc gtgaagttcc tggacgccta tgccgagcac 1680
aagctgcagt tttgggccgt gacagccgag aacgaacctt ctgctggact gctgagcggc 1740
taccccttc agtgcctggg ctttacaccc gagcaccagc gggactttat cgcccgtgat 1800
ctgggaccca cactggccaa tagcacccac cataatgtgc ggctgctgat gctggacgac 1860
cagagactgc ttctgcccca ctgggctaaa gtggtgctga cagatcctga ggccgccaaa 1920

tacgtgcacg gaatcgccgt gcactggtat ctggactttc tggcccctgc caaggccaca 1980
ctgggagaga cacacagact gttccccaac accatgctgt tcgccagcga agcctgtgtg 2040
ggcagcaagt tttgggaaca gagcgtgcgg ctcggcagct gggatagagg catgcagtac 2100
agccacagca tcatcaccaa cctgctgtac cacgtcgtcg gctggaccga ctggaatctg 2160
gccctgaatc ctgaaggcgg ccctaactgg gtccgaaact tcgtggacag ccccatcatc 2220
gtggacatca ccaaggacac cttctacaag cagcccatgt tctaccacct gggacacttc 2280
agcaagttca tccccgaggg ctctcagcgc gttggactgg tggcttccca gaagaacgat 2340

ctggacgccg tggctctgat gcaccctgat ggatctgctg tggtggtggt cctgaaccgc 2400
agcagcaaag atgtgcccct gaccatcaag gatcccgccg tgggattcct ggaaacaatc 2460
agccctggct actccatcca cacctacctg tggcgtagac agtgattgtg gccgaaccgc 2520
cgaactcaga ggccggcccc agaaaacccg agcgagtagg gggcggcgcg caggagggag 2580
gagaactggg ggcgcgggag gctggtgggt gtgggggtg gagatgtaga agatgtgacg 2640
ccgcggcccg gcgggtgcca gattagcgga cgcggtgccc gcggttgcaa cgggatcccg 2700
ggcgctgcag cttgggaggc ggctctcccc aggcggcgtc cgcggagaca cccatccgtg 2760

aaccccaggt cccgggccgc cggctcgccg cgcaccaggg gccggcggac agaagagcgg 2820
ccgagcggct cgaggctggg ggaccgcggg cgcggccgcg cgctgccggg cgggaggctg 2880
gggggccggg gccggggccg tgccccggag cgggtcggag gccggggccg gggccggggg 2940
acggcggctc cccgcgcggc tccagcggct cggggatccc ggccgggccc cgcagggacc 3000

atgatggaga agggccccgt gcgcgccccc gccgagaagc cccgcggcgc ccgctgcagc 3060
aacggcttcc ccgagcgcga cccccccgc cccggcccca gccgccccgc cgagaagccc 3120
ccccgccccg aggccaagag cgcccagccc gccgacggct ggaagggcga gcgcccccgc 3180

agcgaggagg acaacgagct gaacctgccc aacctggccg ccgcctacag cagcatcctg 3240
agcagcctgg gcgagaaccc ccagcgccag ggcctgctga agaccccctg gcgcgccgcc 3300
agcgccatgc agttcttcac caagggctac caggagacca tcagcgacgt gctgaacgac 3360
gccatcttcg acgaggacca cgacgagatg gtgatcgtga aggacatcga catgttcagc 3420
atgtgcgagc accacctggt gcccttcgtg ggcaaggtgc acatcggcta cctgcccaac 3480
aagcaggtgc tgggcctgag caagctggcc cgcatcgtgg agatctacag ccgccgcctg 3540
caggtgcagg agcgcctgac caagcagatc gccgtggcca tcaccgaggc cctgcgcccc 3600

gccggcgtgg gcgtggtggt ggaggccacc cacatgtgca tggtgatgcg cggcgtgcag 3660
aagatgaaca gcaagaccgt gaccagcacc atgctgggcg tgttccgcga ggaccccaag 3720
acccgcgagg agttcctgac cctgatccgc agctgacaat tgttaattaa gtttaaaccc 3780
tcgaggccgc aagccgcatc gataccgtcg actagagctc gctgatcagc ctcgactgtg 3840
ccttctagtt gccagccatc tgttgtttgc ccctcccccg tgccttcctt gaccctggaa 3900
ggtgccactc ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt 3960
aggtgtcatt ctattctggg gggtggggtg gggcaggaca gcaaggggga ggattgggaa 4020

gacaatagca ggcatgctgg ggagagatcc acgataacaa acagcttttt tggggtgaac 4080
atattgactg aattccctgc aggttggcca ctccctctct gcgcgctcgc tcgctcactg 4140
aggccgcccg ggcaaagccc gggcgtcggg cgacctttgg tcgcccggcc tcagtgagcg 4200
agcgagcgcg cagagaggga gtggccaact ccatcactag gggttcctgc ggccgctcgt 4260
acggtctcga ggaattcctg caggataact tgccaacctc attctaaaat gtatatagaa 4320
gcccaaaaga caataacaaa aatattcttg tagaacaaaa tgggaaagaa tgttccacta 4380
aatatcaaga tttagagcaa agcatgagat gtgtggggat agacagtgag gctgataaaa 4440

tagagtagag ctcagaaaca gacccattga tatatgtaag tgacctatga aaaaaatatg 4500
gcattttaca atgggaaaat gatggtcttt ttctttttta gaaaaacagg gaaatatatt 4560
tatatgtaaa aaataaaagg gaacccatat gtcataccat acacacaaaa aaattccagt 4620
gaattataag tctaaatgga gaaggcaaaa ctttaaatct tttagaaaat aatatagaag 4680
catgcagacc agcctggcca acatgatgaa accctctcta ctaataataa aatcagtaga 4740
actactcagg actactttga gtgggaagtc cttttctatg aagacttctt tggccaaaat 4800

taggctctaa atgcaaggag atagtgcatc atgcctggct gcacttactg ataaatgatg 4860

ttatcaccat ctttaaccaa atgcacagga acaagttatg gtactgatgt gctggattga 4920
gaaggagctc tacttccttg acaggacaca tttgtatcaa cttaaaaaag cagatttttg 4980
ccagcagaac tattcattca gaggtaggaa acttagaata gatgatgtca ctgattagca 5040
tggcttcccc atctccacag ctgcttccca cccaggttgc ccacagttga gtttgtccag 5100
tgctcagggc tgcccactct cagtaagaag ccccacacca gcccctctcc aaatatgttg 5160
gctgttcctt ccattaaagt gaccccactt tagagcagca agtggatttc tgtttcttac 5220
agttcaggaa ggaggagtca gctgtgagaa cctggagcct gagatgcttc taagtcccac 5280

tgctactggg gtcagggaag ccagactcca gcatcagcag tcaggagcac taagcccttg 5340
ccaacatcct gtttctcaga gaaactgctt ccattataat ggttgtcctt ttttaagcta 5400
tcaagccaaa caaccagtgt ctaccattat tctcatcacc tgaagccaag ggttctagca 5460
aaagtcaagc tgtcttgtaa tggttgatgt gcctccagct tctgtcttca gtcactccac 5520
tcttagcctg ctctgaatca actctgacca cagttccctg gagcccctgc cacctgctgc 5580
ccctgccacc ttctccatct gcagtgctgt gcagccttct gcactcttgc agagctaata 5640
ggtggagact tgaaggaaga ggaggaaagt ttctcataat agccttgctg caagctcaaa 5700

tgggaggtgg gcactgtgcc caggagcctt ggagcaaagg ctgtgcccaa cctctgactg 5760
catccaggtt tggtcttgac agagataaga agccctggct tttggagcca aaatctaggt 5820
cagacttagg caggattctc aaagtttatc agcagaacat gaggcagaag accctttctg 5880
ctccagcttc ttcaggctca accttcatca gaatagatag aaagagaggc tgtgagggtt 5940
cttaaaacag aagcaaatct gactcagaga ataaacaacc tcctagtaaa ctacagctta 6000
gacagagcat ctggtggtga gtgtgctcag tgtcctactc aactgtctgg tatcagccct 6060
catgaggact tctcttcttt ccctcataga cctccatctc tgttttcctt agcctgcaga 6120

aatctggatg gctattcaca gaatgcctgt gctttcagag ttgcattttt tctctggtat 6180
tctggttcaa gcatttgaag gtaggaaagg ttctccaagt gcaagaaagc cagccctgag 6240
cctcaactgc ctggctagtg tggtcagtag gatgcaaagg ctgttgaatg ccacaaggcc 6300
aaactttaac ctgtgtacca caagcctagc agcagaggca gctctgctca ctggaactct 6360
ctgtcttctt tctcctgagc cttttctttt cctgagtttt ctagctctcc tcaaccttac 6420
ctctgcccta cccaggacaa acccaagagc cactgtttct gtgatgtcct ctccagccct 6480
aattaggcat catgacttca gcctgacctt ccatgctcag aagcagtgct aatccacttc 6540

agatgagctg ctctatgcaa cacaggcaga gcctacaaac ctttgcacca gagccctcca 6600
catatcagtg tttgttcata ctcacttcaa cagcaaatgt gactgctgag attaagattt 6660
tacacaagat ggtctgtaat ttcacagtta gttttatccc attaggtatg aaagaattag 6720
cataattccc cttaaacatg aatgaatctt agatttttta ataaatagtt ttggaagtaa 6780
agacagagac atcaggagca caaggaatag cctgagagga caaacagaac aagaaagagt 6840
ctggaaatac acaggatgtt cttggcctcc tcaaagcaag tgcaagcaga tagtaccagc 6900
agccccaggc tatcagagcc cagtgaagag aagtaccatg aaagccacag ctctaaccac 6960

cctgttccag agtgacagac agtccccaag acaagccagc ctgagccaga gagagaactg 7020
caagagaaag tttctaattt aggttctgtt agattcagac aagtgcaggt catcctctct 7080
ccacagctac tcacctctcc agcctaacaa agcctgcagt ccacactcca accctggtgt 7140
ctcacctcct agcctctccc aacatcctgc tctctgacca tcttctgcat ctctcatctc 7200
accatctccc actgtctaca gcctactctt gcaactacca tctcattttc tgacatcctg 7260
tctacatctt ctgccatact ctgccatcta ccataccacc tcttaccatc taccacacca 7320
tcttttatct ccatccctct cagaagcctc caagctgaat cctgctttat gtgttcatct 7380

cagcccctgc atggaaagct gaccccagag gcagaactat tcccagagag cttggccaag 7440
aaaaacaaaa ctaccagcct ggccaggctc aggagtagta agctgcagtg tctgttgtgt 7500
tctagcttca acagctgcag gagttccact ctcaaatgct ccacatttct cacatcctcc 7560
tgattctggt cactacccat cttcaaagaa cagaatatct cacatcagca tactgtgaag 7620
gactagtcat gggtgcagct gctcagagct gcaaagtcat tctggatggt ggagagctta 7680
caaacatttc atgatgctcc ccccgctctg atggctggag cccaatccct acacagactc 7740
ctgctgtatg tgttttcctt tcactctgag ccacagccag agggcaggca ttcagtctcc 7800

tcttcaggct ggggctgggg cactgagaac tcacccaaca ccttgctctc actccttctg 7860
caaaacaaga aagagctttg tgctgcagta gccatgaaga atgaaaggaa ggctttaact 7920
aaaaaatgtc agagattatt ttcaacccct tactgtggat caccagcaag gaggaaacac 7980
aacacagaga catttttcc cctcaaatta tcaaaagaat cactgcattt gttaaagaga 8040
gcaactgaat caggaagcag agttttgaac atatcagaag ttaggaatct gcatcagaga 8100
caaatgcagt catggttgtt tgctgcatac cagccctaat cattagaagc ctcatggact 8160
tcaaacatca ttccctctga caagatgctc tagcctaact ccatgagata aaataaatct 8220

gcctttcaga gccaaagaag agtccaccag cttcttctca gtgtgaacaa gagctccagt 8280
caggttagtc agtccagtgc agtagaggag accagtctgc atcctctaat tttcaaaggc 8340
aagaagattt gtttaccctg gacaccaggc acaagtgagg tcacagagct cttagatatg 8400

cagtcctcat gagtgaggag actaaagcgc atgccatcaa gacttcagtg tagagaaaac 8460
ctccaaaaaa gcctcctcac tacttctgga atagctcaga ggccgaggcg gcctcggcct 8520
ctgcataaat aaaaaaaatt agtcagccat ggggcggaga atgggcggaa ctgggcggag 8580
ttaggggcgg gatgggcgga gttaggggcg ggactatggt tgctgactaa ttgagatgca 8640

tgctttgcat acttctgcct gctggggagc ctggggactt tccacacctg gttgctgact 8700
aattgagatg catgctttgc atacttctgc ctgctgggga gcctggggac tttccacacc 8760
ctaactgaca cacattccac agctgcatta atgaatcggc caacgcgcgg ggagaggcgg 8820
tttgcgtatt gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg 8880
gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg 8940
ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa 9000
ggccgcgttg ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg 9060

acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc 9120
tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc 9180
ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc 9240
ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg 9300
ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc 9360
actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga 9420
gttcttgaag tggtggccta actacggcta cactagaaga acagtatttg gtatctgcgc 9480

tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac 9540
caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg 9600
atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc 9660
acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa 9720
ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta 9780
ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt 9840
tgcctgactc ctgcaaacca cgttgtgtct caaaatctct gatgttacat tgcacaagat 9900

aaaaatatat catcatgaac aataaaactg tctgcttaca taaacagtaa tacaaggggt 9960
gttatgagcc atattcaacg ggaaacgtct tgctcgaggc cgcgattaaa ttccaacatg 10020
gatgctgatt tatatgggta taaatgggct cgcgataatg tcgggcaatc aggtgcgaca 10080
atctatcgat tgtatgggaa gcccgatgcg ccagagttgt ttctgaaaca tggcaaaggt 10140
agcgttgcca atgatgttac agatgagatg gtcagactaa actggctgac ggaatttatg 10200
cctcttccga ccatcaagca ttttatccgt actcctgatg atgcatggtt actcaccact 10260

gcgatccccg ggaaaacagc attccaggta ttagaagaat atcctgattc aggtgaaaat 10320

attgttgatg cgctggcagt gttcctgcgc cggttgcatt cgattcctgt ttgtaattgt 10380
ccttttaaca gcgatcgcgt atttcgtctc gctcaggcgc aatcacgaat gaataacggt 10440
ttggttgatg cgagtgattt tgatgacgag cgtaatggct ggcctgttga acaagtctgg 10500
aaagaaatgc ataagctttt gccattctca ccggattcag tcgtcactca tggtgatttc 10560
tcacttgata accttatttt tgacgagggg aaattaatag gttgtattga tgttggacga 10620
gtcggaatcg cagaccgata ccaggatctt gccatcctat ggaactgcct cggtgagttt 10680
tctccttcat tacagaaacg gctttttcaa aaatatggta ttgataatcc tgatatgaat 10740

aaattgcagt ttcatttgat gctcgatgag tttttctaag ggcggcctgc caccataccc 10800
acgccgaaac aagcgctcat gagcccgaag tggcgagccc gatcttcccc atcggtgatg 10860
tcggcgatat aggcgccagc aaccgcacct gtggcgccgg tgatgagggc gcgccaagtc 10920
gacgtccggc agtc 10934

<210> 54
<211> 11138
<212> DNA
<213> Artificial sequence
<220><223> Synthetic polynucleotide
<400> 54

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240
tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agtaagtcac 300
tgactgtcta tgcctgggaa agggtgggca ggagatgggg cagtgcagga aaagtggcac 360
tatgaaccct cctggtggcg aggggagggg ggtggtcctc gaacgccttg cagaactggc 420
ctggatacag agtggaccgg ctggccccat ctggaagact tcgagataca ctgttgtctt 480
actgcgctca acagtgtatc tcgaagtctt ccaaatggtg ccagccatcg cagcggggtg 540

caggaaatgg gggcagcccc ccttttggc tatccttcca cgtgttcttt tttgtatctt 600
ttgtgtttcc tagaaaacat ctcagtcacc accgtgatat cacaaggtcc cagggctggg 660
gtcagaaatt ctctcccgag ggaatgaagc cacaggagcc aagagcagga ggaccaaggc 720
cctggcgaag gccgtggcct cgttcaagta aaagatccta gtacagtgca ggtcccaatg 780

