



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2024년08월07일

(11) 등록번호 10-2693318

(24) 등록일자 2024년08월05일

- (51) 국제특허분류(Int. Cl.)
C12N 9/24 (2006.01) *A61K 31/7088* (2006.01)
A61K 48/00 (2006.01) *A61P 25/16* (2006.01)
C07K 14/435 (2006.01) *C12N 15/113* (2010.01)
- (52) CPC특허분류
C12N 9/2402 (2013.01)
A61K 48/005 (2013.01)
- (21) 출원번호 10-2020-7012188
- (22) 출원일자(국제) 2018년10월03일
 심사청구일자 2021년10월01일
- (85) 번역문제출일자 2020년04월27일
- (65) 공개번호 10-2020-0078512
- (43) 공개일자 2020년07월01일
- (86) 국제출원번호 PCT/US2018/054225
- (87) 국제공개번호 WO 2019/070893
 국제공개일자 2019년04월11일
- (30) 우선권주장
 62/567,296 2017년10월03일 미국(US)
- (56) 선행기술조사문헌
 KR1020140130495 A*
 KR1020150079751 A*
 *는 심사관에 의하여 인용된 문헌

- (73) 특허권자
 프리메일 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 39 항

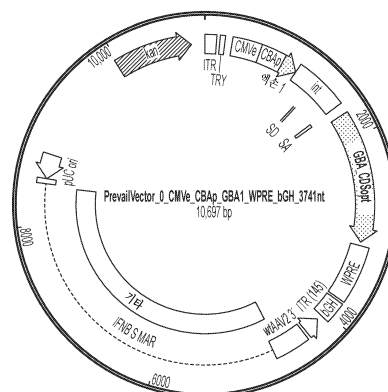
심사관 : 이형곤

(54) 발명의 명칭 리소좀 장애를 위한 유전자 요법

(57) 요약

일부 측면에서, 본 개시내용은 비정상적인 리소좀 기능과 연관된 질환, 예를 들어, 파킨슨병 및 고세병 치료를 위한 조성물 및 방법에 관한 것이다. 일부 실시양태에서, 본 개시내용은 베타-글루코세레브로시다제 (GBA) 또는 그의 일부, 리소좀 막 단백질 2 (LIMP2), 프로사포신, 또는 상기의 임의의 조합을 코딩하는 트랜스진을 포함하는 발현 구축물을 제공한다. 일부 실시양태에서, 본 개시내용은 파킨슨병 방법을 필요로 하는 대상체에게 상기 발현 구축물을 투여함으로써 파킨슨병 방법을 제공한다.

대표도 - 도1



(52) CPC특허분류

A61P 25/16 (2018.01)

C07K 14/435 (2013.01)

C12N 15/113 (2013.01)

C12N 15/86 (2013.01)

C12Y 302/01045 (2013.01)

C12N 2310/14 (2013.01)

C12N 2330/51 (2013.01)

C12N 2750/14143 (2013.01)

명세서

청구범위

청구항 1

(i) 발현 구축물로서,

서열식별번호(SEQ ID NO) 15에 기재된 서열을 포함하는 베타-글루코세레브로시다제 (Gcase) 단백질 코딩 서열을 포함하는 트랜스진;

닭 베타 액틴 (CBA) 프로모터;

CMV 인핸서;

우드척 간염 바이러스 전사후 조절 요소(Woodchuck hepatitis virus Posttranscriptional Regulatory Element: WPRE); 및

소 성장 호르몬 폴리A 신호 테일

을 포함하는 것인 발현 구축물; 및

(ii) 상기 발현 구축물에 플랭킹된 2개의 아테노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR) 서열로서, 제1 ITR 서열은 5' ITR이고, 제2 ITR 서열은 3' ITR인 ITR 서열;

을 포함하는 단리된 핵산이며,

트랜스진은 프로모터에 작동가능하게 연결된 것인, 단리된 핵산.

청구항 2

제1항에 있어서, 트랜스진이 α -시누클레인을 표적화하는 억제성 핵산을 코딩하는 코딩 서열을 추가로 포함하고, 여기서 억제성 핵산은 서열식별번호 20에 기재된 서열에 의해 코딩되는 것인 단리된 핵산.

청구항 3

제1항 또는 제2항에 있어서, 2개의 ITR 서열 각각이 야생형 AAV2 ITR 서열인 단리된 핵산.

청구항 4

제1항 또는 제2항에 있어서, 2개의 ITR 서열 각각이 발현 구축물에 인접한 "D" 영역 (서열식별번호 27)을 포함하는 것인 단리된 핵산.

청구항 5

제1항 또는 제2항에 있어서, ITR 서열 중 적어도 하나가 발현 구축물 기준으로 ITR 서열의 외부에 위치하는 "D" 영역 (서열식별번호 27)을 포함하는 것인 단리된 핵산.

청구항 6

제1항 또는 제2항에 있어서, 발현 구축물 기준으로 5'에 위치하는 ITR 서열이 발현 구축물에 인접한 "D" 영역 (서열식별번호 27)을 포함하고, 발현 구축물 기준으로 3'에 위치하는 ITR 서열이 발현 구축물 기준으로 ITR 서열의 외부에 위치하는 "D" 영역 (서열식별번호 27)을 포함하는 것인 단리된 핵산.

청구항 7

제1항 또는 제2항에 있어서, 5' ITR의 핵산 서열이 서열식별번호 1의 뉴클레오타이드 1 - 145이고, 3' ITR의 핵산 서열이 서열식별번호 1의 뉴클레오타이드 3867 - 4011인 단리된 핵산.

청구항 8

제1항 또는 제2항에 있어서, 5' ITR과 발현 구축물 사이에 TRY 영역을 추가로 포함하고, 여기서 TRY 영역은 서열식별번호 28에 기재된 서열을 갖는 것인 단리된 핵산.

청구항 9

(i) 발현 구축물로서,

서열식별번호 15에 기재된 서열을 포함하는 Gcase 단백질 코딩 서열을 포함하는 트랜스진;

닭 베타 액틴 (CBA) 프로모터;

CMV 인핸서;

우드척 간염 바이러스 전사후 조절 요소; 및

소 성장 호르몬 폴리A 신호 테일

을 포함하는 발현 구축물; 및

(ii) 상기 발현 구축물에 플랭킹된 2개의 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR) 서열로서, 제1 ITR 서열은 5' ITR이고, 제2 ITR 서열은 3' ITR인 ITR 서열;

을 포함하는 재조합 아데노-연관 바이러스 (rAAV) 벡터이며,

트랜스진은 프로모터에 작동가능하게 연결된 것인, rAAV 벡터.

청구항 10

제9항에 있어서, 트랜스진이 α -시누클레인을 표적화하는 억제성 핵산을 코딩하는 코딩 서열을 추가로 포함하고, 여기서 억제성 핵산은 서열식별번호 20에 기재된 서열에 의해 코딩되는 것인 rAAV 벡터.

청구항 11

제9항 또는 제10항에 있어서, 2개의 ITR 서열 각각이 야생형 AAV2 ITR 서열인 rAAV 벡터.

청구항 12

제9항 또는 제10항에 있어서, 2개의 ITR 서열 각각이 발현 구축물에 인접한 "D" 영역 (서열식별번호 27)을 포함하는 것인 rAAV 벡터.

청구항 13

제9항 또는 제10항에 있어서, ITR 서열 중 적어도 하나가 발현 구축물 기준으로 ITR 서열의 외부에 위치하는 "D" 영역 (서열식별번호 27)을 포함하는 것인 rAAV 벡터.

청구항 14

제9항 또는 제10항에 있어서, 발현 구축물 기준으로 5'에 위치하는 ITR 서열이 발현 구축물에 인접한 "D" 영역 (서열식별번호 27)을 포함하고, 발현 구축물 기준으로 3'에 위치하는 ITR 서열이 발현 구축물 기준으로 ITR 서열의 외부에 위치하는 "D" 영역 (서열식별번호 27)을 포함하는 것인 rAAV 벡터.

청구항 15

제9항 또는 제10항에 있어서, 5' ITR의 핵산 서열이 서열식별번호 1의 뉴클레오티드 1 - 145이고, 3' ITR의 핵산 서열이 서열식별번호 1의 뉴클레오티드 3867 - 4011인 rAAV 벡터.

청구항 16

제9항 또는 제10항에 있어서, 5' ITR과 발현 구축물 사이에 TRY 영역을 추가로 포함하고, 여기서 TRY 영역은 서열식별번호 28에 기재된 서열을 갖는 것인 rAAV 벡터.

청구항 17

- (i) 아데노-연관 바이러스 (AAV) 캡시드 단백질; 및
 - (ii) 제9항 또는 제10항의 rAAV 벡터
- 를 포함하는 재조합 아데노-연관 바이러스 (rAAV).

청구항 18

제17항에 있어서, AAV 캡시드 단백질이 AAV9 캡시드 단백질인 rAAV.

청구항 19

5'에서 3' 순으로

- (a) 5' 아데노-연관 바이러스 (AAV) 역전된 말단 반복부 (ITR);
- (b) CMV 인핸서;
- (c) 닭 베타 액틴 (CBA) 프로모터;
- (d) 서열식별번호 15에 기재된 서열을 포함하는 베타-글루코세레브로시다제 (Gcase) 코딩 서열을 포함하는 트랜스진;
- (e) 우드척 간염 바이러스 전사후 조절 요소 (WPRE);
- (f) 소 성장 호르몬 폴리A 신호 테일; 및
- (g) 3' AAV ITR

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

청구항 20

제19항에 있어서, 트랜스진이 α -시누클레인을 표적화하는 억제성 핵산을 코딩하는 코딩 서열을 추가로 포함하고, 여기서 억제성 핵산은 서열식별번호 20에 기재된 서열에 의해 코딩되는 것인 rAAV 벡터.

청구항 21

- (i) 아데노-연관 바이러스 (AAV) 캡시드 단백질; 및
 - (ii) 제19항 또는 제20항의 rAAV 벡터
- 를 포함하는 재조합 아데노-연관 바이러스 (rAAV).

청구항 22

제21항에 있어서, AAV 캡시드 단백질이 AAV9 캡시드 단백질인 rAAV.

청구항 23

제9항, 제10항, 제19항, 및 제20항 중 어느 한 항의 rAAV 벡터를 포함하는 플라스미드.

청구항 24

- (i) 발현 구축물로서,
- 서열식별번호 15에 기재된 서열을 포함하는 베타-글루코세레브로시다제 (Gcase) 단백질 코딩 서열을 포함하는 트랜스진;
- CBA 프로모터;
- CMV 인핸서;
- WPRE; 및
- 소 성장 호르몬 폴리A 신호 테일

을 포함하는 것인 발현 구축물; 및

(ii) 상기 발현 구축물에 플랭킹된 2개의 AAV ITR 서열

을 포함하는 배큘로바이러스 벡터이며,

트랜스진은 프로모터에 작동가능하게 연결된 것인, 배큘로바이러스 벡터.

청구항 25

제24항에 있어서, 트랜스진이 α -시누클레인을 표적화하는 억제성 핵산을 코딩하는 코딩 서열을 추가로 포함하고, 여기서 억제성 핵산은 서열식별번호 20에 기재된 서열에 의해 코딩되는 것인, 배큘로바이러스 벡터.

청구항 26

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

(ii) 제2 벡터로서,

(a) 발현 구축물이며,

서열식별번호 15에 기재된 서열을 포함하는 베타-글루코세레브로시다제 (Gcase) 단백질 코딩 서열을 포함하는 트랜스진;

CBA 프로모터;

CMV 인핸서;

WPRE; 및

소 성장 호르몬 폴리A 신호 테일

을 포함하는 것인 발현 구축물; 및

(b) 상기 발현 구축물에 플랭킹된 2개의 AAV ITR 서열

을 포함하는 제2 벡터

를 포함하는 세포이며,

트랜스진은 프로모터에 작동가능하게 연결된 것인, 세포.

청구항 27

제26항에 있어서, 트랜스진이 α -시누클레인을 표적화하는 억제성 핵산을 코딩하는 코딩 서열을 추가로 포함하고, 여기서 억제성 핵산은 서열식별번호 20에 기재된 서열에 의해 코딩되는 것인 세포.

청구항 28

제26항 또는 제27항에 있어서, 제1 벡터가 플라스미드이고, 제2 벡터가 플라스미드인 세포.

청구항 29

제28항에 있어서, 포유동물 세포인 세포.

청구항 30

제29항에 있어서, 포유동물 세포가 HEK293 세포인 세포.

청구항 31

제26항 또는 제27항에 있어서, 제1 벡터가 배큘로바이러스 벡터이고, 제2 벡터가 배큘로바이러스 벡터인 세포.

청구항 32

제31항에 있어서, 곤충 세포인 세포.

청구항 33

아데노-연관 바이러스 (AAV) 캡시드 단백질 및 제9항, 제10항, 제19항, 및 제20항 중 어느 한 항의 rAAV 벡터를 포함하는 재조합 아데노-연관 바이러스 (rAAV)를 제조하는 방법이며,

- (i) 하나 이상의 AAV rep 단백질 및/또는 하나 이상의 AAV cap 단백질을 코딩하는 제1 벡터, 및 rAAV 벡터를 세포로 전달하는 단계;
- (ii) rAAV의 패키징을 허용하는 조건하에서 상기 (i)의 세포를 배양하는 단계; 및
- (iii) rAAV 수집을 위해, 배양된 숙주 세포 또는 배양 배지를 수거하는 단계를 포함하는 방법.

청구항 34

제1항의 단리된 핵산, 제9항 또는 제19항의 rAAV 벡터, 또는 아데노-연관 바이러스 (AAV) 캡시드 단백질 및 상기 rAAV 벡터를 포함하는 rAAV를 포함하는, 파킨슨병 또는 고세병을 앓거나 또는 앓을 것으로 의심되는 대상체를 치료하기 위한 제약 조성물.

청구항 35

제34항에 있어서, 대상체의 중추 신경계 (CNS)에의 직접 주사용으로 제형화된 제약 조성물.

청구항 36

제35항에 있어서, 대상체의 CNS에의 직접 주사가 대뇌내 주사, 실질내 주사, 경막내 주사, 대수조내 주사, 또는 그의 조합인 제약 조성물.

청구항 37

제35항에 있어서, 대상체의 CNS에의 직접 주사가 컨벡션 강화 전달법 (CED)을 포함하는 것인 제약 조성물.

청구항 38

제34항에 있어서, 말초 주사용으로 제형화된 제약 조성물.

청구항 39

제38항에 있어서, 말초 주사가 정맥내 주사인 제약 조성물.

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

발명의 설명

기술 분야

- [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")의 출원일의 이익을 주장하고, 상기 출원의 전체 내용이 본원에서 참조로 포함된다.

배경 기술

- [0003] 고세병(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 결합성 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 효소 대체 요법, 결합성 Gcase에 결합하고, 안정성을 개선시키는 사페론 유사 소분자 약물, 및 증상 및 병상으로 이어지는, 고세병에서 축적되는 기질의 생산을 차단하는 기질 감소 요법을 포함한다. 그러나, 고세병의 다른 측면 및 치료에 대하여 난치성인 것으로 보인다.

발명의 내용

- [0005] (GBA1 유전자의 두 염색체 대립유전자 모두에 돌연변이를 갖는) 고세병 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, 고세병 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 Gcase (또는 그의 일부), 프로사포신 (또는 그의 일부), LIMP2 (또는 그의 일부), 또는 Gcase (또는 그의 일부) 및 PD 연관 유전자 (예컨대, LIMP2, 프로사포신, 및/또는 α -시누클레인 (α -Syn))로부터의 하나 이상의 추가 유전자 생성물의 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, Gcase를 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, 프로사포신을 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 세포, 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, SCARB2를 코딩하는 핵산 서열에 플랭킹된 AAV ITR을 포함한다.
- [0010] 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.
- [0011] 일부 실시양태에서, 제1 유전자 생성물 또는 제2 유전자 생성물은 Gcase 단백질 또는 그의 일부이다. 일부 실시양태에서, 제1 유전자 생성물 또는 제2 유전자 생성물은 LIMP2 또는 그의 일부, 또는 프로사포신 또는 그의 일부이다. 일부 실시양태에서, 제1 유전자 생성물은 Gcase 단백질이고, 제2 유전자 생성물은 LIMP2 또는 그의 일부, 또는 프로사포신 또는 그의 일부이다.
- [0012] 일부 실시양태에서, 발현 구축물은 간섭 핵산 (예컨대, shRNA, miRNA, dsRNA 등)을 추가로 코딩한다. 일부 실시양태에서, 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).
- [0013] 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.
- [0014] 일부 실시양태에서, 발현 구축물은 내부 리보솜 진입 부위 (IRES)를 추가로 포함한다. 일부 실시양태에서, IRES는 제1 유전자 생성물과 제2 유전자 생성물 사이에 위치한다.
- [0015] 일부 실시양태에서, 발현 구축물은 자기 절단 펩티드 코딩 서열을 추가로 포함한다. 일부 실시양태에서, 자기 절단 펩티드는 T2A 펩티드이다.
- [0016] 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 Δ ITR이다.
- [0017] 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.
- [0018] 일부 실시양태에서, 단리된 핵산 (예컨대, rAAV 벡터)는 TRY 영역을 포함한다. 일부 실시양태에서, TRY 영역은 서열식별번호 28에 기재되어 있는 바와 같은 서열을 포함한다.
- [0019] 일부 실시양태에서, 본 개시내용에 의해 기술된 단리된 핵산은 서열식별번호 1 내지 13, 15, 17, 및 19 중 어느 하나에 기재된 서열을 포함하거나 또는 그로 이루어진다. 일부 실시양태에서, 본 개시내용에 의해 기술된 단리된 핵산은 서열식별번호 14, 16, 및 18 중 어느 하나에 기재된 서열을 포함하거나 또는 그로 이루어진 펩티드를 코딩한다.
- [0020] 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) 벡터이다. 일부 실시양태에서, rAAV 벡터는 단일-가닥 (예컨대, 단일-가닥 DNA)이다.

- [0021] 일부 측면에서, 본 개시내용은 본 개시내용에 의해 기술된 바와 같은 단리된 핵산, 또는 본 개시내용에 의해 기술된 바와 같은 벡터를 포함하는 숙주 세포를 제공한다.
- [0022] 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.
- [0023] 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.
- [0024] 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.
- [0025] 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.
- [0026] 일부 실시양태에서, 투여는 말초 주사를 포함한다. 일부 실시양태에서, 말초 주사는 정맥내 주사이다.

도면의 간단한 설명

- [0027] 도 1은 Gcase (예컨대, *GBA1* 또는 그의 일부)를 코딩하는 발현 구축물을 포함하는 rAAV 벡터를 포함하는 플라스미드의 한 실시양태를 도시한 개략도이다.
- 도 2는 Gcase (예컨대, *GBA1* 또는 그의 일부) 및 LIMP2 (SCARB2) 또는 그의 일부를 코딩하는 발현 구축물을 포함하는 rAAV 벡터를 포함하는 플라스미드의 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 rAAV 벡터를 포함하는 플라스미드의 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 간접 RNA를 코딩하는 발현 구축물을 포함하는 rAAV 벡터를 포함하는 플라스미드의 한 실시양태를 도시한 개략도이다.
- 도 5는 Gcase (예컨대, *GBA1* 또는 그의 일부), 프로사포신 (예컨대, *PSAP* 또는 그의 일부), 및 α -Syn에 대한 간접 RNA를 코딩하는 발현 구축물을 포함하는 rAAV 벡터를 포함하는 플라스미드의 한 실시양태를 도시한 개략도이다.
- 도 6은 Gcase (예컨대, *GBA1* 또는 그의 일부) 및 프로사포신 (예컨대, *PSAP* 또는 그의 일부)을 코딩하는 발현 구축물을 포함하는 rAAV 벡터를 포함하는 플라스미드의 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 rAAV 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 개체 서열은 정확하게 패키징될 수 있다. 5' ITR 서열의 두 변이체 (삽입 박스)를 평가하였다: 상기 변이체들은 야생형 AAV2 ITR의 20-뉴클레오타이드 "D" 영역 내에서 수개의 뉴클레오타이드 차이를 보인다. 일부 실시양태에서, rAAV 벡터는 상단 라인에 제시된 "D" 도메인 뉴클레오타이드 서열을 함유한다. 일부 실시양태에서, rAAV 벡터 돌연변이체 "D" 도메인 (예컨대, "S" 도메인, 여기서 뉴클레오타이드 변이는 하단 라인에 제시)을 보유한다.
- 도 8은 도 7에 기술된 rAAV 벡터를 코딩하는 플라스미드의 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 $n = 8$ 으로 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 GCa6 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 8.8×10^9 vg rAAV로 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 절반은 P38 (3일째)에 희생시키기 전에 3일 동안 CBE를 중단하였다. 모든 처리군 (부형제 + PBS의 경우, $n = 8$, rAAV + PBS의 경우, $n = 7$, 부형제 + CBE의 경우, $n = 8$, 및 rAAV + CBE의 경우, $n = 9$)의 체중을 매일 측정하였고 (좌측 상단), P36에서의 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, $n = 7$ 인 반면, 다른 군은 모두 $n=8$ 이었다. 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, $n = 8$, rAAV + PBS의 경우, $n = 7$, 부형제 + CBE의 경우, $n = 7$, 및 rAAV + CBE의 경우, $n = 9$)의 피질을 사용하여 군에서의 GCa6 활성 (좌측 상단), GluSph 수준 (우측 상단), GluCer 수준 (좌측 하단), 및 벡터 게놈 (우측 하단)을 측정하였다. 생체분포는 게놈 DNA 1 μ g당 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 수집 및 성별은 공변량으로서 보정.

도 13은 부형제 + PBS, 부형제 + CBE, 및 rAAV + 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 GCa6 활성과 음의 상관관계를 가졌다 (B, 선형 회귀에 의해 $p=0.0086$).

도 14는 CBE 마우스 모델에서의 GBA1 rAAV의 생체분포에 대한 대표적인 데이터를 보여주는 것이다. 모든 처리군 (부형제 + PBS의 경우, $n = 8$, rAAV + PBS의 경우, $n = 7$, 부형제 + CBE의 경우, $n = 7$, 및 rAAV + CBE의 경우, $n = 9$)에 대해 간, 비장, 신장, 및 생식선에서 벡터 게놈의 존재를 사정하였다. 생체분포는 게놈 DNA 1 μ g당 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 수집 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 GBA1 rAAV: 3.2×10^9 vg, 1.0×10^{10} vg, 또는 3.2×10^{10} 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. 초기 치사에 기인하여, 행동 검정법에 참여한 마우스의 마리수는 부형제 + PBS의 경우, $n = 10$, 부형제 + CBE의 경우, $n = 9$, 및 3.2×10^9 vg rAAV + CBE의 경우, $n = 6$, 1.0×10^{10} vg rAAV + CBE의 경우, $n = 10$, 3.2×10^{10} vg rAAV + CBE의 경우, $n = 7$ 이었다. 평균이 제시되어 있

다. 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 3.2×10^9 vg rAAV + CBE의 경우, $n = 6$, 1.0×10^{10} vg rAAV + CBE의 경우, $n = 10$, 3.2×10^{10} vg rAAV + 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 μ g당 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 GBA1 rAAV에서의 테이퍼 빔 분석에 대한 대표적인 데이터를 보여주는 것이다. rAAV 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 ($n = 5$)의 운동 수행능을 검정하였다. 총 슬립수 및 활동 시간은 상이한 빔에서의 5회 시도에 대한 총합으로 제시되어 있다. 속도 및 속도당 슬립수는 상이한 빔에서의 5회 시도에 대한 평균으로 제시되어 있다. 평균이 제시되어 있다. 오차 막대는 SEM이다.

도 18은 프로사포신 (PSAP), SCARB2, 및/또는 하나 이상의 억제성 핵산과 함께 조합된, GBA1을 코딩하는 rAAV 구축물의 시험관내 발현에 대한 대표적인 데이터를 보여주는 것이다. 데이터는, HEK293 세포를 각 구축물로 형질감염시켰을 때, GFP 형질감염된 세포와 비교하여 관심 트랜스진이 과다발현되었다는 것을 나타낸다.

도 19는 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 20은 "D" 서열이 야생형으로 (동그라미 표시) 또는 대안적으로 (예컨대, "외부"에; 사각형 표시) 배치되어 있는 ITR을 갖는 rAAV를 이용한 HEK293 세포의 형질도입에 대한 데이터를 보여주는 것이다. "외부"에 배치된 ITR을 갖는 rAAV는 야생형 ITR을 갖는 rAAV만큼 효율적으로 세포를 형질도입시킬 수 있었다.

발명을 실시하기 위한 구체적인 내용

[0028] 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.

[0029] 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) SCNA가 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.

[0030] <표 1>

[0031] PD 연관 유전자의 예

명칭	유전자	기능	NCBI 수탁 번호
리소좀 막 단백질 2	<i>SCARB2/LIMP2</i>	글루코실세라미다제에 대한 리소좀 수용체 (GBA 표적화)	NP_005497.1 (이소형 1), NP_001191184.1 (이소형 2)
프로사포신	<i>PSAP</i>	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	<i>GBA1</i>	글루코세레브로시드의 베타-글리코시드 결합을 절단한다	NP_001005742.1 (이소형 1), NP_001165282.1 (이소형 2), NP_001165283.1 (이소형 3)

[0032]

[0033] 단리된 핵산 및 벡터

[0034] 단리된 핵산은 DNA 또는 RNA일 수 있다. 일부 측면에서, 본 개시내용은 Gcase (예컨대, *GBA1* 유전자의 유전자 생성물) 또는 그의 일부를 코딩하는 발현 구축물을 포함하는 단리된 핵산을 제공한다. β -글루코세레브로시다제 또는 GBA로도 지칭되는 Gcase란, 당지질 대사에서 중간체인 화합물질 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.

[0035] 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 스팅고리피드 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.

[0036] 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.

[0037] 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.

[0038] 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 *SCARB2/LIMP2* 유전자 및/또는 *PSAP* 유전자에 의해 코딩된 단백질 (또는 그의 단편)이다. 그러나, 통상의 기술자는 제1 유전자 생성물 (예컨대, Gcase) 및 제2 유전자 생성물 (예컨대, LIMP2)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.

- [0039] 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).
- [0040] 폴리시스트론성 발현 벡터는 하나 이상의 (예컨대, 1, 2, 3, 4, 5개 또는 그 초과) 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.
- [0041] 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.
- [0042] 병리학상, 장애, 예컨대, PD 및 고세병은 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.
- [0043] 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 뉴클레오타이드에 결합한다 (예컨대, 그와 하이브리드화한다). 일부 실시양태에서, 억제성 핵산 분자는 miRNA 또는 amiRNA, 예를 들어, *SNCA* (α -Syn 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*를 표적화하는 shRNA)이다.
- [0044] 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, 배클로바이러스 벡터 등)일 수 있다. 일부 실시양태에서, 벡터는 플라스미드 (예컨대, 본원에 기술된 바와 같은 단리된 핵산을 포함하는 플라스미드)이다. 일부 실시양태에서, 벡터는 재조합 AAV (rAAV) 벡터이다. 일부 실시양태에서, rAAV 벡터는 단일-가닥 (예컨대, 단일-가닥 DNA)이다. 일부 실시양태에서, 벡터는 배클로바이러스 벡터 (예컨대, 오토그라파 칼리포르니카 핵다각체병 (AcNPV) 벡터)이다.
- [0045] 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, Δ ITR이다.