tgtactagga tcttttactt gaacggggac gccggcatcc gggctcagga ccccccctctc 840

tgccagaggc accaacacca gagttcacaa atcagtctcc tgccctttgc atgtagcaaa 900

gcagccctag gaatgcatct agacaattgt actaaccttc ttctctttcc tctcctgaca 960

gtccggaaag ccaccatgcc caccacccag cagagccccc aggacgagca ggagaagctg 1020

ctggacgagg ccatccaggc cgtgaaggtg cagagcttcc agatgaagcg ctgcctggac 1080

aagaacaagc tgatggacgc cctgaagcac gccagcaaca tgctgggcga gctgcgcacc 1140

agcatgctga gccccaagag ctactacgag ctgtacatgg ccatcagcga cgagctgcac 1200

tacctggagg tgtacctgac cgacgagttc gccaagggcc gcaaggtggc cgacctgtac 1260

gagctggtgc agtacgccgg caacatcatc ccccgcctgt acctgctgat caccgtgggc 1320

gtggtgtacg tgaagagctt cccccagagc cgcaaggaca tcctgaagga cctggtggag 1380

atgtgccgcg gcgtgcagca ccccctgcgc ggcctgttcc tgcgcaacta cctgctgcag 1440

tgcacccgca acatcctgcc cgacgagggc gagcccaccg acgaggagac caccggcgac 1500

atcagcgaca gcatggactt cgtgctgctg aacttcgccg agatgaacaa gctgtgggtg 1560

cgcatgcagc accagggcca cagccgcgac cgcgagaagc gcgagcgcga gcgccaggag 1620

ctgcgcatcc tggtgggcac caacctggtg cgcctgagcc agctggaggg cgtgaacgtg 1680

gagcgctaca agcagatcgt gctgaccggc atcctggagc aggtggtgaa ctgccgcgac 1740

gccctggccc aggagtacct gatggagtgc atcatccagg tgttccccga cgagttccac 1800

ctgcagaccc tgaacccctt cctgcgcgcc tgcgccgagc tgcaccagaa cgtgaacgtg 1860

aagaacatca tcatcgccct gatcgaccgc ctggccctgt tcgcccaccg cgaggacggc 1920

cccggcatcc ccgccgacat caagctgttc gacatcttca gccagcaggt ggccaccgtg 1980

atccagagcc gccaggacat gcccagcgag gacgtggtga gcctgcaggt gagcctgatc 2040

aacctggcca tgaagtgcta ccccgaccgc gtggactacg tggacaaggt gctggagacc 2100

accgtggaga tcttcaacaa gctgaacctg gagcacatcg ccaccagcag cgccgtgagc 2160

aaggagctga cccgcctgct gaagatcccc gtggacacct acaacaacat cctgaccgtg 2220

ctgaagctga agcacttcca ccccctgttc gagtacttcg actacgagag ccgcaagagc 2280

atgagctgct acgtgctgag caacgtgctg gactacaaca ccgagatcgt gagccaggac 2340

caggtggaca gcatcatgaa cctggtgagc accctgatcc aggaccagcc cgaccagccc 2400

gtggaggacc ccgaccccga ggacttcgcc gacgagcaga gcctggtggg ccgcttcatc 2460

cacctgctgc gcagcgagga ccccgaccag cagtacctga tcctgaacac cgcccgcaag 2520

cacttcggcg ccggcggcaa ccagcgcatc cgcttcaccc tgcccccccct ggtgttcgcc 2580

gcctaccagc tggccttccg ctacaaggag aacagcaagg tggacgacaa gtgggagaag 2640

aagtgccaga agatcttcag cttcgcccac cagaccatca gcgccctgat caaggccgag 2700
ctggccgagc tgcccctgcg cctgttcctg cagggcgccc tggccgccgg cgagatcggc 2760
ttcgagaacc acgagaccgt ggcctacgag ttcatgagcc aggccttcag cctgtacgag 2820
gacgagatca gcgacagcaa ggcccagctg gccgccatca ccctgatcat cggcaccttc 2880
gagcgcatga agtgcttcag cgaggagaac cacgagcccc tgcgcaccca gtgcgccctg 2940
gccgccagca agctgctgaa gaagcccgac cagggccgcg ccgtgagcac ctgcgcccac 3000
ctgttctgga gcggccgcaa caccgacaag aacggcgagg agctgcacgg cggcaagcgc 3060

gtgatggagt gcctgaagaa ggccctgaag atcgccaacc agtgcatgga ccccagcctg 3120
caggtgcagc tgttcatcga gatcctgaac cgctacatct acttctacga gaaggagaac 3180
gacgccgtga ccatccaggt gctgaaccag ctgatccaga agatccgcga ggacctgccc 3240
aacctggaga gcagcgagga gaccgagcag atcaacaagc acttccacaa caccctggag 3300
cacctgcgcc tgcgccgcga gagccccgag agcgagggcc ccatctacga gggcctgatc 3360
ctgtgacaat tgttaattaa gtttaaaccc tcgaggccgc aagcttatcg ataatcaacc 3420
tctggattac aaaatttgtg aaagattgac tggtattctt aactatgttg ctccttttac 3480

gctatgtgga tacgctgctt taatgccttt gtatcatgct attgcttccc gtatggcttt 3540
cattttctcc tccttgtata aatcctggtt gctgtctctt tatgaggagt tgtggcccgt 3600
tgtcaggcaa cgtggcgtgg tgtgcactgt gtttgctgac gcaacccccca ctggttgggg 3660
cattgccacc acctgtcagc tcctttccgg gactttcgct ttccccctcc ctattgccac 3720
ggcggaactc atcgccgcct gccttgcccg ctgctggaca ggggctcggc tgttgggcac 3780
tgacaattcc gtggtgttgt cggggaaatc atcgtccttt ccttggctgc tcgcctgtgt 3840
tgccacctgg attctgcgcg ggacgtcctt ctgctacgtc ccttcggccc tcaatccagc 3900

ggaccttcct tcccgcggcc tgctgccggc tctgcggcct cttccgcgtc ttcgccttcg 3960
ccctcagacg agtcggatct ccctttgggc cgcctccccg catcgatacc gtcgactaga 4020
gctcgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt ttgcccctcc 4080
cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta ataaaatgag 4140
gaaattgcat cgcattgtct gagtaggtgt cattctattc tggggggtgg ggtggggcag 4200
gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggagag atccacgata 4260
acaaacagct tttttggggt gaacatattg actgaattcc ctgcaggttg gccactccct 4320

ctctgcgcgc tcgctcgctc actgaggccg cccgggcaaa gcccgggcgt cgggcgacct 4380
ttggtcgccc ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aactccatca 4440
ctaggggttc ctgcggccgc tcgtacggtc tcgaggaatt cctgcaggat aacttgccaa 4500
cctcattcta aaatgtatat agaagcccaa aagacaataa caaaaatatt cttgtagaac 4560
aaaatgggaa agaatgttcc actaaatatc aagatttaga gcaaagcatg agatgtgtgg 4620
ggatagacag tgaggctgat aaaatagagt agagctcaga aacagaccca ttgatatatg 4680
taagtgacct atgaaaaaaa tatggcattt tacaatggga aaatgatggt ctttttcttt 4740

tttagaaaaa cagggaaata tatttatatg taaaaaataa aagggaaccc atatgtcata 4800
ccatacacac aaaaaaattc cagtgaatta taagtctaaa tggagaaggc aaaactttaa 4860
atcttttaga aaataatata gaagcatgca gaccagcctg gccaacatga tgaaaccctc 4920
tctactaata ataaaatcag tagaactact caggactact ttgagtggga agtccttttc 4980
tatgaagact tctttggcca aaattaggct ctaaatgcaa ggagatagtg catcatgcct 5040
ggctgcactt actgataaat gatgttatca ccatctttaa ccaaatgcac aggaacaagt 5100
tatggtactg atgtgctgga ttgagaagga gctctacttc cttgacagga cacatttgta 5160

tcaacttaaa aaagcagatt tttgccagca gaactattca ttcagaggta ggaaacttag 5220
aatagatgat gtcactgatt agcatggctt ccccatctcc acagctgctt cccacccagg 5280
ttgcccacag ttgagtttgt ccagtgctca gggctgccca ctctcagtaa gaagccccac 5340
accagcccct ctccaaatat gttggctgtt ccttccatta aagtgacccc actttagagc 5400
agcaagtgga tttctgtttc ttacagttca ggaaggagga gtcagctgtg agaacctgga 5460
gcctgagatg cttctaagtc ccactgctac tggggtcagg gaagccagac tccagcatca 5520
gcagtcagga gcactaagcc cttgccaaca tcctgtttct cagagaaact gcttccatta 5580

taatggttgt cctttttaa gctatcaagc caaacaacca gtgtctacca ttattctcat 5640
cacctgaagc caagggttct agcaaaagtc aagctgtctt gtaatggttg atgtgcctcc 5700
agcttctgtc ttcagtcact ccactcttag cctgctctga atcaactctg accacagttc 5760
cctggagccc ctgccacctg ctgcccctgc caccttctcc atctgcagtg ctgtgcagcc 5820
ttctgcactc ttgcagagct aataggtgga gacttgaagg aagaggagga aagtttctca 5880
taatagcctt gctgcaagct caaatgggag gtgggcactg tgcccaggag ccttggagca 5940
aaggctgtgc ccaacctctg actgcatcca ggtttggtct tgacagagat aagaagccct 6000

ggcttttgga gccaaaatct aggtcagact taggcaggat tctcaaagtt tatcagcaga 6060
acatgaggca gaagaccctt tctgctccag cttcttcagg ctcaaccttc atcagaatag 6120
atagaaagag aggctgtgag ggttcttaaa acagaagcaa atctgactca gagaataaac 6180

aacctcctag taaactacag cttagacaga gcatctggtg gtgagtgtgc tcagtgtcct 6240
actcaactgt ctggtatcag ccctcatgag gacttctctt ctttccctca tagacctcca 6300
tctctgtttt ccttagcctg cagaaatctg gatggctatt cacagaatgc ctgtgctttc 6360
agagttgcat tttttctctg gtattctggt tcaagcattt gaaggtagga aaggttctcc 6420

aagtgcaaga aagccagccc tgagcctcaa ctgcctggct agtgtggtca gtaggatgca 6480
aaggctgttg aatgccacaa ggccaaactt taacctgtgt accacaagcc tagcagcaga 6540
ggcagctctg ctcactggaa ctctctgtct tctttctcct gagccttttc ttttcctgag 6600
ttttctagct ctcctcaacc ttacctctgc cctacccagg acaaacccaa gagccactgt 6660
ttctgtgatg tcctctccag ccctaattag gcatcatgac ttcagcctga ccttccatgc 6720
tcagaagcag tgctaatcca cttcagatga gctgctctat gcaacacagg cagagcctac 6780
aaacctttgc accagagccc tccacatatc agtgtttgtt catactcact tcaacagcaa 6840

atgtgactgc tgagattaag attttacaca agatggtctg taatttcaca gttagtttta 6900
tcccattagg tatgaaagaa ttagcataat tccccttaaa catgaatgaa tcttagattt 6960
tttaataaat agttttggaa gtaaagacag agacatcagg agcacaagga atagcctgag 7020
aggacaaaca gaacaagaaa gagtctggaa atacacagga tgttcttggc ctcctcaaag 7080
caagtgcaag cagatagtac cagcagcccc aggctatcag agcccagtga agagaagtac 7140
catgaaagcc acagctctaa ccaccctgtt ccagagtgac agacagtccc caagacaagc 7200
cagcctgagc cagagagaga actgcaagag aaagtttcta atttaggttc tgttagattc 7260

agacaagtgc aggtcatcct ctctccacag ctactcacct ctccagccta acaaagcctg 7320
cagtccacac tccaaccctg gtgtctcacc tcctagcctc tcccaacatc ctgctctctg 7380
accatcttct gcatctctca tctcaccatc tcccactgtc tacagcctac tcttgcaact 7440
accatctcat tttctgacat cctgtctaca tcttctgcca tactctgcca tctaccatac 7500
cacctcttac catctaccac accatctttt atctccatcc ctctcagaag cctccaagct 7560
gaatcctgct ttatgtgttc atctcagccc ctgcatggaa agctgacccc agaggcagaa 7620
ctattcccag agagcttggc caagaaaaac aaaactacca gcctggccag gctcaggagt 7680

agtaagctgc agtgtctgtt gtgttctagc ttcaacagct gcaggagttc cactctcaaa 7740
tgctccacat ttctcacatc ctcctgattc tggtcactac ccatcttcaa agaacagaat 7800
atctcacatc agcatactgt gaaggactag tcatgggtgc agctgctcag agctgcaaag 7860
tcattctgga tggtggagag cttacaaaca tttcatgatg ctccccccgc tctgatggct 7920
ggagcccaat ccctacacag actcctgctg tatgtgtttt cctttcactc tgagccacag 7980
ccagagggca ggcattcagt ctcctcttca ggctggggct ggggcactga gaactcaccc 8040

aacaccttgc tctcactcct tctgcaaaac aagaaagagc tttgtgctgc agtagccatg 8100

aagaatgaaa ggaaggcttt aactaaaaaa tgtcagagat tattttcaac cccttactgt 8160
ggatcaccag caaggaggaa acacaacaca gagacatttt ttcccctcaa attatcaaaa 8220
gaatcactgc atttgttaaa gagagcaact gaatcaggaa gcagagtttt gaacatatca 8280
gaagttagga atctgcatca gagacaaatg cagtcatggt tgtttgctgc ataccagccc 8340
taatcattag aagcctcatg gacttcaaac atcattccct ctgacaagat gctctagcct 8400
aactccatga gataaaataa atctgccttt cagagccaaa gaagagtcca ccagcttctt 8460
ctcagtgtga acaagagctc cagtcaggtt agtcagtcca gtgcagtaga ggagaccagt 8520

ctgcatcctc taattttcaa aggcaagaag atttgtttac cctggacacc aggcacaagt 8580
gaggtcacag agctcttaga tatgcagtcc tcatgagtga ggagactaaa gcgcatgcca 8640
tcaagacttc agtgtagaga aaacctccaa aaaagcctcc tcactacttc tggaatagct 8700
cagaggccga ggcggcctcg gcctctgcat aaataaaaaa aattagtcag ccatggggcg 8760
gagaatgggc ggaactgggc ggagttaggg gcgggatggg cggagttagg ggcgggacta 8820
tggttgctga ctaattgaga tgcatgcttt gcatacttct gcctgctggg gagcctgggg 8880
actttccaca cctggttgct gactaattga gatgcatgct ttgcatactt ctgcctgctg 8940

gggagcctgg ggactttcca caccctaact gacacacatt ccacagctgc attaatgaat 9000
cggccaacgc gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac 9060
tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt 9120
aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca 9180
gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc 9240
ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact 9300
ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct 9360

gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag 9420
ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca 9480
cgaacccccc gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa 9540
cccggtaaga cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc 9600
gaggtatgta ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag 9660
aagaacagta tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg 9720
tagctcttga tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca 9780

gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc 9840

tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag 9900

gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata 9960

tgagtaaact tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat 10020

ctgtctattt cgttcatcca tagttgcctg actcctgcaa accacgttgt gtctcaaaat 10080

ctctgatgtt acattgcaca agataaaaat atatcatcat gaacaataaa actgtctgct 10140

tacataaaca gtaatacaag gggtgttatg agccatattc aacgggaaac gtcttgctcg 10200

aggccgcgat taaattccaa catggatgct gatttatatg ggtataaatg ggctcgcgat 10260

aatgtcgggc aatcaggtgc gacaatctat cgattgtatg ggaagcccga tgcgccagag 10320

ttgtttctga aacatggcaa aggtagcgtt gccaatgatg ttacagatga gatggtcaga 10380

ctaaactggc tgacggaatt tatgcctctt ccgaccatca agcattttat ccgtactcct 10440

gatgatgcat ggttactcac cactgcgatc cccgggaaaa cagcattcca ggtattagaa 10500

gaatatcctg attcaggtga aaatattgtt gatgcgctgg cagtgttcct gcgccggttg 10560

cattcgattc ctgtttgtaa ttgtcctttt aacagcgatc gcgtatttcg tctcgctcag 10620

gcgcaatcac gaatgaataa cggtttggtt gatgcgagtg attttgatga cgagcgtaat 10680

ggctggcctg ttgaacaagt ctggaaagaa atgcataagc ttttgccatt ctcaccggat 10740

tcagtcgtca ctcatggtga tttctcactt gataacctta tttttgacga ggggaaatta 10800

ataggttgta ttgatgttgg acgagtcgga atcgcagacc gataccagga tcttgccatc 10860

ctatggaact gcctcggtga gttttctcct tcattacaga aacggctttt tcaaaaatat 10920

ggtattgata atcctgatat gaataaattg cagtttcatt tgatgctcga tgagtttttc 10980

taagggcggc ctgccaccat acccacgccg aaacaagcgc tcatgagccc gaagtggcga 11040

gcccgatctt ccccatcggt gatgtcggcg atataggcgc cagcaaccgc acctgtggcg 11100

ccggtgatga gggcgcgcca agtcgacgtc cggcagtc 11138

<210> 55

<211> 242

<212> PRT

<213> Artificial sequence

<220><223> Synthetic polypeptide

<400> 55

Met Pro Arg Gly Phe Thr Trp Leu Arg Tyr Leu Gly Ile Phe Leu Gly

1 5 10 15

Val Ala Leu Gly Asn Glu Pro Leu Glu Met Trp Pro Leu Thr Gln Asn

20 25 30

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

Val Ser Leu Ser Ala Thr Glu Ser Val Gln Asp Val Leu Leu Glu Gly
100 105 110
His Pro Ser Trp Lys Tyr Leu Gln Glu Val Glu Thr Leu Leu Leu Asn
115 120 125
Val Gln Gln Gly Leu Thr Asp Val Glu Val Ser Pro Lys Val Glu Ser
130 135 140
Val Leu Ser Leu Leu Asn Ala Pro Gly Pro Asn Leu Lys Leu Val Arg
145 150 155 160

Pro Lys Ala Leu Leu Asp Asn Cys Phe Arg Val Met Glu Leu Leu Tyr
165 170 175
Cys Ser Cys Cys Lys Gln Ser Ser Val Leu Asn Trp Gln Asp Cys Glu
180 185 190
Val Pro Ser Pro Gln Ser Cys Ser Pro Glu Pro Ser Leu Gln Tyr Ala
195 200 205
Ala Thr Gln Leu Tyr Pro Pro Pro Pro Trp Ser Pro Ser Ser Pro Pro
210 215 220

His Ser Thr Gly Ser Val Arg Pro Val Arg Ala Gln Gly Glu Gly Leu
225 230 235 240
Leu Pro