[0046] 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 19에 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 19). 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.

[0047] 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.

[0048] 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 초과 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.

[0049] 본 개시내용에 의해 기술된 바와 같은 단리된 핵산 또는 rAAV 벡터는 예를 들어, 서열식별번호 28에 기재된 바와 같은, 또는 문헌 [Francois, et al. The Cellular TATA Binding Protein Is Required for Rep-Dependent Replication of a Minimal Adeno-Associated Virus Type 2 p5 Element. *J Virol.* 2005]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.

[0050] 일부 측면에서, 본 개시내용은 본 개시내용에 의해 기술된 바와 같은 단리된 핵산 또는 rAAV 벡터를 포함하는 배칼로바이러스 벡터에 관한 것이다. 일부 실시양태에서, 배칼로바이러스 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.

[0051] 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.

[0052] rAAV

[0053] 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 야생

형 AAV 캡시드 단백질 기준으로 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개, 또는 10개 초과 (예컨대, 15, 20, 25, 50, 100개 등) 아미노산 치환 (예컨대, 돌연변이)을 포함하는 캡시드 단백질 변이체를 포함한다.

[0054] 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.

[0055] 일부 실시양태에서, 본 개시내용에 의해 기술된 바와 같은 rAAV (예컨대, AAV 캡시드 단백질에 의해 캡시드화되어 rAAV 캡시드 입자를 형성하는 재조합 rAAV 계통을 포함하는 것)는 배컬로바이러스 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.

[0056] 제약 조성물

[0057] 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 비바람직할 생물학적 효과를 유발하지 않거나, 또는 그와 함께 함유되어 있는 조성물의 성분 중 어느 것이라도 유해한 방식으로 상호작용하지 않으면서, 개체에게 투여될 수 있다는 것을 지칭한다.

[0058] 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.

[0059] 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 환부의 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.

[0060] 방법

[0061] 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.

[0062] 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.

[0063] 일부 실시양태에서, 조성물은 대상체의 CNS로 직접, 예를 들어, 대상체의 뇌 및/또는 척수 내로의 직접 주사에 의해 투여된다. CNS에의 직접 주사 방식의 예로는 대뇌내 주사, 뇌실내 주사, 수조내 주사, 실질내 주사, 경막

내 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).

[0064] 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.

[0065] 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.

[0066] 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.

[0067] 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.

[0068] 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.

[0069] 실시예

[0070] 실시예 1: rAAV 벡터

[0071] 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 캡티드 링커, 예컨대, T2A 또는 P2A 사용으로 캡티드 결합이 생성되지 못하게 방해됨에 따라 아미노산이 부가된 2개의 캡티드 단편이 유도됨으로써; IRES 요소를 사용하여; 또는 2개의 별개의 발현 카세트들을 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 rAAV 벡터를 포함하는 플라스미드의 예는 도 1-6 및 하기 표 2에 제시되어 있다.

[0072] <표 2>

명칭	프로모터 1	shRNA	CDS1	폴리A 1	비시스트론성 요소	프로모터 2	CDS2	폴리A 2	ITR 사이의 길이
CMVc_CBAp_GBA1_WPRE_bGH	CBA		GBA1	WPRE-bGH					3741
LT1s_JetLong_mRNAi aSYn_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_mRNAi aSYn_P SAP-T2A-GBA1_bGH_4353nt	JetLong	aSyn	PSAP	bGH	T2A	-	GBA1	-	4353
PrevailVector_LI2_JetLong_PSAP_IRES-GBA1_SyntheticpolyA_4337nt	JetLong	-	PSAP	합성 pA	IRES	-	GBA1	-	4337

[0073]

[0074] 실시예 2: GBA 결핍 세포로의 바이러스 형질도입에 관한 세포 기반 검정법

[0075] 예를 들어, GD 환자 단핵구, 또는 hES 세포로부터의 섬유모세포, 또는 환자 유래의 유도 만능 줄기 세포 (iPS C)로서 GBA1 결핍 세포를 획득한다. 이들 세포는 기질, 예컨대, 글루코실세라미드 및 글루코실스핑고신 (GluCer 및 GluSph)을 축적한다. 야생형 또는 돌연변이체 배양 세포주를 Gcase 억제제, 예컨대, CBE로 처리하는 것 또한 GBA 결핍 세포를 획득하는 데 사용된다.

[0076] 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 텍스트란 또는 다른 마커의 세포내이입 구획을 통한 흡수에 의해 리소좀 이상에 대한 영상화 또한 수행한다. 리소좀과의 결합성 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.

[0077] 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.

[0078] 실시예 3: 돌연변이체 마우스를 이용한 생체내 검정법

[0079] 본 실시예는 돌연변이체 마우스를 이용한 AAV 벡터에 관한 생체내 검정법을 기술한다. 돌연변이체 마우스에서의 상기과 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.

[0080] 예를 들어, 5-10 μl의 주사 부피로 진한 AAV 스톱을 사용하여 (예컨대, 2x10¹¹ vg/마우스 용량의) 비히클 대조군 및 AAV 벡터의 경막내 또는 뇌실내 전달을 수행한다. 컨택션 강화 전달법에 의한 실질내 전달을 수행한다.

[0081] 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.

[0082] 실시예 4: 질환의 화학적 모델

[0083] 본 실시예는 화학적으로 유도된 고세병을 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 고세병을 앓는 마우스 모델에서 상기 AAV 벡터에 관한 생체내 연구를 수행한다.

- [0084] 예를 들어, 5-10 μ l의 주사 부피로 진한 AAV 스톱을 사용하여 (예컨대, 2×10^{11} vg/마우스의) 비히클 대조군 및 AAV 벡터의 경막내 또는 뇌실내 전달을 수행한다. 컨벡션 강화 전달법에 의한 실질내 전달을 수행한다. 꼬리 정맥 주사에 의해 말초 전달을 달성한다.
- [0085] 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.
- [0086] 실시예 5: PD, LBD, 고세병 환자에서의 임상 시험
- [0087] 일부 실시양태에서, 특정 형태의 고세병 (예컨대, GD1)을 앓는 환자는 파킨슨병 (PD) 또는 루이소체 치매 (LBD)가 발생할 위험이 증가되어 있다. 본 실시예는 고세병, PD 및/또는 LBD를 앓는 환자에서 본 개시내용에 의해 기술된 바와 같은 rAAV의 안전성 및 효능을 사정하는 임상 시험을 기술한다.
- [0088] 문헌 [Grabowski *et al.* (1995) *Ann. Intern. Med.* 122(1):33-39]에 기술된 것과 유사한 연구 디자인을 사용하여 고세병, PD 및/또는 LBD 치료를 위한 상기 벡터의 임상 시험을 수행한다.
- [0089] 실시예 6: 말초 질환 치료
- [0090] 일부 실시양태에서, 예를 들어, 문헌 [Biegstraaten *et al.* (2010) *Brain* 133(10):2909-2919]에 기술된 바와 같이, 특정 형태의 고세병을 앓는 환자는 말초 신경병증의 증상을 보인다.
- [0091] 본 실시예는 고세병 (예컨대, 타입 1 고세병)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 고세병 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.
- [0092] 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 검정법, 또는 영상화 연구에 의해 검정한다.
- [0093] 실시예 7: CNS 형태의 치료
- [0094] 본 실시예는 CNS 형태의 고세병의 치료를 위한 본원에 기술된 바와 같은 rAAV 벡터에 관한 생체내 검정법을 기술한다. 간략하면, CNS 형태의 고세병 (예컨대, 타입 2 또는 타입 3 고세병)을 앓는 것으로 확인된 고세병 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.
- [0095] 실시예 8: GBA1에 돌연변이를 갖는 대상체에서 파킨슨병의 유전자 요법
- [0096] 본 실시예는 GBA1 중의 돌연변이를 특징으로 하는 파킨슨병을 앓는 대상체에게 GBA1을 코딩하는 재조합 아테노-연관 바이러스 (rAAV)를 투여하는 것을 기술한다.
- [0097] rAAV 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 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 막

는 스테퍼(stuffer) 서열을 함유한다. rAAV 벡터를 도시하는 개략도는 도 8에 제시되어 있다. rAAV 벡터는 AAV9 혈청형 캡시드 단백질을 이용하여 rAAV로 패키징된다.

- [0098] 대수조 내로의 (대수조내; ICM) 투시 조영 유도하의 후두하 주사를 통해 GBA1-rAAV를 단일 용량으로 대상체에게 투여한다. 투약 요법 연구의 한 실시양태는 하기와 같다:
- [0099] 단일 용량의 rAAV를, 비임상 약리학 및 독성 연구의 결과에 기초하여 결정된 두 용량 수준 (3×10^{13} vg (저용량); 1×10^{14} vg (고용량) 등) 중 하나로 환자 (N=12)에게 투여한다.
- [0100] 초기 연구는 rAAV 벡터 및 변이체 rAAV 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.
- [0101] 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 이외에, 돌연변이체 D 도메인 (본원에서 "S" 도메인으로 명명)을 포함하는 제2 벡터 형태 또한 평가하였다. rAAV 및 변이체 rAAV, 둘 모두 동일한 트랜스진을 발현한다. 하기에서 상세하게 설명되는 바와 같이, 상기 두 벡터 모두 생체내에서 효과적인 바이러스를 생산한 반면, 야생형 "D" 도메인을 함유하는 rAAV는 추가 개발을 위해 선택하였다.
- [0102] 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 주사받은 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.
- [0103] 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.
- [0104] 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 또는 부형제를 전달한 후, 이어서, P8에 IP CBE 또는 PBS의 매일 처리를 개시하였다 (도 10).
- [0105] rAAV를 받은, CBE로 처리된 마우스는 로타로드에서 부형제를 받은 것보다 통계학상 유의적으로 더욱 우수한 수행능을 보였다 (도 11). 변이체 벡터 처리군의 마우스는 다른 행동 척도, 예컨대, 검사 동안 이동한 총 거리에 있어서는 부형제로 처리된 마우스와 다르지 않았다 (도 11).
- [0106] 생존 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. GBA1 rAAV로 처리된 마우스에서는 GCase 활성이 증가한 반면, CBE 처리는 GCase 활성을 감소시켰다. 추가로, CBE 및 GBA1-rAAV, 둘 모두를 받은 마우스는 PBS 처리군과 유사한 GCase 활성 수준을 보였는데, 이는 rAAV 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 처리는 기질 축적을 유의적으로 감소시켰다.
- [0107] 지질 수준은 처리군 간에 GCase 활성 및 로타로드에서의 수행능, 둘 모두와 음의 상관관계를 가졌다. rAAV 투

여 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 μ g당 >100 벡터 게놈인 것을 양성인 것으로 정의한다). CBE로 처리 및 비처리된, GBA1-rAAV를 받은 마우스, 둘 모두 피질에서 rAAV 벡터 게놈에 대해 양성이었다고, 이는 ICV 전달을 통해 rAAV가 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.

[0108] CBE 모델에서의 좀더 큰 대규모 연구를 통해 CBE 모델에서 GBA1-rAAV의 유효 용량을 추가로 조사하였다. 25 mg/kg CBE 용량 모델을 이용하여, 부형제 또는 GBA1-rAAV를 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 용량의 효과를 사정하였다:

[0109] 부형제 ICV + PBS IP

[0110] 부형제 ICV + 25 mg/kg CBE IP

[0111] 3.2e9 vg (2.13e10 vg/g 뇌) rAAV ICV + 25 mg/kg CBE IP

[0112] 1.0e10 vg (6.67e10 vg/g 뇌) rAAV ICV + 25 mg/kg CBE IP

[0113] 3.2e10 vg (2.13e11 vg/g 뇌) rAAV ICV + 25 mg/kg CBE IP.

[0114] 최고 용량의 rAAV는 P37에 CBE 처리와 관련된 체중 증가 불능을 구제하였다. 추가로, 부형제 + CBE 처리군과 비교하였을 때, 상기 용량은 로타로드 및 테이퍼 빔(tapered beam)에서의 수행능을 통계학상 유의적으로 증가시켰다 (도 15). 부형제 처리군 및 rAAV 처리군, 둘 모두를 포함하는, 여러 군에서 치사가 관찰되었다 (부형제 + PBS: 0; 부형제 + 25 mg/kg CBE: 1; 3.2e9 vg rAAV + 25 mg/kg CBE: 4; 1.0e10 vg rAAV + 25 mg/kg CBE: 0; 3.2e10 vg rAAV + 25 mg/kg CBE: 3).

[0115] 생전 연구 완료시, 생화학적 분석을 위해, 마우스를 희생시켰다 (도 16). 형광분석 검정법에 의해 생물학적으로 삼중으로 피질에서 GCase 활성을 사정하였다. CBE로 처리된 마우스는 감소된 GCase 활성을 보인 반면, 고용량의 rAAV를 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 투여에 의해 구제되었다.

[0116] 확립된 화학적 CBE 모델에 추가로, Gba1에서의 V394L GD 돌연변이에 대해 동형접합성이고, 또한 GCCase 국제화 및 활성에 영향을 주는 사포신이 부분적으로 결핍되어 있는 4L/PS-NA 유전적 모델에서도 또한 GBA1-rAAV를 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 μ l의 최대 역가 바이러스를 ICV에 의해 전달하였고, 여기서 최종 용량은 2.4e10 vg (6.0e10 vg/g 뇌)였다. 군당 마우스 6마리씩이며, 처리군은 하기와 같았다:

[0117] WT + 부형제 ICV

[0118] 4L/PS-NA + 부형제 ICV

[0119] 4L/PS-NA + 2.4e10 vg (6.0e10 vg/g 뇌) rAAV ICV.

[0120] rAAV 전달 후 4주째에 빔 워크 검사에 의한 운동 수행능을 사정하였다. GBA1-rAAV를 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, GCCase 활성, 및 생체분포를 사정한다.

[0121] 0.03x, 0.1x, 및 1x 제안된 1상의 임상 고용량에 상응하는, 추가의 더 낮은 용량의 rAAV를 CBE 모델을 이용하여 현재 시험하고 있다. 각 군은 군당 마우스 10마리씩 (5M/5F) 포함한다:

[0122] 부형제 ICV

[0123] 부형제 ICV + 25 mg/kg CBE IP

- [0124] 3.2e8 vg (2.13e9 vg/g 뇌) rAAV ICV + 25 mg/kg CBE IP
- [0125] 1.0e9 vg (6.67e9 vg/g 뇌) rAAV ICV + 25 mg/kg CBE IP
- [0126] 1.0e10 vg (6.67e10 vg/g 뇌) rAAV ICV + 25 mg/kg CBE IP.
- [0127] 운동 표현형 이외에도, 피질에서 지질 수준 및 GCase 활성을 사정한다. 시간 경과에 따른 처리 및 분석 또한 수행한다.
- [0128] 효능 및 안전성 데이터를 평가하기 위해 더 큰 용량 범위의 연구를 개시하였다. 10마리의 4L/PS-NA 마우스 (군당 5M/5F)에 10 μ l의 rAAV를 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:
- [0129] WT + 부형제 ICV
- [0130] 4L/PS-NA + 부형제 ICV
- [0131] 4L/PS-NA + 4.3e9 vg (1.1e10 vg/g 뇌) rAAV ICV
- [0132] 4L/PS-NA + 4.3e10 vg (1.1e11 vg/g/ 뇌) rAAV ICV
- [0133] 4L/PS-NA + 1.3e11 vg (3.2e11 vg/g 뇌) rAAV ICV
- [0134] 4L/PS-NA + 4.3e11 vg (1.1e12 vg/g 뇌) rAAV ICV.
- [0135] CBE 모델에서의 비임상 연구에 관한 요약은 하기 표 3에 제시되어 있다.
- [0136] <표 3>
- [0137] CBE 마우스 모델에서의 결과 요약

시험 물질	연구 번호	용량 코호트	행동 변화			지질	효소	BD	
			비 교 적 행 동	행 동 변 화	비 교 적 행 동			+	-
GBA1- rAAV	CBE 모델에서의 PRV-2018-	3.2e9 vg (2.13e10 vg/g 뇌)	NS	NS	NS	NS	NS	+	-
	005 용량 범위	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 GBA1- rAAV	CBE 모델에서의 PRV-2018- 005 용량 범위 rAAV	8.8e9 vg (5.9e10 vg/g 뇌)	S	N/A	NS	S	S	+	+

- [0138]
- [0139] 양성 생체분포는 >100 vg/1 μ g 계층 DNA인 것으로 정의된다는 점에 주의한다.
- [0140] 약어: BD = 생체분포; NS = 비유의적; T = 경향; S = 유의적; N/A = 적용불가능; + = 양성; - = 음성.
- [0141] 실시예 9: rAAV 벡터의 시험관내 분석
- [0142] 프로사포신 (PSAP) 및 SCARB2를 단독으로, 또는 GBA1 및/또는 하나 이상의 억제성 RNA와 함께 조합하여 코딩하는 rAAV 벡터의 시험관내 활성을 사정하는 예비 연구를 수행하였다. PSAP 및 프로그래놀린 (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 18은 HEK293 세포를 각 구축물로 형질감염시켰을 때, 모의 형질감염된 세포와 비교하여 상응한 유전자 생성물이 과다발현되었다는 것을 나타내는 대표적인 데이터를 보여주는 것이다.

<표 4>

ID	프로모터	억제성 RNA	프로모터	트랜스진
I00015	JL_인트론	SCNA	JetLong	Opt-PSAP_GBA1
I00039	-	-	JetLong	Opt-PSAP-GRN
I00046	-	-		Opt-PSAP
I00014	JetLong	SCNA	JetLong	Opt-SCARB2_GBA1

실시예 10: ITR "D" 서열 배치 및 세포 형질도입

ITR "D" 서열의 배치가 rAAV 벡터의 세포 형질도입에 미치는 효과를 조사하였다. 도 19에 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 20).

실시예 11: 시험관내 독성 연구

오십 (50) 마리의 마우스에 생후 3일째 GBA1 코딩 rAAV를 4 μ l 뇌실내 (ICV) 주사를 통해 투여하였다. 모든 마우스는 처리군에 따라 콘두리틀 B-에폭시드 (CBE) 또는 PBS를 생후 8일째부터 연구 종료시까지 매일 복강내 (IP) 주사로 받았다. 마우스를 그의 최종 IP 투약 후 24시간째에 안락사시켰다. 안락사시킨 후, 표적 조직을 수거하고, 냉각 4% 파라포름알데히드 중에 점적 고정시키고, 4°C에서 보관한 후, 조직병리학 프로세싱 및 평가를 위해 보냈다. 연구 진행 동안 여덟 (8)마리의 동물이 사망하였고, 이들은 보내거나, 분석하거나 하지 않았다.

38-40일째에 안락사된 마흔 두 (42) 마리의 동물로부터 얻은 조직을 트리밍하고, 프로세싱하고, 파라핀 블록 중에 포매시켰다. 이어서, ~5 μ m 절편을 제작하고, 헤마톡실린 및 에오신 (H&E)으로 염색하고, 평가를 위해 슬라이드에 부착시켰다.

rAAV 처리에 기인하여, 조직병리학적 소견 또는 독성의 증거는 없었다. 콘두리틀 B-에폭시드 (CBE)로 처리된 마우스에서는 중추 신경계 (CNS)에 대뇌 피질 중 신경교 흉터 및 뉴런 괴사, 및 뇌간 및 흉부 척수 중 뉴런 괴사를 포함하는 소견이 있었다. 고용량의 rAAV 처리는 상기 CNS 소견의 발병률을 현저히 감소시킨 반면, 저용량 및 중간 용량의 바이러스는 용량에 의존하는 방식으로 대뇌 피질 중 신경교 흉터의 발병률을 감소시켰고, 다른 CNS 소견에 미치는 효과는 모호하였다.

등가물

본 출원은 하기 문헌의 내용, 그 전문을 참조로 포함한다: 2018년 10월 3일 출원된, 대리인 사건 번호 P1094.70003W000으로 언급되는 국제 PCT 출원; 2018년 10월 3일 출원된, 대리인 사건 번호 P1094.70004W000으로 언급되는 국제 PCT 출원; 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 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").

따라서, 본 발명의 적어도 하나의 실시양태의 여러 측면을 기술하였는 바, 관련 기술분야의 통상의 기술자는 다

양한 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.

[0154] 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 구조를 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.

[0155] 명세서 중 및 청구범위 중 본원에서 사용되는 바, "단수형태"는 달리 명확하게 그 반대로 명시되지 않는 한, "적어도 하나"를 의미하는 것으로 이해되어야 한다.

[0156] 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 경우에는 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.

[0157] 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 초과를 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 해석되어야 한다. 청구범위에서 사용될 때, "본질적으로 ~으로 이루어진"이라는 것은 특허법 분야에서 사용되는 바와 같이, 그의 일반적인 의미를 가져야 한다.

[0158] 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 초과를 A를 포함하는 적어도 하나의 A를 지칭할 수 있고; 또 다른 실시양태에서는, (임의적으로 A 이외의 다른 요소를 포함하면서) A는 존재하지 않고, 임의적으로 1개 초과를 B를 포함하는 적어도 하나의 B를 지칭할 수 있고; 추가의 또 다른 실시양태에서, 임의적으로 1개 초과를 A를 포함하는 적어도 하나의 A, 및 임의적으로 1개 초과를 B를 포함하는 적어도 하나의 B (및 다른 요소를 포함하는

것) 등을 칭할 수 있다.

[0159] 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 또는 반-폐쇄형 전환 어구여야 한다.

[0160] 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.

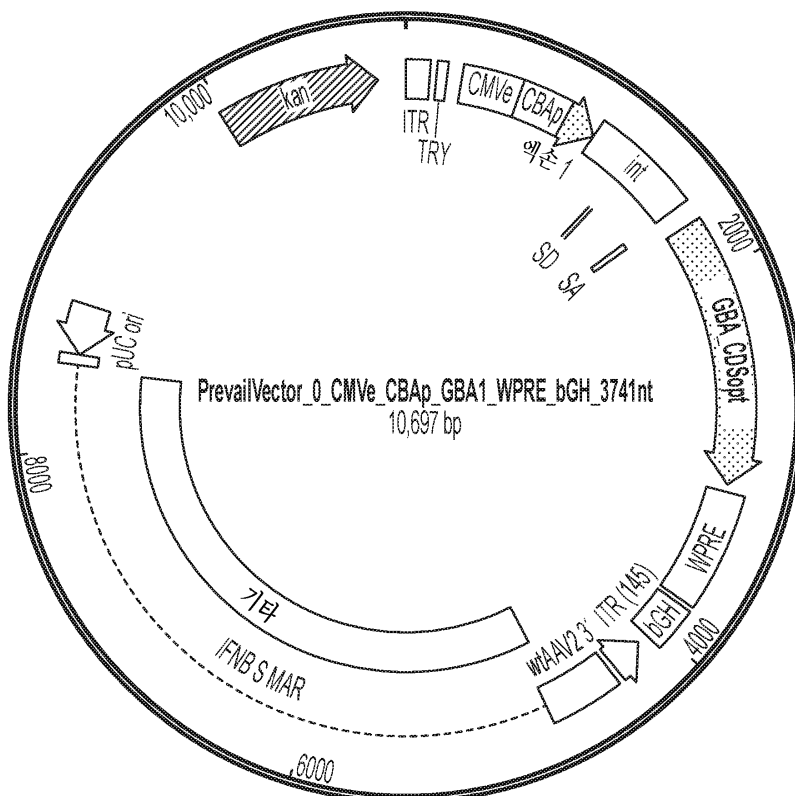
[0161] 명백하게 반대로 명시되지 않는 한, 1개 초과 단계 또는 조치를 포함하는 본원에서 청구되는 임의의 방법에서, 본 방법의 단계 또는 조치의 순서가 반드시 본 방법의 단계 또는 조치가 언급된 순서로 제한되는 것은 아니라는 것도 또한 이해하여야 한다.

[0162] 서열

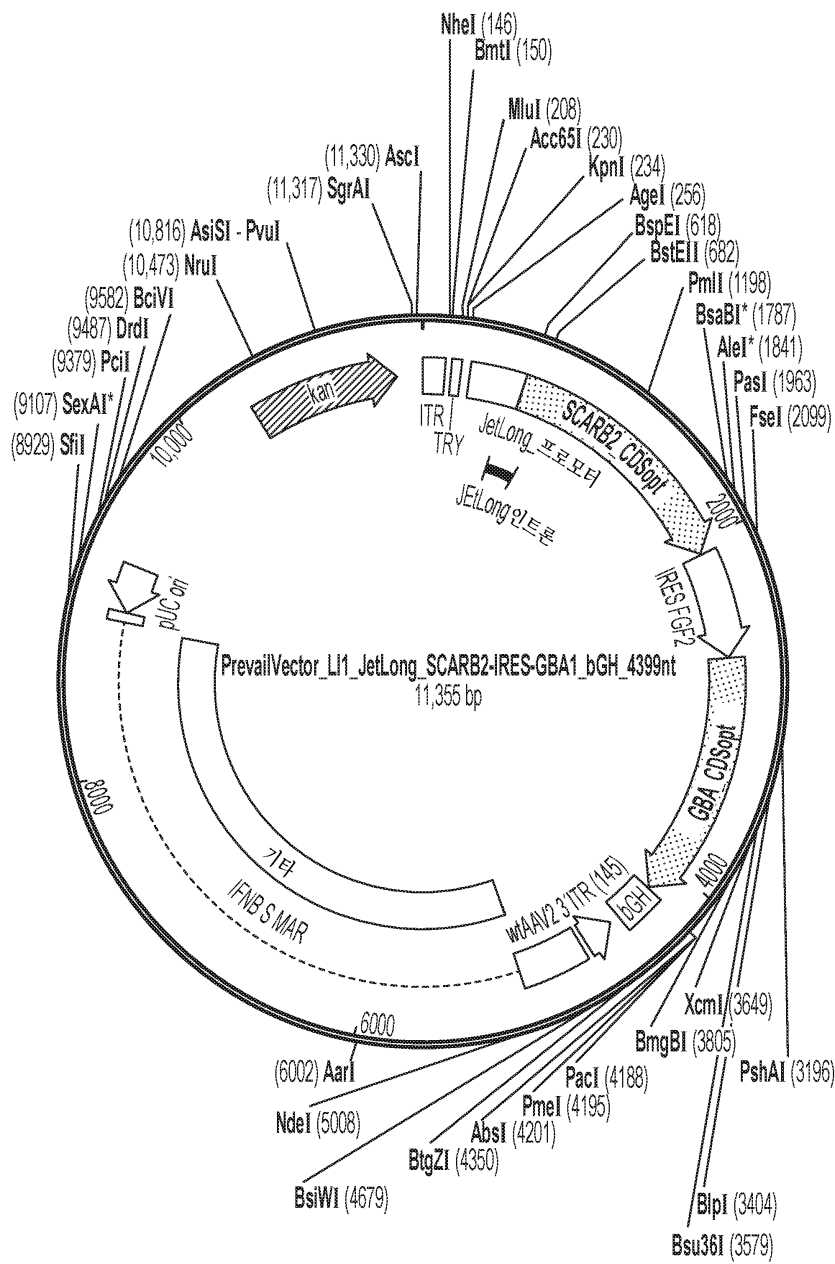
[0163] 일부 실시양태에서, 하나 이상의 유전자 생성물 (예컨대, 제1, 제2 및/또는 제3 유전자 생성물)을 코딩하는 발현 카세트는 서열식별번호 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 또는 25 중 어느 하나에 기재된 서열을 포함하거나 또는 그로 이루어진다 (또는 상기 서열을 갖는 펩티드를 코딩한다). 일부 실시양태에서, 유전자 생성물은 서열식별번호 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 또는 25 중 어느 하나의 일부 (예컨대, 단편)에 의해 코딩된다.

도면

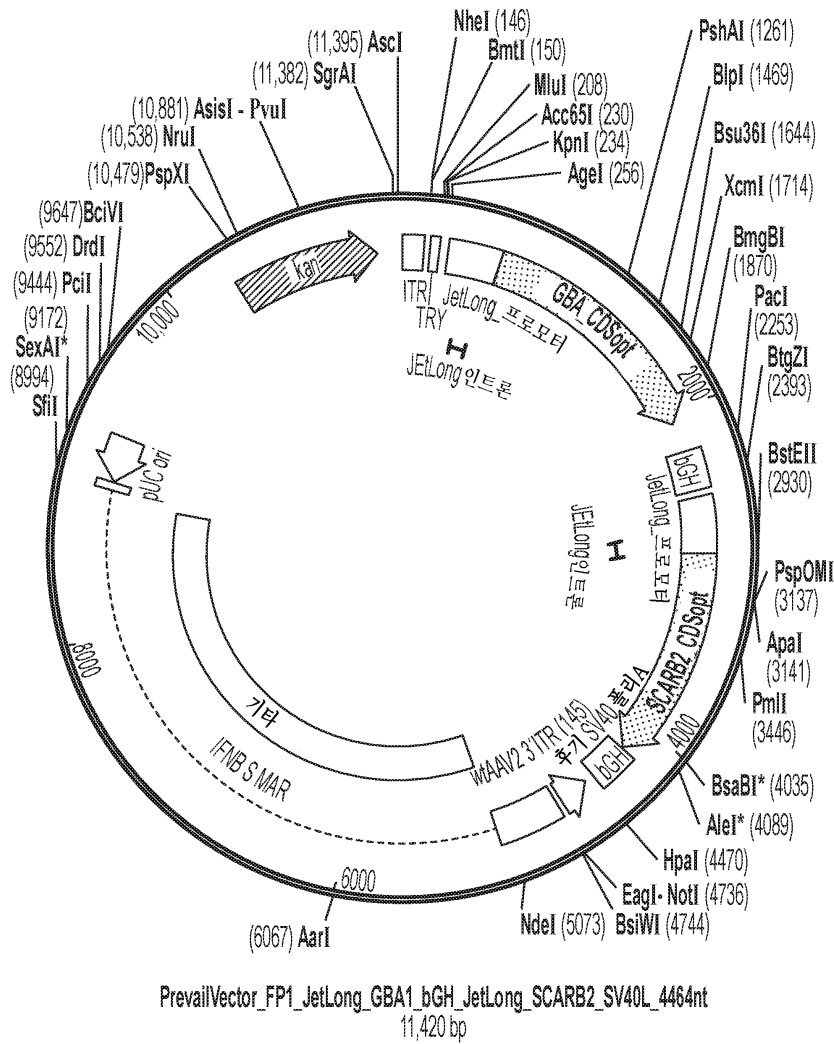
도면1



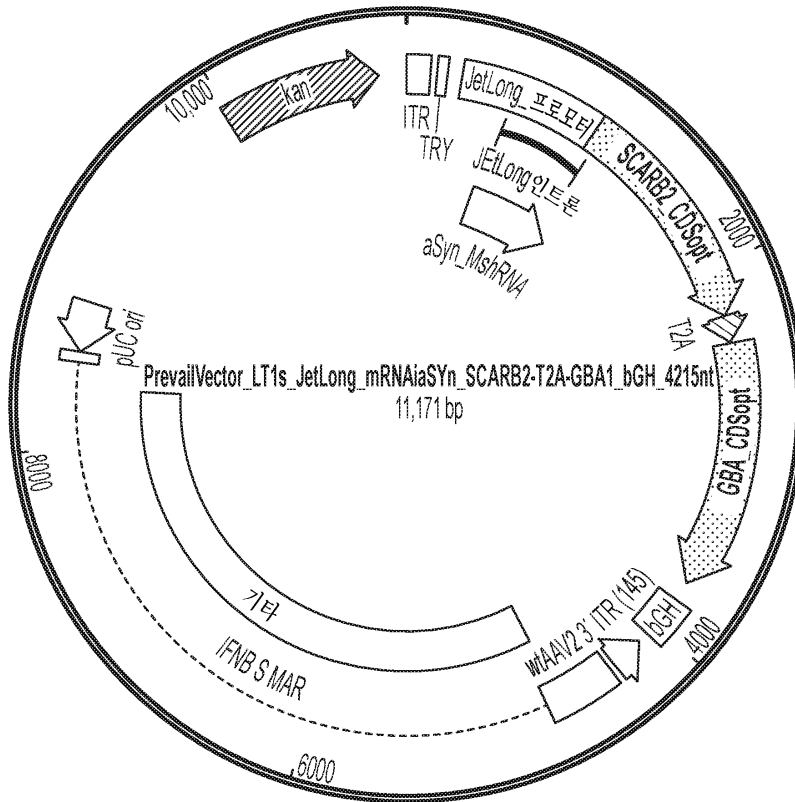
도면2



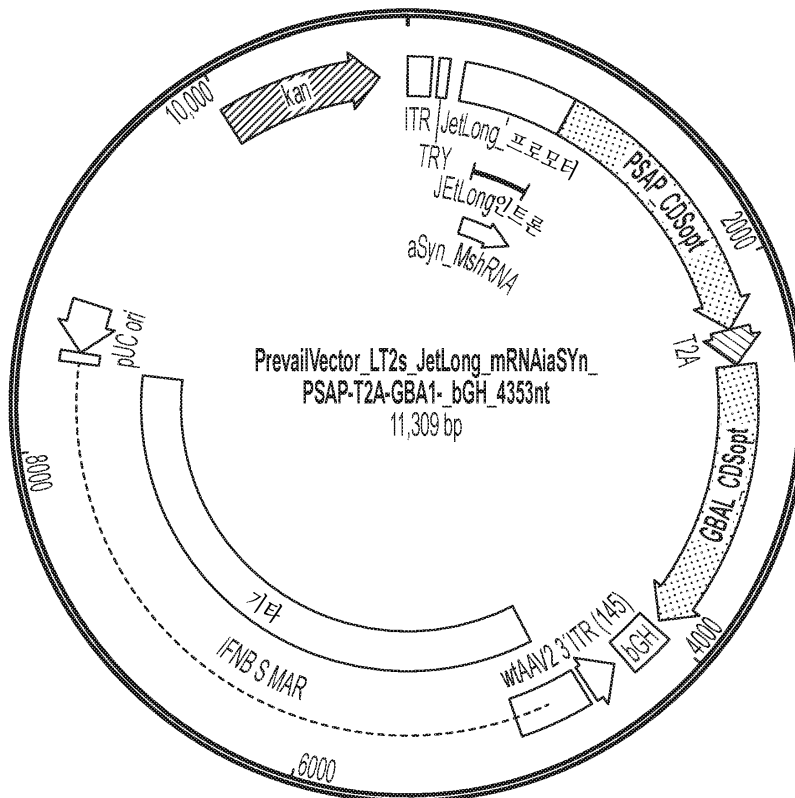
도면3



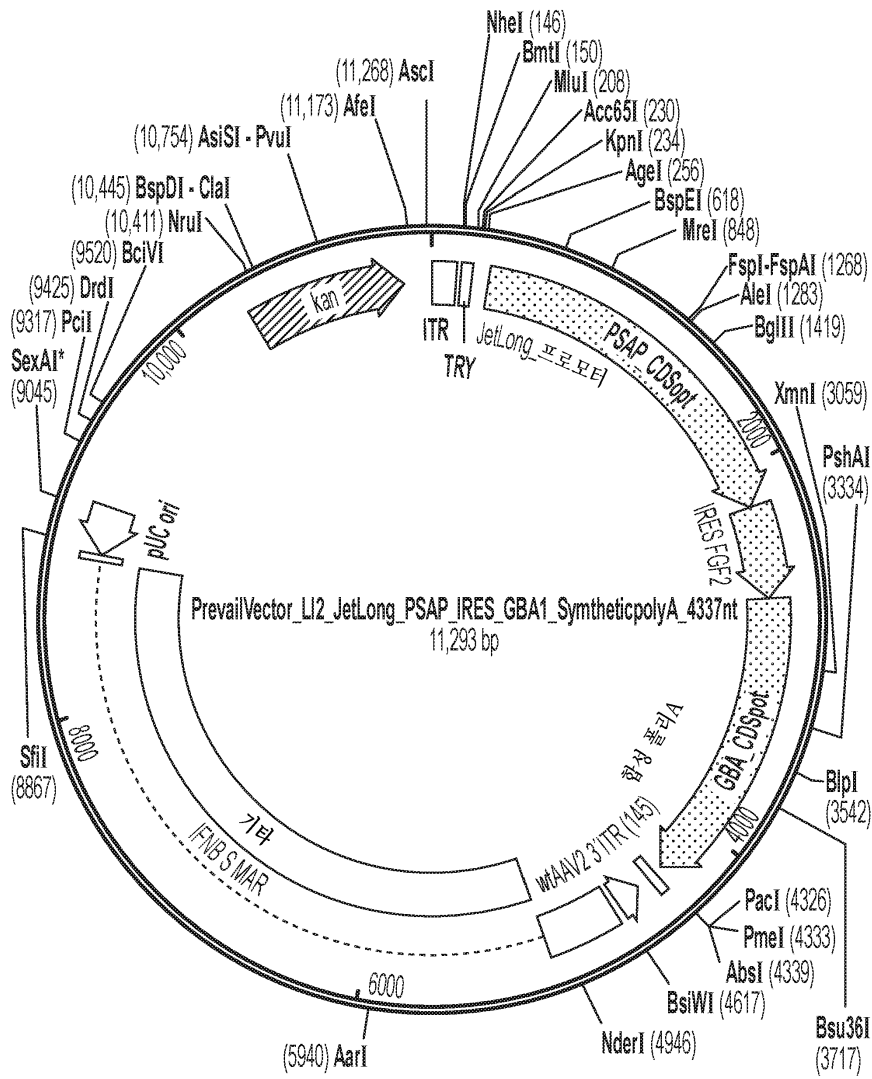
도면4



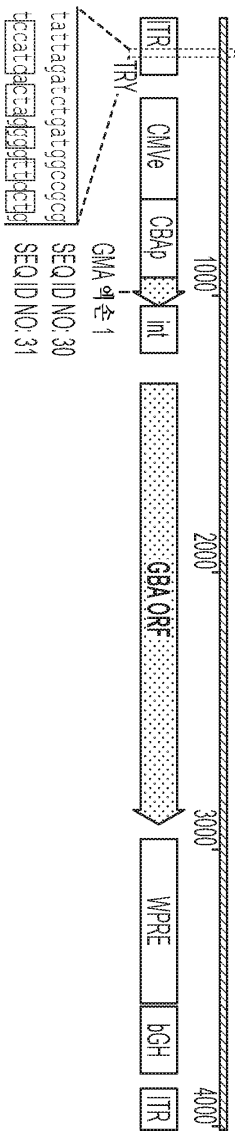
도면5



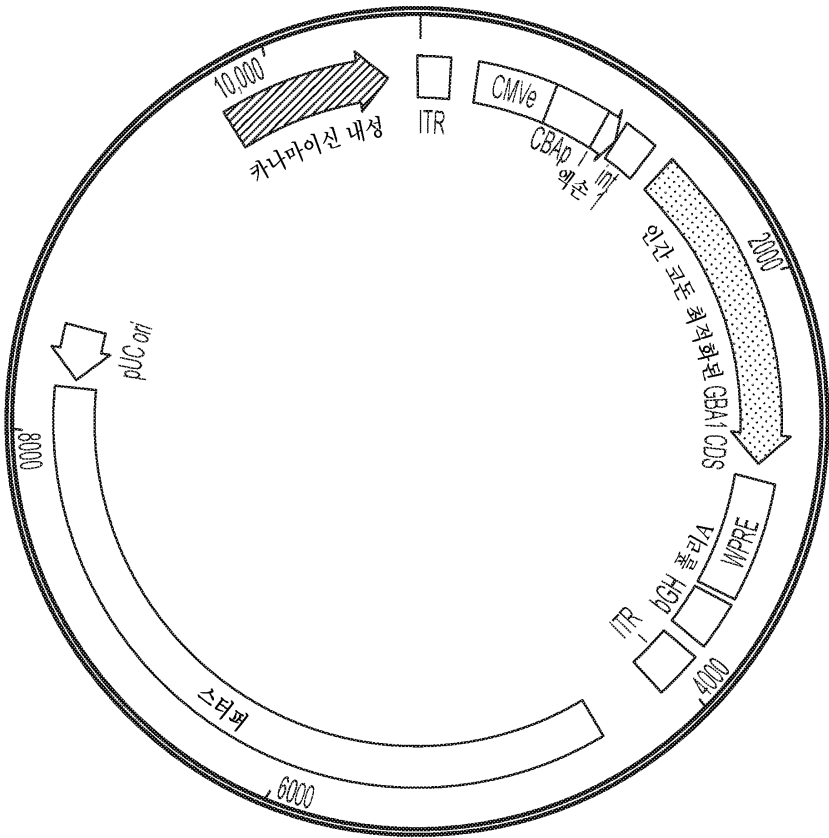
도면6



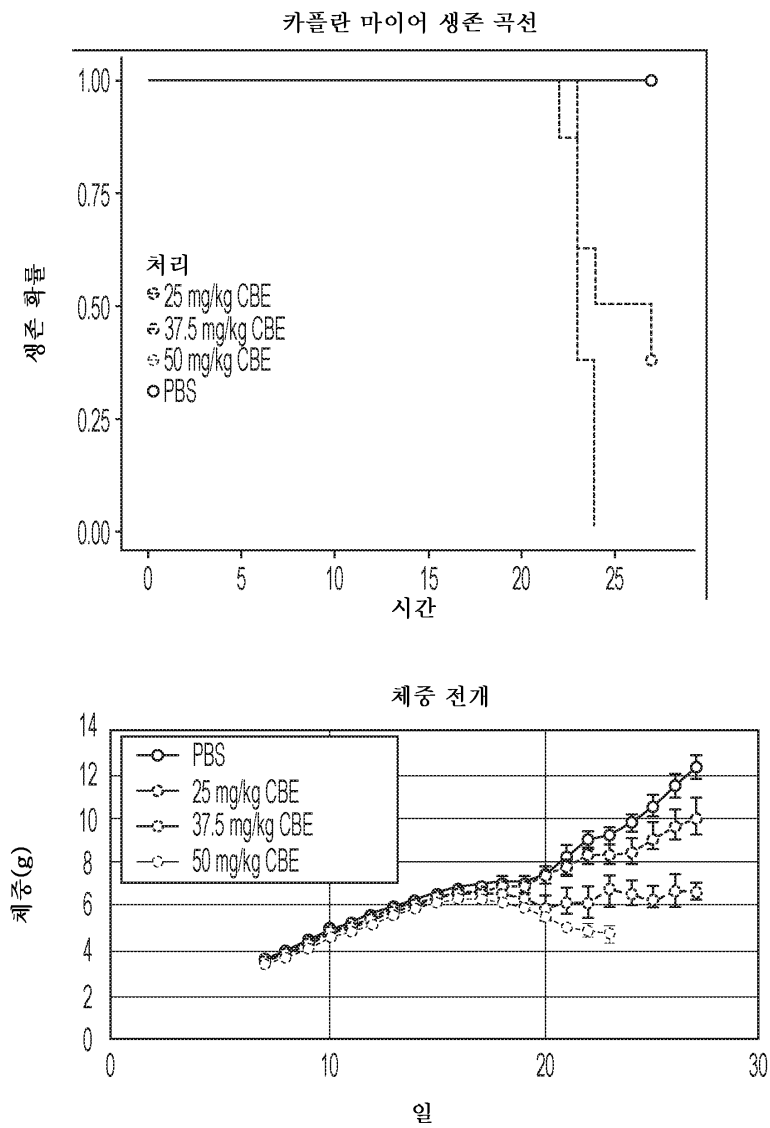
도면7



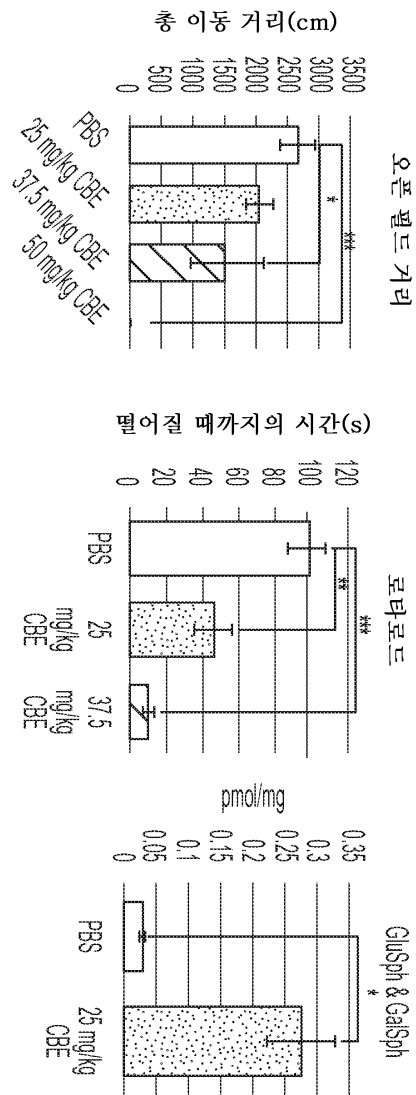
도면8



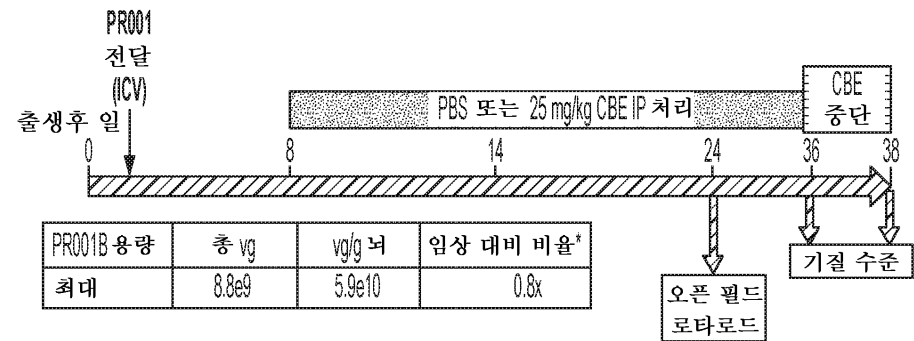
도면9i



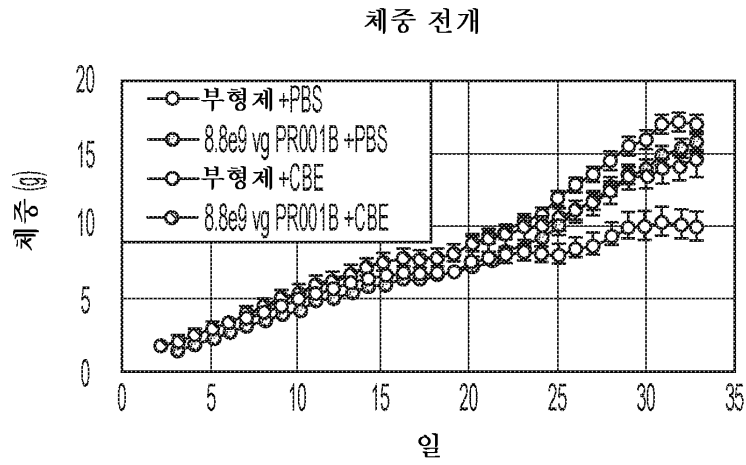
도면9i



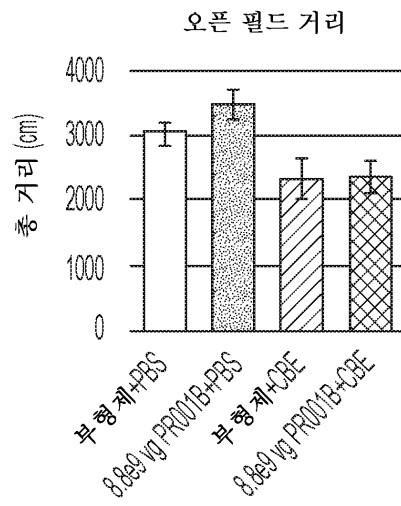
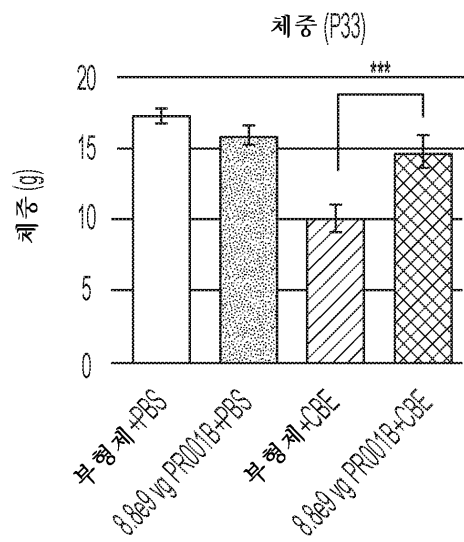
도면10