<210> 56
<211> 729

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 56

atgccccgcg gcttcacctg gctgcgctac ctgggcatct tcctgggcgt ggccctgggc 60

aacgagcccc tggagatgtg gcccctgacc cagaacgagg agtgcaccgt gaccggcttc 120

ctgcgcgaca agctgcagta ccgcagccgc ctgcagtaca tgaagcacta cttccccatc 180

aactacaaga tcagcgtgcc ctacgagggc gtgttccgca tcgccaacgt gacccgcctg 240

cagcgcgccc aggtgagcga gcgcgagctg cgctacctgt gggtgctggt gagcctgagc 300

gccaccgaga gcgtgcagga cgtgctgctg gagggccacc ccagctggaa gtacctgcag 360

gaggtggaga ccctgctgct gaacgtgcag cagggcctga ccgacgtgga ggtgagcccc 420

aaggtggaga gcgtgctgag cctgctgaac gccccggcc ccaacctgaa gctggtgcgc 480

cccaaggccc tgctggacaa ctgcttccgc gtgatggagc tgctgtactg cagctgctgc 540

aagcagagca gcgtgctgaa ctggcaggac tgcgaggtgc ccagccccca gagctgcagc 600

cccgagccca gcctgcagta cgccgccacc cagctgtacc cccccccccc ctggagcccc 660

agcagccccc cccacagcac cggcagcgtg cgccccgtgc gcgcccaggg cgagggcctg 720

ctgccctaa 729

<210> 57

<211> 230

<212> PRT

<213> Artificial sequence

<220><223> Synthetic polypeptide

<400> 57

Met Glu Pro Leu Arg Leu Leu Ile Leu Leu Phe Val Thr Glu Leu Ser

1 5 10 15

Gly Ala His Asn Thr Thr Val Phe Gln Gly Val Ala Gly Gln Ser Leu

20 25 30

Gln Val Ser Cys Pro Tyr Asp Ser Met Lys His Trp Gly Arg Arg Lys

35 40 45

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

50 55 60

Ser Thr His Asn Leu Trp Leu Leu Ser Phe Leu Arg Arg Trp Asn Gly

65 70 75 80

Ser Thr Ala Ile Thr Asp Asp Thr Leu Gly Gly Thr Leu Thr Ile Thr

85 90 95

Leu Arg Asn Leu Gln Pro His Asp Ala Gly Leu Tyr Gln Cys Gln Ser

100 105 110

Leu His Gly Ser Glu Ala Asp Thr Leu Arg Lys Val Leu Val Glu Val

115 120 125

Leu Ala Asp Pro Leu Asp His Arg Asp Ala Gly Asp Leu Trp Phe Pro

130 135 140

Gly Glu Ser Glu Ser Phe Glu Asp Ala His Val Glu His Ser Ile Ser

145 150 155 160

Arg Ser Leu Leu Glu Gly Glu Ile Pro Phe Pro Pro Thr Ser Ile Leu

165 170 175

Leu Leu Leu Ala Cys Ile Phe Leu Ile Lys Ile Leu Ala Ala Ser Ala

180 185 190

Leu Trp Ala Ala Ala Trp His Gly Gln Lys Pro Gly Thr His Pro Pro

195 200 205

Ser Glu Leu Asp Cys Gly His Asp Pro Gly Tyr Gln Leu Gln Thr Leu

210 215 220

Pro Gly Leu Arg Asp Thr

225 230

<210> 58

<211> 690

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 58

atggagcccc tgcgcctgct gatcctgctg ttcgtgaccg agctgagcgg cgcccacaac 60

accaccgtgt tccagggcgt ggccggccag agcctgcagg tgagctgccc ctacgacagc 120

atgaagcact ggggccgccg caaggcctgg tgccgccagc tgggcgagaa gggcccctgc 180

cagcgcgtgg tgagcaccca caacctgtgg ctgctgagct tcctgcgccg ctggaacggc 240

agcaccgcca tcaccgacga caccctgggc ggcaccctga ccatcaccct gcgcaacctg 300

cagccccacg acgccggcct gtaccagtgc cagagcctgc acggcagcga ggccgacacc 360

ctgcgcaagg tgctggtgga ggtgctggcc gaccccctgg accaccgcga cgccggcgac 420

ctgtggttcc ccggcgagag cgagagcttc gaggacgccc acgtggagca cagcatcagc 480

cgcagcctgc tggagggcga gatccccttc ccccccacca gcatcctgct gctgctggcc 540

tgcatcttcc tgatcaagat cctggccgcc agcgccctgt gggccgccgc ctggcacggc 600

cagaagcccg gcacccaccc ccccagcgag ctggactgcg gccacgaccc cggctaccag 660

ctgcagaccc tgcccggcct gcgcgacacc 690

<210> 59

<211> 11060

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 59

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420

tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctcctg gtggcgaggg gagggggggtg gtcctcgaac gccttgcaga 600

actggcctgg atacagagtg gaccggctgg ccccatctgg aagacttcga gatacactgt 660

tgtcttactg cgctcaacag tgtatctcga agtcttccaa atggtgccag ccatcgcagc 720

ggggtgcagg aaatgggggc agcccccctt tttggctatc cttccacgtg ttctttttg 780

tatcttttgt gtttcctaga aaacatctca gtcaccaccg cagccctagg aatgcatcta 840

gacaattgta ctaaccttct tctctttcct ctcctgacag tccggaaagc caccatggaa 900

ttcagcagcc ccagcagaga ggaatgcccc aagcctctga gccgggtgtc aatcatggcc 960

ggatctctga caggactgct gctgcttcag gccgtgtctt gggcttctgg cgctagacct 1020

tgcatcccca agagcttcgg ctacagcagc gtcgtgtgcg tgtgcaatgc cacctactgc 1080
gacagcttcg accctcctac ctttcctgct ctgggcacct tcagcagata cgagagcacc 1140
agatccggca gacggatgga actgagcatg ggacccatcc aggccaatca cacaggcact 1200

ggcctgctgc tgacactgca gcctgagcag aaattccaga aagtgaaagg cttcggcgga 1260
gccatgacag atgccgccgc tctgaatatc ctggctctgt ctccaccagc tcagaacctg 1320
ctgctcaaga gctacttcag cgaggaaggc atcggctaca acatcatcag agtgcccatg 1380
gccagctgcg acttcagcat caggacctac acctacgccg acacacccga cgatttccag 1440
ctgcacaact tcagcctgcc tgaagaggac accaagctga agatccctct gatccacaga 1500
gccctgcagc tggcacaaag acccgtgtca ctgctggcct ctccatggac atctcccacc 1560
tggctgaaaa caaatggcgc cgtgaatggc aagggcagcc tgaaaggcca acctggcgac 1620

atctaccacc agacctgggc cagatacttc gtgaagttcc tggacgccta tgccgagcac 1680
aagctgcagt tttgggccgt gacagccgag aacgaacctt ctgctggact gctgagcggc 1740
taccccttc agtgcctggg ctttacaccc gagcaccagc gggactttat cgcccgtgat 1800
ctgggaccca cactggccaa tagcacccac cataatgtgc ggctgctgat gctggacgac 1860
cagagactgc ttctgcccca ctgggctaaa gtggtgctga cagatcctga ggccgccaaa 1920
tacgtgcacg gaatcgccgt gcactggtat ctggactttc tggcccctgc caaggccaca 1980
ctgggagaga cacacagact gttccccaac accatgctgt tcgccagcga agcctgtgtg 2040

ggcagcaagt tttgggaaca gagcgtgcgg ctcggcagct gggatagagg catgcagtac 2100
agccacagca tcatcaccaa cctgctgtac cacgtcgtcg gctggaccga ctggaatctg 2160
gccctgaatc ctgaaggcgg ccctaactgg gtccgaaact tcgtggacag ccccatcatc 2220
gtggacatca ccaaggacac cttctacaag cagcccatgt tctaccacct gggacacttc 2280
agcaagttca tccccgaggg ctctcagcgc gttggactgg tggcttccca gaagaacgat 2340
ctggacgccg tggctctgat gcaccctgat ggatctgctg tggtggtggt cctgaaccgc 2400
agcagcaaag atgtgcccct gaccatcaag gatcccgccg tgggattcct ggaaacaatc 2460

agccctggct actccatcca cacctacctg tggcgtagac aggagggcag aggaagtctt 2520
ctgacatgcg gagacgtgga agagaatccc ggccctatgc cccgcggctt cacctggctg 2580
cgctacctgg gcatcttcct gggcgtggcc ctgggcaacg agcccctgga gatgtggccc 2640
ctgacccaga acgaggagtg caccgtgacc ggcttcctgc gcgacaagct gcagtaccgc 2700
agccgcctgc agtacatgaa gcactacttc cccatcaact acaagatcag cgtgccctac 2760
gagggcgtgt tccgcatcgc caacgtgacc cgcctgcagc gcgcccaggt gagcgagcgc 2820

gagctgcgct acctgtgggt gctggtgagc ctgagcgcca ccgagagcgt gcaggacgtg 2880

ctgctggagg gccaccccag ctggaagtac ctgcaggagg tggagaccct gctgctgaac 2940
gtgcagcagg gcctgaccga cgtggaggtg agccccaagg tggagagcgt gctgagcctg 3000
ctgaacgccc ccggccccaa cctgaagctg gtgcgcccca aggccctgct ggacaactgc 3060
ttccgcgtga tggagctgct gtactgcagc tgctgcaagc agagcagcgt gctgaactgg 3120
caggactgcg aggtgcccag cccccagagc tgcagccccg agcccagcct gcagtacgcc 3180
gccacccagc tgtacccccc ccccccctgg agccccagca gcccccccca cagcaccggc 3240
agcgtgcgcc ccgtgcgcgc ccagggcgag ggcctgctgc cctaatgaca attgttaatt 3300

aagtttaaac cctcgaggcc gcaagcttat cgataatcaa cctctggatt acaaaatttg 3360
tgaaagattg actggtattc ttaactatgt tgctcctttt acgctatgtg gatacgctgc 3420
tttaatgcct ttgtatcatg ctattgcttc ccgtatggct ttcattttct cctccttgta 3480
taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtcaggc aacgtggcgt 3540
ggtgtgcact gtgtttgctg acgcaacccc cactggttgg ggcattgcca ccacctgtca 3600
gctcctttcc gggactttcg ctttccccct ccctattgcc acggcggaac tcatcgccgc 3660
ctgccttgcc cgctgctgga caggggctcg gctgttgggc actgacaatt ccgtggtgtt 3720

gtcggggaaa tcatcgtcct ttccttggct gctcgcctgt gttgccacct ggattctgcg 3780
cgggacgtcc ttctgctacg tcccttcggc cctcaatcca gcggaccttc cttcccgcgg 3840
cctgctgccg gctctgcggc ctcttccgcg tcttcgcctt cgccctcaga cgagtcggat 3900
ctccctttgg gccgcctccc cgcatcgata ccgtcgacta gagctcgctg atcagcctcg 3960
actgtgcctt ctagttgcca gccatctgtt gtttgcccct cccccgtgcc ttccttgacc 4020
ctggaaggtg ccactcccac tgtcctttcc taataaaatg aggaaattgc atcgcattgt 4080
ctgagtaggt gtcattctat tctggggggt ggggtggggc aggacagcaa gggggaggat 4140

tgggaagaca atagcaggca tgctggggag agatccacga taacaaacag ctttttttggg 4200
gtgaacatat tgactgaatt ccctgcaggt tggccactcc ctctctgcgc gctcgctcgc 4260
tcactgaggc cgcccgggca aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag 4320
tgagcgagcg agcgcgcaga gagggagtgg ccaactccat cactaggggt tcctgcggcc 4380
gctcgtacgg tctcgaggaa ttcctgcagg ataacttgcc aacctcattc taaaatgtat 4440
atagaagccc aaaagacaat aacaaaaata ttcttgtaga acaaaatggg aaagaatgtt 4500
ccactaaata tcaagattta gagcaaagca tgagatgtgt ggggatagac agtgaggctg 4560

ataaaataga gtagagctca gaaacagacc cattgatata tgtaagtgac ctatgaaaaa 4620
aatatggcat tttacaatgg gaaaatgatg gtctttttct tttttagaaa aacagggaaa 4680
tatatttata tgtaaaaaat aaaagggaac ccatatgtca taccatacac acaaaaaaat 4740
tccagtgaat tataagtcta aatggagaag gcaaaacttt aaatctttta gaaaataata 4800
tagaagcatg cagaccagcc tggccaacat gatgaaaccc tctctactaa taataaaatc 4860
agtagaacta ctcaggacta ctttgagtgg gaagtccttt tctatgaaga cttctttggc 4920
caaaattagg ctctaaatgc aaggagatag tgcatcatgc ctggctgcac ttactgataa 4980

atgatgttat caccatcttt aaccaaatgc acaggaacaa gttatggtac tgatgtgctg 5040
gattgagaag gagctctact tccttgacag gacacatttg tatcaactta aaaaagcaga 5100
tttttgccag cagaactatt cattcagagg taggaaactt agaatagatg atgtcactga 5160
ttagcatggc ttccccatct ccacagctgc ttcccaccca ggttgcccac agttgagttt 5220
gtccagtgct cagggctgcc cactctcagt aagaagcccc acaccagccc ctctccaaat 5280
atgttggctg ttccttccat taaagtgacc ccactttaga gcagcaagtg gatttctgtt 5340
tcttacagtt caggaaggag gagtcagctg tgagaacctg gagcctgaga tgcttctaag 5400

tcccactgct actggggtca gggaagccag actccagcat cagcagtcag gagcactaag 5460
cccttgccaa catcctgttt ctcagagaaa ctgcttccat tataatggtt gtccttttt 5520
aagctatcaa gccaaacaac cagtgtctac cattattctc atcacctgaa gccaagggtt 5580
ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct ccagcttctg tcttcagtca 5640
ctccactctt agcctgctct gaatcaactc tgaccacagt tccctggagc ccctgccacc 5700
tgctgcccct gccaccttct ccatctgcag tgctgtgcag ccttctgcac tcttgcagag 5760
ctaataggtg gagacttgaa ggaagaggag gaaagtttct cataatagcc ttgctgcaag 5820

ctcaaatggg aggtgggcac tgtgcccagg agccttggag caaaggctgt gcccaacctc 5880
tgactgcatc caggtttggt cttgacagag ataagaagcc ctggcttttg gagccaaaat 5940
ctaggtcaga cttaggcagg attctcaaag tttatcagca gaacatgagg cagaagaccc 6000
tttctgctcc agcttcttca ggctcaacct tcatcagaat agatagaaag agaggctgtg 6060
agggttctta aaacagaagc aaatctgact cagagaataa acaacctcct agtaaactac 6120
agcttagaca gagcatctgg tggtgagtgt gctcagtgtc ctactcaact gtctggtatc 6180
agccctcatg aggacttctc ttctttccct catagacctc catctctgtt ttccttagcc 6240

tgcagaaatc tggatggcta ttcacagaat gcctgtgctt tcagagttgc atttttctc 6300
tggtattctg gttcaagcat ttgaaggtag gaaaggttct ccaagtgcaa gaaagccagc 6360
cctgagcctc aactgcctgg ctagtgtggt cagtaggatg caaaggctgt tgaatgccac 6420

aaggccaaac tttaacctgt gtaccacaag cctagcagca gaggcagctc tgctcactgg 6480
aactctctgt cttctttctc ctgagccttt tcttttcctg agttttctag ctctcctcaa 6540
ccttacctct gccctaccca ggacaaaccc aagagccact gtttctgtga tgtcctctcc 6600
agccctaatt aggcatcatg acttcagcct gaccttccat gctcagaagc agtgctaatc 6660

cacttcagat gagctgctct atgcaacaca ggcagagcct acaaaccttt gcaccagagc 6720
cctccacata tcagtgtttg ttcatactca cttcaacagc aaatgtgact gctgagatta 6780
agattttaca caagatggtc tgtaatttca cagttagttt tatcccatta ggtatgaaag 6840
aattagcata attcccctta aacatgaatg aatcttagat tttttaataa atagttttgg 6900
aagtaaagac agagacatca ggagcacaag gaatagcctg agaggacaaa cagaacaaga 6960
aagagtctgg aaatacacag gatgttcttg gcctcctcaa agcaagtgca agcagatagt 7020
accagcagcc ccaggctatc agagcccagt gaagagaagt accatgaaag ccacagctct 7080

aaccaccctg ttccagagtg acagacagtc cccaagacaa gccagcctga gccagagaga 7140
gaactgcaag agaaagtttc taatttaggt tctgttagat tcagacaagt gcaggtcatc 7200
ctctctccac agctactcac ctctccagcc taacaaagcc tgcagtccac actccaaccc 7260
tggtgtctca cctcctagcc tctcccaaca tcctgctctc tgaccatctt ctgcatctct 7320
catctcacca tctcccactg tctacagcct actcttgcaa ctaccatctc attttctgac 7380
atcctgtcta catcttctgc catactctgc catctaccat accacctctt accatctacc 7440
acaccatctt ttatctccat ccctctcaga agcctccaag ctgaatcctg ctttatgtgt 7500

tcatctcagc ccctgcatgg aaagctgacc ccagaggcag aactattccc agagagcttg 7560
gccaagaaaa acaaaactac cagcctggcc aggctcagga gtagtaagct gcagtgtctg 7620
ttgtgttcta gcttcaacag ctgcaggagt tccactctca aatgctccac atttctcaca 7680
tcctcctgat tctggtcact acccatcttc aaagaacaga atatctcaca tcagcatact 7740
gtgaaggact agtcatgggt gcagctgctc agagctgcaa agtcattctg gatggtggag 7800
agcttacaaa catttcatga tgctcccccc gctctgatgg ctggagccca atccctacac 7860
agactcctgc tgtatgtgtt ttcctttcac tctgagccac agccagaggg caggcattca 7920

gtctcctctt caggctgggg ctggggcact gagaactcac ccaacacctt gctctcactc 7980
cttctgcaaa acaagaaaga gctttgtgct gcagtagcca tgaagaatga aaggaaggct 8040
ttaactaaaa aatgtcagag attattttca acccccttact gtggatcacc agcaaggagg 8100
aaacacaaca cagagacatt ttttcccctc aaattatcaa aagaatcact gcatttgtta 8160
aagagagcaa ctgaatcagg aagcagagtt ttgaacatat cagaagttag gaatctgcat 8220
cagagacaaa tgcagtcatg gttgtttgct gcataccagc cctaatcatt agaagcctca 8280

tggacttcaa acatcattcc ctctgacaag atgctctagc ctaactccat gagataaaat 8340

aaatctgcct ttcagagcca aagaagagtc caccagcttc ttctcagtgt gaacaagagc 8400
tccagtcagg ttagtcagtc cagtgcagta gaggagacca gtctgcatcc tctaattttc 8460
aaaggcaaga agatttgttt accctggaca ccaggcacaa gtgaggtcac agagctctta 8520
gatatgcagt cctcatgagt gaggagacta aagcgcatgc catcaagact tcagtgtaga 8580
gaaaacctcc aaaaaagcct cctcactact tctggaatag ctcagaggcc gaggcggcct 8640
cggcctctgc ataaataaaa aaaattagtc agccatgggg cggagaatgg gcggaactgg 8700
gcggagttag gggcgggatg ggcggagtta ggggcgggac tatggttgct gactaattga 8760

gatgcatgct ttgcatactt ctgcctgctg gggagcctgg ggactttcca cacctggttg 8820
ctgactaatt gagatgcatg ctttgcatac ttctgcctgc tggggagcct ggggactttc 8880
cacaccctaa ctgacacaca ttccacagct gcattaatga atcggccaac gcgcggggag 8940
aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt 9000
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga 9060
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 9120
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc ccccctgacg agcatcacaa 9180

aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 9240
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 9300
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 9360
cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 9420
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 9480
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 9540
tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat 9600

ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 9660
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 9720
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 9780
aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct 9840
tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga 9900
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc 9960
catagttgcc tgactcctgc aaaccacgtt gtgtctcaaa atctctgatg ttacattgca 10020

caagataaaa atatatcatc atgaacaata aaactgtctg cttacataaa cagtaataca 10080

aggggtgtta tgagccatat tcaacgggaa acgtcttgct cgaggccgcg attaaattcc 10140

aacatggatg ctgatttata tgggtataaa tgggctcgcg ataatgtcgg gcaatcaggt 10200

gcgacaatct atcgattgta tgggaagccc gatgcgccag agttgtttct gaaacatggc 10260

aaaggtagcg ttgccaatga tgttacagat gagatggtca gactaaactg gctgacggaa 10320

tttatgcctc ttccgaccat caagcatttt atccgtactc ctgatgatgc atggttactc 10380

accactgcga tccccgggaa aacagcattc caggtattag aagaatatcc tgattcaggt 10440

gaaaatattg ttgatgcgct ggcagtgttc ctgcgccggt tgcattcgat tcctgtttgt 10500

aattgtcctt ttaacagcga tcgcgtattt cgtctcgctc aggcgcaatc acgaatgaat 10560

aacggtttgg ttgatgcgag tgattttgat gacgagcgta atggctggcc tgttgaacaa 10620

gtctggaaag aaatgcataa gcttttgcca ttctcaccgg attcagtcgt cactcatggt 10680

gatttctcac ttgataacct tatttttgac gaggggaaat taataggttg tattgatgtt 10740

ggacgagtcg gaatcgcaga ccgataccag gatcttgcca tcctatggaa ctgcctcggt 10800

gagttttctc cttcattaca gaaacggctt tttcaaaaat atggtattga taatcctgat 10860

atgaataaat tgcagtttca tttgatgctc gatgagtttt tctaagggcg gcctgccacc 10920

atacccacgc cgaaacaagc gctcatgagc ccgaagtggc gagcccgatc ttccccatcg 10980

gtgatgtcgg cgatataggc gccagcaacc gcacctgtgg cgccggtgat gagggcgcgc 11040

caagtcgacg tccggcagtc 11060

<210> 60

<211> 10913

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 60

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420

tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctcctg gtggcgaggg gagggggtg gtcctcgaac gccttgcaga 600
actggcctgg atacagagtg gaccggctgg ccccatctgg aagacttcga gatacactgt 660
tgtcttactg cgctcaacag tgtatctcga agtcttccaa atggtgccag ccatcgcagc 720
ggggtgcagg aaatgggggc agcccccctt tttggctatc cttccacgtg ttctttttg 780
tatcttttgt gtttcctaga aaacatctca gtcaccaccg cagccctagg aatgcatcta 840
gacaattgta ctaaccttct tctctttcct ctcctgacag tccggaaagc caccatggaa 900
ttcagcagcc ccagcagaga ggaatgcccc aagcctctga gccgggtgtc aatcatggcc 960

ggatctctga caggactgct gctgcttcag gccgtgtctt gggcttctgg cgctagacct 1020
tgcatcccca agagcttcgg ctacagcagc gtcgtgtgcg tgtgcaatgc cacctactgc 1080
gacagcttcg accctcctac ctttcctgct ctgggcacct tcagcagata cgagagcacc 1140
agatccggca gacggatgga actgagcatg ggacccatcc aggccaatca cacaggcact 1200
ggcctgctgc tgacactgca gcctgagcag aaattccaga aagtgaaagg cttcggcgga 1260
gccatgacag atgccgccgc tctgaatatc ctggctctgt ctccaccagc tcagaacctg 1320
ctgctcaaga gctacttcag cgaggaaggc atcggctaca acatcatcag agtgcccatg 1380

gccagctgcg acttcagcat caggacctac acctacgccg acacacccga cgatttccag 1440
ctgcacaact tcagcctgcc tgaagaggac accaagctga agatccctct gatccacaga 1500
gccctgcagc tggcacaaag acccgtgtca ctgctggcct ctccatggac atctcccacc 1560
tggctgaaaa caaatggcgc cgtgaatggc aagggcagcc tgaaaggcca acctggcgac 1620
atctaccacc agacctgggc cagatacttc gtgaagttcc tggacgccta tgccgagcac 1680
aagctgcagt tttgggccgt gacagccgag aacgaacctt ctgctggact gctgagcggc 1740
taccccttc agtgcctggg ctttacaccc gagcaccagc gggactttat cgcccgtgat 1800

ctgggaccca cactggccaa tagcacccac cataatgtgc ggctgctgat gctggacgac 1860
cagagactgc ttctgcccca ctgggctaaa gtggtgctga cagatcctga ggccgccaaa 1920
tacgtgcacg gaatcgccgt gcactggtat ctggactttc tggcccctgc caaggccaca 1980
ctgggagaga cacacagact gttccccaac accatgctgt tcgccagcga agcctgtgtg 2040
ggcagcaagt tttgggaaca gagcgtgcgg ctcggcagct gggatagagg catgcagtac 2100
agccacagca tcatcaccaa cctgctgtac cacgtcgtcg gctggaccga ctggaatctg 2160
gccctgaatc ctgaaggcgg ccctaactgg gtccgaaact tcgtggacag ccccatcatc 2220

gtggacatca ccaaggacac cttctacaag cagcccatgt tctaccacct gggacacttc 2280
agcaagttca tccccgaggg ctctcagcgc gttggactgg tggcttccca gaagaacgat 2340
ctggacgccg tggctctgat gcaccctgat ggatctgctg tggtggtggt cctgaaccgc 2400
agcagcaaag atgtgcccct gaccatcaag gatcccgccg tgggattcct ggaaacaatc 2460
agccctggct actccatcca cacctacctg tggcgtagac agtgattgtg gccgaaccgc 2520
cgaactcaga ggccggcccc agaaaacccg agcgagtagg gggcggcgcg caggagggag 2580
gagaactggg ggcgcgggag gctggtgggt gtgggggtg gagatgtaga agatgtgacg 2640

ccgcggcccg gcgggtgcca gattagcgga cgcggtgccc gcggttgcaa cgggatcccg 2700
ggcgctgcag cttgggaggc ggctctcccc aggcggcgtc cgcggagaca cccatccgtg 2760
aaccccaggt cccgggccgc cggctcgccg cgcaccaggg gccggcggac agaagagcgg 2820
ccgagcggct cgaggctggg ggaccgcggg cgcggccgcg cgctgccggg cgggaggctg 2880
gggggccggg gccggggccg tgccccggag cgggtcggag gccggggccg gggccggggg 2940
acggcggctc cccgcgcggc tccagcggct cggggatccc ggccgggccc cgcagggacc 3000
atgatgcccc gcggcttcac ctggctgcgc tacctgggca tcttcctggg cgtggccctg 3060

ggcaacgagc ccctggagat gtggcccctg acccagaacg aggagtgcac cgtgaccggc 3120
ttcctgcgcg acaagctgca gtaccgcagc cgcctgcagt acatgaagca ctacttcccc 3180
atcaactaca agatcagcgt gccctacgag ggcgtgttcc gcatcgccaa cgtgacccgc 3240
ctgcagcgcg cccaggtgag cgagcgcgag ctgcgctacc tgtgggtgct ggtgagcctg 3300
agcgccaccg agagcgtgca ggacgtgctg ctggagggcc accccagctg gaagtacctg 3360
caggaggtgg agaccctgct gctgaacgtg cagcagggcc tgaccgacgt ggaggtgagc 3420
cccaaggtgg agagcgtgct gagcctgctg aacgcccccg gccccaacct gaagctggtg 3480

cgccccaagg ccctgctgga caactgcttc cgcgtgatgg agctgctgta ctgcagctgc 3540
tgcaagcaga gcagcgtgct gaactggcag gactgcgagg tgcccagccc ccagagctgc 3600
agccccgagc ccagcctgca gtacgccgcc acccagctgt acccccccc cccctggagc 3660
cccagcagcc ccccccacag caccggcagc gtgcgccccg tgcgcgccca gggcgagggc 3720
ctgctgccct aatgacaatt gttaattaag tttaaaccct cgaggccgca agccgcatcg 3780
ataccgtcga ctagagctcg ctgatcagcc tcgactgtgc cttctagttg ccagccatct 3840
gttgtttgcc cctcccccgt gccttccttg accctggaag gtgccactcc cactgtcctt 3900

tcctaataaa atgaggaaat tgcatcgcat tgtctgagta ggtgtcattc tattctgggg 3960
ggtggggtgg ggcaggacag caagggggag gattgggaag acaatagcag gcatgctggg 4020
gagagatcca cgataacaaa cagctttttt ggggtgaaca tattgactga attccctgca 4080

ggttggccac tccctctctg cgcgctcgct cgctcactga ggccgcccgg gcaaagcccg 4140
ggcgtcgggc gacctttggt cgcccggcct cagtgagcga gcgagcgcgc agagagggag 4200
tggccaactc catcactagg ggttcctgcg gccgctcgta cggtctcgag gaattcctgc 4260
aggataactt gccaacctca ttctaaaatg tatatagaag cccaaaagac aataacaaaa 4320

atattcttgt agaacaaaat gggaaagaat gttccactaa atatcaagat ttagagcaaa 4380
gcatgagatg tgtggggata gacagtgagg ctgataaaat agagtagagc tcagaaacag 4440
acccattgat atatgtaagt gacctatgaa aaaaatatgg cattttacaa tgggaaaatg 4500
atggtctttt tctttttag aaaaacaggg aaatatattt atatgtaaaa aataaaaggg 4560
aacccatatg tcataccata cacacaaaaa aattccagtg aattataagt ctaaatggag 4620
aaggcaaaac tttaaatctt ttagaaaata atatagaagc atgcagacca gcctggccaa 4680
catgatgaaa ccctctctac taataataaa atcagtagaa ctactcagga ctactttgag 4740

tgggaagtcc ttttctatga agacttcttt ggccaaaatt aggctctaaa tgcaaggaga 4800
tagtgcatca tgcctggctg cacttactga taaatgatgt tatcaccatc tttaaccaaa 4860
tgcacaggaa caagttatgg tactgatgtg ctggattgag aaggagctct acttccttga 4920
caggacacat ttgtatcaac ttaaaaaagc agatttttgc cagcagaact attcattcag 4980
aggtaggaaa cttagaatag atgatgtcac tgattagcat ggcttcccca tctccacagc 5040
tgcttcccac ccaggttgcc cacagttgag tttgtccagt gctcagggct gcccactctc 5100
agtaagaagc cccacaccag cccctctcca aatatgttgg ctgttccttc cattaaagtg 5160

accccacttt agagcagcaa gtggatttct gtttcttaca gttcaggaag gaggagtcag 5220
ctgtgagaac ctggagcctg agatgcttct aagtcccact gctactgggg tcagggaagc 5280
cagactccag catcagcagt caggagcact aagcccttgc caacatcctg tttctcagag 5340
aaactgcttc cattataatg gttgtccttt tttaagctat caagccaaac aaccagtgtc 5400
taccattatt ctcatcacct gaagccaagg gttctagcaa aagtcaagct gtcttgtaat 5460
ggttgatgtg cctccagctt ctgtcttcag tcactccact cttagcctgc tctgaatcaa 5520
ctctgaccac agttccctgg agcccctgcc acctgctgcc cctgccacct tctccatctg 5580

cagtgctgtg cagccttctg cactcttgca gagctaatag gtggagactt gaaggaagag 5640
gaggaaagtt tctcataata gccttgctgc aagctcaaat gggaggtggg cactgtgccc 5700
aggagccttg gagcaaaggc tgtgcccaac ctctgactgc atccaggttt ggtcttgaca 5760
gagataagaa gccctggctt ttggagccaa aatctaggtc agacttaggc aggattctca 5820
aagtttatca gcagaacatg aggcagaaga ccctttctgc tccagcttct tcaggctcaa 5880
ccttcatcag aatagataga aagagaggct gtgagggttc ttaaaacaga agcaaatctg 5940

actcagagaa taaacaacct cctagtaaac tacagcttag acagagcatc tggtggtgag 6000

tgtgctcagt gtcctactca actgtctggt atcagccctc atgaggactt ctcttctttc 6060
cctcatagac ctccatctct gttttcctta gcctgcagaa atctggatgg ctattcacag 6120
aatgcctgtg ctttcagagt tgcatttttt ctctggtatt ctggttcaag catttgaagg 6180
taggaaaggt tctccaagtg caagaaagcc agccctgagc ctcaactgcc tggctagtgt 6240
ggtcagtagg atgcaaaggc tgttgaatgc cacaaggcca aactttaacc tgtgtaccac 6300
aagcctagca gcagaggcag ctctgctcac tggaactctc tgtcttcttt ctcctgagcc 6360
ttttcttttc ctgagttttc tagctctcct caaccttacc tctgccctac ccaggacaaa 6420

cccaagagcc actgtttctg tgatgtcctc tccagcccta attaggcatc atgacttcag 6480
cctgaccttc catgctcaga agcagtgcta atccacttca gatgagctgc tctatgcaac 6540
acaggcagag cctacaaacc tttgcaccag agccctccac atatcagtgt ttgttcatac 6600
tcacttcaac agcaaatgtg actgctgaga ttaagatttt acacaagatg gtctgtaatt 6660
tcacagttag ttttatccca ttaggtatga aagaattagc ataattcccc ttaaacatga 6720
atgaatctta gatttttaa taaatagttt tggaagtaaa gacagagaca tcaggagcac 6780
aaggaatagc ctgagaggac aaacagaaca agaaagagtc tggaaataca caggatgttc 6840

ttggcctcct caaagcaagt gcaagcagat agtaccagca gccccaggct atcagagccc 6900
agtgaagaga agtaccatga aagccacagc tctaaccacc ctgttccaga gtgacagaca 6960
gtccccaaga caagccagcc tgagccagag agagaactgc aagagaaagt ttctaattta 7020
ggttctgtta gattcagaca agtgcaggtc atcctctctc cacagctact cacctctcca 7080
gcctaacaaa gcctgcagtc cacactccaa ccctggtgtc tcacctccta gcctctccca 7140
acatcctgct ctctgaccat cttctgcatc tctcatctca ccatctccca ctgtctacag 7200
cctactcttg caactaccat ctcattttct gacatcctgt ctacatcttc tgccatactc 7260

tgccatctac cataccacct cttaccatct accacaccat cttttatctc catccctctc 7320
agaagcctcc aagctgaatc ctgctttatg tgttcatctc agcccctgca tggaaagctg 7380
accccagagg cagaactatt cccagagagc ttggccaaga aaaacaaaac taccagcctg 7440
gccaggctca ggagtagtaa gctgcagtgt ctgttgtgtt ctagcttcaa cagctgcagg 7500
agttccactc tcaaatgctc cacatttctc acatcctcct gattctggtc actacccatc 7560
ttcaaagaac agaatatctc acatcagcat actgtgaagg actagtcatg ggtgcagctg 7620
ctcagagctg caaagtcatt ctggatggtg gagagcttac aaacatttca tgatgctccc 7680

cccgctctga tggctggagc ccaatcccta cacagactcc tgctgtatgt gttttccttt 7740
cactctgagc cacagccaga gggcaggcat tcagtctcct cttcaggctg gggctggggc 7800
actgagaact cacccaacac cttgctctca ctccttctgc aaaacaagaa agagctttgt 7860
gctgcagtag ccatgaagaa tgaaaggaag gctttaacta aaaaatgtca gagattattt 7920
tcaacccctt actgtggatc accagcaagg aggaaacaca acacagagac atttttccc 7980
ctcaaattat caaaagaatc actgcatttg ttaaagagag caactgaatc aggaagcaga 8040
gttttgaaca tatcagaagt taggaatctg catcagagac aaatgcagtc atggttgttt 8100

gctgcatacc agccctaatc attagaagcc tcatggactt caaacatcat tccctctgac 8160
aagatgctct agcctaactc catgagataa aataaatctg cctttcagag ccaaagaaga 8220
gtccaccagc ttcttctcag tgtgaacaag agctccagtc aggttagtca gtccagtgca 8280
gtagaggaga ccagtctgca tcctctaatt ttcaaaggca agaagatttg tttaccctgg 8340
acaccaggca caagtgaggt cacagagctc ttagatatgc agtcctcatg agtgaggaga 8400
ctaaagcgca tgccatcaag acttcagtgt agagaaaacc tccaaaaaag cctcctcact 8460
acttctggaa tagctcagag gccgaggcgg cctcggcctc tgcataaata aaaaaaatta 8520

gtcagccatg gggcggagaa tgggcggaac tgggcggagt taggggcggg atgggcggag 8580
ttaggggcgg gactatggtt gctgactaat tgagatgcat gctttgcata cttctgcctg 8640
ctggggagcc tggggacttt ccacacctgg ttgctgacta attgagatgc atgctttgca 8700
tacttctgcc tgctggggag cctggggact ttccacaccc taactgacac acattccaca 8760
gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg ggcgctcttc 8820
cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc 8880
tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag gaaagaacat 8940

gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt 9000
ccataggctc cgcccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg 9060
aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc 9120
tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt 9180
ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa 9240
gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaacta 9300
tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa 9360

caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa 9420
ctacggctac actagaagaa cagtatttgg tatctgcgct ctgctgaagc cagttacctt 9480
cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta gcggtggttt 9540

ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat 9600

cttttctacg gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat 9660

gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc 9720

aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc 9780

acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc tgcaaaccac 9840

gttgtgtctc aaaatctctg atgttacatt gcacaagata aaaatatatc atcatgaaca 9900

ataaaactgt ctgcttacat aaacagtaat acaaggggtg ttatgagcca tattcaacgg 9960

gaaacgtctt gctcgaggcc gcgattaaat tccaacatgg atgctgattt atatgggtat 10020

aaatgggctc gcgataatgt cgggcaatca ggtgcgacaa tctatcgatt gtatgggaag 10080

cccgatgcgc cagagttgtt tctgaaacat ggcaaaggta gcgttgccaa tgatgttaca 10140

gatgagatgg tcagactaaa ctggctgacg gaatttatgc ctcttccgac catcaagcat 10200

tttatccgta ctcctgatga tgcatggtta ctcaccactg cgatccccgg gaaaacagca 10260

ttccaggtat tagaagaata tcctgattca ggtgaaaata ttgttgatgc gctggcagtg 10320

ttcctgcgcc ggttgcattc gattcctgtt tgtaattgtc cttttaacag cgatcgcgta 10380

tttcgtctcg ctcaggcgca atcacgaatg aataacggtt tggttgatgc gagtgatttt 10440

gatgacgagc gtaatggctg gcctgttgaa caagtctgga aagaaatgca taagcttttg 10500

ccattctcac cggattcagt cgtcactcat ggtgatttct cacttgataa ccttattttt 10560

gacgagggga aattaatagg ttgtattgat gttggacgag tcggaatcgc agaccgatac 10620

caggatcttg ccatcctatg gaactgcctc ggtgagtttt ctccttcatt acagaaacgg 10680

ctttttcaaa aatatggtat tgataatcct gatatgaata aattgcagtt tcatttgatg 10740

ctcgatgagt ttttctaagg gcggcctgcc accataccca cgccgaaaca agcgctcatg 10800

agcccgaagt ggcgagcccg atcttcccca tcggtgatgt cggcgatata ggcgccagca 10860

accgcacctg tggcgccggt gatgagggcg cgccaagtcg acgtccggca gtc 10913

<210> 61

<211> 11209

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 61

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240
tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300
tagggcggag ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga 360
atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420
tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta 480

agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540
tggcactatg aaccctgcag ccctaggaat gcatctagac aattgtacta accttcttct 600
ctttcctctc ctgacagtcc ggaaagccac catggaattc agcagcccca gcagagagga 660
atgccccaag cctctgagcc gggtgtcaat catggccgga tctctgacag gactgctgct 720
gcttcaggcc gtgtcttggg cttctggcgc tagaccttgc atccccaaga gcttcggcta 780
cagcagcgtc gtgtgcgtgt gcaatgccac ctactgcgac agcttcgacc ctcctacctt 840
tcctgctctg ggcaccttca gcagatacga gagcaccaga tccggcagac ggatggaact 900

gagcatggga cccatccagg ccaatcacac aggcactggc ctgctgctga cactgcagcc 960
tgagcagaaa ttccagaaag tgaaaggctt cggcggagcc atgacagatg ccgccgctct 1020
gaatatcctg gctctgtctc caccagctca gaacctgctg ctcaagagct acttcagcga 1080
ggaaggcatc ggctacaaca tcatcagagt gcccatggcc agctgcgact tcagcatcag 1140
gacctacacc tacgccgaca cacccgacga tttccagctg cacaacttca gcctgcctga 1200
agaggacacc aagctgaaga tccctctgat ccacagagcc ctgcagctgg cacaaagacc 1260
cgtgtcactg ctggcctctc catggacatc tcccacctgg ctgaaaacaa atggcgccgt 1320

gaatggcaag ggcagcctga aaggccaacc tggcgacatc taccaccaga cctgggccag 1380
atacttcgtg aagttcctgg acgcctatgc cgagcacaag ctgcagtttt gggccgtgac 1440
agccgagaac gaaccttctg ctggactgct gagcggctac ccctttcagt gcctgggctt 1500
tacacccgag caccagcggg actttatcgc ccgtgatctg ggacccacac tggccaatag 1560
cacccaccat aatgtgcggc tgctgatgct ggacgaccag agactgcttc tgccccactg 1620
ggctaaagtg gtgctgacag atcctgaggc cgccaaatac gtgcacggaa tcgccgtgca 1680
ctggtatctg gactttctgg cccctgccaa ggccacactg ggagagacac acagactgtt 1740

ccccaacacc atgctgttcg ccagcgaagc ctgtgtgggc agcaagtttt gggaacagag 1800
cgtgcggctc ggcagctggg atagaggcat gcagtacagc cacagcatca tcaccaacct 1860
gctgtaccac gtcgtcggct ggaccgactg gaatctggcc ctgaatcctg aaggcggccc 1920
taactgggtc cgaaacttcg tggacagccc catcatcgtg gacatcacca aggacacctt 1980

ctacaagcag cccatgttct accacctggg acacttcagc aagttcatcc ccgagggctc 2040

tcagcgcgtt ggactggtgg cttcccagaa gaacgatctg gacgccgtgg ctctgatgca 2100

ccctgatgga tctgctgtgg tggtggtcct gaaccgcagc agcaaagatg tgcccctgac 2160

catcaaggat cccgccgtgg gattcctgga aacaatcagc cctggctact ccatccacac 2220

ctacctgtgg cgtagacagt gacaattgtt aattaagttt aaaccctcga ggccgcaagc 2280

cgcatcgata ccgtcgacta gagctcgctg atcagcctcg actgtgcctt ctagttgcca 2340

gccatctgtt gtttgcccct cccccgtgcc ttccttgacc ctggaaggtg ccactcccac 2400

tgtcctttcc taataaaatg aggaaattgc atcgcattgt ctgagtaggt gtcattctat 2460

tctgggggt ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca 2520

tgctggggag agatccacga taacaaacag ctttttggg ggatatcaaa ctgcctgttt 2580

gggcttctca tttcttacct ccccttccct ctcccacctg ctactgggtg catctctgct 2640

ccccccttcc ccagcagatg gttacctttg ggctgttgct ttcttgtcac catctgagtt 2700

ctcagacgct ggaaagccat gttctcggct ctgtgaatga caatgctgac tggagtgctg 2760

cccctctgta aagggctggg tgtggatggt cacaagcccc tcacatgcct cagccaagag 2820

gaagtagtac aggggtcagc ccagaggtcc aggggaaagg agtggaaacc gatttcccca 2880

ccaagggagg ggcctgtacc tcagctgttc ccatagctta cttgccacaa ctgccaagca 2940

agtttcgctg agtttgacac atggatccct gtggatcaac tgccctagga ctccgtttgc 3000

acccatgtga cactgttgac tttgccctga cgaagcaggg ccaacagtcc cctaacttaa 3060

ttacaaaaac taatgactaa gagagaggtg gctagagctg aggcccctga gtcaggctgt 3120

gggtgggatc atctccagta caggaagtga gactttcatt tcctcctttc caagagaggg 3180

ctgagggagc agggttgagc aactggtgca gacagcctag ctggactttg ggtgaggcgg 3240

ttcagccata tcgaattctg ctggggctac tggcaggtaa ggaggaagga ggctgagggg 3300

agggggcccc tgggagggag cctgccctgg gttgctaacc atctcctctc tgccaaaagt 3360

ccggaaagcc accatggagc ccctgcgcct gctgatcctg ctgttcgtga ccgagctgag 3420

cggcgcccac aacaccaccg tgttccaggg cgtggccggc cagagcctgc aggtgagctg 3480

cccctacgac agcatgaagc actggggccg ccgcaaggcc tggtgccgcc agctgggcga 3540

gaagggcccc tgccagcgcg tggtgagcac ccacaacctg tggctgctga gcttcctgcg 3600

ccgctggaac ggcagcaccg ccatcaccga cgacaccctg ggcggcaccc tgaccatcac 3660

cctgcgcaac ctgcagcccc acgacgccgg cctgtaccag tgccagagcc tgcacggcag 3720

cgaggccgac accctgcgca aggtgctggt ggaggtgctg gccgaccccc tggaccaccg 3780

cgacgccggc gacctgtggt tccccggcga gagcgagagc ttcgaggacg cccacgtgga 3840

gcacagcatc agccgcagcc tgctggaggg cgagatcccc ttccccccca ccagcatcct 3900
gctgctgctg gcctgcatct tcctgatcaa gatcctggcc gccagcgccc tgtgggccgc 3960
cgcctggcac ggccagaagc ccggcaccca ccccccccagc gagctggact gcggccacga 4020
ccccggctac cagctgcaga ccctgcccgg cctgcgcgac acctgaccca ggggactcag 4080
cggccgctcg agtctagagg gcccgtttaa acccgctgat cagcctcgaa gacatgataa 4140
gatacattga tgagtttgga caaaccacaa caagaatgca gtgaaaaaaa tgctttattt 4200
gtgaaatttg tgatgctatt gctttatttg taaccattat aagctgcaat aaacaagtta 4260

acaacaacaa ttgcattcat tttatgtttc aggttcaggg ggagatgtgg gaggtttttt 4320
aaagcaagta aaacctctac aaatgtggta tgaacatatt gactgaattc cctgcaggtt 4380
ggccactccc tctctgcgcg ctcgctcgct cactgaggcc gcccgggcaa agcccgggcg 4440
tcgggcgacc tttggtcgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc 4500
caactccatc actaggggtt cctgcggccg ctcgtacggt ctcgaggaat tcctgcagga 4560
taacttgcca acctcattct aaaatgtata tagaagccca aaagacaata acaaaaatat 4620
tcttgtagaa caaaatggga aagaatgttc cactaaatat caagatttag agcaaagcat 4680

gagatgtgtg gggatagaca gtgaggctga taaaatagag tagagctcag aaacagaccc 4740
attgatatat gtaagtgacc tatgaaaaaa atatggcatt ttacaatggg aaaatgatgg 4800
tctttttctt ttttagaaaa acagggaaat atatttatat gtaaaaaata aaagggaacc 4860
catatgtcat accatacaca caaaaaaatt ccagtgaatt ataagtctaa atggagaagg 4920
caaaacttta aatcttttag aaaataatat agaagcatgc agaccagcct ggccaacatg 4980
atgaaaccct ctctactaat aataaaatca gtagaactac tcaggactac tttgagtggg 5040
aagtcctttt ctatgaagac ttctttggcc aaaattaggc tctaaatgca aggagatagt 5100

gcatcatgcc tggctgcact tactgataaa tgatgttatc accatcttta accaaatgca 5160
caggaacaag ttatggtact gatgtgctgg attgagaagg agctctactt ccttgacagg 5220
acacatttgt atcaacttaa aaaagcagat ttttgccagc agaactattc attcagaggt 5280
aggaaactta gaatagatga tgtcactgat tagcatggct tccccatctc cacagctgct 5340
tcccacccag gttgcccaca gttgagtttg tccagtgctc agggctgccc actctcagta 5400
agaagcccca caccagcccc tctccaaata tgttggctgt tccttccatt aaagtgaccc 5460
cactttagag cagcaagtgg atttctgttt cttacagttc aggaaggagg agtcagctgt 5520

gagaacctgg agcctgagat gcttctaagt cccactgcta ctggggtcag ggaagccaga 5580
ctccagcatc agcagtcagg agcactaagc ccttgccaac atcctgtttc tcagagaaac 5640
tgcttccatt ataatggttg tccttttta agctatcaag ccaaacaacc agtgtctacc 5700
attattctca tcacctgaag ccaagggttc tagcaaaagt caagctgtct tgtaatggtt 5760
gatgtgcctc cagcttctgt cttcagtcac tccactctta gcctgctctg aatcaactct 5820
gaccacagtt ccctggagcc cctgccacct gctgcccctg ccaccttctc catctgcagt 5880
gctgtgcagc cttctgcact cttgcagagc taataggtgg agacttgaag gaagaggagg 5940

aaagtttctc ataatagcct tgctgcaagc tcaaatggga ggtgggcact gtgcccagga 6000
gccttggagc aaaggctgtg cccaacctct gactgcatcc aggtttggtc ttgacagaga 6060
taagaagccc tggcttttgg agccaaaatc taggtcagac ttaggcagga ttctcaaagt 6120
ttatcagcag aacatgaggc agaagaccct ttctgctcca gcttcttcag gctcaacctt 6180
catcagaata gatagaaaga gaggctgtga gggttcttaa aacagaagca aatctgactc 6240
agagaataaa caacctccta gtaaactaca gcttagacag agcatctggt ggtgagtgtg 6300
ctcagtgtcc tactcaactg tctggtatca gccctcatga ggacttctct tctttccctc 6360

atagacctcc atctctgttt tccttagcct gcagaaatct ggatggctat tcacagaatg 6420
cctgtgcttt cagagttgca ttttttctct ggtattctgg ttcaagcatt tgaaggtagg 6480
aaaggttctc caagtgcaag aaagccagcc ctgagcctca actgcctggc tagtgtggtc 6540
agtaggatgc aaaggctgtt gaatgccaca aggccaaact ttaacctgtg taccacaagc 6600
ctagcagcag aggcagctct gctcactgga actctctgtc ttctttctcc tgagcctttt 6660
cttttcctga gttttctagc tctcctcaac cttacctctg ccctacccag gacaaaccca 6720
agagccactg tttctgtgat gtcctctcca gccctaatta ggcatcatga cttcagcctg 6780

accttccatg ctcagaagca gtgctaatcc acttcagatg agctgctcta tgcaacacag 6840
gcagagccta caaacctttg caccagagcc ctccacatat cagtgtttgt tcatactcac 6900
ttcaacagca aatgtgactg ctgagattaa gattttacac aagatggtct gtaatttcac 6960
agttagtttt atcccattag gtatgaaaga attagcataa ttccccttaa acatgaatga 7020
atcttagatt ttttaataaa tagttttgga agtaaagaca gagacatcag gagcacaagg 7080
aatagcctga gaggacaaac agaacaagaa agagtctgga aatacacagg atgttcttgg 7140
cctcctcaaa gcaagtgcaa gcagatagta ccagcagccc caggctatca gagcccagtg 7200

aagagaagta ccatgaaagc cacagctcta accaccctgt tccagagtga cagacagtcc 7260
ccaagacaag ccagcctgag ccagagagag aactgcaaga gaaagtttct aatttaggtt 7320
ctgttagatt cagacaagtg caggtcatcc tctctccaca gctactcacc tctccagcct 7380

aacaaagcct gcagtccaca ctccaaccct ggtgtctcac ctcctagcct ctcccaacat 7440
cctgctctct gaccatcttc tgcatctctc atctcaccat ctcccactgt ctacagccta 7500
ctcttgcaac taccatctca ttttctgaca tcctgtctac atcttctgcc atactctgcc 7560
atctaccata ccacctctta ccatctacca caccatcttt tatctccatc cctctcagaa 7620

gcctccaagc tgaatcctgc tttatgtgtt catctcagcc cctgcatgga aagctgaccc 7680
cagaggcaga actattccca gagagcttgg ccaagaaaaa caaaactacc agcctggcca 7740
ggctcaggag tagtaagctg cagtgtctgt tgtgttctag cttcaacagc tgcaggagtt 7800
ccactctcaa atgctccaca tttctcacat cctcctgatt ctggtcacta cccatcttca 7860
aagaacagaa tatctcacat cagcatactg tgaaggacta gtcatgggtg cagctgctca 7920
gagctgcaaa gtcattctgg atggtggaga gcttacaaac atttcatgat gctccccccg 7980
ctctgatggc tggagcccaa tccctacaca gactcctgct gtatgtgttt tcctttcact 8040

ctgagccaca gccagagggc aggcattcag tctcctcttc aggctggggc tggggcactg 8100
agaactcacc caacaccttg ctctcactcc ttctgcaaaa caagaaagag ctttgtgctg 8160
cagtagccat gaagaatgaa aggaaggctt taactaaaaa atgtcagaga ttattttcaa 8220
ccccttactg tggatcacca gcaaggagga aacacaacac agagacattt tttcccctca 8280
aattatcaaa agaatcactg catttgttaa agagagcaac tgaatcagga agcagagttt 8340
tgaacatatc agaagttagg aatctgcatc agagacaaat gcagtcatgg ttgtttgctg 8400
cataccagcc ctaatcatta gaagcctcat ggacttcaaa catcattccc tctgacaaga 8460

tgctctagcc taactccatg agataaaata aatctgcctt tcagagccaa agaagagtcc 8520
accagcttct tctcagtgtg aacaagagct ccagtcaggt tagtcagtcc agtgcagtag 8580
aggagaccag tctgcatcct ctaattttca aaggcaagaa gatttgttta ccctggacac 8640
caggcacaag tgaggtcaca gagctcttag atatgcagtc ctcatgagtg aggagactaa 8700
agcgcatgcc atcaagactt cagtgtagag aaaacctcca aaaaagcctc ctcactactt 8760
ctggaatagc tcagaggccg aggcggcctc ggcctctgca taaataaaaa aaattagtca 8820
gccatggggc ggagaatggg cggaactggg cggagttagg ggcgggatgg gcggagttag 8880

gggcgggact atggttgctg actaattgag atgcatgctt tgcatacttc tgcctgctgg 8940
ggagcctggg gactttccac acctggttgc tgactaattg agatgcatgc tttgcatact 9000
tctgcctgct ggggagcctg gggactttcc acaccctaac tgacacacat tccacagctg 9060
cattaatgaa tcggccaacg cgcggggaga ggcggtttgc gtattgggcg ctcttccgct 9120
tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac 9180
tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga 9240

gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat 9300

aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac 9360
ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct 9420
gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg 9480
ctttctcata gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg 9540
ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt 9600
cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg 9660
attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac 9720

ggctacacta gaagaacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga 9780
aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggtttttt 9840
gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt 9900
tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga 9960
ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc 10020
taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct 10080
atctcagcga tctgtctatt tcgttcatcc atagttgcct gactcctgca aaccacgttg 10140

tgtctcaaaa tctctgatgt tacattgcac aagataaaaa tatatcatca tgaacaataa 10200
aactgtctgc ttacataaac agtaatacaa ggggtgttat gagccatatt caacgggaaa 10260
cgtcttgctc gaggccgcga ttaaattcca acatggatgc tgatttatat gggtataaat 10320
gggctcgcga taatgtcggg caatcaggtg cgacaatcta tcgattgtat gggaagcccg 10380
atgcgccaga gttgtttctg aaacatggca aaggtagcgt tgccaatgat gttacagatg 10440
agatggtcag actaaactgg ctgacggaat ttatgcctct tccgaccatc aagcatttta 10500
tccgtactcc tgatgatgca tggttactca ccactgcgat ccccgggaaa acagcattcc 10560