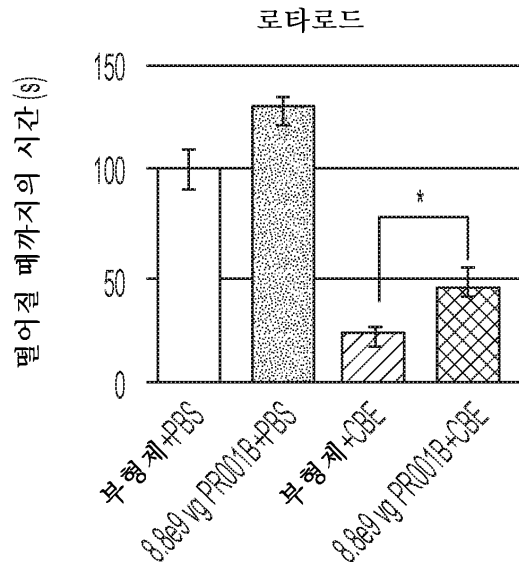
도면11i



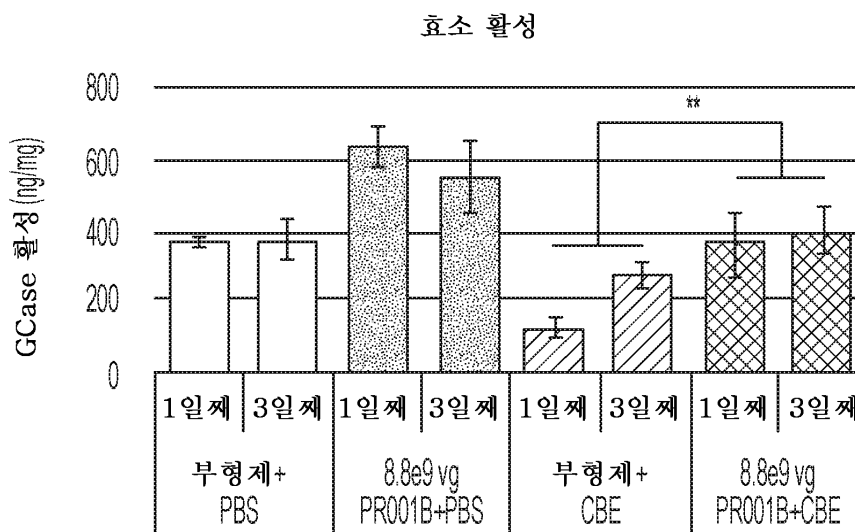
도면11ii



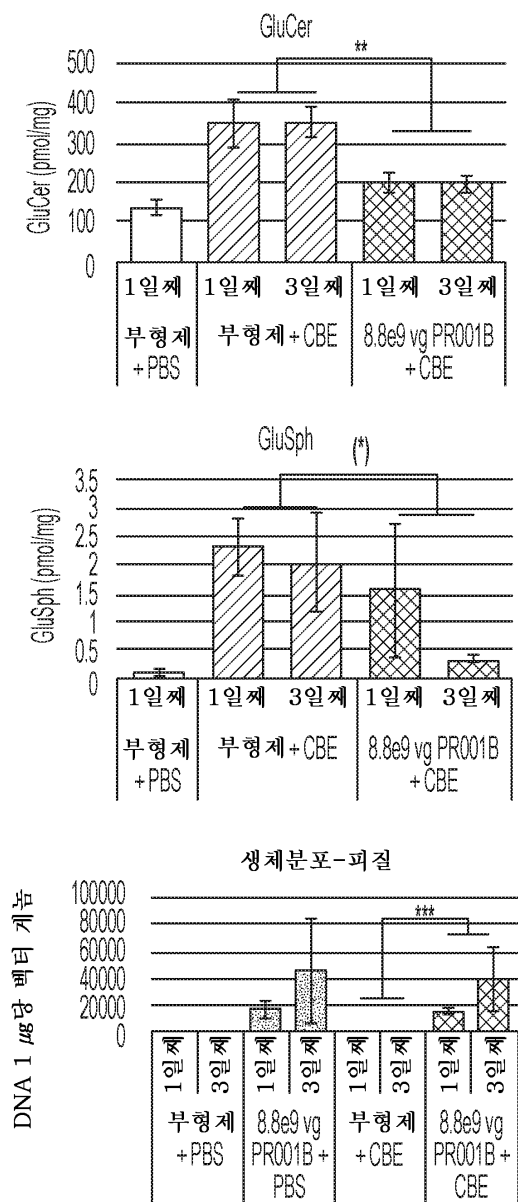
도면11iii



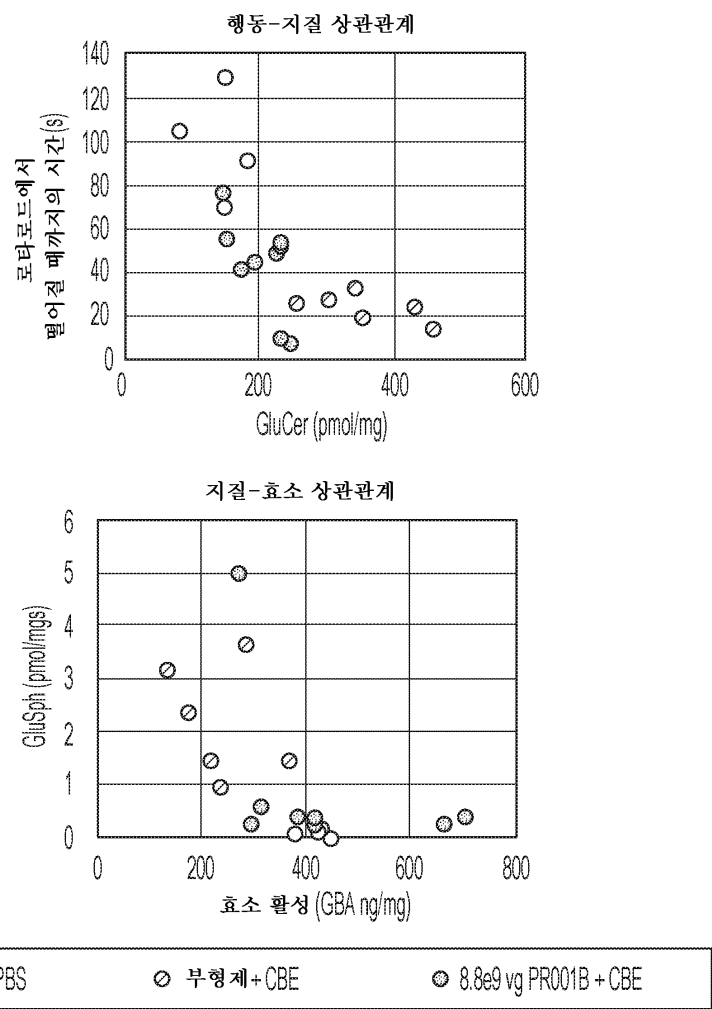
도면12i



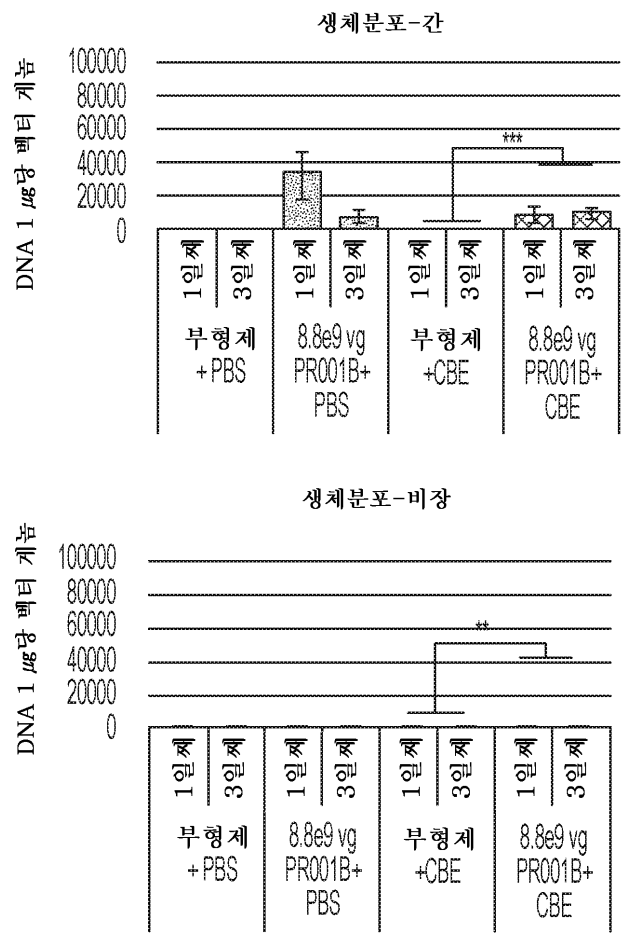
도면12i



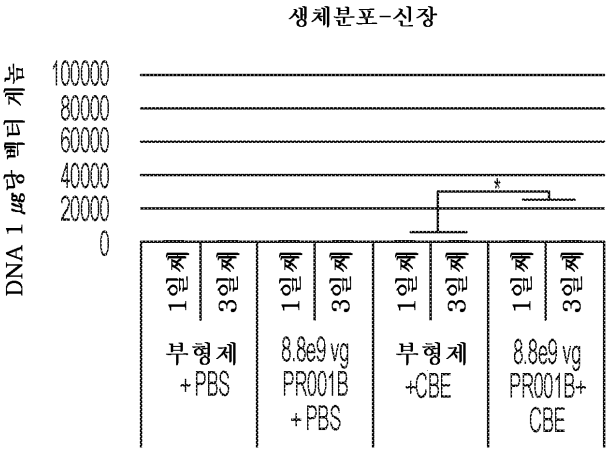
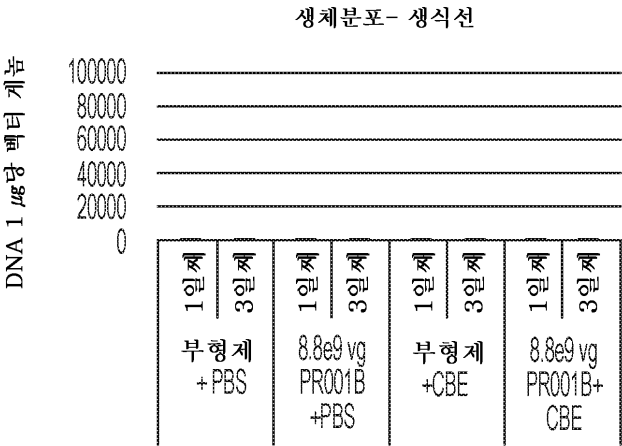
도면13



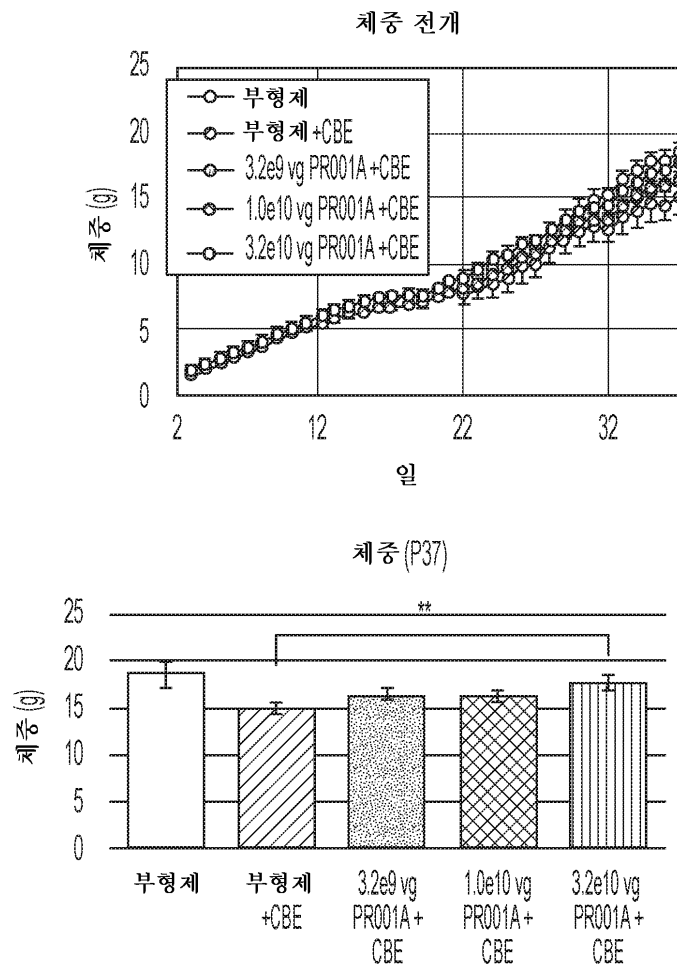
도면14i



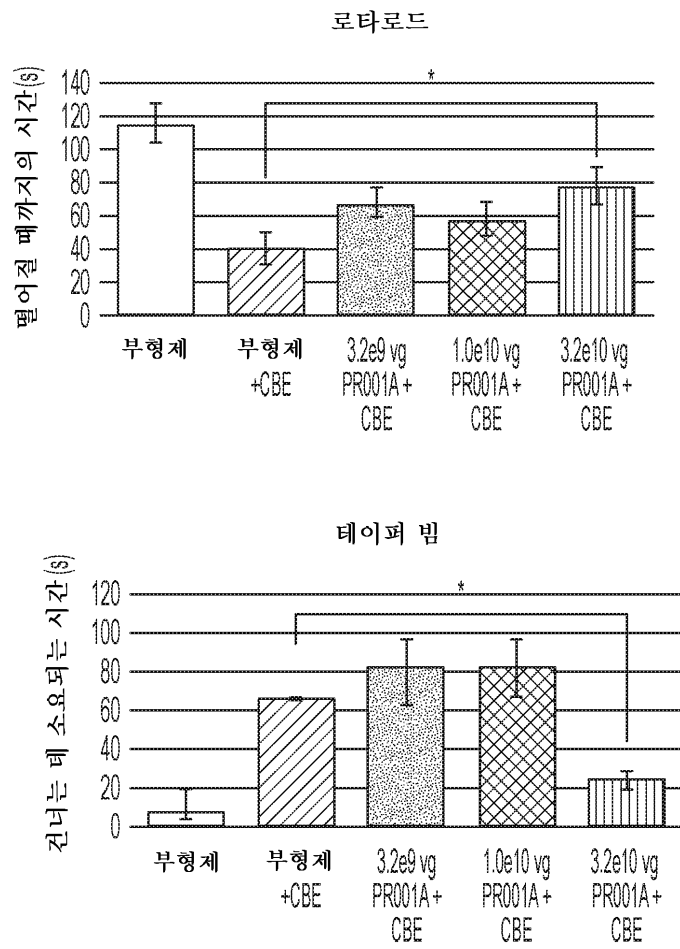
도면14ii



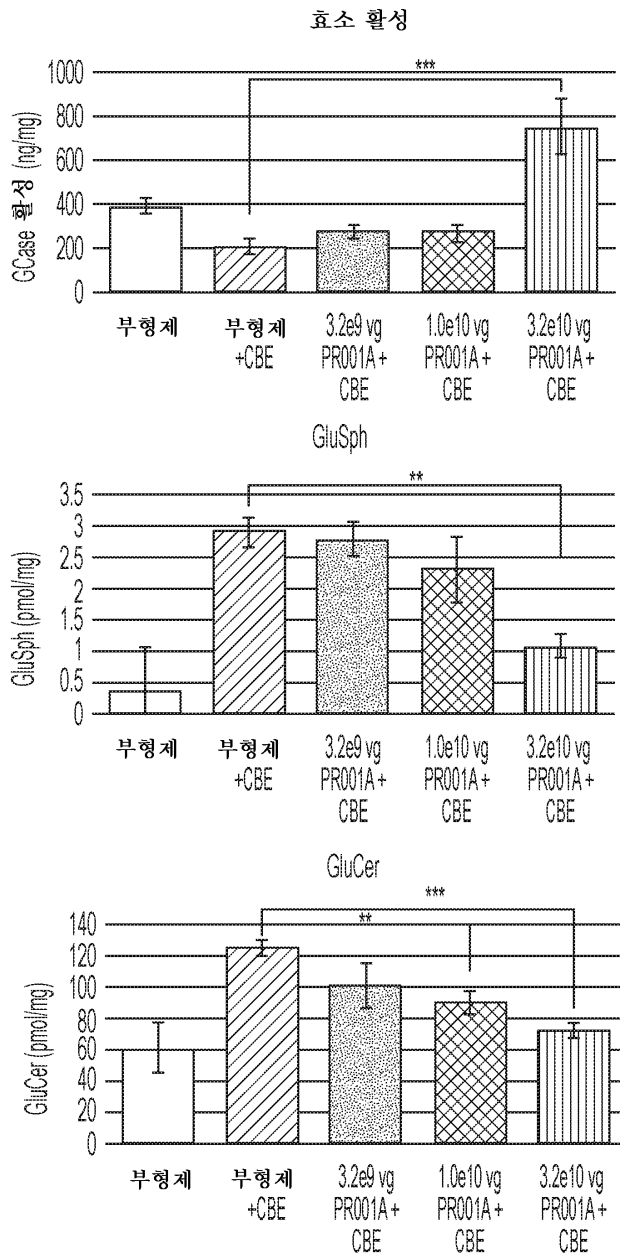
도면15i



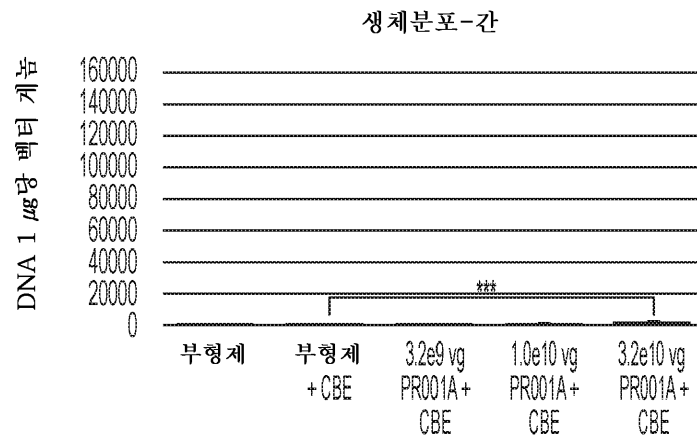
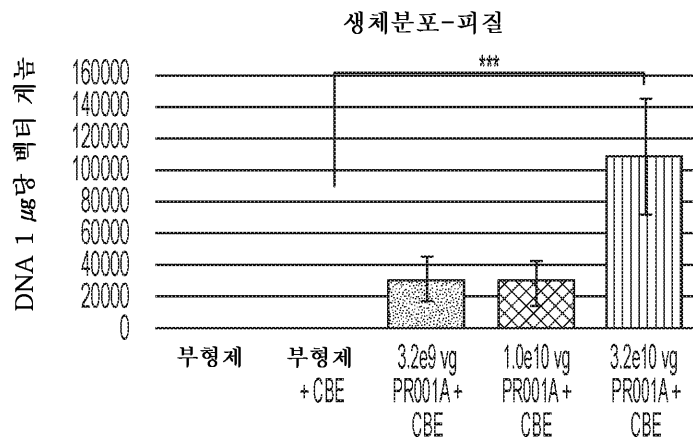
도면15i



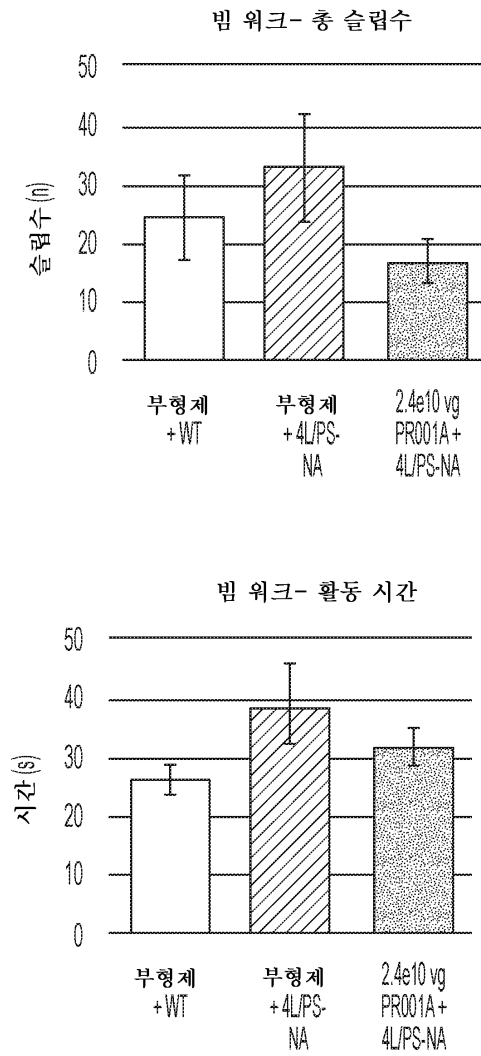
도면16i



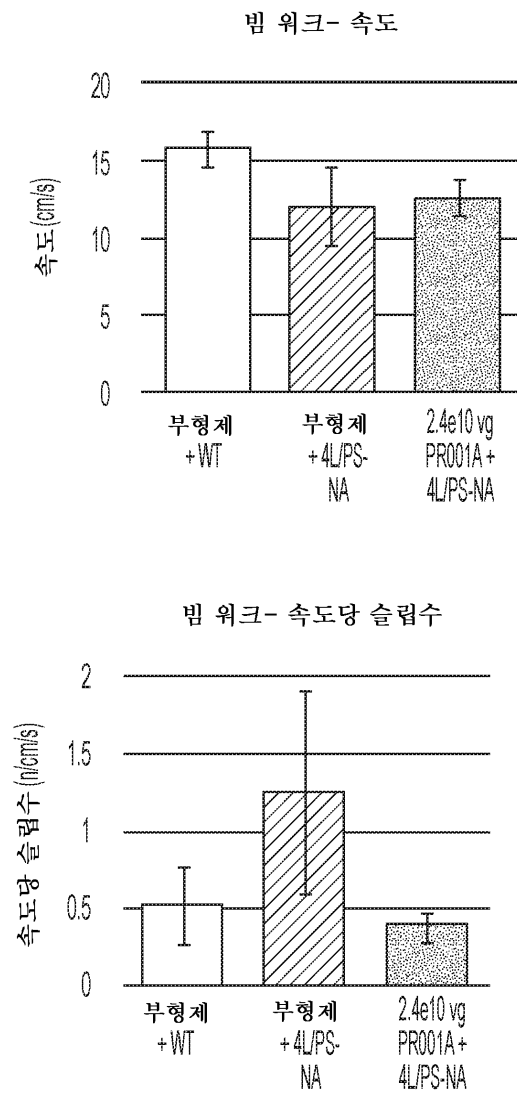
도면16ii



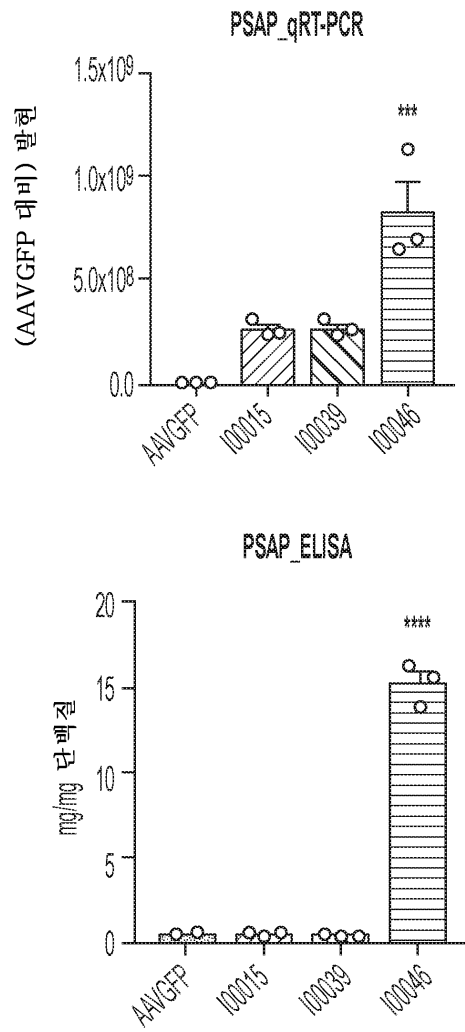
도면17i



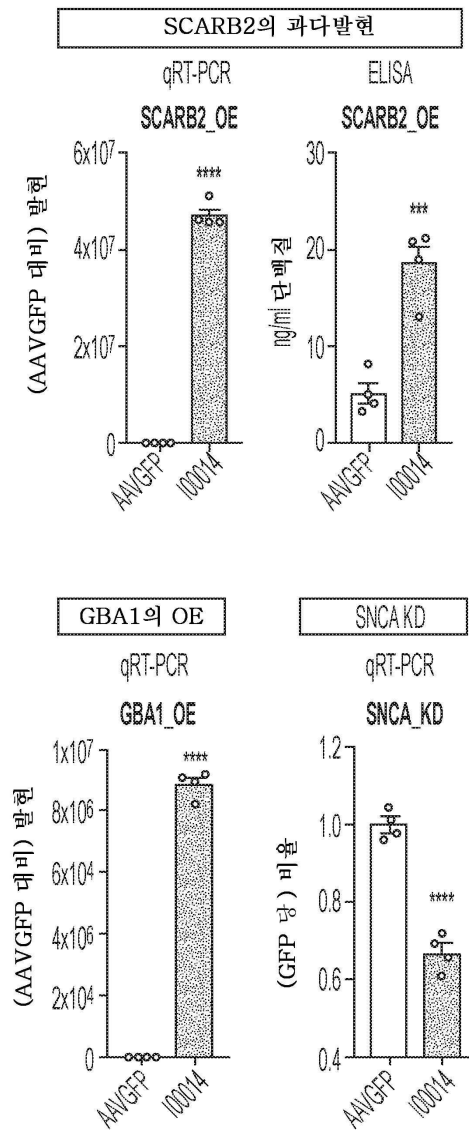
도면17ii



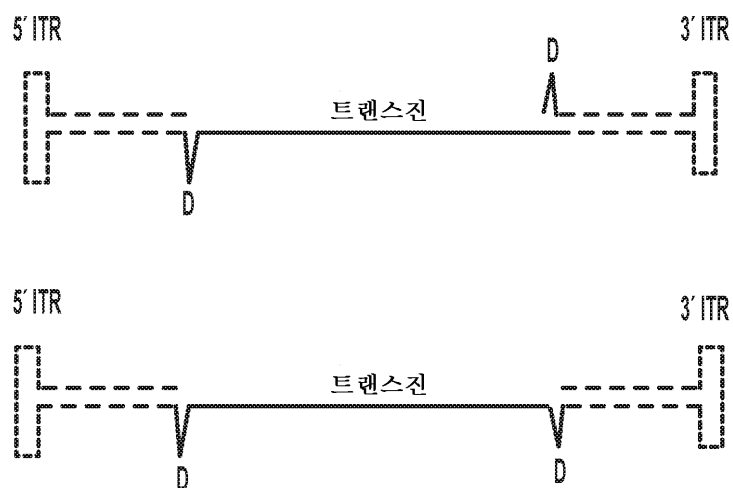
도면18i



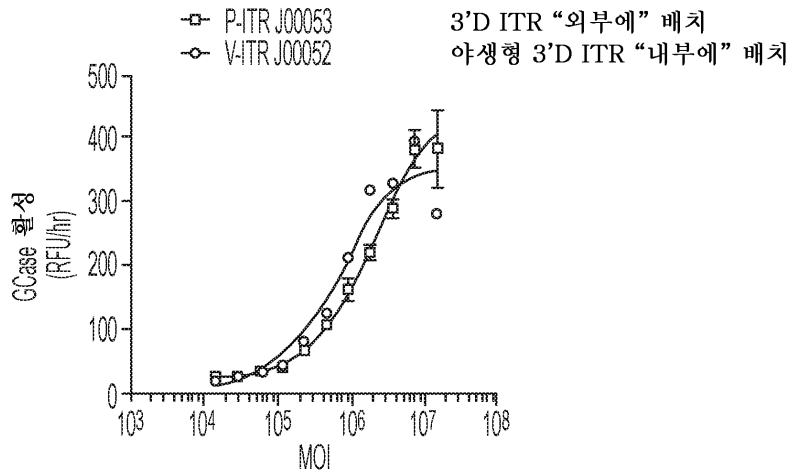
도면18i



도면19



도면20



서열목록

SEQUENCE LISTING

<110> Prevail Therapeutics, Inc.

<120> GENE THERAPIES FOR LYSOSOMAL DISORDERS

<130> P1094.70002W000

<140> Not Yet Assigned

<141> Concurrently herewith

<150> US 62/567,296

<151> 2017-10-03

<160> 31

<170> PatentIn version 3.5

<210> 1

<211> 10697

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 1

ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg 120

gccaaactcca tctactagggg ttcttgctag 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 cccccgcca 420
ttgacgtcaa taatgacgta tgttcccata gtaacgcaa tagggacttt ccattgacgt 480
caatgggtgg aciattttacg gtaaaactgcc cacttggcag tacatcaagt gtatcatatg 540

ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgccag 600
tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgtatt 660
accatggtcg aggtgagccc cacgttctgc ttactctcc ccatctcccc cccctccca 720
cccccaattt tgtattttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg 780
gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg 840
agaggtgctgg cggcagccaa tcagagcggc gcgctccgaa agtttcttt tatggcgagg 900
cggcggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg 960

ctgccttcgc cccgtgcccc gctccgccgc cgctcgcgc cgcccgcccc ggctctgact 1020
gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctcctcgg gctgtaatta 1080
gcgcttgggt taatgacggc ttgttttctg tggctgcgtg aaagccttga ggggctcgg 1140
gagctagagc ctctgctaac catgttcatg ctttcttctt tttctacag ctctgggca 1200
acgtgctggt tattgtgctg tctcatcatt ttggcaaaga attcctcgaa gatccgaagg 1260
gaaagtcctt cagcactgtg ggatccgttc gaagatatca ccggttgagc caccatggaa 1320
ttcagcagcc ccagcagaga ggaatgcccc aagcctctga gccgggtgtc aatcatggcc 1380

ggatctctga caggactgct gctgcttcag gccgtgtctt gggttcttgg cgctagacct 1440
tgcatcccca agagcttcgg ctacagcagc gtcgtgtgcg tgtgcaatgc cacctactgc 1500
gacagcttcg accctctac ctttctgct ctgggcacct tcagcagata cgagagcacc 1560
agatccggca gacggatgga actgagcatg ggacccatcc aggccaatca cacaggcact 1620
ggcctgctgc tgacactgca gcctgagcag aaattccaga aagtgaagg cttcggcgga 1680
gccatgacag atgccgccg tctgaatatc ctggctctgt ctccaccagc tcagaacctg 1740
ctgctcaaga gctacttcag cgaggaaggc atcggtctaca acatcatcag agtgcceatg 1800

gccagctgcg acttcagcat caggacctac acctacgcc acacaccga cgatttcag 1860
ctgcacaact tcagcctgcc tgaagaggac accaagctga agatccctct gatccacaga 1920
gccctgcagc tggcacaaag acccgtgtca ctgctggcct ctccatggac atctcccacc 1980
tggtgaaaa caaatggcgc cgtgaatggc aagggcagcc tgaaaggcca acctggcgac 2040
atctaccacc agacctgggc cagatacttc gtgaagtcc tggacgccta tgccgagcac 2100
aagctgcagt ttggggcgt gacagccgag aacgaacctt ctgctggact gctgagcggc 2160

tacccctttc agtgcctggg ctttacaccc gagcaccagc gggactttat cgcccgtgat	2220
ctgggaccca cactggccaa tagcaccac cataatgtgc ggctgctgat gctggacgac	2280
cagagactgc ttctgcccc ctgggctaaa gtgggtgctga cagatcctga ggccgcaaaa	2340
tacgtgcacg gaatcgccgt gcactgggtat ctggactttc tggcccctgc caaggccaca	2400
ctgggagaga cacacagact gttccccaac accatgctgt tgcgcagcga agcctgtgtg	2460
ggcagcaagt tttgggaaca gagcgtgcgg ctcggcagct gggatagagg catgcagtac	2520
agccacagca tcatcacaa cctgctgtac cacgtcgtcg gctggaccga ctggaatctg	2580
gccctgaatc ctgaaggcgg ccctaactgg gtccgaaact tcgtggacag ccccatcatc	2640
gtggacatca ccaaggacac cttctacaag cagcccatgt tctaccacct gggacacttc	2700
agcaagtcca tccccaggg ctctcagcgc gttagactgg tggcttcca gaagaacgat	2760
ctggacgccg tggtctgat gcacctgat ggatctgctg tgggtggtgt cctgaaccgc	2820
agcagcaaag atgtgccct gaccatcaag gatcccgccg tgggattcct ggaaacaatc	2880
agccctggct actccatcca cacctacctg tggcgtagac agtgacaatt gtaattaag	2940
tttaaaccct cgaggccga agcttatcga taatcaacct ctggattaca aaatttgtga	3000
aagattgact ggtattctta actatgttgc tccttttacg ctatgtggat acgtctcttt	3060
aatgcctttg tatcatgcta ttgcttcccg tatggcttcc attttctcct ccttgtataa	3120
atcctggttg ctgtctcttt atgaggagtt gtggcccggt gtcaggcaac gtggcgtggt	3180
gtgcactgtg tttgctgacg caacccccac tggttggggc attgccacca cctgtcagct	3240
cctttccggg actttcgtt tccccctccc tatigccacg gcggaactca tcgccgctg	3300
ccttgccgcg tgctggacag gggctcggct gttaggcact gacaattccg tgggtgtgtc	3360
ggggaaatca tcgtctttc cttggtgctc cgctgtgtt gccacctgga ttctgcgcgg	3420
gacgtccttc tgctacgtcc cttcgccct caatccagcg gaccttctt cccgcggcct	3480
gctgccggtc ctgcggcctc ttccgctct tgccttcgc cctcagacga gtcggatctc	3540
cctttgggcc gcctccccgc atcgataccg tcgactagag ctgctgac agcctcgact	3600
gtgccttcta gttgccagcc atctgttgtt tgccccctcc ccgtgccttc cttgaccctg	3660
gaaggtgcc accccactgt cctttcctaa taaaatgagg aaattgcac gcattgtctg	3720
agtaggtgtc attctattct ggggggtggg gtggggcagg acagcaagg ggaggattgg	3780
gaagacaata gcaggcatgc tggggagaga tccacgataa caaacagctt ttttggggtg	3840
aacatattga ctgaattccc tgcaggttgg ccactccctc tctgcgcgt cgctcgtca	3900

ctgaggccgc ccgggcaaag cccgggctc gggcgacctt tggtcgccc gcctcagtga	3960
gcgagcgagc gcgcagagag ggagtggcca actccatcac taggggttcc tgcggccgct	4020
cgtacggtct cgaggaattc ctgcaggata acttgccaac ctcatctaa aatgtatata	4080
gaagcccaaa agacaataac aaaaatattc ttgtagaaca aaatgggaaa gaatgttcca	4140
ctaaatatca agatttagag caaagcatga gatgtgtggg gatagacagt gaggctgata	4200
aaatagagta gagctcagaa acagacccat tgatatatgt aagtaccta tgaaaaaat	4260
atggcatttt acaatgggaa aatgatgttc tttttctttt ttagaaaaac agggaaatat	4320
atttatatgt aaaaaataaa agggaacca 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 ctctacttc ttgacaggac acatttgtat caacttaaaa aagcagattt	4740
ttgccagcag aactattcat tcagaggtag gaaacttaga atagatgatg tcaactgatta	4800
gcatggcttc cccatctcca cagctgcttc ccaccagggt tgcccacagt tgagttgtc	4860
cagtgtcag ggctgcccc tctcagtaag aagccccaca ccagccctc tccaaatatg	4920
ttggctgttc ctccattaa agtgaccca ctttagagca gcaagtggat ttctgtttct	4980
tacagttcag gaaggaggag tcagctgtga gaacctggag cctgagatgc ttctaagtcc	5040
cactgctact ggggtcaggg aagccagact ccagcatcag cagtcaggag cactaagccc	5100
ttgccaacat cctgtttctc agagaaactg ctccattat aatggttgtc cttttttaag	5160
ctatcaagcc aaacaaccag tgtctacat tattctcatc acctgaagcc aagggttcta	5220
gcaaaagtca agctgtcttg taatggttga tgtgcctcca gcttctgtct tcagtcactc	5280
cactcttagc ctgctctgaa tcaactctga ccacagtcc ctggagcccc tgccacctgc	5340
tgccctgcc accttctcca tctgcagtgc tgtgcagcct tctgcactct tgcagagcta	5400
ataggtggag acttgaagga agaggaggaa agtttctcat aatagccttg ctgcaagctc	5460
aaatgggagg tgggcactgt gccaggagc cttggagcaa aggtgtgtcc caacctctga	5520
ctgcatccag gtttggctct gacagagata agaagccctg gcttttgag ccaaaatcta	5580
ggtcagactt aggcaggatt ctcaaagttt atcagcagaa catgaggcag aagacccttt	5640
ctgctccagc ttcttcagc tcaaccttca tcagaataga tagaaagaga ggctgtgagg	5700
gttcttaaaa cagaagcaaa tctgactcag agaataaaca acctcctagt aaactacagc	5760

ttagacagag catctgggtg tgagtgtgct cagtgtccta ctcaactgtc tggatcagc	5820
cctcatgagg acttctcttc tttccctcat agacctccat ctctgttttc cttagcctgc	5880
agaaatctgg atggctattc acagaatgcc tgtgttttca gaggttgcatt ttttctctgg	5940
tattctgggt caagcatttg aaggtaggaa aggtttctcca agtgcaagaa agccagccct	6000
gagcctcaac tgcctggcta gtgtggtcag taggatgcaa aggcgtgtga atgccacaag	6060
gccaaacttt aacctgtgta ccacaagcct agcagcagag gcagctctgc tcaactggaac	6120
tctctgtctt ctttctcctg agccttttct tttcctgagt tttctagctc tcctcaacct	6180
tacctctgcc ctaccagga caaacccaag agccactgtt tctgtgatgt cctctccagc	6240
cctaattagg catcatgact tcagcctgac cttccatgct cagaagcagt gctaattcac	6300
ttcagatgag ctgctctatg caacacaggc agagcctaca aacctttgca ccagagccct	6360
ccacatatca gtgtttgttc atactcactt caacagcaaa tgtgactgct gagattaaga	6420
ttttacacaa gatggctctgt aatttcacag ttagttttat cccattaggt atgaaagaat	6480
tagcataatt ccccttaaac atgaatgaat cttagatttt ttaataaata gttttggaag	6540
taaagacaga gacatcagga gcacaaggaa tagcctgaga ggacaaacag aacaagaaag	6600
agtctggaaa tacacaggat gttcttggcc tcctcaaagc aagtgcagc agatagtagc	6660
agcagcccca ggctatcaga gccagtgaa gagaagtagc atgaaagcca cagctctaac	6720
cacctgttc cagagtgaca gacagtcccc aagacaagcc agcctgagcc agagagagaa	6780
ctgcaagaga aagtttctaa tttaggttct gttagattca gacaagtgca ggtcatcctc	6840
tctccacagc tactcacctc tccagcctaa caaagcctgc agtccacact ccaaccctgg	6900
tgtctcacct cctagcctct cccaacatcc tgcctctga ccatcttctg catctctcat	6960
ctcaccatct cccactgtct acagcctact ctgcaacta ccatctcatt tcttgacatc	7020
ctgtctacat cttctgcat actctgcat ctaccatacc acctcttacc atctaccaca	7080
ccatctttta tctccatccc tctcagaagc ctccaagctg aatcctgctt tatgtgttca	7140
tctcagcccc tgcatggaaa gctgaccca gaggcagaac tattcccaga gagcttggcc	7200
aagaaaaaca aaactaccag cctggccagg ctgaggagta gtaagctgca gtgtctgttg	7260
tgttctagct tcaacagctg caggagtcc actctcaaact gctccacatt tctcacatcc	7320
tctgtattct ggtcactacc catcttcaaa gaacagaata tctcacatca gcatactgtg	7380
aaggactagt catgggtgca gctgctcaga gctgcaaagt cattctggat ggtggagagc	7440
ttacaaacat ttcatgatgc tcccccgct ctgatggctg gagcccaatc cctacacaga	7500
ctcctgctgt atgtgttttc ctttactct gagccacagc cagagggcag gcattcagtc	7560
tcctcttcag gctggggctg gggcactgag aactcaccca acaccttgct ctcactcctt	7620

ctgcaaaaca agaaagagct ttgtgctgca gtagccatga agaataaaag gaaggcttta	7680
actaaaaaat gtcagagatt attttcaacc ctttactgtg gatcaccagc aaggaggaaa	7740
cacaacacag agacattttt tccctcaaaa ttatcaaaag aatcactgca ttgtttaaag	7800
agagcaactg aatcaggaag cagagttttg aacatatcag aagttaggaa tctgcatcag	7860
agacaaatgc agtcatggtt gtttctgca taccagccct aatcattaga agcctcatgg	7920
acttcaaaca tcattccctc tgacaagatg ctctagccta actccatgag ataaaataaa	7980
tctgcctttc agagccaaag aagagtccac cagcttcttc tcagtgtgaa caagagctcc	8040
agtcaggtta gtcagtccag tgcagtagag gagaccagtc tgcctcctct aattttcaaa	8100
ggcaagaaga ttgttttacc ctggacacca ggcacaagtg aggtcacaga gctcttagat	8160
atgcagtcct catgagttag gagactaaag cgcatgccat caagacttca gtgtagagaa	8220
aacctcaaaa aaagcctcct cactacttct ggaatagctc agaggccgag gcggcctcgg	8280
cctctgcata aataaaaaaa attagtgcgc catggggcgg agaattggcg gaactggcg	8340
gagttagggg cgggatgggc ggagttaggg gcgggactat ggttctgac taattgagat	8400
gcatgctttg catacttctg cctgtggggg agcctgggga ctttccacac ctggttctg	8460
actaatttag atgcatgctt tgcatacttc tgctgtctgg ggagcctggg gactttccac	8520
accctaactg acacacattc cacagctgca ttaatgaatc ggccaacgcg cggggagagg	8580
cggtttgcgt attgggcgct cttccgcttc ctgcctcact gactcgctgc gctcggctgt	8640
tcggctgcgg cgagcggat cagctcactc aaaggcggta atacggttat ccacagaatc	8700
aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa	8760
aaaggccgcg ttgtcggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa	8820
tcgacgtca agtcagaggt ggcgaaacc gacaggacta taaagatacc aggcgtttcc	8880
ccctggaagc tcctcgtgc gctctcctgt tccgacctg ccgcttacg gatacctgtc	8940
cgctttctc cttcgggaa gcgtggcgct ttctcatagc tcacgtgta ggtatctcag	9000
ttcgggttag gtcgttcgt ccaagctggg ctgtgtgcac gaacccccg ttcagcccga	9060
ccgtgcgcc ttatccgta actatcgtct tgagtccaac ccgtaagac acgacttacc	9120
gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac	9180
agagtctttg aagtgggtgc ctaactacgg ctacactaga agaacagtat ttggtatctg	9240
cgctctgctg aagccagtta cttcggaaa aagagtgggt agctcttgat ccggcaaaaa	9300
aaccaccgct ggtagcgtg gttttttgt ttgcaagcag cagattacgc gcagaaaaaa	9360

aggatctcaa gaagatcctt tgatcttttc tacggggtct gacgtcagtg ggaacgaaaa 9420

ctcacgttaa gggatttttg tcatgagatt atcaaaaagg atcttcacct agatcctttt 9480

aaattaaaaa tgaagtttta aatcaatcta aagtatatat gagtaaacctt ggtctgacag 9540

ttaccaatgc ttaatcagtg aggcacctat ctacgcatc tgtctatttc gttcatccat 9600

agttgcctga ctctgcaaa ccacgttggtg 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 atgatgcag gttactcacc 10020

actgcgatcc ccgggaaaaac agcattccag gtattagaag aatatcctga ttcaggtgaa 10080

aatattgttg atgcgtggc agtgttctg cgccggttg attcgattcc tgtttgtaat 10140

tgctctttta acagcgatcg cgtatttcgt ctgcgtcagg cgcaatcacg aatgaataac 10200

ggttttggtg atgcgagtga ttttgatgac gagcgtaatg gctggcctgt tgaacaagtc 10260

tggaaagaaa tgcataagct tttgccattc tcaccgatt cagtcgtcac tcatggtgat 10320

ttctcacttg ataaccttat ttttgacgag gggaaattaa taggttgat tgatgttgga 10380

cgagtcggaa tcgcagaccg ataccaggat cttgccatcc tatggaactg cctcggtag 10440

ttttctcctt cattacagaa acggcttttt caaaaatatg gtattgataa tcttgatatg 10500

aataaattgc agtttcattt gatgctgat gagtttttct aagggcggcc tgccaccata 10560

cccacgccga aacaagcgt catgagcccg aagtggcgag cccgatcttc cccatcggtg 10620

atgtcggcga tataggcgcc agcaaccgca cctgtggcgc cggatgatgag ggcgcgccaa 10680

gtcgacgtcc ggcagtc 10697

<210> 2

<211> 11355

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 2

ttggccactc cctctctgcg cgctcgtcgtg ctactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccc ggccgcctca gtgagcgagc gagcgcgag agaggagtg 120

gccaaactca tctactagggg ttctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaaagtgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagtccg 420

tgcagccggg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcaacttgta 480

agtactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctgcag ccctaggaat gcacttagac aattgtacta accttcttct 600

ctttcctctc ctgacagtcc ggaaagccac catgggccgc tgcgtcttct acaccgccgg 660

caccctgagc ctgctgctgc tggtgaccag cgtgaccctg ctggtggccc gcgtgttcca 720

gaaggccgtg gaccagagca tcgagaagaa gatcgtgctg cgcaacggca ccgaggcctt 780

cgacagctgg gagaagcccc ccctgcccgt gtacaccag ttctacttct tcaacgtgac 840

caaccccgag gagatcctgc gcggcgagac cccccgctg gaggaggtgg gcccctacac 900

ctaccgcgag ctgcgcaaca aggccaacat ccagttcggc gacaacggca ccaccatcag 960

cgccgtgagc aacaaggcct acgtgttcga gcgcgaccag agcgtgggcg accccaagat 1020

cgacctgac cgcaccctga acatccccgt gctgaccgtg atcgagtgga gccaggtgca 1080

cttctgccc gagatcatcg aggccatgct gaaggcctac cagcagaagc tgttcgtgac 1140

ccacaccgtg gacgagctgc tgtggggctc caaggacgag atcctgagcc tgatccacgt 1200

gttccgccc gacatcagcc cctacttcgg cctgttctac gagaagaacg gcaccaacga 1260

cggcgactac gtgttctga ccggcgagga cagctacctg aacttcacca agatcgtgga 1320

gtggaacggc aagaccagcc tggactgggt gatcaccgac aagtgaaca tgatcaacgg 1380

caccgacggc gacagcttcc acccctgat caccaaggac gaggtagctgt acgtgttccc 1440

cagcgacttc tgccgcagcg tgtacatcac cttcagcgac tacgagagcg tgcagggcct 1500

gcccgccttc cgctacaagg tgcccgcga gatcctggcc aacaccagcg acaacgccgg 1560

cttctgcatc cccgagggca actgcctggg cagcggcgtg ctgaacgtga gcatctgcaa 1620

gaacggcgcc cccatcatca tgagcttccc ccacttctac caggccgacg agcgcttcgt 1680

gagcgccatc gagggcattg accccaacca ggaggaccac gagaccttcg tggacatcaa 1740

ccccctgacc ggcatcatcc tgaaggccgc caagccttc 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
gaggctgggtg ggtgtggggg gtggagatgt agaagatgtg acgcccgggc ccggcgggtg 2220
ccagattagc ggacgcgggtg cccgcggttg caacgggatc ccgggcgctg cagcttggga 2280
ggcggctctc ccagggcggc gtccgcggag acacccatcc gtgaaccca ggtcccgggc 2340

cgccggtcgc ccgcgcacca ggggcccggc gacagaagag cggccgagcg gctcagggt 2400
gggggaccgc gggcgcggcc gcgcgtgcc gggcgggagg ctggggggcc ggggcccggg 2460
ccgtgccccg gagcgggtcg gaggccgggg ccggggccgg gggacggcgg ctccccgcgc 2520
ggctccagcg gctcggggat cccgccggg cccgcaggg accatgatgg aattcagcag 2580
ccccagcaga gaggaatgcc ccaagcctct gagccgggtg tcaatcatgg ccgatctct 2640
gacaggactg ctgctgttc aggccgtgtc ttgggttct ggcgctagac cttgcatccc 2700
caagagcttc ggctacagca gcgtcgtgtg cgtgtgcaat gccacctact gcgacagctt 2760

cgacctcct acctttctg ctctgggcac cttcagcaga tacgagagca ccagatccgg 2820
cagacggatg gaactgagca tgggacccat ccaggccaat cacacaggca ctggcctgct 2880
gtgacactg cagcctgagc agaaattcca gaaagtgaag ggcttcggcg gagccatgac 2940
agatgccgcc gctctgaata tcttggtct gtctccacca gctcagaacc tgctgtcaa 3000
gagctacttc agcgaggaag gcatcggtc caacatcatc agagtgccca tggccagctg 3060
cgacttcagc atcaggacct acacctacgc cgacacaccc gacgatttcc agctgcacaa 3120
cttcagcctg cctgaagagg acaccaagct gaagatccct ctgatccaca gagccctgca 3180

gtcggcacia agaccctgt cactgctggc ctctccatgg acatctccca cctggtgaa 3240
aacaatggc gccgtgaatg gcaagggcag cctgaaaggc caacctggcg acatctacca 3300
ccagacctgg gccagatact tcgtgaagt cctggacgcc tatgccgagc acaagctgca 3360
gttttgggc gtgacagccg agaacgaacc ttctgctgga ctgctgagcg getaccctt 3420
tcagtgcctg ggctttacac ccgagcacca gcgggacttt atgccccgtg atctgggacc 3480
cacactggcc aatagcacc accataatgt gcggctgctg atgctggacg accagagact 3540
gtttctgccc cactgggcta aagtgtgtg gacagatcct gaggccgcca aatacgtgca 3600

cggaatgcc gtgcactggt atctggactt tctggccct gccaaggcca cactgggaga 3660
gacacacaga ctgttccca acacatgct gtctgccagc gaagcctgtg tgggcagcaa 3720
gttttgggaa cagagcgtgc ggctcggcag ctgggataga ggcatgcagt acagccacag 3780

catcatcacc aacctgctgt accacgtcgt cggctggacc gactggaatc tggccctgaa	3840
tcctgaaggc ggcctaact gggccgaaa ctctgtggac agcccatca tcgtggacat	3900
caccaaggac accttctaca agcagcccat gttctaccac ctgggacact tcagcaagtt	3960
catccccgag ggctctcagc gcgttggact ggtggcttcc cagaagaacg atctggacgc	4020
cgtaggctctg atgcacctg atggatctgc tgtggtggtg gtcctgaacc gcagcagcaa	4080
agatgtgccc ctgacatca aggatccgc cgtgggattc ctggaacaa tcagccctgg	4140
ctactccatc cacacctacc tgtggcgtag acagtgacaa ttgttaatta agtttaaac	4200
ctcagaggcc caagccgcat cgataccgtc gactagagct cgctgacag cctcgactgt	4260
gccttctagt tgccagccat ctgttgtttg cccctcccc gtgccttctt tgaccctgga	4320
aggtgccact cccactgtcc ttctctaata aatgaggaa attgcatcgc attgtctgag	4380
taggtgtcat tctattctgg ggggtgggt ggggcaggac agcaagggg aggattggga	4440
agacaatagc aggcattctg gggagagatc cacgataaca aacagctttt ttggggtgaa	4500
catattgact gaattccctg caggttggcc actccctctc tgcgcgtcg ctgcctcact	4560
gaggccgccc gggcaagcc cgggcgtcgg gcgaccttg gtcgcccgcc ctgagtgcg	4620
gagcgagcgc gcagagagg agtgccaac tccatcacta ggggttctg cggccgctcg	4680
tacggtctcg aggaattctt gcaggataac ttgccaacct cattctaaaa tgtatataga	4740
agcccaaaag acaataacaa aaatatctt gtagaacaaa atgggaaaga atgttcact	4800
aaatatcaag atttagagca aagcatgaga tgtgtgggga tagacagtga ggctgataaa	4860
atagagtaga gtcagaaac agaccattg atatatgtaa gtgacctatg aaaaaatat	4920
ggcattttac aatgggaaaa tgatggtctt ttctttttt agaaaaacag ggaaatatat	4980
ttatatgtaa aaaaataaag ggaaccata tgcatacca tacacacaaa aaaattccag	5040
tgaattataa gtctaaatgg agaaggcaaa actttaaatc ttttagaaaa taatatagaa	5100
gcatgcagac cagcctggcc aacatgatga aacctctct actaataata aaatcagtag	5160
aactactcag gactactttg agtgggaagt cttttctat gaagacttct ttggccaaaa	5220
ttaggtctta aatgcaagga gatagtgc atgcctggc tgcacttact gataaatgat	5280
gttatcacca tctttaacca aatgcacagg aacaagtat ggtactgatg tgctggattg	5340
agaaggagct ctacttctt gacaggacac atttgtatca acttaaaaa gcagattttt	5400
gccagcagaa ctattcttc agaggtagga aacttagaat agatgatgic actgattagc	5460
atggcttccc catctccaca gctgcttccc acccaggttg cccacagttg agtttgtcca	5520
gtgctcaggg ctgcccactc tcagtaagaa gccccacac agccctctc caaatatgtt	5580
ggctgttctt tccattaaag tgacccact ttagagcagc aagtggattt ctgtttctta	5640

cagttcagga aggaggagtc agctgtgaga acctggagcc tgagatgctt ctaagtccea 5700

ctgctactgg ggtcaggga gccagactcc agcatcagca gtcaggagca ctaagccctt 5760

gccaacatcc tgtttctcag agaaactgct tccattataa tggttgtcct tttttaagct 5820

atcaagccaa acaaccagtg tctaccattt tttctatcac ctgaagccaa gggttctagc 5880

aaaagtcaag ctgtcttgta atgggtgatg tgcctccagc tttctgtctt agtactcca 5940

ctcttagcct gctctgaatc aactctgacc acagttccct ggagccctg ccacctgctg 6000

cccctgccac ctctccatc tgcagtgtg tgcagcctt tgcactcttg cagagctaat 6060

aggtggagac ttgaaggaag aggaggaaag tttctcataa tagccttgct gcaagctcaa 6120

atgggagggtg ggcactgtgc ccaggagcct tggagcaaag gctgtgcca acctctgact 6180

gcatccaggt ttggtcttga cagagataag aagccctggc ttttgagcc aaaatctagg 6240

tcagacttag gcagattct caaagtttat cagcagaaca tgaggcagaa gaccctttct 6300

gctccagctt ctccaggctc aaccttcac agaataagata gaaagagagg ctgtgagggt 6360

tcttaaaaca gaagcaaatc tgactcagag aataaacaac ctctagtaa actacagctt 6420

agacagagca tctgggtgtg agtgtgtca gtgtcctact caactgtctg gtatcagccc 6480

tcatgaggac ttctcttctt tccctcatag acctccatct ctgttttct tagcctgcag 6540

aaatctggat ggctattcac agaatgctg tgtttcaga gttgcatttt ttctctggta 6600

ttctgttca agcatttgaa ggtaggaaag gttctccaag tgcaagaaag ccagccctga 6660

gcctcaactg cctggctagt gtggtcagta ggatgcaaag gctgttgaat gccacaaggc 6720

caaaacttta cctgtgtacc acaagcctag cagcagaggc agctctgtc actggaactc 6780

tctgtcttct ttctcctgag cttttcttt tctgagttt tctagctctc ctcaacctta 6840

cctctgccct acccaggaca aaccaagag ccaactgttc tgtgatgtcc tctccagccc 6900

taattaggca tcatgaactc agcctgacct tccatgtca gaagcagtgc taatccactt 6960

cagatgagct gctctatgca acacaggcag agcctacaaa cctttgcacc agagccctcc 7020

acatatcagt gtttgttcat actcacttca acagcaaatg tgactgtga gattaagatt 7080

ttacacaaga tggctctgta tttcacagtt agttttatcc cattaggtat gaaagaatta 7140

gcataattcc ctttaaacat gaatgaatct tagatttttt aataaatagt ttggaagta 7200

aagacagaga catcaggagc acaaggaata gctgagagg acaaacagaa caagaaagag 7260

tctggaata cacaggatgt tcttgccctc ctcaaagcaa gtgcaagcag atagtaccag 7320

cagccccagg ctatcagagc ccagtgaaga gaagtacat gaaagccaca gctctaacca 7380

ccctgttcca gaggtagaca cagtcaccaa gacaagccag cctgagccag agagagaact 7440
gcaagagaaa gtttctaatt taggttctgt tagattcaga caagtgcagg tcatcctctc 7500
tccacagcta ctcacctctc cagcctaaca aagcctgcag tccacactcc aaccctggtg 7560
tctacactcc tagcctctcc caacatcctg ctctctgacc atcttctgca tctctcatct 7620
caccatctcc cactgtctac agcctactct tgcaactacc atctcatttt ctgacatcct 7680
gtctacatct tctgccatac tctgccatct accataccac ctcttaccat ctaccacacc 7740
atcttttacc tccatccctc tcagaagcct ccaagctgaa tccgtcttta tgtgttcatc 7800

tcagcccttg catggaaagc tgaccccaaga ggacagaacta ttccacagaga gcttggccaa 7860
gaaaaacaaa actaccagcc tggccaggct caggagtagt aagctgcagt gtctgttgtg 7920
ttctagcttc aacagctgca ggagttccac tctcaaatgc tccacatttc tcacatcctc 7980
ctgattctgg tcaactacca tcttcaaga acagaatata tcacatcagc atactgtgaa 8040
ggactagtca tgggtgcagc tgctcagagc tgcaaagtca ttctggatgg tggagagctt 8100
acaaacattt catgatgctc cccccgtct gatggctgga gcccaatccc tacacagact 8160
cctgctgtat gtgttttctt ttcactctga gccacagcca gagggcaggc attcagtctc 8220

ctcttcaggc tggggctggg gcactgagaa ctcaccaac acctgtctct cactccttct 8280
gcaaaacaag aaagagcttt gtgctgcagt agccatgaag aatgaaagga aggccttaac 8340
taaaaaatgt cagagattat tttaacccc ttactgtgga tcaccagcaa ggaggaaaca 8400
caacacagag acattttttc ccctcaaat atcaaaagaa tcactgcatt tgttaaagag 8460
agcaactgaa tcaggaagca gagttttgaa catatcagaa gtttaggaatc tgcacagag 8520
acaaatgcag tcatggttgt ttgctgcata ccagccctaa tcattagaag cctcatggac 8580
ttcaaacatc attccctctg acaagatgct ctacgctaac tccatgagat aaaataaatc 8640

tgcccttcag agccaaagaa gaggccacca gcttcttctc agtgtgaaca agagctccag 8700
tcaggtagt cagtcagtg cagtagagga gaccagtctg catcctctaa ttttcaaagg 8760
caagaagatt tgtttaccct ggacaccagg cacaagttag gtcacagagc tcttagatat 8820
gcagtcctca tgagttagga gactaaagcg catgcatca agacttcagt gtagagaaaa 8880
cctcaaaaa agcctcctca ctacttctgg aatagctcag aggccgaggc ggcctcggcc 8940
tctgcataaa taataaaat tagtcagca tggggcggag aatgggcgga actgggcgga 9000
gttagggcg ggatggcg agttagggc gggactatgg ttgctgacta attgagatgc 9060