aggtattaga agaatatcct gattcaggtg aaaatattgt tgatgcgctg gcagtgttcc 10620
tgcgccggtt gcattcgatt cctgtttgta attgtccttt taacagcgat cgcgtatttc 10680
gtctcgctca ggcgcaatca cgaatgaata acggtttggt tgatgcgagt gattttgatg 10740
acgagcgtaa tggctggcct gttgaacaag tctggaaaga aatgcataag cttttgccat 10800
tctcaccgga ttcagtcgtc actcatggtg atttctcact tgataacctt atttttgacg 10860
aggggaaatt aataggttgt attgatgttg gacgagtcgg aatcgcagac cgataccagg 10920
atcttgccat cctatggaac tgcctcggtg agttttctcc ttcattacag aaacggcttt 10980

ttcaaaaata tggtattgat aatcctgata tgaataaatt gcagtttcat ttgatgctcg 11040

atgagttttt ctaagggcgg cctgccacca tacccacgcc gaaacaagcg ctcatgagcc 11100

cgaagtggcg agcccgatct tccccatcgg tgatgtcggc gatataggcg ccagcaaccg 11160

cacctgtggc gccggtgatg agggcgcgcc aagtcgacgt ccggcagtc 11209

<210> 62

<211> 11459

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 62

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300

cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc 360

cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga cccccgccca 420

ttgacgtcaa taatgacgta tgttcccata gtaacgccaa tagggacttt ccattgacgt 480

caatgggtgg actatttacg gtaaactgcc cacttggcag tacatcaagt gtatcatatg 540

ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgcccag 600

tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt 660

accatggtcg aggtgagccc cacgttctgc ttcactctcc ccatctcccc cccctcccca 720

cccccaattt tgtatttatt tatttttaa ttattttgtg cagcgatggg ggcgggggg 780

gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg 840

agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg 900

cggcggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg 960

ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cgcccgcccc ggctctgact 1020

gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctcctccgg gctgtaatta 1080

gcgcttggtt taatgacggc ttgttttctg tggctgcgtg aaagccttga ggggctccgg 1140

gagctagagc ctctgctaac catgttcatg ccttcttctt tttcctacag ctcctgggca 1200

acgtgctggt tattgtgctg tctcatcatt ttggcaaaga attcctcgaa gatccgaagg 1260

gaaagtcttc cacgactgtg ggatccgttc gaagatatca ccggttgagc caccatggaa 1320

ttcagcagcc ccagcagaga ggaatgcccc aagcctctga gccgggtgtc aatcatggcc 1380

ggatctctga caggactgct gctgcttcag gccgtgtctt gggcttctgg cgctagacct 1440
tgcatcccca agagcttcgg ctacagcagc gtcgtgtgcg tgtgcaatgc cacctactgc 1500
gacagcttcg accctcctac ctttcctgct ctgggcacct tcagcagata cgagagcacc 1560
agatccggca gacggatgga actgagcatg ggacccatcc aggccaatca cacaggcact 1620
ggcctgctgc tgacactgca gcctgagcag aaattccaga aagtgaaagg cttcggcgga 1680
gccatgacag atgccgccgc tctgaatatc ctggctctgt ctccaccagc tcagaacctg 1740
ctgctcaaga gctacttcag cgaggaaggc atcggctaca acatcatcag agtgcccatg 1800

gccagctgcg acttcagcat caggacctac acctacgccg acacacccga cgatttccag 1860
ctgcacaact tcagcctgcc tgaagaggac accaagctga agatccctct gatccacaga 1920
gccctgcagc tggcacaaag acccgtgtca ctgctggcct ctccatggac atctcccacc 1980
tggctgaaaa caaatggcgc cgtgaatggc aagggcagcc tgaaaggcca acctggcgac 2040
atctaccacc agacctgggc cagatacttc gtgaagttcc tggacgccta tgccgagcac 2100
aagctgcagt tttgggccgt gacagccgag aacgaacctt ctgctggact gctgagcggc 2160
taccccttc agtgcctggg ctttacaccc gagcaccagc gggactttat cgcccgtgat 2220

ctgggaccca cactggccaa tagcacccac cataatgtgc ggctgctgat gctggacgac 2280
cagagactgc ttctgcccca ctgggctaaa gtggtgctga cagatcctga ggccgccaaa 2340
tacgtgcacg gaatcgccgt gcactggtat ctggactttc tggcccctgc caaggccaca 2400
ctgggagaga cacacagact gttccccaac accatgctgt tcgccagcga agcctgtgtg 2460
ggcagcaagt tttgggaaca gagcgtgcgg ctcggcagct gggatagagg catgcagtac 2520
agccacagca tcatcaccaa cctgctgtac cacgtcgtcg gctggaccga ctggaatctg 2580
gccctgaatc ctgaaggcgg ccctaactgg gtccgaaact tcgtggacag ccccatcatc 2640

gtggacatca ccaaggacac cttctacaag cagcccatgt tctaccacct gggacacttc 2700
agcaagttca tccccgaggg ctctcagcgc gttggactgg tggcttccca gaagaacgat 2760
ctggacgccg tggctctgat gcaccctgat ggatctgctg tggtggtggt cctgaaccgc 2820
agcagcaaag atgtgcccct gaccatcaag gatcccgccg tgggattcct ggaaacaatc 2880
agccctggct actccatcca cacctacctg tggcgtagac agtgacaatt gttaattaag 2940
tttaaaccct cgaggccgca agccgcatcg ataccgtcga ctagagctcg ctgatcagcc 3000
tcgactgtgc cttctagttg ccagccatct gttgtttgcc cctcccccgt gccttccttg 3060

accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcatcgcat 3120
tgtctgagta ggtgtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag 3180
gattgggaag acaatagcag gcatgctggg gagagatcca cgataacaaa cagctttttt 3240
gggggggcgg agttagggcg gagccaatca gcgtgcgccg ttccgaaagt tgccttttat 3300
ggctgggcgg agaatgggcg gtgaacgccg atgattatat aaggacgcgc cgggtgtggc 3360
acagctagtt ccgtcgcagc cgggatttgg gtcgcggttc ttgtttgtgg atccctgtga 3420
tcgtcacttg gtaagtcact gactgtctat gcctgggaaa gggtgggcag gagatggggc 3480

agtgcaggaa aagtggcact atgaaccctg cagccctagg aatgcatcta gacaattgta 3540
ctaaccttct tctctttcct ctcctgacag tccggaaagc caccatgccc cgcggcttca 3600
cctggctgcg ctacctgggc atcttcctgg gcgtggccct gggcaacgag cccctggaga 3660
tgtggcccct gacccagaac gaggagtgca ccgtgaccgg cttcctgcgc gacaagctgc 3720
agtaccgcag ccgcctgcag tacatgaagc actacttccc catcaactac aagatcagcg 3780
tgccctacga gggcgtgttc cgcatcgcca acgtgacccg cctgcagcgc gcccaggtga 3840
gcgagcgcga gctgcgctac ctgtgggtgc tggtgagcct gagcgccacc gagagcgtgc 3900

aggacgtgct gctggagggc caccccagct ggaagtacct gcaggaggtg gagaccctgc 3960
tgctgaacgt gcagcagggc ctgaccgacg tggaggtgag ccccaaggtg gagagcgtgc 4020
tgagcctgct gaacgccccc ggccccaacc tgaagctggt gcgccccaag gccctgctgg 4080
acaactgctt ccgcgtgatg gagctgctgt actgcagctg ctgcaagcag agcagcgtgc 4140
tgaactggca ggactgcgag gtgcccagcc cccagagctg cagccccgag cccagcctgc 4200
agtacgccgc cacccagctg taccccccc ccccctggag ccccagcagc cccccccaca 4260
gcaccggcag cgtgcgcccc gtgcgcgccc agggcgaggg cctgctgccc taatgaccca 4320

ggggactcag cggccgctcg agtctagagg gcccgtttaa acccgctgat cagcctcgaa 4380
gacatgataa gatacattga tgagtttgga caaaccacaa caagaatgca gtgaaaaaaa 4440
tgctttattt gtgaaatttg tgatgctatt gctttatttg taaccattat aagctgcaat 4500
aaacaagtta acaacaacaa ttgcattcat tttatgtttc aggttcaggg ggagatgtgg 4560
gaggtttttt aaagcaagta aaacctctac aaatgtggta tgaacatatt gactgaattc 4620
cctgcaggtt ggccactccc tctctgcgcg ctcgctcgct cactgaggcc gcccgggcaa 4680
agcccgggcg tcgggcgacc tttggtcgcc cggcctcagt gagcgagcga gcgcgcagag 4740

agggagtggc caactccatc actaggggtt cctgcggccg ctcgtacggt ctcgaggaat 4800
tcctgcagga taacttgcca acctcattct aaaatgtata tagaagccca aaagacaata 4860
acaaaaatat tcttgtagaa caaaatggga aagaatgttc cactaaatat caagatttag 4920

agcaaagcat gagatgtgtg gggatagaca gtgaggctga taaaatagag tagagctcag 4980
aaacagaccc attgatatat gtaagtgacc tatgaaaaaa atatggcatt ttacaatggg 5040
aaaatgatgg tctttttctt ttttagaaaa acagggaaat atatttatat gtaaaaaata 5100
aaagggaacc catatgtcat accatacaca caaaaaaatt ccagtgaatt ataagtctaa 5160

atggagaagg caaaacttta aatcttttag aaaataatat agaagcatgc agaccagcct 5220
ggccaacatg atgaaaccct ctctactaat aataaaatca gtagaactac tcaggactac 5280
tttgagtggg aagtcctttt ctatgaagac ttctttggcc aaaattaggc tctaaatgca 5340
aggagatagt gcatcatgcc tggctgcact tactgataaa tgatgttatc accatcttta 5400
accaaatgca caggaacaag ttatggtact gatgtgctgg attgagaagg agctctactt 5460
ccttgacagg acacatttgt atcaacttaa aaaagcagat ttttgccagc agaactattc 5520
attcagaggt aggaaactta gaatagatga tgtcactgat tagcatggct tccccatctc 5580

cacagctgct tcccacccag gttgcccaca gttgagtttg tccagtgctc agggctgccc 5640
actctcagta agaagcccca caccagcccc tctccaaata tgttggctgt tccttccatt 5700
aaagtgaccc cactttagag cagcaagtgg atttctgttt cttacagttc aggaaggagg 5760
agtcagctgt gagaacctgg agcctgagat gcttctaagt cccactgcta ctggggtcag 5820
ggaagccaga ctccagcatc agcagtcagg agcactaagc ccttgccaac atcctgtttc 5880
tcagagaaac tgcttccatt ataatggttg tccttttta agctatcaag ccaaacaacc 5940
agtgtctacc attattctca tcacctgaag ccaagggttc tagcaaaagt caagctgtct 6000

tgtaatggtt gatgtgcctc cagcttctgt cttcagtcac tccactctta gcctgctctg 6060
aatcaactct gaccacagtt ccctggagcc cctgccacct gctgcccctg ccaccttctc 6120
catctgcagt gctgtgcagc cttctgcact cttgcagagc taataggtgg agacttgaag 6180
gaagaggagg aaagtttctc ataatagcct tgctgcaagc tcaaatggga ggtgggcact 6240
gtgcccagga gccttggagc aaaggctgtg cccaacctct gactgcatcc aggtttggtc 6300
ttgacagaga taagaagccc tggcttttgg agccaaaatc taggtcagac ttaggcagga 6360
ttctcaaagt ttatcagcag aacatgaggc agaagaccct ttctgctcca gcttcttcag 6420

gctcaacctt catcagaata gatagaaaga gaggctgtga gggttcttaa aacagaagca 6480
aatctgactc agagaataaa caacctccta gtaaactaca gcttagacag agcatctggt 6540
ggtgagtgtg ctcagtgtcc tactcaactg tctggtatca gccctcatga ggacttctct 6600
tctttccctc atagacctcc atctctgttt tccttagcct gcagaaatct ggatggctat 6660
tcacagaatg cctgtgcttt cagagttgca ttttttctct ggtattctgg ttcaagcatt 6720
tgaaggtagg aaaggttctc caagtgcaag aaagccagcc ctgagcctca actgcctggc 6780

tagtgtggtc agtaggatgc aaaggctgtt gaatgccaca aggccaaact ttaacctgtg 6840

taccacaagc ctagcagcag aggcagctct gctcactgga actctctgtc ttctttctcc 6900
tgagcctttt cttttcctga gttttctagc tctcctcaac cttacctctg ccctacccag 6960
gacaaaccca agagccactg tttctgtgat gtcctctcca gccctaatta ggcatcatga 7020
cttcagcctg accttccatg ctcagaagca gtgctaatcc acttcagatg agctgctcta 7080
tgcaacacag gcagagccta caaacctttg caccagagcc ctccacatat cagtgtttgt 7140
tcatactcac ttcaacagca aatgtgactg ctgagattaa gattttacac aagatggtct 7200
gtaatttcac agttagtttt atcccattag gtatgaaaga attagcataa ttccccttaa 7260

acatgaatga atcttagatt ttttaataaa tagttttgga agtaaagaca gagacatcag 7320
gagcacaagg aatagcctga gaggacaaac agaacaagaa agagtctgga aatacacagg 7380
atgttcttgg cctcctcaaa gcaagtgcaa gcagatagta ccagcagccc caggctatca 7440
gagcccagtg aagagaagta ccatgaaagc cacagctcta accaccctgt tccagagtga 7500
cagacagtcc ccaagacaag ccagcctgag ccagagagag aactgcaaga gaaagtttct 7560
aatttaggtt ctgttagatt cagacaagtg caggtcatcc tctctccaca gctactcacc 7620
tctccagcct aacaaagcct gcagtccaca ctccaaccct ggtgtctcac ctcctagcct 7680

ctcccaacat cctgctctct gaccatcttc tgcatctctc atctcaccat ctcccactgt 7740
ctacagccta ctcttgcaac taccatctca ttttctgaca tcctgtctac atcttctgcc 7800
atactctgcc atctaccata ccacctctta ccatctacca caccatcttt tatctccatc 7860
cctctcagaa gcctccaagc tgaatcctgc tttatgtgtt catctcagcc cctgcatgga 7920
aagctgaccc cagaggcaga actattccca gagagcttgg ccaagaaaaa caaaactacc 7980
agcctggcca ggctcaggag tagtaagctg cagtgtctgt tgtgttctag cttcaacagc 8040
tgcaggagtt ccactctcaa atgctccaca tttctcacat cctcctgatt ctggtcacta 8100

cccatcttca aagaacagaa tatctcacat cagcatactg tgaaggacta gtcatgggtg 8160
cagctgctca gagctgcaaa gtcattctgg atggtggaga gcttacaaac atttcatgat 8220
gctccccccg ctctgatggc tggagcccaa tccctacaca gactcctgct gtatgtgttt 8280
tcctttcact ctgagccaca gccagagggc aggcattcag tctcctcttc aggctggggc 8340
tggggcactg agaactcacc caacaccttg ctctcactcc ttctgcaaaa caagaaagag 8400
ctttgtgctg cagtagccat gaagaatgaa aggaaggctt taactaaaaa atgtcagaga 8460
ttattttcaa ccccttactg tggatcacca gcaaggagga aacacaacac agagacattt 8520

tttcccctca aattatcaaa agaatcactg catttgttaa agagagcaac tgaatcagga 8580
agcagagttt tgaacatatc agaagttagg aatctgcatc agagacaaat gcagtcatgg 8640
ttgtttgctg cataccagcc ctaatcatta gaagcctcat ggacttcaaa catcattccc 8700
tctgacaaga tgctctagcc taactccatg agataaaata aatctgcctt tcagagccaa 8760
agaagagtcc accagcttct tctcagtgtg aacaagagct ccagtcaggt tagtcagtcc 8820
agtgcagtag aggagaccag tctgcatcct ctaattttca aaggcaagaa gatttgttta 8880
ccctggacac caggcacaag tgaggtcaca gagctcttag atatgcagtc ctcatgagtg 8940

aggagactaa agcgcatgcc atcaagactt cagtgtagag aaaacctcca aaaaagcctc 9000
ctcactactt ctggaatagc tcagaggccg aggcggcctc ggcctctgca taaataaaaa 9060
aaattagtca gccatggggc ggagaatggg cggaactggg cggagttagg ggcgggatgg 9120
gcggagttag gggcgggact atggttgctg actaattgag atgcatgctt tgcatacttc 9180
tgcctgctgg ggagcctggg gactttccac acctggttgc tgactaattg agatgcatgc 9240
tttgcatact tctgcctgct ggggagcctg gggactttcc acaccctaac tgacacacat 9300
tccacagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc gtattgggcg 9360

ctcttccgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt 9420
atcagctcac tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa 9480
gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc 9540
gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag 9600
gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt 9660
gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg 9720
aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt aggtcgttcg 9780

ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg 9840
taactatcgt cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac 9900
tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg 9960
gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc tgaagccagt 10020
taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg 10080
tggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc 10140
tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt 10200

ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt 10260
taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag 10320
tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct gactcctgca 10380

aaccacgttg tgtctcaaaa tctctgatgt tacattgcac aagataaaaa tatatcatca 10440

tgaacaataa aactgtctgc ttacataaac agtaatacaa ggggtgttat gagccatatt 10500

caacgggaaa cgtcttgctc gaggccgcga ttaaattcca acatggatgc tgatttatat 10560

gggtataaat gggctcgcga taatgtcggg caatcaggtg cgacaatcta tcgattgtat 10620

gggaagcccg atgcgccaga gttgtttctg aaacatggca aaggtagcgt tgccaatgat 10680

gttacagatg agatggtcag actaaactgg ctgacggaat ttatgcctct tccgaccatc 10740

aagcatttta tccgtactcc tgatgatgca tggttactca ccactgcgat ccccgggaaa 10800

acagcattcc aggtattaga agaatatcct gattcaggtg aaaatattgt tgatgcgctg 10860

gcagtgttcc tgcgccggtt gcattcgatt cctgtttgta attgtccttt taacagcgat 10920

cgcgtatttc gtctcgctca ggcgcaatca cgaatgaata acggtttggt tgatgcgagt 10980

gattttgatg acgagcgtaa tggctggcct gttgaacaag tctggaaaga aatgcataag 11040

cttttgccat tctcaccgga ttcagtcgtc actcatggtg atttctcact tgataacctt 11100

atttttgacg aggggaaatt aataggttgt attgatgttg gacgagtcgg aatcgcagac 11160

cgataccagg atcttgccat cctatggaac tgcctcggtg agttttctcc ttcattacag 11220

aaacggcttt ttcaaaaata tggtattgat aatcctgata tgaataaatt gcagtttcat 11280

ttgatgctcg atgagttttt ctaagggcgg cctgccacca tacccacgcc gaaacaagcg 11340

ctcatgagcc cgaagtggcg agcccgatct tccccatcgg tgatgtcggc gatataggcg 11400

ccagcaaccg cacctgtggc gccggtgatg agggcgcgcc aagtcgacgt ccggcagtc 11459

<210> 63

<211> 274

<212> PRT

<213> Artificial sequence

<220><223> Synthetic polypeptide

<400> 63

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

145 150 155 160

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 Thr Tyr Phe Gly His Ser

225 230 235 240

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> 64

<211> 825

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 64

atgggcaaga gcctgagcca cctgcccctg cacagcagca aggaggacgc ctacgacggc 60

gtgaccagcg agaacatgcg caacggcctg gtgaacagcg aggtgcacaa cgaggacggc 120

cgcaacggcg acgtgagcca gttcccctac gtggagttca ccggccgcga cagcgtgacc 180

tgccccacct gccagggcac cggccgcatc ccccgcggcc aggagaacca gctggtggcc 240

ctgatcccct acagcgacca gcgcctgcgc ccccgccgca ccaagctgta cgtgatggcc 300

agcgtgttcg tgtgcctgct gctgagcggc ctggccgtgt tcttcctgtt cccccgcagc 360

atcgacgtga agtacatcgg cgtgaagagc gcctacgtga gctacgacgt gcagaagcgc 420

accatctacc tgaacatcac caacaccctg aacatcacca acaacaacta ctacagcgtg 480

gaggtggaga acatcaccgc ccaggtgcag ttcagcaaga ccgtgatcgg caaggcccgc 540

ctgaacaaca tcaccatcat cggccccctg gacatgaagc agatcgacta caccgtgccc 600

accgtgatcg ccgaggagat gagctacatg tacgacttct gcaccctgat cagcatcaag 660

gtgcacaaca tcgtgctgat gatgcaggtg accgtgacca ccacctactt cggccacagc 720

gagcagatca gccaggagcg ctaccagtac gtggactgcg gccgcaacac cacctaccag 780

ctgggccaga gcgagtacct gaacgtgctg cagccccagc agtaa 825

<210> 65

<211> 267

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 65

gtgatatcac aaggtcccag ggctggggtc agaaattctc tcccgaggga atgaagccac 60

aggagccaag agcaggagga ccaaggccct ggcgaaggcc gtggcctcgt tcaagtaaaa 120

gatcctagta cagtgcaggt cccaatgtgt actaggatct tttacttgaa cggggacgcc 180

ggcatccggg ctcaggaccc ccctctctgc cagaggcacc aacaccagag ttcacaaatc 240

agtctcctgc cctttgcatg tagcaaa 267

<210> 66

<211> 267

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 66

tttgctacat gcaaagggca ggagactgat ttgtgaactc tggtgttggt gcctctggca 60

gagagggggg tcctgagccc ggatgccggc gtccccgttc aagtaaaaga tcctagtaca 120

cattgggacc tgcactgtac taggatcttt tacttgaacg aggccacggc cttcgccagg 180

gccttggtcc tcctgctctt ggctcctgtg gcttcattcc ctcgggagag aatttctgac 240

cccagccctg ggaccttgtg atatcac 267

<210> 67

<211> 593

<212> PRT

<213> Artificial sequence

<220><223> Synthetic polypeptide

<400> 67

Met Trp Thr Leu Val Ser Trp Val Ala Leu Thr Ala Gly Leu Val Ala

1 5 10 15

Gly Thr Arg Cys Pro Asp Gly Gln Phe Cys Pro Val Ala Cys Cys Leu

20 25 30

Asp Pro Gly Gly Ala Ser Tyr Ser Cys Cys Arg Pro Leu Leu Asp Lys

35 40 45

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

50 55 60

Ala His Cys Ser Ala Gly His Ser Cys Ile Phe Thr Val Ser Gly Thr

65 70 75 80

Ser Ser Cys Cys Pro Phe Pro Glu Ala Val Ala Cys Gly Asp Gly His

85 90 95

His Cys Cys Pro Arg Gly Phe His Cys Ser Ala Asp Gly Arg Ser Cys

100 105 110

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

115 120 125

Ser Gln Phe Glu Cys Pro Asp Phe Ser Thr Cys Cys Val Met Val Asp

130 135 140

Gly Ser Trp Gly Cys Cys Pro Met Pro Gln Ala Ser Cys Cys Glu Asp

145 150 155 160

Arg Val His Cys Cys Pro His Gly Ala Phe Cys Asp Leu Val His Thr

165 170 175

Arg Cys Ile Thr Pro Thr Gly Thr His Pro Leu Ala Lys Lys Leu Pro

180 185 190

Ala Gln Arg Thr Asn Arg Ala Val Ala Leu Ser Ser Ser Val Met Cys

195 200 205

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

210 215 220

Pro Ser Gly Lys Tyr Gly Cys Cys Pro Met Pro Asn Ala Thr Cys Cys

225 230 235 240

Ser Asp His Leu His Cys Cys Pro Gln Asp Thr Val Cys Asp Leu Ile

245 250 255

Gln Ser Lys Cys Leu Ser Lys Glu Asn Ala Thr Thr Asp Leu Leu Thr

260 265 270

Lys Leu Pro Ala His Thr Val Gly Asp Val Lys Cys Asp Met Glu Val

275 280 285

Ser Cys Pro Asp Gly Tyr Thr Cys Cys Arg Leu Gln Ser Gly Ala Trp

290 295 300

Gly Cys Cys Pro Phe Thr Gln Ala Val Cys Cys Glu Asp His Ile His

305 310 315 320

Cys Cys Pro Ala Gly Phe Thr Cys Asp Thr Gln Lys Gly Thr Cys Glu

325 330 335

Gln Gly Pro His Gln Val Pro Trp Met Glu Lys Ala Pro Ala His Leu

340 345 350

Ser Leu Pro Asp Pro Gln Ala Leu Lys Arg Asp Val Pro Cys Asp Asn

355 360 365

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

370 375 380

Glu Trp Gly Cys Cys Pro Ile Pro Glu Ala Val Cys Cys Ser Asp His

385 390 395 400

Gln His Cys Cys Pro Gln Gly Tyr Thr Cys Val Ala Glu Gly Gln Cys

405 410 415

Gln Arg Gly Ser Glu Ile Val Ala Gly Leu Glu Lys Met Pro Ala Arg

420 425 430

Arg Ala Ser Leu Ser His Pro Arg Asp Ile Gly Cys Asp Gln His Thr

435 440 445

Ser Cys Pro Val Gly Gln Thr Cys Cys Pro Ser Leu Gly Gly Ser Trp

450 455 460

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

465 470 475 480

Cys Cys Pro Ala Gly Tyr Thr Cys Asn Val Lys Ala Arg Ser Cys Glu

485 490 495

Lys Glu Val Val Ser Ala Gln Pro Ala Thr Phe Leu Ala Arg Ser Pro

500 505 510

His Val Gly Val Lys Asp Val Glu Cys Gly Glu Gly His Phe Cys His

515 520 525

Asp Asn Gln Thr Cys Cys Arg Asp Asn Arg Gln Gly Trp Ala Cys Cys

530 535 540

Pro Tyr Arg Gln Gly Val Cys Cys Ala Asp Arg Arg His Cys Cys Pro

545 550 555 560

Ala Gly Phe Arg Cys Ala Ala Arg Gly Thr Lys Cys Leu Arg Arg Glu

565 570 575

Ala Pro Arg Trp Asp Ala Pro Leu Arg Asp Pro Ala Leu Arg Gln Leu

580 585 590

Leu

<210> 68

<211> 1779

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 68

atgtggaccc tggtgagctg ggtggccctg accgccggcc tggtggccgg cacccgctgc 60

cccgacggcc agttctgccc cgtggcctgc tgcctggacc ccggcggcgc cagctacagc 120

tgctgccgcc ccctgctgga caagtggccc accaccctga gccgccacct gggcggcccc 180

tgccaggtgg acgcccactg cagcgccggc cacagctgca tcttcaccgt gagcggcacc 240

agcagctgct gccccttccc cgaggccgtg gcctgcggcg acggccacca ctgctgcccc 300

cgcggcttcc actgcagcgc cgacggccgc agctgcttcc agcgcagcgg caacaacagc 360

gtgggcgcca tccagtgccc cgacagccag ttcgagtgcc ccgacttcag cacctgctgc 420

gtgatggtgg acggcagctg gggctgctgc cccatgcccc aggccagctg ctgcgaggac 480

cgcgtgcact gctgccccca cggcgccttc tgcgacctgg tgcacacccg ctgcatcacc 540

cccaccggca cccacccccct ggccaagaag ctgcccgccc agcgcaccaa ccgcgccgtg 600

gccctgagca gcagcgtgat gtgccccgac gcccgcagcc gctgccccga cggcagcacc 660

tgctgcgagc tgcccagcgg caagtacggc tgctgcccca tgcccaacgc cacctgctgc 720

agcgaccacc tgcactgctg cccccaggac accgtgtgcg acctgatcca gagcaagtgc 780

ctgagcaagg agaacgccac caccgacctg ctgaccaagc tgcccgccca caccgtgggc 840

gacgtgaagt gcgacatgga ggtgagctgc cccgacggct acacctgctg ccgcctgcag 900

agcggcgcct ggggctgctg ccccttcacc caggccgtgt gctgcgagga ccacatccac 960

tgctgccccg ccggcttcac ctgcgacacc cagaagggca cctgcgagca gggcccccac 1020

caggtgccct ggatggagaa ggcccccgcc cacctgagcc tgcccgaccc ccaggccctg 1080

aagcgcgacg tgccctgcga caacgtgagc agctgcccca gcagcgacac ctgctgccag 1140

ctgaccagcg gcgagtgggg ctgctgcccc atccccgagg ccgtgtgctg cagcgaccac 1200

cagcactgct gcccccaggg ctacacctgc gtggccgagg gccagtgcca gcgcggcagc 1260

gagatcgtgg ccggcctgga gaagatgccc gcccgccgcg ccagcctgag ccacccccgc 1320

gacatcggct gcgaccagca caccagctgc cccgtgggcc agacctgctg ccccagcctg 1380

ggcggcagct gggcctgctg ccagctgccc cacgccgtgt gctgcgagga ccgccagcac 1440

tgctgccccg ccggctacac ctgcaacgtg aaggcccgca gctgcgagaa ggaggtggtg 1500

agcgcccagc ccgccacctt cctggcccgc agcccccacg tgggcgtgaa ggacgtggag 1560

tgcggcgagg gccacttctg ccacgacaac cagacctgct gccgcgacaa ccgccagggc 1620

tgggcctgct gcccctaccg ccagggcgtg tgctgcgccg accgccgcca ctgctgcccc 1680

gccggcttcc gctgcgccgc ccgcggcacc aagtgcctgc gccgcgaggc cccccgctgg 1740

gacgcccccc tgcgcgaccc cgccctgcgc cagctgctg 1779

<210> 69

<211> 10871

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 69

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300

ctagttatta atagtaatca attacggggt cattagttca tagcccatat atggagttcc 360

gcgttacata acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat 420

tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc 480

aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc 540

caagtacgcc ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt 600

acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta 660

ccatggtcga ggtgagcccc acgttctgct tcactctccc catctccccc ccctccccac 720

ccccaatttt gtatttattt atttttaat tattttgtgc agcgatgggg gcgggggggg 780

gggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga 840

gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa gtttcctttt atggcgaggc 900

ggcggcggcg gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgacgc 960

tgccttcgcc ccgtgccccg ctccgccgcc gcctcgcgcc gcccgccccg gctctgactg 1020

accgcgttac tcccacaggt gagcgggcgg gacggccctt ctcctccggg ctgtaattag 1080

cgcttggttt aatgacggct tgtttctttt ctgtggctgc gtgaaagcct tgaggggctc 1140

cgggagctag agcctctgct aaccatgttc atgccttctt ctttttccta cagctcctgg 1200

gcaacgtgct ggttattgtg ctgtctcatc attttggcaa agaattcctc gaagatccga 1260

agggaaagtc ttccacgact gtgggatccg ttcgaagata tcaccggttg agccaccatg 1320

tggaccctgg tgagctgggt ggccctgacc gccggcctgg tggccggcac ccgctgcccc 1380

gacggccagt tctgccccgt ggcctgctgc ctggaccccg gcggcgccag ctacagctgc 1440

tgccgccccc tgctggacaa gtggcccacc accctgagcc gccacctggg cggcccctgc 1500

caggtggacg cccactgcag cgccggccac agctgcatct tcaccgtgag cggcaccagc 1560

agctgctgcc ccttccccga ggccgtggcc tgcggcgacg gccaccactg ctgcccccgc 1620

ggcttccact gcagcgccga cggccgcagc tgcttccagc gcagcggcaa caacagcgtg 1680

ggcgccatcc agtgccccga cagccagttc gagtgccccg acttcagcac ctgctgcgtg 1740

atggtggacg gcagctgggg ctgctgcccc atgccccagg ccagctgctg cgaggaccgc 1800

gtgcactgct gcccccacgg cgccttctgc gacctggtgc acacccgctg catcaccccc 1860

accggcaccc acccccctggc caagaagctg cccgcccagc gcaccaaccg cgccgtggcc 1920

ctgagcagca gcgtgatgtg ccccgacgcc cgcagccgct gccccgacgg cagcacctgc 1980

tgcgagctgc ccagcggcaa gtacggctgc tgccccatgc ccaacgccac ctgctgcagc 2040

gaccacctgc actgctgccc ccaggacacc gtgtgcgacc tgatccagag caagtgcctg 2100

agcaaggaga acgccaccac cgacctgctg accaagctgc ccgcccacac cgtgggcgac 2160

gtgaagtgcg acatggaggt gagctgcccc gacggctaca cctgctgccg cctgcagagc 2220

ggcgcctggg gctgctgccc cttcacccag gccgtgtgct gcgaggacca catccactgc 2280

tgccccgccg gcttcacctg cgacacccag aagggcacct gcgagcaggg cccccaccag 2340

gtgccctgga tggagaaggc ccccgcccac ctgagcctgc ccgacccccca ggccctgaag 2400

cgcgacgtgc cctgcgacaa cgtgagcagc tgccccagca gcgacacctg ctgccagctg 2460

accagcggcg agtggggctg ctgccccatc cccgaggccg tgtgctgcag cgaccaccag 2520

cactgctgcc cccagggcta cacctgcgtg gccgagggcc agtgccagcg cggcagcgag 2580

atcgtggccg gcctggagaa gatgcccgcc cgccgcgcca gcctgagcca cccccgcgac 2640

atcggctgcg accagcacac cagctgcccc gtgggccaga cctgctgccc cagcctgggc 2700

ggcagctggg cctgctgcca gctgccccac gccgtgtgct gcgaggaccg ccagcactgc 2760

tgccccgccg gctacacctg caacgtgaag gcccgcagct gcgagaagga ggtggtgagc 2820

gcccagcccg ccaccttcct ggcccgcagc ccccacgtgg gcgtgaagga cgtggagtgc 2880

ggcgagggcc acttctgcca cgacaaccag acctgctgcc gcgacaaccg ccagggctgg 2940

gcctgctgcc cctaccgcca gggcgtgtgc tgcgccgacc gccgccactg ctgccccgcc 3000

ggcttccgct gcgccgcccg cggcaccaag tgcctgcgcc gcgaggcccc ccgctgggac 3060

gcccccctgc gcgaccccgc cctgcgccag ctgctgtgac aattgttaat taagtttaaa 3120

ccctcgaggc cgcaagctta tcgataatca acctctggat tacaaaattt gtgaaagatt 3180

gactggtatt cttaactatg ttgctccttt tacgctatgt ggatacgctg ctttaatgcc 3240

tttgtatcat gctattgctt cccgtatggc tttcattttc tcctccttgt ataaatcctg 3300

gttgctgtct ctttatgagg agttgtggcc cgttgtcagg caacgtggcg tggtgtgcac 3360
tgtgtttgct gacgcaaccc ccactggttg gggcattgcc accacctgtc agctcctttc 3420
cgggactttc gctttccccc tccctattgc cacggcggaa ctcatcgccg cctgccttgc 3480
ccgctgctgg acaggggctc ggctgttggg cactgacaat tccgtggtgt tgtcggggaa 3540

atcatcgtcc tttccttggc tgctcgcctg tgttgccacc tggattctgc gcgggacgtc 3600
cttctgctac gtcccttcgg ccctcaatcc agcggacctt ccttcccgcg gcctgctgcc 3660
ggctctgcgg cctcttccgc gtcttcgcct tcgccctcag acgagtcgga tctccctttg 3720
ggccgcctcc ccgcatcgat accgtcgact agagctcgct gatcagcctc gactgtgcct 3780
tctagttgcc agccatctgt tgtttgcccc tcccccgtgc cttccttgac cctggaaggt 3840
gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgcattg tctgagtagg 3900
tgtcattcta ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac 3960

aatagcaggc atgctgggga gagatccacg ataacaaaca gctttttgg ggtgaacata 4020
ttgactgaat tccctgcagg ttggccactc cctctctgcg cgctcgctcg ctcactgagg 4080
ccgcccgggc aaagcccggg cgtcgggcga cctttggtcg cccggcctca gtgagcgagc 4140
gagcgcgcag agagggagtg gccaactcca tcactagggg ttcctgcggc cgctcgtacg 4200
gtctcgagga attcctgcag gataacttgc caacctcatt ctaaaatgta tatagaagcc 4260
caaaagacaa taacaaaaat attcttgtag aacaaaatgg gaaagaatgt tccactaaat 4320
atcaagattt agagcaaagc atgagatgtg tggggataga cagtgaggct gataaaatag 4380