atgctttgca tacttctgcc tgctggggag cctggggact ttccacacct ggttgctgac 9120
taattgagat gcatgcttg catacttctg cctgctgggg agcctgggga ctttccacac 9180
cctaactgac acacattcca cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg 9240

gtttgcgtat tgggcgctct tccgcttcct cgtcactga ctcgctgcgc tcggtcgttc	9300
ggctgcggcg agcggatatca gctcactcaa aggcggtaat acggttatcc acagaatcag	9360
gggataacgc aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa	9420
aggccgcgtt gctggcgttt ttccataggc tccgcccccc tgacgagcat cacaaaaatc	9480
gacgctcaag tcagaggtag cgaaacccga caggactata aagataaccag gcgtttcccc	9540
ctggaagctc cctcgctgcg tctctgttgc cgaccctgcc gcttaccgga tacctgtccg	9600
cctttctccc ttcggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt	9660
cgggtgtaggt cgttcgctcc aagctgggct gtgtgcacga accccccgtt cagcccagcc	9720
gtcgcgcctt atccggtaac tatcgtcttg agtccaaccc ggtaagacac gacttatcgc	9780
cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag	9840
agtctctgaa gtggtggcct aactacggct aactagaag aacagtatct ggtatctgcg	9900
ctctgctgaa gccagttacc ttcgaaaaa gagttggtag ctcttgatcc ggcaaaaaa	9960
ccaccgctgg tagcggtagt tttttgttt gcaagcagca gattacgcgc agaaaaaaag	10020
gatctcaaga agatcctttg atcttttcta cggggtctga cgtcagtagg aacgaaaact	10080
cacgttaagg gatcttggtc atgagattat caaaaaggat cttcacctag atccttttaa	10140
attaaaaatg aagttttaa tcaatctaaa gtatatatga gtaaaacttg tctgacagtt	10200
accaatgctt aatcagtagg gcacctatct cagcgatctg tctatctcgt tcatccatag	10260
ttgcctgact cctgcaaacc acgttgtgtc tcaaaatctc tgatgttaca ttgcacaaga	10320
taaaaatata tcatcatgaa caataaaact gtctgcttac ataaacagta atacaagggg	10380
tgttatgagc catattcaac gggaacgctc ttgctcgagg ccgcgattaa attccaacat	10440
ggatgctgat ttatatgggt ataaatgggc tcgcgataat gtcgggcaat caggtgcgac	10500
aatctatcga ttgtatggga agcccgatgc gccagagttg tttctgaaac atggcaaagg	10560
tagcgttgcc aatgatgta cagatgagat ggtcagacta aactggctga cggaatttat	10620
gcctcttccg accatcaagc attttatccg tactcctgat gatgcatggt tactcaccac	10680
tgcatcccc gggaaaacag cattccaggt attagaagaa tatcctgatt caggtgaaaa	10740
tattgttgat gcgctggcag tgttctgcg ccggttgcat tcgattcctg tttgtaattg	10800
tccttttaac agcgatcgcg tatttcgtct cgtcaggcg caatcacgaa tgaataacgg	10860
tttggttgat gcgagtgatt ttgatgacga gcgtaatggc tggcctgttg aacaagtctg	10920
gaaagaaatg cataagcttt tgccattctc accggattca gtcgtcactc atggtgattt	10980
ctcacttgat aaccttatct ttgacgaggg gaaattaata ggttgatttg atgttgagc	11040
agtcggaatc gcagaccgat accaggatct tgccatccta tggaactgcc tcggtgagtt	11100

ttctccttca ttacagaaac ggctttttca aaaatatggt attgataatc ctgatatgaa 11160

taaattgcag ttccatttga tgctcgatga gtttttctaa gggcggcctg ccaccatacc 11220

cacgcgaaa caagcgtca 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 ccctctctcg cgctcgctcg ctactgagg ccggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg 120

gccaaactcca tcaactagggg ttcttgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300

tagggcggag ccaatcagcg tgcgccgttc cgaagttgc cttttatggc tgggcggaga 360

atgggcgggt aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagtccg 420

tgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcaacttgta 480

agtcactgac tgctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540

tggcactatg aaccctgcag ccttaggaat gcacttagac aattgtacta accttcttct 600

ctttctctc ctgacagtcc ggaaagccac catggaattc agcagcccca gcagagagga 660

atgcccgaag cccttgagcc ggggtgtcaat catggccgga tctctgacag gactgtgct 720

gcttcaggcc gtgtcttggg cttctggcgc tagacctgc atccccaaga gcttcggcta 780

cagcagcgtc gtgtgcgtgt gcaatgccac ctactgcgac agcttcgacc ctctacctt 840

tcctgtcttg ggcaccttca gcagatacga gagcaccaga tccggcagac ggatggaact 900

gagcatggga cccatccagg ccaatcacac aggcactggc ctgctgctga cactgcagcc 960

tgagcagaaa ttccagaaag tgaaaggctt cggcggagcc atgacagatg ccgccgtct 1020

gaatatcctg gctctgtctc caccagctca gaacctgctg ctcaagagct acttcagcga 1080

ggaaggcatc ggtacaaca tcatcagagt gccatggcc agctgcgact tcagcatcag 1140

gacctacacc tacgccgaca caccgacga ttccagctg 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
 caccaccat 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 gtcgtcggt ggaccgactg gaatctggcc ctgaatcctg aaggcgcccc 1920
 taactgggtc cgaaacttcg tggacagccc catcatctg gacatcacca aggacacctt 1980
 ctacaagcag cccatgttct accacctggg acacttcagc aagttcatcc ccgagggctc 2040
 tcagcgcgtt ggactgggtg cttcccagaa gaacgatctg gacgccgtgg ctctgatgca 2100
 ccctgatgga tctgctgtgg tgggtgtcct gaaccgcagc agcaaagatg tgccccgtac 2160
 catcaaggat cccgccgtgg gattcctgga aacaatcagc cctggctact ccatecacac 2220

 ctacctgtgg cgtagacagt gacaattgtt aattaagttt catcgatacc gtcgactaga 2280
 gtcgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt ttgccctcc 2340
 cccgtgcctt ccttgacctt ggaagggtgc actccactg tcctttccta ataaaatgag 2400
 gaaattgcat cgcattgtct gtagtaggtgt cattctattc tgggggggtgg ggtggggcag 2460
 gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggagag atccacgata 2520
 acaaacagct tttttggggg ggcggagtta gggcggagcc aatcagcgtg cgccgttccg 2580
 aaagttgcct tttatggctg ggcggagaat gggcgggtgaa cgccgatgat tatataagga 2640

 cgcgccgggt gtggcacagc tagttccgtc gcagccggga tttgggtcgc ggttcttgtt 2700
 tgtggatccc tgtgatctc acttggttaag tctactactg tctatgcctg ggaaagggtg 2760
 ggcaggagat ggggcagtgc aggaaaagtg gcactatgaa cctgcagcc ctaggaatgc 2820
 atctagacaa ttgtactaac cttcttctct ttctctcct gacagtccgg aaagccacca 2880
 tgggccgtg ctgcttctac accgccgga cctgagcct gctgctgctg gtgaccagcg 2940
 tgacctgct ggtggccccg gtgttcacga aggccgtgga ccagagcatc gagaagaaga 3000

tcgtgctgcg caacggcacc gaggccttcg acagctggga gaagccccc ctgcccgtgt	3060
acaccagtt 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 ttcgtgacct acaccgtgga cgagctgctg tggggctaca	3420
aggacgagat cctgagcctg atccacgtgt tccgccccga catcagcccc tacttcggcc	3480
tgttctacga gaagaacggc accaacgacg gcgactacgt gttcctgacc ggcgaggaca	3540
gctacctgaa cttaccaag atcgtggagt ggaacggcaa gaccagcctg gactggtgga	3600
tcaccgacaa gtgcaacatg atcaacggca ccgacggcga cagcttcac cccctgatca	3660
ccaaggacga ggtgctgtac gtgttcccca gcgacttctg ccgcagcgtg tacatcacct	3720
tcagcgacta cgagagcgtg cagggcctgc ccgccttccg ctacaagggtg cccgccgaga	3780
tcctggccaa caccagcgac aacgccggt tctgcatccc cgagggcaac tgcctgggca	3840
gcggcgtgct gaacgtgagc atctgcaaga acggcgcccc catcatcatg agcttcccc	3900
acttctacca ggccgacgag cgcttcgtga gcgccatcga gggcatgcac cccaaccagg	3960
aggaccacga gaccttcgtg gacatcaacc cctgaccgg catcatcctg aaggccgcca	4020
agcgcttcca gatcaacatc tacgtgaaga agctggacga cttcgtggag accggcgaca	4080
tccgcacat ggtgttcccc gtgatgtacc tgaacgagag cgtgcacatc gacaaggaga	4140
ccgccagccg cctgaagagc atgatcaaca ccacctgat catcaccaac atcccctaca	4200
tcacatggc cctgggcgtg ttcttcggcc tgggtgttac ctggctggcc tgcaagggcc	4260
agggcagcat ggacgagggc accgccgacg agcgcgcccc cctgatccgc acctgacce	4320
ggggactcaa tcagcctcga agacatgata agataattg atgagtttg acaaaccaca	4380
acaagaatgc agtgaaaaaa atgctttatt tgtgaaattt gtgatgctat tgctttattt	4440
gtaaccatta taagctgcaa taaacaagt aacaacaaca attgcattca tttatgttt	4500
caggttcagg gggagatgtg ggaggttttt taaagcaagt aaaacctcta caaatgtggt	4560
atgaacatat tgactgaatt cctgcaggt tggccactcc ctctctgcgc gctcgtcgc	4620
tcactgaggc cgcccgggca aagcccgggc gtcggcgac ctttggctgc cggcctcag	4680
tgagcgagcg agcgcgaga gagggagtgg ccaactccat cactaggggt tcctgcggcc	4740

gctcgtacgg tctcgaggaa ttctgcagg ataacttgcc aacctcattc taaaatgtat	4800
atagaagccc aaaagacaat aacaaaaata ttcttgtaga acaaatggg aaagaatgtt	4860
ccactaaata tcaagattta gagcaaagca tgagatgtgt ggggatagac agtgaggctg	4920
ataaaataga gtagagctca gaaacagacc cattgatata tgtaagtac ctatgaaaaa	4980
aatatggcat ttacaatgg gaaatgatg gtctttttct ttttagaaa aacagggaaa	5040
tatatattata tgtaaaaaat aaaagggaa ccatatgtca taccatacac aaaaaaaat	5100
tccagtgaat tataagtcta aatggagaag gcaaaacttt aaatcttta 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 aaccaaagtc acaggaacaa gttatggtac tgatgtgctg	5400
gattgagaag gagctctact tccttgacag gacacatttg tatcaactta aaaaagcaga	5460
tttttgccag cagaactatt cattcagagg taggaaactt agaatagatg atgtcactga	5520
ttagcatggc ttccccatct ccacagctgc ttcccacca ggttgccac agttgagttt	5580
gtccagtgtc cagggtgcc cactctcagt aagaagcccc acaccagccc ctctccaaat	5640
atgttggtg ttcttccat taaagtacc ccactttaga gcagcaagtg gatttctgtt	5700
tcttacagtt caggaaggag gagtcagctg tgagaacctg gagcctgaga tgcttctaag	5760
tcccactgct actggggta gggaagccag actccagcat cagcagtcag gagcactaag	5820
cccttgccaa catcctgttt ctcagagaaa ctgcttccat tataatggtt gtctttttt	5880
aagctatcaa gccaaacaac cagtgtctac cattattctc atcacctgaa gccaaagggtt	5940
ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct ccagcttctg tcttcagtca	6000
ctccactctt agcctgtctt gaatcaactc tgaccacagt tccttgagc cctgcccacc	6060
tgctgcccc gccaccttct ccatctgcag tgctgtgcag ccttctgcac tcttgagag	6120
ctaatagggtg gagacttgaa ggaagaggag gaaagtttct cataatagcc ttgctgcaag	6180
ctcaaatggg aggtgggcac tgtgcccagg agccttgag caaaggctgt gcccaacctc	6240
tgactgcatc caggtttggc cttgacagag ataagaagcc ctggcttttg gagccaaat	6300
ctaggtcaga cttaggcagg attctcaaag ttatcagca gaacatgagg cagaagacct	6360
tttctgctcc agcttcttca ggctcaacct tcatcagaat agatagaaag agaggctgtg	6420
agggttctta aaacagaagc aaatctgact cagagaataa acaacctcct agtaaaactac	6480
agcttagaca gagcatctgg tggtagtgt gctcagtgct ctaactcaact gtctggtatc	6540
agccctcatg aggacttctc ttctttccct catagacctc catctctgtt ttccttagcc	6600

tgcagaaatc tggatggcta ttcacagaat gcctgtgctt tcagagttgc attttttctc	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 tcttttctg agttttctag ctctctctca	6900
ccttacctct gccctacca ggacaaacc aagagccact gtttctgtga tgtctctcc	6960
agccctaatt aggcacatg acttcagcct gaccttccat gctcagaagc agtgctaate	7020
cacttcagat gagctgctct atgcaacaca ggcagagcct acaaaccttt gcaccagagc	7080
cctccacata tcagtgtttg ttcatactca cttcaacagc aaatgtgact gctgagatta	7140
agattttaca caagatggtc tgtaatttca cagttagttt tatccatta ggtatgaaag	7200
aattagcata attccccta aacatgaatg aatcttagat tttttaataa atagttttgg	7260
aagtaaagac agagacatca ggagcacaag gaatagcctg agaggacaaa cagaacaaga	7320
aagagctcgg aaatacacag gatgttcttg gcctctctca agcaagtgca agcagatagt	7380
accagcagcc ccaggtatc agagcccagt gaagagaagt accatgaaag ccacagctct	7440
aaccaccctg ttccagagtg acagacagtc cccaagacaa gccagcctga gccagagaga	7500
gaactgcaag agaaagtctc taatttaggt tctgttagat tcagacaagt gcaggtcatc	7560
ctctctccac agctactcac ctctccagcc taacaaagcc tgcagtcac actccaacc	7620
tggtgtctca cctcctagcc tctcccaaca tctgtctctc tgaccatctt ctgcactct	7680
catctacca tctcccactg tctacagcct actcttgcaa ctaccatctc atttctgac	7740
atcctgtcta catcttctgc catactctgc catctaccat accacctctt accatctacc	7800
acaccatctt ttatctccat cctctcaga agcctccaag ctgaatcctg ctttatgtgt	7860
tcactcagc cctgcatgg aaagetgacc ccagaggcag aactattccc agagagcttg	7920
gccaaagaaa aaaaaactac cagcctggcc aggctcagga gtagtaagct gcagtgtctg	7980
ttgtgttcta gttcaacag ctgcaggagt tccactctca aatgtccac atttctcaca	8040
tctctctgat tctggtcact acccatcttc aaagaacaga atatctcaca tcagcatact	8100
gigaaggact agtcatgggt gcagctgctc agagctgcaa agtcattctg gatggtggag	8160
agcttacaaa catttcatga tgcctcccc gcctctgatg ctggagccca atccctacac	8220
agactcctgc tgiatgtgtt ttcctttcac tctgagccac agccagaggg caggcattca	8280
gtctctctt caggctgggg ctggggcact gagaactcac ccaacacctt gctctcactc	8340
cttctgcaaa acaagaaaga gctttgtgct gcagtagcca tgaagaatga aaggaaggct	8400
ttaactaaaa aatgtcagag attattttca accccttact gtggatcacc agcaaggagg	8460

aaacacaaca cagagacatt ttttccctc aaattatcaa aagaatcact gcatttgta	8520
aagagagcaa ctgaatcagg aagcagagt ttgaacatat cagaagttag gaatctgcat	8580
cagagacaaa tgcagtcag gttgtttgct gcataccagc cctaattcatt agaagcctca	8640
tggacttcaa acatcattcc ctctgacaag atgctctagc ctaactccat gagataaaat	8700
aaatctgcct ttcagagcca aagaagagtc caccagcttc ttctcagtg gaacaagagc	8760
tccagtcagg ttagtcagtc cagtcagta gaggagacca gtctgcatcc tctaatttc	8820
aaaggcaaga agatttggtt accctggaca ccaggcaca gtgaggtcac agagcttta	8880
gatatgcagt cctcatgagt gaggagacta aagcgcatgc catcaagact tcagtgtaga	8940
gaaaacctcc aaaaaagcct cctcactact tctggaatag ctgagaggcc gaggcggcct	9000
cggcctctgc ataaataaaa aaaattagtc agccatgggg cggagaatgg gcggaactgg	9060
gcggagttag ggccgggatg ggccggagtta gggcggggac tatggttgct gactaatga	9120
gatgcatgct ttgcatactt ctgcctgctg gggagcctgg ggactttcca cacctggttg	9180
ctgactaatt gagatgcatg ctttgcatc ttctgcctgc tggggagcct ggggactttc	9240
cacacctaa ctgacacaca ttccacagct gcattaatga atcgccaac gcgcggggag	9300
aggcggtttg cgtattgggc gctcttcgc ttctcgcctc actgactgc tgcgctcgg	9360
cgttcggtg cgccgagcgg tatcagctca ctcaaaggcg gtaatacgt tatccacaga	9420
atcaggggat aacgcaggaa agaacatgt agcaaaaggc cagcaaaagg ccaggaaaccg	9480
taaaaaggcc gcgttgctgg cggttttcca taggtccgc cccctgacg agcatcaca	9540
aatcgacgc tcaagtcaa ggtggcgaaa ccgacagga ctataaagat accagcggt	9600
tccccctgga agtccctcg tgcgtctcc tgtccgacc ctgccgtta ccgatacct	9660
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agtcacgct gtaggtatct	9720
cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc	9780
cgaccgtgc gccttatccg gtaactatcg tcttgagtcc aaccggtaa gacacgactt	9840
atgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc	9900
tacagagttc ttgaagtgg ggcctaacta cggtacact agaagaacag tatttggtat	9960
ctgcgtctg ctgaagccag ttaccttcgg aaaaagagt gtagctctt gatccggcaa	10020
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa	10080
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga	10140
aaactcacgt taagggtttt tggatcatg attatcaaaa aggatcttca cctagatcct	10200

tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttgggtctga	10260
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcac	10320
catagtggcc tgactcctgc aaaccacgtt gtgtctcaaa atctctgatg ttacattgca	10380
caagataaaa atatatcatc atgaacaata aaactgtctg cttacataaa cagtaataca	10440
aggggtgtta tgagccatct tcaacgggaa acgtcttgct cgaggccgag attaaattcc	10500
aacatggatg ctgatttata tgggtataaa tgggctcgcg ataatgtcgg gcaatcaggt	10560
gcgacaatct atcgattgta tgggaagccc gatgcgccag agttgtttct gaaacatggc	10620
aaaggtagcg ttgccaatga ttttacagat gagatggta gactaaactg gctgacggaa	10680
tttatgcctc ttcgacat caagcatttt atccgtactc ctgatgatgc atggttactc	10740
accactgcga tccccgggaa aacagcattc caggtattag aagaatatcc tgattcaggt	10800
gaaaatattg ttgatgcgtt ggcagtgttc ctgcgccggt tgcattcgat tcctgtttgt	10860
aattgtcctt ttaacagcga tcgcgtattt cgtctcgtc aggcgcaatc acgaatgaat	10920
aacggtttgg ttgatgcgag tgattttgat gacgagcgta atggctggcc tgttgaacaa	10980
gtctggaaag aaatgataa gcttttgcca ttctcaccgg attcagtcgt cactcatggt	11040
gattttctac ttgataacct tatttttgac gaggggaaat taataggttg tattgatgtt	11100
ggacgagtcg gaatcgaga ccgataccag gatcttgcca tcctatggaa ctgcctcgg	11160
gagttttctc cttcattaca gaaacggctt tttcaaaaat atggtattga taatcctgat	11220
atgaataaat tgcagtttca tttgatgtc gatgagtttt tctaaggcg gctgccacc	11280
ataccacgc cgaacaagc gctcatgagc ccgaagtggc gagcccgatc ttccccatcg	11340
gtgatgtcgg cgatatagcg gccagcaacc gcacctgtgg cgccggtgat gagggcgcg	11400
caagtcgacg tccggcagtc	11420

<210> 4

<211> 11171

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 4

ttggccactc ccctctctcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc	60
cgacgcccgg gctttgccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg	120
gccaactcca tctactagggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc	180
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc	240
tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt	300

tagggcggag ccaatcagcg tgcgccgttc cgaaagtgc cttttatggc tgggcggaga	360
atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagtccg	420
tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcaacttgta	480
agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag	540
tggcactatg aacctctctg gtggcgaggg gaggggggtg gtcctcgaa gccttgacaga	600
actggcctgg atacagagtg gaccggctgg ccccatctgg aagacttcga gataactgt	660
tgtcttactg cgctcaacag tgtatctcga agtcttccaa atggcgccag ccatcgacgc	720
ggggtgcagg aaatgggggc agccccctt tttggctatc cttccacgtg ttcttttttg	780
tatcttttgt gtttcctaga aaacatctca gtcaccaccg cagccctagg aatgcatcta	840
gacaattgta ctaaccttct tctctttcct ctctgacag tccggaaagc caccatgggc	900
cgctgctgct tctacaccgc cggcaccctg agcctgctgc tgctgggtgac cagcgtgacc	960
ctgctggtgg cccgcgtgtt ccagaaggcc gtggaccaga gcatcgagaa gaagatcgtg	1020
ctgcgcaacg gcaccgaggc cttcgacagc tgggagaagc cccccctgcc cgtgtacacc	1080
cagttctact tcttcaacgt gaccaacccc gaggagatcc tgcgcggcga gacccccgc	1140
gtggaggagg tgggccccta cacctaccgc gagctgcgca acaaggccaa catccagttc	1200
ggcgacaacg gcaccacat cagcgccgtg agcaacaagg cctacgtgtt cgagcgcgac	1260
cagagcgtgg gcgaccccaa gatcgacctg atccgcaccc tgaacatccc cgtgctgacc	1320
gtgatcgagt ggagccaggt gcaattcctg cgcgagatca tcgaggccat gctgaaggcc	1380
taccagcaga agctgttcgt gaccacacc 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 tccaccctt gatcaccaag	1680
gacgaggtgc tgtactgtt cccagcgac tttgtccgca gcgtgtacat caccttcagc	1740
gactacgaga gcgtgcaggg cctgcccgc ttcgctaca aggtgcccgc cgagatcctg	1800
gccaacacca gcgacaacgc cggcttctgc atccccgagg gcaactgcct gggcagcggc	1860
gtgctgaacg tgagcatctg caagaacggc gccccatca tcatgagctt ccccaacttc	1920
taccaggccg acgagcgctt cgtgagcgcc atcgaggga tgcaccccaa ccaggaggac	1980
cacgagacct tcgtggacat caacccttg 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 cggcctgggtg ttcacctggc tggcctgcaa gggccagggc	2280
agcatggacg agggcaccgc cgacgagcgc gccccctga tccgcaccga gggcagagga	2340
agtcttctga catgcggaag cgtggaagag aatcccggcc ctatggaatt cagcagcccc	2400
agcagagagg aatgccccaa gcctctgagc cgggtgtcaa tcatggccgg atctctgaca	2460
ggactgctgc tgcttcaggc cgtgtcttgg gcttctggcg ctagacctg catccccaag	2520
agcttcggct acagcagcgt cgtgtgcgtg tgcaatgcca cctactgca cagcttcgac	2580
cctcctacct ttctgtctt 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 ctacccgac acacccgacg atttcagct gcacaacttc	2940
agcctgcctg aagaggacac caagctgaag atccctctga tccacagagc cctgcagctg	3000
gcacaaagac ccgtgtcact gctggcctct ccatggacat ctccacactg gctgaaaaca	3060
aatggcgccg tgaatggcaa gggcagcctg aaaggccaac ctggcgacat ctaccaccag	3120
acctgggcca gatacttcgt gaagtctctg gacgcctatg ccgagcaca gctgcagttt	3180
tgggccgtga cagccgagaa cgaaccttct gctggactgc tgagcggcta cccctttcag	3240
tgctgggct ttacaccga gcaccagcgg gactttatcg cccgtgatct gggaccaca	3300
ctggccaata gcaccacca taatgtgcgg ctgctgatgc tggacgacca gagactgctt	3360
ctgccccact gggctaaagt ggtgttgaca gatcctgagg ccgccaata cgtgcacgga	3420
atcgccgtgc actggtatct ggactttctg gccctgcca aggccacact gggagagaca	3480
cacagactgt tcccaaacac catgctgttc gccagcgaag cctgtgtggg cagcaagttt	3540
tgggaacaga gcgtgcggct cggcagctgg gatagaggca tgcagtacag ccacagcatc	3600
atcaccaacc tgcgtgacca cgtcgtcggc tggaccgact ggaatctggc cctgaatcct	3660
gaaggcggcc ctaactgggt ccgaaacttc gtggacagcc ccatcatcgt ggacatcacc	3720
aaggacacct tctacaagca gcccatgttc taccacctgg gacacttcag caagttcatc	3780
cccaggggct ctcagcgcgt tggactgggtg gcttcccaga agaacgatct ggacgccgtg	3840
gctctgatgc accctgatgg atctgctgtg gtggtggtcc tgaaccgcag cagcaaagat	3900

gtgccccctga ccatcaagga tccccccgtg ggatttcttg aaacaatcag ccctggctac 3960
tccatccaca cctacctgtg gcgtagacag tgacaattgt taattaagtt taaaccctcg 4020
aggccgcaag ccgcatcgat accgtcgact agagctcgct gatcagcctc gactgtgcct 4080
tctagttgcc agccatctgt tgtttgcccc tccccctgtc cttccttgac cctggaaggt 4140

gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgattg tctgagtagg 4200
tgtcattcta ttctgggggg tgggtgggg caggacagca agggggagga ttgggaagac 4260
aatagcaggc atgctgggga gagatccacg ataacaaca gcttttttgg ggtgaacata 4320
ttgactgaat tccctgcagg ttggccactc cctctctgcg cgctcgctcg ctactgagg 4380
ccgccccggc aaagccggg cgctgggcga cctttggtcg cccggcctca gtgagcgagc 4440
gagcgcgag agaggagtg gccaaactca tcactagggg ttcctgcggc cgctcgtacg 4500
gtctcgagga attcctgcag gataacttgc caacctcatt ctaaaatgta tatagaagcc 4560

caaaagacaa taacaaaaat attctttag aacaaaatgg gaaagaatgt tccactaat 4620
atcaagattt agagcaaac atgagatgtg tgggataga cagtgaggct gataaaatag 4680
agtagagctc agaaacagac ccattgatat atgtaagtga cctatgaaaa aaatatggca 4740
ttttacaatg ggaaaatgat ggtctttttc ttttttagaa aaacaggga atatatatt 4800
atgtaaaaaa taaaaggga cccatatgtc ataccataca cacaaaaaa ttccagtga 4860
ttataagtct aaatggagaa ggcaaaactt taaatctttt agaaaataat atagaagcat 4920
gcagaccagc ctggccaaca tgatgaaacc ctcttacta ataataaaat cagtagaact 4980

actcaggact actttgagt ggaagtcctt ttctatgaag acttttttg ccaaaattag 5040
gctctaaatg caaggagata gtgcatcatg cctggctgca cttactgata aatgatgtta 5100
tcaccatctt taaccaaag cacaggaaca agttatggta ctgatgtgct ggattgagaa 5160
ggagctctac ttcttgaca ggacacattt gtatcaactt aaaaaagcag atttttgcca 5220
gcagaactat tcattcagag gtaggaaact tagaatagat gatgtcactg attagcatgg 5280
cttccccatc tccacagctg ctccccacc aggttgccca cagttgagtt tgtccagtgc 5340
tcagggtgc ccactctcag taagaagccc cacaccagcc cctctccaaa tatgttggt 5400

gttccttcca ttaaagtac cccactttag agcagcaagt ggatttctgt ttcttacagt 5460
tcaggaagga ggagtcagt gtgagaacct ggagcctgag atgcttctaa gtcccactgc 5520
tactggggtc agggaagcca gactccagca tcagcagtca ggagcactaa gcccttgcca 5580
acatcctgtt tctcagagaa actgcttcca ttataatgg tgtccttttt taagctatca 5640
agccaaacaa ccagtgtcta ccattattct catcacctga agccaagggt tctagcaaaa 5700
gtcaagctgt ctgtaatgg ttgatgtgcc tccagcttct gtcttcagtc actccactct 5760

tagcctgctc tgaatcaact ctgaccacag ttccttggag cccctgccac ctgctgcccc	5820
tgccaccttc tccatctgca gtgctgtgca gccttctgca ctcttgca gctaataagg	5880
ggagacttga aggaagagga ggaaagtttc tcataatagc cttgctgcaa gctcaaatgg	5940
gaggtgggca ctgtgccag gagccttga gcaaaggctg tgcccaacct ctgactgcat	6000
ccaggtttgg tcttgacaga gataagaagc cctggctttt ggagccaaaa tctaggtcag	6060
acttaggcag gatttctaaa gtttatcagc agaacatgag gcagaagacc ctttctgctc	6120
cagcttcttc aggtctcaacc ttcacagaa tagatagaaa gagaggctgt gagggttctt	6180
aaaacagaag caaatctgac tcagagaata aacaacctcc tagtaaaacta cagcttagac	6240
agagcatctg gtggtgagtg tgctcagtgt cctactcaac tgtctggtat cagccctcat	6300
gaggacttct cttctttccc tcatagacct ccactctgt tttccttagc ctgcagaaat	6360
ctggatggct attcacagaa tgctgtgtct ttcagagtig cattttttct ctggtattct	6420
ggttcaagca ttgaaggta ggaaagggtc tccaagtga agaaaggcag cctgagcct	6480
caactgcctg gctagtgtgg tcagttagat gcaaaggctg ttgaatgcca caaggccaaa	6540
ctttaacctg tgtaccaca gcctagcagc agaggcagct ctgctcactg gaactctctg	6600
tcttctttct cctgagcctt tcttttctt gagttttcta gctctctca accttacctc	6660
tgccttacc aggacaaacc caagagccac tgtttctgtg atgtctctc cagccctaat	6720
taggcatcat gacttcagcc tgaccttcca tgctcagaag cagtgtctaat ccacttcaga	6780
tgagctgctc tatgcaacac aggcagagcc tacaaacctt tgcaccagag cctccacat	6840
atcagtgttt gttcatactc acttcaacag caaatgtgac tgctgagatt aagattttac	6900
acaagatggt ctgtaatttc acagttagtt ttatccatt aggtatgaaa gaattagcat	6960
aattccctt aaacatgaat gaactttaga ttttttaata aatagttttg gaagtaaaga	7020
cagagacatc aggagcaca ggaatagcct gagaggacaa acagaacaag aaagagtctg	7080
gaaatacaca ggatgttctt ggctctctca aagcaagtgc aagcagatag taccagcagc	7140
cccaggtat cagagcccag tgaagagaag taccatgaaa gccacagctc taaccacct	7200
gttccagagt gacagacagt cccaagaca agccagcctg agccagagag agaactgcaa	7260
gagaaagttt ctaatttagg ttctgttaga ttcagacaag tgcaggatcat cctctctcca	7320
cagctactca cctctccagc ctaacaaagc ctgcagtcca cactccaacc ctggtgtctc	7380
acctctagc ctctcccaac atctgtctct ctgacctct tctgcatctc tcatctcacc	7440
atctcccact gtctacagcc tactcttgca actaccatct cattttctga catctgtct	7500

acatcttctg ccatactctg ccatactacca taccacctct taccatctac cacaccatct 7560
 tttatctcca tccctctcag aagcctccaa gctgaatcct gctttatgtg ttcactcag 7620
 cccctgcatg gaaagctgac cccagaggca gaactattcc cagagagctt ggccaagaaa 7680
 aacaaaacta ccagcctggc caggctcagg agtagtaagc tgcagtgtct gttgtgttct 7740
 agcttcaaca gctgcaggag ttccactctc aaatgctcca catttctcac atcctcctga 7800
 ttctggtcac taccatctt caaagaacag aatctctcac atcagcatac tgtgaaggac 7860
 tagtcatggg tgcagctgct cagagctgca aagtcattct ggatgggtgga gagcttaciaa 7920

 acatttcatg atgctccccc cgctctgatg gctggagccc aatccctaca cagactcctg 7980
 ctgtatgtgt tttcctttca ctctgagcca cagccagagg gcaggcattc agtctcctct 8040
 tcaggctggg gctggggcac tgagaactca cccaacacct tgctctcact cttctgcaa 8100
 aacaagaaag agctttgtgc tgcagtagcc atgaagaatg aaaggaaggc tttaactaaa 8160
 aaatgtcaga gattattttc aaccctttac tgtggatcac cagcaaggag gaaacacaac 8220
 acagagacat tttttccct caaattatca aaagaatcac tgcatttggt aaagagagca 8280
 actgaatcag gaagcagagt ttggaacata tcagaagtta ggaatctgca tcagagacaa 8340

 atgcagtcac ggttgtttgc tgcataccag ccctaactcat tagaagcctc atggacttca 8400
 aacatcattc cctctgacaa gatgctctag cctaactcca tgagataaaa taaatctgcc 8460
 tttcagagcc aaagaagagt ccaccagctt ctctcagtg tgaacaagag ctccagtcag 8520
 gttagtcagt ccagtgcagt agaggagacc agtctgcatc ctctaatttt caaaggcaag 8580
 aagatttggt taccctggac accaggcaca agtgaggtea 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 gggcggagt aggggcggga ctatggttgc tgactaattg agatgcatgc 8880
 tttgcatact tctgctgct ggggagcctg gggactttcc acacctggtt gctgactaat 8940
 tgagatgcat gctttgcata ctctgctctg ctggggagcc tggggacttt ccacacccta 9000
 actgacacac attccacagc tgcattaatg aatcgcccaa cgcgcgggga gaggcggttt 9060
 gcgtattggg cgctcttccg ctctctctg cactgactcg ctgcgctcgg tcgttcggct 9120
 gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 9180

 taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 9240
 cggttgctg gcgtttttcc ataggctccg cccctctgac gagcatcaca aaaatcgacg 9300
 ctcaagtcat aggtggcgaa acccgacagg actataaaga taccaggcgt tccccctgg 9360

aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt	9420
tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt	9480
gtaggtcggt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg	9540
cgctttatcc ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact	9600
ggcagcagcc actggttaaca ggattagcag agcgaggtat gtaggcgggtg ctacagagtt	9660
cttgaagtgg tggcctaact acggtacac tagaagaaca gtatttggta tctgcgctct	9720
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac	9780
cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgagaa aaaaaggatc	9840
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaaacg aaaactcacg	9900
ttaagggttt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta	9960
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttgggtctg acagttacca	10020
atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagtgc	10080
ctgactctg caaaccacgt tgtgtctcaa aatctctgat gttacattgc acaagataaa	10140
aatatatcat catgaacaat aaaactgtct gcttacataa acagtaatac aagggtgtt	10200
atgagccata ttcaacggga aacgtcttgc tcgaggccgc gattaaattc caacatggat	10260
gctgatttat atgggtataa atgggctcgc gataatgtcg ggcaatcagg tgcgacaatc	10320
tatcgattgt atgggaagcc cgatgcgcca gattgtttc tgaaacatgg caaaggtagc	10380
gttgccaatg atgttacaga tgagatggtc agactaaact ggctgacgga atttatgcct	10440
cttccgacca tcaagcattt tatecgtaact cctgatgatg catggttact caccactgcg	10500
atccccggga aaacagcatt ccaggtattt gaagaatatc ctgattcagg tgaaaatatt	10560
gttgatgcgc tggcagtgtt cctgcgccgg ttgcattcga ttcctgtttg taattgtcct	10620
tttaacagcg atcggtatt tcgtctcgt caggcgcaat cacgaatgaa taacggtttg	10680
gttgatgcga gtgattttga tgacgagcgt aatggctggc ctgttgaaca agtctggaaa	10740
gaaatgcata agcttttgcc attctcaccg gattcagtcg tcaactcatgg tgattttcct	10800
cttgataacc ttatttttga cgaggggaaa ttaatagggt gtattgatgt tggacgagtc	10860
ggaatcgcag accgatacca ggatcttgcc atcctatgga actgcctcgg tgagttttct	10920
ccttcattac agaaacggct ttttcaaaaa tatggtattg ataactctga tatgaataaa	10980
ttgcagtttc atttgatgct cgatgagttt ttctaagggc ggcctgccac cataccacg	11040
ccgaacaag cgctcatgag cccgaagtgg cgagcccgat cttcccatc ggtgatgtcg	11100
gcgatatagg cgccagcaac cgcacctgtg gcgccgtga tgaggcgcg ccaagtcgac	11160
gtccggcagt c	11171

<210> 5

<211> 11309

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 5

ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc	60
cgacgcccgg gctttgcccg ggcgccctca gtgagcgagc gagcgcgag agagggagtg	120
gccaaactcca tctactagggg ttcttgctag ctctgggtat ttaagcccga gtgagcacgc	180
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc	240
tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt	300
tagggcggag ccaatcacgc tgcgcggttc cgaaagtgc cttttatggc tggcgggaga	360
atggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagtccg	420
tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcaattggtg	480
agtcactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag	540
tggcactatg aaccctcctg gtggcgaggg gaggggggtg gtctctgaac gccttgacga	600
actggcctgg atacagagtg gaccggctgg ccccatctgg aagacttcga gataactgt	660
tgtcttactg cgctcaacag tgtatctcga agtcttccaa atggtgccag ccatcgagc	720
ggggtgcagg aaatgggggc agccccctt tttggctatc cttccacgtg ttcttttttg	780
tatcttttgt gtttcctaga aaacatctca gtcaccaccg cagccctagg aatgcatcta	840
gacaattgta ctaaccttct tctctttcct ctctgacag tccggaagc caccatgtac	900
gccctgttcc tgctggccag cctgtctggc gccgccctgg ccggccccgt gctgggcctg	960
aaggagtgca cccgcggcag cgccgtgtgg tgccagaac tgaagaccgc cagcgactgc	1020
ggcgccgtga agcactgcct gcagaccgtg tggaacaagc ccaccgtgaa gaggcctgcc	1080
tgcgacatct gcaaggacgt ggtgaccgcc gccggcgaca tgctgaagga caacgccacc	1140
gaggaggaga tcctggtgta cctggagaag acctgcgact ggctgccccaa gcccaacatg	1200
agcgccagct gcaaggagat cgtggacagc tacctgcccg tgatcctgga catcatcaag	1260
ggcgagatga gccgccccgg cgaggtgtgc agcgccctga acctgtgcga gagcctgcag	1320
aagcacctgg ccgagctgaa ccaccagaag cagctggaga gcaacaagat ccccagctg	1380
gacatgaccg aggtggtggc ccccttcatg gccaacatcc cctgtctgct gtacccccag	1440
gacggcccc gcagcaagcc ccagcccaag gacaacggcg acgtgtgcca ggactgcatc	1500

cagatggtga ccgacatcca gaccgccgtg cgcaccaaca gcaccttcgt gcaggccctg 1560
gtggagcacg tgaaggagga gtgcgaccgc ctggggcccg gcatggccga catctgcaag 1620
aactacatca gccagtacag cgagatcgcc atccagatga tgatgcacat gcagcccaag 1680

gagatctgcg ccctggtggg cttctgcgac gaggtgaagg agatgccc at gcagaccctg 1740
gtgcccgcga aggtggccag caagaacgtg atccccgcc 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 ctccccgcc tgaccgtgca cgtgaccag 2100

cccaaggacg gcggcttctg cgaggtgtgc aagaagctgg tgggctacct ggaccgcaac 2160
ctggagaaga acagcaccaa gcaggagatc ctggccgcc tggagaaggg ctgcagcttc 2220
ctgcccgaac cctaccagaa gcagtgcgac cagtctgtgg ccgagtacga gcccgtgctg 2280
atcgagatcc tgggtggaggt gatggacccc agcttctgtg gcctgaagat cggcgccctg 2340
cccagcgccc acaagcccct gctgggcacc gagaagtga tctggggccc cagctactgg 2400
tgccagaaca ccgagaccgc cgcccagtgc aacgccgtgg agcactgcaa gcgccacgtg 2460
tggaacgagg gcagaggaag tcttctgaca tgcggagacg tggaagagaa tcccggccct 2520

atggaattca gcagccccag cagagaggaa tgcccaagc ctctgagccg ggtgtcaatc 2580
atggccggat ctctgacagg actgctgtg cttcaggccg tgtcttgggc ttctggcgt 2640
agaccttgca tcccaagag cttcggtac agcagcgtcg tgtgctgtg caatgccacc 2700
tactgcgaca gttcgacc tctaccttt cctgctctgg gcaccttcag cagatacag 2760
agcaccagat ccggcagacg gatggaactg agcatgggac ccatccaggc caatcacaca 2820
ggcactggcc tgctgctgac actgcagcct gagcagaaat tccagaaagt gaaaggcttc 2880
ggcggagcca tgacagatgc cgccgctctg aatatactgg ctctgtctcc accagctcag 2940

aacctgctgc tcaagagcta cttcagcgag gaaggcatcg gctacaacat catcagagtg 3000
cccatggcca gctgcgactt cagcatcagg acctacacct acgccgacac acccgacgat 3060
ttccagctgc acaacttcag cctgcctgaa gaggacacca agctgaagat ccctctgac 3120
cacagagccc tgcagctggc acaaagacc gtgtcactgc tggcctctcc atggacatct 3180
cccacctggc tgaatacaaa tggcgccgtg aatggcaagg gcagcctgaa aggccaacct 3240
ggcgacatct accaccagac ctgggccaga tacttctgta agttcctgga cgcctatgcc 3300

gagcacaagc tgcagttttg ggccgtgaca gccgagaacg aaccttctgc tggactgctg 3360

agcggctacc cttttcagtg cctgggcttt acacccgagc accagcggga ctttatcgcc 3420

cgtgatctgg gacccacact ggccaatagc acccaccata atgtgcggct gctgatgctg 3480

gacgaccaga gactgcttct gcccactgg gctaaagtgg tgctgacaga tctgaggcc 3540

gccaataacg tgcacggaat cgccgtgcac tggatatctg 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 cgagggtctt cagcgcgttg gactggtggc ttcccagaag 3960

aacgatctgg acgccgtggc tctgatgcac cctgatggat ctgctgtggt ggtggtcctg 4020

aaccgcagca gcaaagatgt gcccctgacc atcaaggatc ccgccgtggg attcctggaa 4080

acaatcagcc ctggctactc catecacacc tacctgtggc gtagacagtg acaattgtta 4140

attaagttaa aaccctcgag gccgcaagcc gcatcgatac cgtcgactag agctcgctga 4200

tcagcctega ctgtgccttc tagttgccag ccactctgtt tttgcccctc ccccgtgcct 4260

tccttgacc tggaaggctg cactccact gtcttttct aataaaatga ggaaattgca 4320

tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg ggggtgggca ggacagcaag 4380

ggggaggatt gggaagacaa tagcaggcat gctggggaga gatccacgat aacaaacagc 4440

ttttttgggg tgaacatatt gactgaattc cctgcagggt ggccactccc tctctgcgcg 4500

ctcgtcgtct cactgaggcc gcccgggcaa agccccggcg tcgggcgacc tttggtcgcc 4560

cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc actaggggtt 4620

cctgcggccg ctctacggt ctgaggaat tctgcagga taacttgcca acctcattct 4680

aaaatgtata tagaagccca aaagacaata acaaaaatat tctttagaa caaatggga 4740

aagaatgttc cactaaatat caagatttag agcaaagcat gagatgtgtg gggatagaca 4800

gtgaggctga taaaatagag tagagctcag aaacagacc attgatatat gtaagtgacc 4860

tatgaaaaaa atatggcatt ttacaatggg aaaatgatgg tctttttctt ttttagaaaa 4920

acagggaaat atattttat gtaaaaaata aaagggaacc catatgtcat accatacaca 4980

caaaaaaatt ccagtgaatt ataagtctaa atggagaagg caaaacttta aatcttttag 5040

aaaataatat agaagcatgc agaccagcct ggccaacatg atgaaaccct ctctactaat	5100
aataaaatca gtagaactac tcaggactac tttagtgagg aagtcctttt ctatgaagac	5160
ttctttggcc aaaattaggc tctaaatgca aggagatagt gcatcatgcc tggctgcact	5220
tactgataaa tgatgttata accatcttta accaaatgca caggaacaag ttatggtact	5280
gatgtgctgg attgagaagg agctctactt ccttgacagg acacatttgt atcaacttaa	5340
aaaagcagat ttttgccagc agaactattc attcagaggt aggaaactta gaatagatga	5400
tgctactgat tagcatggct tccccatctc cacagctgct tcccaccag gttgccaca	5460
gttgagtttg tccagtgtc agggtgccc actctcagta agaagcccca caccagcccc	5520
tctccaaata tgttggctgt tccttcatt aaagtgacc cacttagag cagcaagtgg	5580
atttctgttt cttacagttc aggaaggagg agtcagctgt gagaacctgg agcctgagat	5640
gcttctaagt cccactgcta ctggggtcag ggaagccaga ctccagcatc agcagtcagg	5700
agcactaagc ccttgccaac atcctgtttc tcagagaaac tgcttcatt ataatggttg	5760
tcctttttta agctatcaag ccaacaacc agtgtctacc attattctca tcacctgaag	5820
ccaagggttc tagcaaaagt caagctgtct tgtaatggtt gatgtgctc cagcttctgt	5880
cttcagtcac tccactctta gcctgctctg aatcaactct gaccacagtt cctggagcc	5940
cctgccacct gctgccctg ccaccttctc catctgcagt gctgtgcagc cttctgcact	6000
cttgccagagc taataggtgg agacttgaag gaagaggagg aaagtttctc ataatagcct	6060
tgctgcaagc tcaaatggga ggtgggcact gtgccagga gccttgagc aaaggctgtg	6120
cccaacctct gactgcatcc aggtttgtc ttgacagaga taagaagccc tggcttttgg	6180
agccaaaatc taggtcagac ttaggcagga ttctcaaagt ttatcagcag aacatgaggc	6240
agaagaccct ttctgtcca gcttcttcag gctcaacctt catcagaata gatagaaaga	6300
gaggctgtga gggttcttaa aacagaagca aatctgactc agagaataaa caacctccta	6360
glaaactaca gcttagacag agcatctggt ggtgagtggt ctcagtgtcc tactcaactg	6420
tctggatca gccctcatga ggacttctct tctttccctc atagacctcc atctctgttt	6480
tccttagcct gcagaaatct ggatggctat tcacagaatg cctgtgcttt cagagttgca	6540
ttttttctct ggtattctgg ttcaagcatt tgaaggtagg aaaggttctc caagtcaag	6600
aaagccagcc ctgagctca actgcctggc tagtgtgtc agtaggatgc aaaggctgtt	6660
gaatgccaca aggccaaact ttaacctgtg taccacaagc ctgacagcag aggcagctct	6720
gctcactgga actctctgtc ttctttctcc tgagcctttt cttttctga gttttctagc	6780
tctctcaac cttacctctg ccctaccag gacaaacca agagccactg tttctgtgat	6840
gtcctctcca gccctaatta ggcatcatga cttcagcctg acctccatg ctgagaagca	6900

gtgctaattcc acttcagatg agctgctctta tgcaacacag gcagagccta caaacctttg	6960
caccagagcc ctccacatat cagtgtttgt tcatactcac ttcaacagca aatgtgactg	7020
ctgagattaa gattttacac aagatggctct gtaatttcac agtttagttt atcccattag	7080
gtatgaaaga attagcataa ttccccitaa 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 ctgittagatt cagacaagtg	7440
caggtcatcc tctctccaca gctactcacc tctccagcct aacaaagcct gcagtccaca	7500
ctccaacctt ggtgtctcac ctctagcct ctcccaacat cctgctctct gaccatcttc	7560
tgcattctct atctcaccat ctcccactgt ctacagccta ctcttgcaac taccatctca	7620
ttttctgaca tctgtcttac atcttctgcc atactctgcc atctaccata ccacctctta	7680
ccattaccac caccatcttt tatctccatc ctctcagaa gcctccaagc tgaatcctgc	7740
tttatgtgtt catctcagcc cctgcatgga aagctgacct cagaggcaga actattccca	7800
gagagcttgg ccaagaaaaa caaaactacc agcctggcca ggctcaggag tagtaagctg	7860
cagtgtctgt tgtgttctag ctccaacagc tgcaggagtt ccactctcaa atgctccaca	7920
tttctcacat cctcctgatt ctggctacta cccatcttca aagaacagaa tatctcacat	7980
cagcactatg tgaaggacta gtcatgggtg cagctgctca gagctgcaaa gtcattcttg	8040
atggtggaga gcttacaac atttcatgat gctcccccg ctctgatggc tggagcccaa	8100
tccttacaca 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 tttccctca aattatcaaa agaactcactg	8400
catttgtaa agagagcaac tgaatcagga agcagagttt tgaacataac agaagttagg	8460
aatctgcatc agagacaaat gcagtcattg ttgtttgctg cataccagcc ctaatcatta	8520
gaagcctcat ggacttcaaa catcattccc tctgacaaga tgctctagcc taactccatg	8580
agataaaata aatctgcctt tcagagccaa agaagagtcc accagcttct tctcagtgtg	8640
aacaagagct ccagtcaggt tagtcagtcc agtgagtag aggagaccag tctgcatcct	8700
ctaattttca aaggcaagaa gatttgttta ccctggacac caggcacaag tgaggtcaca	8760

gagctcttag atatgcagtc ctcatgagtg aggagactaa agcgcatgcc atcaagactt 8820

cagtgtagag aaaacctcca aaaaagcctc ctactacttt ctggaatagc tcagaggccg 8880

aggcgccctc ggctctgca taaataaaaa aaattagtca gccatggggc ggagaatggg 8940

cggaactggg cggagttagg ggcgggatgg gcggagttag gggcgggact atggttgctg 9000

actaatlgag atgcatgctt tgcatacttc tgcctgctgg ggagcctggg gactttccac 9060

acctggttgc tgactaatlg agatgcatgc tttgcatact tctgcctgct ggggagcctg 9120

gggactttcc acacctaac tgacacacat tccacagctg cattaatgaa tcggccaacg 9180

cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tectcgctca ctgactcgct 9240

gcgctcggtc gticggctgc ggcgagcggg atcagctcac tcaaaggcgg taatacggtt 9300

atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 9360

caggaacctg aaaaaggccg cgttgctggc gtttttccat aggtccgccc cccctgacga 9420

gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 9480

ccaggcgttt ccccttgaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac 9540

cggatacctg tccgccttcc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 9600

taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 9660

cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccgtaag 9720

acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatgt 9780

aggcggtgct acagagtctt tgaagtgtg gcctaactac ggctacacta gaagaacagt 9840

atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 9900

atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 9960

gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca 10020

gtggaacgaa aactcacgtt aagggaattt ggatcatgaga ttatcaaaaa ggatcttcac 10080

ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 10140

ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 10200

tcgttcatcc atagtgcct gactcctgca aaccacgttg tgtctcaaaa tctctgatgt 10260

tacattgcac aagataaaaa tataatcatca tgaacaataa aactgtctgc ttacataaac 10320

agtaatacaa ggggtgttat gagccatatt caacgggaaa cgtcttgctc gaggccgca 10380

ttaaattcca acatggatgc tgatttatat gggtataaat gggctcgca taatgtcggg 10440

caatcaggtg cgacaatcta tcgattgtat gggaagcccc atgcgccaga gttgtttctg 10500

aaacatggca aaggtagcgt tgccaatgat gttacagatg agatggtcag actaaactgg 10560
ctgacggaat ttatgcctct tccgaccatc aagcatttta tccgtactcc tgatgatgca 10620
tggttactca ccaactgcgat ccccgggaaa acagcattcc aggtattaga agaatacct 10680
gattcagggtg aaaaatttgt tgatgcgctg gcagtgttcc tgcgccggtt gcattcgatt 10740
cctgtttgta attgtccttt taacagcgat cgcgtatttc gtctcgtca ggcgcaatca 10800
cgaatgaata acggttttgt tgatgcgagt gatattgatg acgagcgtaa tggctggcct 10860
gttgaacaag tctggaaaga aatgcataag cttttgcat tctcacgga ttcagtcgtc 10920

actcatggtg atttctcact tgataacctt atttttgacg aggggaaatt aataggttgt 10980
attgatgttg gacgagtcgg aatcgagac cgataccagg atcttgccat cctatggaac 11040
tgcctcgggtg agttttctcc ttcattacag aaacggcttt ttcaaaaata tggattgat 11100
aatctgata tgaataaatt gcagtttcat ttgatgctcg atgagttttt ctaagggcgg 11160
cctgccacca taccacgcc 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 cgctcgtcgt ctactgagg ccgggcgacc aaaggtcgcc 60
cgacgcccg gctttgcccg ggccgcctca gtgagcgagc gagcgcgag agagggagtg 120
gccaaactcca tctactagggg ttcttgctag ctctgggtat ttaagcccga gtgagcacgc 180
agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240
tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg agggcggagt 300
tagggcggag ccaatcagcg tgcgccgttc cgaaggtgc cttttatggc tgggcggaga 360

atgggcggtg aacgccgatg attatataag gacgcgccgg gtgtggcaca gctagttccg 420
tcgcagccgg gatttgggtc gcggttcttg tttgtggatc cctgtgatcg tcaattggta 480
agtactgac tgtctatgcc tgggaaaggg tgggcaggag atggggcagt gcaggaaaag 540
tggcactatg aaccctgcag ccttaggaat gcatctagac aattgtacta accttcttct 600
ctttctctc ctgacagtc ggaaagccac catgtacgcc ctgttctgc tggccagcct 660
gctgggcgcc gccctggccg gcccgtgct gggcctgaag gagtgcaccc gcggcagcgc 720