agtagagctc agaaacagac ccattgatat atgtaagtga cctatgaaaa aaatatggca 4440
ttttacaatg ggaaaatgat ggtctttttc ttttttagaa aaacagggaa atatatttat 4500
atgtaaaaaa taaaagggaa cccatatgtc ataccataca cacaaaaaaa ttccagtgaa 4560
ttataagtct aaatggagaa ggcaaaactt taaatctttt agaaaataat atagaagcat 4620
gcagaccagc ctggccaaca tgatgaaacc ctctctacta ataataaaat cagtagaact 4680
actcaggact actttgagtg ggaagtcctt ttctatgaag acttctttgg ccaaaattag 4740
gctctaaatg caaggagata gtgcatcatg cctggctgca cttactgata aatgatgtta 4800

tcaccatctt taaccaaatg cacaggaaca agttatggta ctgatgtgct ggattgagaa 4860
ggagctctac ttccttgaca ggacacattt gtatcaactt aaaaaagcag atttttgcca 4920
gcagaactat tcattcagag gtaggaaact tagaatagat gatgtcactg attagcatgg 4980
cttccccatc tccacagctg cttcccaccc aggttgccca cagttgagtt tgtccagtgc 5040
tcagggctgc ccactctcag taagaagccc cacaccagcc cctctccaaa tatgttggct 5100
gttccttcca ttaaagtgac cccactttag agcagcaagt ggatttctgt ttcttacagt 5160

tcaggaagga ggagtcagct gtgagaacct ggagcctgag atgcttctaa gtcccactgc 5220

tactggggtc agggaagcca gactccagca tcagcagtca ggagcactaa gcccttgcca 5280
acatcctgtt tctcagagaa actgcttcca ttataatggt tgtccttttt taagctatca 5340
agccaaacaa ccagtgtcta ccattattct catcacctga agccaagggt tctagcaaaa 5400
gtcaagctgt cttgtaatgg ttgatgtgcc tccagcttct gtcttcagtc actccactct 5460
tagcctgctc tgaatcaact ctgaccacag ttccctggag cccctgccac ctgctgcccc 5520
tgccaccttc tccatctgca gtgctgtgca gccttctgca ctcttgcaga gctaataggt 5580
ggagacttga aggaagagga ggaaagtttc tcataatagc cttgctgcaa gctcaaatgg 5640

gaggtgggca ctgtgcccag gagccttgga gcaaaggctg tgcccaacct ctgactgcat 5700
ccaggtttgg tcttgacaga gataagaagc cctggctttt ggagccaaaa tctaggtcag 5760
acttaggcag gattctcaaa gtttatcagc agaacatgag gcagaagacc ctttctgctc 5820
cagcttcttc aggctcaacc ttcatcagaa tagatagaaa gagaggctgt gagggttctt 5880
aaaacagaag caaatctgac tcagagaata aacaacctcc tagtaaacta cagcttagac 5940
agagcatctg gtggtgagtg tgctcagtgt cctactcaac tgtctggtat cagccctcat 6000
gaggacttct cttctttccc tcatagacct ccatctctgt tttccttagc ctgcagaaat 6060

ctggatggct attcacagaa tgcctgtgct ttcagagttg catttttct ctggtattct 6120
ggttcaagca tttgaaggta ggaaaggttc tccaagtgca agaaagccag ccctgagcct 6180
caactgcctg gctagtgtgg tcagtaggat gcaaaggctg ttgaatgcca caaggccaaa 6240
ctttaacctg tgtaccacaa gcctagcagc agaggcagct ctgctcactg gaactctctg 6300
tcttctttct cctgagcctt ttcttttcct gagttttcta gctctcctca accttacctc 6360
tgccctaccc aggacaaacc caagagccac tgtttctgtg atgtcctctc cagccctaat 6420
taggcatcat gacttcagcc tgaccttcca tgctcagaag cagtgctaat ccacttcaga 6480

tgagctgctc tatgcaacac aggcagagcc tacaaacctt tgcaccagag ccctccacat 6540
atcagtgttt gttcatactc acttcaacag caaatgtgac tgctgagatt aagattttac 6600
acaagatggt ctgtaatttc acagttagtt ttatcccatt aggtatgaaa gaattagcat 6660
aattcccctt aaacatgaat gaatcttaga ttttttaata aatagttttg gaagtaaaga 6720
cagagacatc aggagcacaa ggaatagcct gagaggacaa acagaacaag aaagagtctg 6780
gaaatacaca ggatgttctt ggcctcctca aagcaagtgc aagcagatag taccagcagc 6840
cccaggctat cagagcccag tgaagagaag taccatgaaa gccacagctc taaccaccct 6900

gttccagagt gacagacagt ccccaagaca agccagcctg agccagagag agaactgcaa 6960
gagaaagttt ctaatttagg ttctgttaga ttcagacaag tgcaggtcat cctctctcca 7020
cagctactca cctctccagc ctaacaaagc ctgcagtcca cactccaacc ctggtgtctc 7080
acctcctagc ctctcccaac atcctgctct ctgaccatct tctgcatctc tcatctcacc 7140
atctcccact gtctacagcc tactcttgca actaccatct cattttctga catcctgtct 7200
acatcttctg ccatactctg ccatctacca taccacctct taccatctac cacaccatct 7260
tttatctcca tccctctcag aagcctccaa gctgaatcct gctttatgtg ttcatctcag 7320

cccctgcatg gaaagctgac cccagaggca gaactattcc cagagagctt ggccaagaaa 7380
aacaaaacta ccagcctggc caggctcagg agtagtaagc tgcagtgtct gttgtgttct 7440
agcttcaaca gctgcaggag ttccactctc aaatgctcca catttctcac atcctcctga 7500
ttctggtcac tacccatctt caaagaacag aatatctcac atcagcatac tgtgaaggac 7560
tagtcatggg tgcagctgct cagagctgca aagtcattct ggatggtgga gagcttacaa 7620
acatttcatg atgctccccc cgctctgatg gctggagccc aatccctaca cagactcctg 7680
ctgtatgtgt tttcctttca ctctgagcca cagccagagg gcaggcattc agtctcctct 7740

tcaggctggg gctggggcac tgagaactca cccaacacct tgctctcact ccttctgcaa 7800
aacaagaaag agctttgtgc tgcagtagcc atgaagaatg aaaggaaggc tttaactaaa 7860
aaatgtcaga gattattttc aacccccttac tgtggatcac cagcaaggag gaaacacaac 7920
acagagacat tttttcccct caaattatca aaagaatcac tgcatttgtt aaagagagca 7980
actgaatcag gaagcagagt tttgaacata tcagaagtta ggaatctgca tcagagacaa 8040
atgcagtcat ggttgtttgc tgcataccag ccctaatcat tagaagcctc atggacttca 8100
aacatcattc cctctgacaa gatgctctag cctaactcca tgagataaaa taaatctgcc 8160

tttcagagcc aaagaagagt ccaccagctt cttctcagtg tgaacaagag ctccagtcag 8220
gttagtcagt ccagtgcagt agaggagacc agtctgcatc ctctaatttt caaaggcaag 8280
aagatttgtt taccctggac accaggcaca agtgaggtca cagagctctt agatatgcag 8340
tcctcatgag tgaggagact aaagcgcatg ccatcaagac ttcagtgtag agaaaacctc 8400
caaaaaagcc tcctcactac ttctggaata gctcagaggc cgaggcggcc tcggcctctg 8460
cataaataaa aaaaattagt cagccatggg gcggagaatg ggcggaactg ggcggagtta 8520
ggggcgggat gggcggagtt aggggcggga ctatggttgc tgactaattg agatgcatgc 8580

tttgcatact tctgcctgct ggggagcctg gggactttcc acacctggtt gctgactaat 8640
tgagatgcat gctttgcata cttctgcctg ctggggagcc tggggacttt ccacacccta 8700
actgacacac attccacagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt 8760

gcgtattggg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct 8820
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 8880
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 8940
cgcgttgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg 9000

ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 9060
aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 9120
tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt 9180
gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg 9240
cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact 9300
ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt 9360
cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggta tctgcgctct 9420

gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 9480
cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 9540
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg 9600
ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 9660
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca 9720
atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc 9780
ctgactcctg caaaccacgt tgtgtctcaa aatctctgat gttacattgc acaagataaa 9840

aatatatcat catgaacaat aaaactgtct gcttacataa acagtaatac aaggggtgtt 9900
atgagccata ttcaacggga aacgtcttgc tcgaggccgc gattaaattc caacatggat 9960
gctgatttat atgggtataa atgggctcgc gataatgtcg ggcaatcagg tgcgacaatc 10020
tatcgattgt atgggaagcc cgatgcgcca gagttgtttc tgaaacatgg caaaggtagc 10080
gttgccaatg atgttacaga tgagatggtc agactaaact ggctgacgga atttatgcct 10140
cttccgacca tcaagcattt tatccgtact cctgatgatg catggttact caccactgcg 10200
atccccggga aaacagcatt ccaggtatta gaagaatatc ctgattcagg tgaaaatatt 10260

gttgatgcgc tggcagtgtt cctgcgccgg ttgcattcga ttcctgtttg taattgtcct 10320
tttaacagcg atcgcgtatt tcgtctcgct caggcgcaat cacgaatgaa taacggtttg 10380
gttgatgcga gtgattttga tgacgagcgt aatggctggc ctgttgaaca agtctggaaa 10440
gaaatgcata agcttttgcc attctcaccg gattcagtcg tcactcatgg tgatttctca 10500
cttgataacc ttatttttga cgaggggaaa ttaataggtt gtattgatgt tggacgagtc 10560
ggaatcgcag accgatacca ggatcttgcc atcctatgga actgcctcgg tgagttttct 10620

ccttcattac agaaacggct ttttcaaaaa tatggtattg ataatcctga tatgaataaa 10680

ttgcagtttc atttgatgct cgatgagttt ttctaagggc ggcctgccac catacccacg 10740

ccgaaacaag cgctcatgag cccgaagtgg cgagcccgat cttccccatc ggtgatgtcg 10800

gcgatatagg cgccagcaac cgcacctgtg gcgccggtga tgagggcgcg ccaagtcgac 10860

gtccggcagt c 10871

<210> 70

<211> 4151

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 70

gggaggttac gcgttcgtcg actactagtg ggtaccagag cgggcggagt tagggcggag 60

ccaatcagcg tgcgccgttc cgaaagttgc cttttatggc tgggcggaga atgggcggtg 120

aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg tcgcagccgg 180

gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcacttggta agtcactgac 240

tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag tggcactatg 300

aaccctgcag ccctaggaat gcatctagac aattgtacta accttcttct ctttcctctc 360

ctgacagtcc ggaaagccac catgtacgcc ctgttcctgc tggccagcct gctgggcgcc 420

gccctggccg gccccgtgct gggcctgaag gagtgcaccc gcggcagcgc cgtgtggtgc 480

cagaacgtga agaccgccag cgactgcggc gccgtgaagc actgcctgca gaccgtgtgg 540

aacaagccca ccgtgaagag cctgccctgc gacatctgca aggacgtggt gaccgccgcc 600

ggcgacatgc tgaaggacaa cgccaccgag gaggagatcc tggtgtacct ggagaagacc 660

tgcgactggc tgcccaagcc caacatgagc gccagctgca aggagatcgt ggacagctac 720

ctgcccgtga tcctggacat catcaagggc gagatgagcc gccccggcga ggtgtgcagc 780

gccctgaacc tgtgcgagag cctgcagaag cacctggccg agctgaacca ccagaagcag 840

ctggagagca acaagatccc cgagctggac atgaccgagg tggtggcccc cttcatggcc 900

aacatccccc tgctgctgta cccccaggac ggcccccgca gcaagcccca gcccaaggac 960

aacggcgacg tgtgccagga ctgcatccag atggtgaccg acatccagac cgccgtgcgc 1020

accaacagca ccttcgtgca ggccctggtg gagcacgtga aggaggagtg cgaccgcctg 1080

ggccccggca tggccgacat ctgcaagaac tacatcagcc agtacagcga gatcgccatc 1140

cagatgatga tgcacatgca gcccaaggag atctgcgccc tggtgggctt ctgcgacgag 1200

gtgaaggaga tgcccatgca gaccctggtg cccgccaagg tggccagcaa gaacgtgatc 1260
cccgccctgg agctggtgga gcccatcaag aagcacgagg tgcccgccaa gagcgacgtg 1320
tactgcgagg tgtgcgagtt cctggtgaag gaggtgacca agctgatcga caacaacaag 1380

accgagaagg agatcctgga cgccttcgac aagatgtgca gcaagctgcc caagagcctg 1440
agcgaggagt gccaggaggt ggtggacacc tacggcagca gcatcctgag catcctgctg 1500
gaggaggtga gccccgagct ggtgtgcagc atgctgcacc tgtgcagcgg cacccgcctg 1560
cccgccctga ccgtgcacgt gacccagccc aaggacggcg gcttctgcga ggtgtgcaag 1620
aagctggtgg gctacctgga ccgcaacctg gagaagaaca gcaccaagca ggagatcctg 1680
gccgccctgg agaagggctg cagcttcctg cccgacccct accagaagca gtgcgaccag 1740
ttcgtggccg agtacgagcc cgtgctgatc gagatcctgg tggaggtgat ggaccccagc 1800

ttcgtgtgcc tgaagatcgg cgcctgcccc agcgcccaca agcccctgct gggcaccgag 1860
aagtgcatct ggggccccag ctactggtgc cagaacaccg agaccgccgc ccagtgcaac 1920
gccgtggagc actgcaagcg ccacgtgtgg aacagaagaa agagaggaag tggagagggc 1980
agaggaagtc ttctgacatg cggagacgtg gaagagaatc ccggccctat gtggaccctg 2040
gtgagctggg tggccctgac cgccggcctg gtggccggca cccgctgccc cgacggccag 2100
ttctgccccg tggcctgctg cctggacccc ggcggcgcca gctacagctg ctgccgcccc 2160
ctgctggaca agtggcccac caccctgagc cgccacctgg gcggcccctg ccaggtggac 2220

gcccactgca gcgccggcca cagctgcatc ttcaccgtga gcggcaccag cagctgctgc 2280
cccttccccg aggccgtggc ctgcggcgac ggccaccact gctgcccccg cggcttccac 2340
tgcagcgccg acggccgcag ctgcttccag cgcagcggca acaacagcgt gggcgccatc 2400
cagtgccccg acagccagtt cgagtgcccc gacttcagca cctgctgcgt gatggtggac 2460
ggcagctggg gctgctgccc catgccccag gccagctgct gcgaggaccg cgtgcactgc 2520
tgcccccacg gcgccttctg cgacctggtg cacacccgct gcatcacccc caccggcacc 2580
cacccccctgg ccaagaagct gcccgcccag cgcaccaacc gcgccgtggc cctgagcagc 2640

agcgtgatgt gccccgacgc ccgcagccgc tgccccgacg gcagcacctg ctgcgagctg 2700
cccagcggca agtacggctg ctgccccatg cccaacgcca cctgctgcag cgaccacctg 2760
cactgctgcc cccaggacac cgtgtgcgac ctgatccaga gcaagtgcct gagcaaggag 2820
aacgccacca ccgacctgct gaccaagctg cccgcccaca ccgtgggcga cgtgaagtgc 2880
gacatggagg tgagctgccc cgacggctac acctgctgcc gcctgcagag cggcgcctgg 2940
ggctgctgcc ccttcaccca ggccgtgtgc tgcgaggacc acatccactg ctgccccgcc 3000

ggcttcacct gcgacaccca gaagggcacc tgcgagcagg gcccccacca ggtgccctgg 3060

atggagaagg cccccgccca cctgagcctg cccgaccccc aggccctgaa gcgcgacgtg 3120
ccctgcgaca acgtgagcag ctgccccagc agcgacacct gctgccagct gaccagcggc 3180
gagtggggct gctgccccat ccccgaggcc gtgtgctgca gcgaccacca gcactgctgc 3240
ccccagggct acacctgcgt ggccgagggc cagtgccagc gcggcagcga gatcgtggcc 3300
ggcctggaga agatgcccgc ccgccgcgcc agcctgagcc acccccgcga catcggctgc 3360
gaccagcaca ccagctgccc cgtgggccag acctgctgcc ccagcctggg cggcagctgg 3420
gcctgctgcc agctgcccca cgccgtgtgc tgcgaggacc gccagcactg ctgccccgcc 3480

ggctacacct gcaacgtgaa ggcccgcagc tgcgagaagg aggtggtgag cgcccagccc 3540
gccaccttcc tggcccgcag cccccacgtg ggcgtgaagg acgtggagtg cggcgagggc 3600
cacttctgcc acgacaacca gacctgctgc cgcgacaacc gccagggctg ggcctgctgc 3660
ccctaccgcc agggcgtgtg ctgcgccgac cgccgccact gctgccccgc cggcttccgc 3720
tgcgccgccc gcggcaccaa gtgcctgcgc cgcgaggccc cccgctggga cgcccccctg 3780
cgcgaccccg ccctgcgcca gctgctgtga caattgttaa ttaagtttaa accctcgagg 3840
ccgcaagcaa taaaatatct ttattttcat tacatctgtg tgttggtttt ttgtgtgaca 3900

attgttaatt aagtttaaac gttcgaggcc gcaagcgaga tccacgataa caaacagctt 3960
ttttggggtg aacatattga ctgaattccc tgcaggttgg ccactccctc tctgcgcgct 4020
cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc gggcgacctt tggtcgcccg 4080
gcctcagtga gcgagcgagc gcgcagagag ggagtggcca actccatcac taggggttcc 4140
tgcggccgct c 4151

<210> 71
<211> 23
<212> DNA
<213> Artificial sequence
<220><223> Synthetic polynucleotide
<400> 71
aagagggtgt tctctatgta ggc 23

<210> 72
<211> 22
<212> DNA
<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 72

gctcctccaa catttgtcac tt 22

<210> 73

<211> 23

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 73

acacagtacc taccgttata gca 23

<210> 74

<211> 23

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 74

tgttgtcaca gtaacttgca tca 23

<210> 75

<211> 19

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 75

ctgggctaca ctgagcacc 19

<210> 76

<211> 21

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 76

aagtggtcgt tgagggcaat g 21

<210> 77

<211> 20

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 77

tattagatct gatggccgcg 20

<210> 78

<211> 20

<212> DNA

<213> Artificial sequence

<220><223> Synthetic polynucleotide

<400> 78

tccatcacta ggggttcctg 20