cggtgtggtgc cagaacgtga agaccgccag cgactgcggc gccgtgaagc actgcctgca	780
gaccgtgtgg aacaagccca ccgtgaagag cctgccctgc gacatctgca aggacgtggt	840
gaccgccgcc ggcgacatgc tgaaggacaa cgccaccgag gaggagatcc tgggtgtacct	900
ggagaagacc tgcgactggc tgcccaagcc caacatgagc gccagctgca aggagatcgt	960
ggacagctac ctgcccgtga tcttggacat catcaagggc gagatgagcc gccccggcga	1020
ggtgtgcagc gccctgaacc tgtgagagag cctgcagaag cacctggcgg agctgaacca	1080
ccagaagcag ctggagagca acaagatccc cgagctggac atgaccgagg tgggtggcccc	1140
cttcatggcc aacatccccc tctgtctgta cccccaggac ggcccccgca gcaagcccca	1200
gcccgaaggac aacggcgacg tgtgccagga ctgcatccag atggtgaccg acatccagac	1260
cgccgtgcgc accaacagca ccttcgtgca ggccctggtg gagcacgtga aggaggagtg	1320
cgaccgcctg ggccccggca tggccgacat ctgcaagaac tacatcagcc agtacagcga	1380
gatcgccatc cagatgatga tgcacatgca gcccgaaggag atctgcgccc tgggtgggctt	1440
ctgcgacgag gtgaaggaga tgcccatgca gaccctggtg cccgccaagg tggccagcaa	1500
gaacgtgac cccgccctgg agctgggtgga gcccatcaag aagcacgagg tgcccccaa	1560
gagcgacgtg tactgcgagg tgtgagagtt cctggtgaag gaggtgacca agctgatcga	1620
caacaacaag accgagaagg agatcctgga cgccttcgac aagatgtgca gcaagctgcc	1680
caagagcctg agcgaggagt gccaggaggt ggtggacacc tacggcagca gcatcctgag	1740
catcctgctg gaggaggatga gccccgagct ggtgtgcagc atgtgcacc tgtgcagcgg	1800
caccgcctg cccgccctga ccgtgcacgt gaccagccc aaggacggcg gcttctgcga	1860
ggtgtgcaag aagctggtgg gctacctgga ccgcaacctg gagaagaaca gcaccaagca	1920
ggagatcctg gccgccctgg agaagggtg cagcttctg cccgaccct accagaagca	1980
gtgcgaccag ttcgtggccg agtacgagcc cgtgctgac gagatcctgg tggaggtgat	2040
ggaccccagc ttcgtgtgcc tgaagatcgg cgctgcccc agcgcccaca agcccctgct	2100
gggcaccgag aagtgcactt ggggccccag ctactggtgc cagaacaccg agaccgccgc	2160
ccagtgaac gccgtggagc actgcaagcg ccacgtgtgg aactgattgt ggccgaaccg	2220
ccgaactcag aggccggccc cagaaaacc gagcgagtag gggcgggcgc gcaggaggga	2280
ggagaactgg gggcgcgga ggctggtggg tgtggggggg ggagatgtag aagatgtgac	2340
gccgcggccc ggcgggtgcc agattagcgg acgcggtgcc cgcggttgca acgggatccc	2400
gggcgctgca gcttgggagg cggctctccc caggcgcgct ccgcggagac acccatccgt	2460

gaaccccagg tcccgggccc ccggctcgcc gcgcaccagg ggccggcgga cagaagagcg 2520
gccgagcggc tcgaggctgg gggaccgcgg gcgcggccgc gcgctgccgg gcgggaggct 2580
ggggggccgg ggccggggcc gtgccccgga gcgggtcgga ggccggggcc ggggcccggg 2640
gacggcggtt cccgcgcgg ctccagcggc tcggggatcc cggccgggcc ccgcagggac 2700
catgatgga ttcagcagcc ccagcagaga ggaatgcccc aagcctctga gccgggtgtc 2760
aatcatggcc ggatctctga caggactgct gctgcttcag gccgtgtctt gggcttctgg 2820
cgctagacct tgcatccca agagcttcgg ctacagcagc gtcgtgtgcg tgtgcaatgc 2880

cacctactgc gacagcttcg accctcctac ctttctgct ctgggcacct tcagcagata 2940
cgagagcacc agatccggca gacggatgga actgagcatg ggacccatcc aggccaatca 3000
cacaggcact ggcttgcgtc tgacactgca gcctgagcag aaattccaga aagtgaagg 3060
cttcggcgga gccatgacag atgccgccgc tctgaatcc ctggctctgt ctccaccagc 3120
tcagaacctg ctgctcaaga gctacttcag cgaggaaggc atcggtctaca acatcatcag 3180
agtgcccatg gccagctgcg acttcagcat caggacctac acctacgccg acacaccga 3240
cgatttccag ctgcacaact tcagcctgcc tgaagaggac accaagctga agatccctct 3300

gatccacaga gccctgcagc tggcacaag acccgtgtca ctgctggcct ctccatggac 3360
atctcccacc tggctgaaaa caaatggcgc cgtgaatggc aagggcagcc tgaaaggcca 3420
acctggcgac atctaccacc agacctgggc cagatacttc gtgaagtcc tggacgccta 3480
tgccgagcac aagctgcagt tttgggccgt gacagccgag aacgaacctt ctgctggact 3540
gctgagcggc taccctttc agtgccctggg ctttacacc gagcaccagc gggactttat 3600
cgcccgtgat ctgggaccca cactggccaa tagcacccac cataatgtgc ggctgctgat 3660
gctggacgac cagagactgc ttctgcccc ctgggctaaa gtggtgctga cagatcctga 3720

ggccgcaaaa tacgtgcagc gaatgcctgt gactggtat ctggactttc tggcccctgc 3780
caaggccaca ctgggagaga cacacagact gtccccaac accatgctgt tcgccagcga 3840
agcctgtgtg ggcagcaagt tttgggaaca gacgtgcgg ctcggcagct gggatagagg 3900
catgcagtac agccacagca tcatcaccaa cctgctgtac cagtcgtcg gctggaccga 3960
ctggaatctg gccctgaatc ctgaaggcgg ccctaactgg gtccgaaact tcgtggacag 4020
ccccatcgc gtggacatca ccaaggacac cttctacaag cagcccatgt tctaccacct 4080
gggacacttc agcaagtcca tccccgagg ctctcagcgc gttggactgg tggcttccca 4140

gaagaacgat ctggacgccg tggctctgat gcaccctgat ggatctgctg tgggtgtggt 4200
cctgaaccgc agcagcaaat atgtgccct gaccatcaag gatcccgcgg tgggattcct 4260
ggaaacaatc agccctggct actccatcca cacctacctg tggcgtagac agtgacaatt 4320

gttaattaag tttaaaccct cgaggccgca agcaataaaa tatctttatt ttcattacat 4380
ctgtgtgttg gttttttgtg tggagatcca cgataacaaa cagctttttt ggggtgaaca 4440
tattgactga attccctgca ggttggccac tccctctctg cgcgctcgct cgctcactga 4500
ggccgcccgg gcaaagcccg ggcgtcgggc gaccttttgt cgcccggcct cagtgagcga 4560

gcgagcgcg agagaggag tggccaactc catcactagg ggttcctgcg gccgctcgta 4620
cggctctcag gaattcctgc aggataactt gccaacctca ttctaaaatg tatatagaag 4680
cccaaaagac aataacaaaa atattcttgi agaacaaaat gggaaagaat gttccactaa 4740
atatcaagat ttagagcaaa gcatgagatg tgtggggata gacagtgagg ctgataaaat 4800
agagtagagc tcagaaacag acccattgat atatgtaagt gacctatgaa aaaaatatgg 4860
cattttacaa tgggaaaatg atggtctttt tcttttttag aaaaacaggg aaatatattt 4920
atatgtaaaa aataaaaggg aacctatatg tcataccata cacacaaaaa aattccagtg 4980

aattataagt ctaaatggag aaggcaaac tttaaactt ttagaaaata atatagaagc 5040
atgcagacca gcctggccaa catgatgaaa cctctctac taataataaa atcagtagaa 5100
ctactcagga ctactttgag tgggaagtcc ttttctatga agacttctt ggccaaaatt 5160
aggctctaaa tgcaaggaga tagtgcacat tgccctggctg cacttactga taaatgatgt 5220
taccaccatc tttaacaaa tgcacaggaa caagttatgg tactgatgtg ctggattgag 5280
aaggagctct acttcttga caggacacat ttgtatcaac ttaaaaaagc agatttttgc 5340
cagcagaact attcattcag aggtaggaaa cttagaatag atgatgtcac tgattagcat 5400

ggcttcccca tctccacagc tgcttccac ccaggttgcc cacagttgag tttgtccagt 5460
gctcagggtc gccactctc agtaagaagc cccacaccag cccctctcca aatatgttgg 5520
ctgttcttc cattaaagt accccactt agagcagcaa gtggatttct gtttcttaca 5580
gttcaggaag gaggagtcat ctgtgagaac ctggagcctg agatgcttct aagteccact 5640
gctactgggg tcaggaagc cagactccag catcagcagt caggagcact aagcccttgc 5700
caacatctg tttctcagaa aaactgctc cattataatg gttgtcttt tttaagctat 5760
caagccaaac aaccagtgc taccattatt ctcacacct gaagccaagg gttctagcaa 5820

aagtcaagct gtcttgtaat ggttgatgtg cctccagctt ctgtcttcag tcactccact 5880
cttagcctgc tctgaatcaa ctctgaccac agttccctgg agccctgcc acctgtgcc 5940
cctgccacct tctccatctg cagtgtgtg cagccttctg cactcttgca gagctaatag 6000
gtggagactt gaaggaagag gaggaaagt tctcataata gccttgctgc aagctcaaat 6060
gggaggtggg cactgtgcc aggagcctt gagcaaaggc tgtgcccaac ctctgactgc 6120
atccaggttt ggtcttgaca gagataagaa gccttgctt ttggagcaa aatctaggtc 6180

agacttaggc aggattctca aagtttatca gcagaacatg aggagaaga ccctttctgc 6240

tccagcttct tcaggctcaa ccttcacag aatagataga aagagaggct gtgagggttc 6300

ttaaaacaga agcaaactcg actcagagaa taaacaacct cctagtaaac tacagcttag 6360

acagagcatc tggtaggtgag tgtgctcag gtctactca actgtctggt atcagccctc 6420

atgaggactt ctctcttttc cctcatagac ctccatctct gttttctta gcctgcagaa 6480

atctggatgg ctattcacag aatgcctgtg ctttcagagt tgcatTTTT ctctggtatt 6540

ctggttcaag catttgaagg taggaaaggt tctcaagtg caagaaagcc agccctgagc 6600

ctcaactgcc tggctagtgt ggtcagtagg atgcaaaggc tgttgaatgc cacaaggcca 6660

aactttaacc tgtgtaccac aagcctagca gcagaggcag ctctgctcac tggaaactctc 6720

tgtcttcttt ctctgagcc tttcttttc ctgagttttc tagctctcct caaccttacc 6780

tctgccctac ccaggacaaa cccaagagcc actgtttctg tgatgtctc tccagcccta 6840

attaggcatc atgacttcag cctgaccttc catgctcaga agcagtgtca atccacttca 6900

gatgagctgc tctatgcaac acaggcagag cctacaaacc tttgcaccag agccctccac 6960

atatcagtgt ttgttcatac tcaattcaac agcaaatgtg actgtctgaga ttaagatTTT 7020

acacaagatg gtctgttaatt tcacagttag ttttatccca ttaggtatga aagaattagc 7080

ataattcccc ttaaacatga atgaatctta gatTTTTaa taaatagttt tggaaagtaa 7140

gacagagaca tcaggagcac aaggaatagc ctgagaggac aaacagaaca agaaagagtc 7200

tggaaataca caggatgttc ttggctcct caaagcaagt gcaagcagat agtaccagca 7260

gccccaggct atcagagccc agtgaagaga agtaccatga aagccacagc tctaaccacc 7320

ctgttccaga gtgacagaca gtcccaaga caagccagcc tgagccagag agagaactgc 7380

aagagaaagt ttctaattta ggttctgtta gattcagaca agtgcaggtc atctctctc 7440

cacagctact cacctctcca gcctaacaaa gcctgcagtc cacactcaa ccctggtgtc 7500

tcacctcta gcctctccca acatctgtct ctctgacct cttctgcac tctcatctca 7560

ccatctccca ctgtctacag cctactcttg caactacct ctcatTTTct gacatcctgt 7620

ctacatcttc tgccatactc tgccatctac cataccacct cttaccatct accacacct 7680

cttttatctc catccctctc agaagcctcc aagctgaatc ctgctttatg tgttcatctc 7740

agccctgca tggaaagctg accccagagg cagaactatt cccagagagc ttggccaaga 7800

aaaacaaaac taccagcctg gccaggctca ggagtagtaa gctgcagtgt ctgttgtgtt 7860

ctagcttcaa cagctgcagg agttccactc tcaaatgtc cacatttctc acatctctct 7920

gattctggtc actacccatc ttcaaagaac agaatacttc acatcagcat actgtgaagg 7980
 actagtcacg ggtgcagctg ctacagctg caaagtcatt ctggatgggt gagagcttac 8040
 aaacatttca tgatgctccc cccgctctga tggctggagc ccaatcccta cacagactcc 8100
 tgctgtatgt gttttccttt cactctgagc cacagccaga gggcaggcat tcagtctcct 8160
 cttcaggctg gggctggggc actgagaact cacccaacac cttgtctctca ctcttctgc 8220
 aaaacaagaa agagctttgt gctgcagtag ccatgaagaa tgaaaggaag gctttaacta 8280
 aaaaatgtca gagattatct tcaacccctt actgtggatc accagcaagg aggaaacaca 8340

 acacagagac attttttccc ctcaaattat caaaagaatc actgcatttg ttaaagagag 8400
 caactgaatc aggaagcaga gttttgaaca tatcagaagt taggaatctg catcagagac 8460
 aaatgcagtc atggttgttt gctgcatacc agccctaatac 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 ttaccctgg 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
 tagggggcgg atggggcggg ttaggggcgg gactatggtt gctgactaat tgagatgcat 9000
 gctttgcata ctctgcctg ctggggagcc tggggacttt ccacacctgg ttgtgacta 9060
 attgagatgc atgctttgca tacttctgcc tctggggag cctggggact ttccacaccc 9120
 taactgacac acattccaca gctgcattaa tgaatcggcc aacgcgcggg gagaggcgg 9180

 ttgcgtattg ggcgctcttc cgcttcctcg ctactgact cgctgcgctc ggtcgttcgg 9240
 ctgcggcgag cggtatcagc tcactcaaag gcggtaatc gggtatccac agaatcaggg 9300
 gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag 9360
 gccgcgttgc tggcgttttt ccataggtc cgccccctg acgagcatca caaaaatcga 9420
 cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct 9480
 ggaagctccc tcgtgcgctc tctgttccg accctgccg ttaccggata cctgtccgcc 9540
 tttctccctt cggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg 9600

 gtgtaggtcg ttgcgtccaa gctgggctgt gtgcacgaac cccccgttca gcccaccgc 9660
 tgcgccttat ccggttaacta tcgtcttgag tccaacccgg taagacacga ctatcgcca 9720
 ctggcagcag cactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag 9780

ttcttgaagt ggtagcctaa ctacggctac actagaagaa cagtatttgg tatctgcgct	9840
ctgctgaagc cagttacctt cggaataaga gttgtagct cttgatccgg caaacaacc	9900
accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaagga	9960
tctcaagaag atcctttgat cttttctacg gggctctgacg ctacgtggaa cgaaaaactca	10020
cgtaaaggga ttttggcat gagattatca aaaaggatct tcacctagat ctttttaaat	10080
taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagtac	10140
caatgcttaa tcagttaggc acctatctca gcgactctgc tatttcgttc atccatagtt	10200
gcctgactcc tgcaaacacc gttgtgtctc aaaatctctg atgttacatt gcacaagata	10260
aaaatatatc atcatgaaca ataaaactgt ctgcttacat aaacagtaat acaaggggtg	10320
ttatgagcca taitcaacgg 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
ctcttcgac catcaagcat tttatccgta ctctgatga tgcatggtta ctaccactg	10620
cgatccccgg gaaaacagca ttccaggtat tagaagaata tcttgattca ggtgaaaata	10680
ttgttgatgc gctggcagtg ttctgcgcc ggttgcatc gattcctgtt tgtaattgtc	10740
cttttaacag cgatcgcta tttcgtctcg ctacggcgca atcacgaatg aataacggtt	10800
tggttgatgc gattgatttt gatgacgagc gtaatggctg gcctgttgaa caagtctgga	10860
aagaaatgca taagcttttg ccattctcac cggattcagt cgtcactcat ggtgatttct	10920
cacttgataa cctatTTTT gacgagggga aattaatagg ttgtattgat gttggacgag	10980
tcggaatcgc agaccgatac caggatcttg ccacctatg gaactgcctc ggtgagtttt	11040
ctcttcatt acagaaacgg ctttttcaaa aatatggtat tgataatcct gatatgaata	11100
aattgcagtt tcatttgatg ctcatgagc ttttctaagg gcggcctgcc accataccca	11160
cgccgaaca agcgtcatg agcccgaagt ggcgagcccg atcttccca tcggtgatgt	11220
cgcgcatata ggcgcagca 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 ctactgagg cgcgccgggc aaagcccggg	60
cgtcgggcga cctttggtcg cccggcctca gtgagcgagc gagcgcgag agagggagt	120
gccaaactcca tctactagggg ttccctgctag ctctgggtat ttaagcccga gtgagcacgc	180
agggctctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc	240
tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcgggtac	300
ctagttatta atagtaatca attacggggg cattagtcca tagcccatat atggagtacc	360
gcgttacata acttacgcta aatggccgcg ctggctgacc gccaacgac ccccgcccat	420
tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc	480
aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc	540
caagtacgcc ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgccagtc	600
acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta	660
ccatggctga ggtgagcccc acgttctgct tctctctccc catctcccc cctctccac	720
ccccaat ttt gtattttat ttttttaaat tttttgtgc agcgatgggg gcgggggggg	780
ggggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga	840
gaggtgcggc ggcagccaat cagagcgggc cgctccgaaa gtttcctttt atggcgaggc	900
ggcgggcgcg gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgacgc	960
tgcttcgcc ccgtgcccc ctcccgccc gcctcgccc gcccggccc gctctgactg	1020
accgcgttac tcccacaggt gagcgggcgg gacggccctt ctctccggg ctgtaattag	1080
cgcttggttt aatgacggct tgtttctttt ctgtggctgc gtgaaagcct tgaggggctc	1140
cgggagctag agcctctgct aacctgttc atgccttctt ctttttcta cagctcctgg	1200
gcaacgtgct ggttattgtg ctgtctcatc attttgcaa agaattctc gaagatccga	1260
agggaaagtc ttccacgact gtgggatccg ttcaagata tcaccggttg agccaccatg	1320
gaattcagca gcccagcag agaggaatgc cccaagctc tgagccgggt gtcaatcatg	1380
gccgatctc tgacaggact gctgtgtctt caggccgtgt cttgggcttc tggcgctaga	1440
ccttgcatcc ccaagagctt cggctacagc agcgtcgtgt gcgtgtgcaa tgccacctac	1500
tgcgacagct tcgacctcc tacctttctt gctctgggca cttcagcag atacgagagc	1560
accagatccg gcagacggat ggaactgagc atgggacca tccaggcca tcacacaggc	1620
actggcctgc tgctgacct gcagcctgag cagaaattcc agaaagtga aggcttcggc	1680
ggagccatga catatgccg cgctctgaat atcctggctc tgtctccacc agctcagaac	1740
ctgctgctca agagctactt cagcgaggaa ggcatcggct acaacatcat cagagtgcgc	1800

atggccagct gcgacttcag catcaggacc tacacctacg ccgacacacc cgacgatttc	1860
cagctgcaca acttcagcct gcctgaagag gacaccaagc tgaagatccc tctgatccac	1920
agagccctgc agctggcaca aagaccctg tctactgctgg cctctccatg gacatctccc	1980
acctggctga aaacaaatgg cgccgtgaat ggcaaggcca gcctgaaagg ccaacctggc	2040
gacatctacc accagacctg ggccagatac ttcgtgaagt tcctggacgc ctatgccgag	2100
cacaagctgc agttttgggc cgtgacagcc gagaacgaac cttctgctgg actgctgagc	2160
ggctacccct ttcagtgcct gggcittaca cccgagcacc agcgggactt tatcgcccgt	2220
gatctgggac ccacactggc caatagcacc caccataatg tgcggctgct gatgctggac	2280
gaccagagac tgcttctgcc ccactgggct aaagtgggtg 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 cagcccccac	2640
atcgtggaca tcaccaagga caccittctac aagcagccca tgtttacca cctgggacac	2700
ttcagcaagt tcatccccga gggctctcag cgcgttggac tgggtggcttc ccagaagaac	2760
gatctggacg ccgtggctct gatgcacct gatggatctg ctgtggtggt ggtcctgaac	2820
cgcagcagca aagatgtgcc cctgaccatc aaggatcccg ccgtgggatt cctggaacaa	2880
atcagccctg gctactccat ccacacctac ctgtggcgta gacagtgaca atgtttaatt	2940
aagtttaaac cctcagagcc gcaagcttat cgataatcaa cctctggatt acaaaatttg	3000
tgaaagattg actggatttc ttaactatgt tgctcctttt acgctatgtg gatacgtgc	3060
tttaatgcct ttgtatcatg ctattgcttc ccgtatggct ttcattttct cctccttgta	3120
taaatcctgg ttgctgtctc ttatgagga gttgtggccc gttgtcaggc aacgtggcgt	3180
ggtgtgcact gtgtttgtg acgcaacccc cactggttgg ggcatgcca ccacctgtca	3240
gtctctttcc gggactttcg ctttccccct cctattgcc acggcggaac tcatgccgc	3300
ctgccttgcc cgtctgtgga caggggctcg gctgttgggc actgacaatt ccgtggtgtt	3360
gtcggggaaa tcatgtctt ttcttggct gctgcctgt gttgccacct ggattctgcg	3420
cgggacgtcc ttctgtacg tcccttcggc cctcaatcca gcggaccttc cttcccgcgg	3480
cctgtgcgg gctctgcggc ctcttcgcg tcttcgctt cgcctcaga cgagtcggat	3540
ctccctttgg gccgcctccc cgcacgata ccgtcgacta gagtcgtg atcagcctcg	3600

actgtgcctt ctagttgcc a gccatctgtt gtttggccct ccccggtgcc ttccttgacc 3660

ctggaagggtg ccactccac tgccttttcc taataaaatg aggaattgc atcgattgt 3720

ctgagtaggt gtcattctat tctgggggt ggggtgggc aggacagcaa gggggaggat 3780

tgggaagaca atagcaggca tgctggggag agatccacga taacaacag ctttttggg 3840

gtgaacatat tgactgaatt cctgcagggt tggccactcc ctctctgcgc gctcgtcgc 3900

tcactgaggc cgccgggca aagccgggc gtcgggcgac ctttggcgc cgggcctcag 3960

tgagcgagcg agcgcgaga gagggagtgg ccaactccat cactaggggt tctgcggcc 4020

gctcgtacgg tctcaggaa ttctgcagg ataactgcc aacctcattc taaaatgtat 4080

atagaagccc aaaagacaat aacaaaaata ttctttaga acaaatggg aaagaatgtt 4140

ccactaaata tcaagattta gagcaaagca tgagatgtgt ggggatagac agtgaggctg 4200

ataaataga gtagagctca gaaacagacc cattgatata tgtaagtac ctatgaaaaa 4260

aatatggcat ttacaatgg gaaatgatg gtcttttct ttttagaaa aacagggaaa 4320

tatatattata tgtaaaaaat aaaagggaac ccataatgtca taccatacac aaaaaaaat 4380

tccagtgaat tataagtcta aatggagaag gaaaaacttt aaatctttta gaaaataata 4440

tagaagcatg cagaccagcc tggccaacat gatgaaaccc tctctactaa taataaaatc 4500

agtagaacta ctcaggacta ctttgagtgg gaagtccttt tctatgaaga ctcttttggc 4560

caaaattagg ctctaaatgc aaggagatag tgcattatgc ctggctgcac ttactgataa 4620

atgatgttat caccatcttt aaccaaagtc acaggaacaa gttatggtac tgatgtgctg 4680

gattgagaag gagctctact tccttgacag gacacatttg tatcaactta aaaaagcaga 4740

tttttgcag cagaactatt cattcagagg taggaaactt agaatagatg atgtcactga 4800

ttagcatggc ttcccatct ccacagctgc ttccacca ggttggccac agttgagttt 4860

gtccagtgtc cagggtgcc cactctcagt aagaagcccc acaccagccc ctctccaaat 4920

atgttggctg ttccttccat taaagtgacc ccactttaga gcagcaagtg gatttctgtt 4980

tcttacagtt caggaaggag gactcagctg tgagaacctg gagcctgaga tgcttctaag 5040

tccactgct actggggctca gggaagccag actccagcat cagcagtcag gagcactaag 5100

cccttgccaa catcctgttt ctgagagaaa ctgcttccat tataatggtt gtcctttttt 5160

aagctatcaa gccaaacaac cagtgtctac cattattctc atcacctgaa gccagggtt 5220

ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct ccagcttctg tcttcagtca 5280

ctccactctt agcctgctct gaatcaactc tgaccacagt tccctggagc cctgccacc 5340

tgctgccct gccaccttct ccatctgcag tgcgtgcag cttctgcac tcttcagag	5400
ctaatagggt gagacttgaa ggaagaggag gaaagtttct cataatagcc ttgctgaag	5460
ctcaaatggg aggtgggcac tgtgcccagg agccttggag caaaggctgt gcccacctc	5520
tgactgcatc caggtttgggt cttgacagag ataagaagcc ctggcttttg gagccaaaat	5580
ctaggtcaga cttaggcagg attctcaaag ttatcagca gaacatgagg cagaagaccc	5640
tttctgctcc agcttcttca ggctcaacct tcatcagaat agatagaaag agaggctgtg	5700
agggttctta aaacagaagc aaatctgact cagagaataa acaacctcct agtaaacctac	5760
agcttagaca gagcatctgg tggtagtgt gctcagtgtc ctactcaact gtctggtatc	5820
agccctcatg aggacttctc ttctttccct catagacctc catctctgtt ttccttagcc	5880
tgcagaaatc tggatggcta ttcacagaat gcctgtgctt tcagagtgc atttttctc	5940
tggttattctg gttcaagcat ttgaaggtag gaaagtttct ccaagtgcaa gaaagccagc	6000
cctgagcctc aactgcctgg ctagtgtgtt cagtaggatg caaaggctgt tgaatgccac	6060
aaggccaaac tttaacctgt gtaccacaag cctagcagca gaggcagctc tgctcactgg	6120
aactctctgt cttctttctc ctgagccttt tcttttctg agttttctag ctctctcaa	6180
ccttacctct gccctacca ggacaaacc aagagccact gtttctgtga tgcctctcc	6240
agccctaatt aggcacatg acttcagcct gaccttccat gctcagaagc agtgctaate	6300
cacttcagat gagctgctct atgcaacaca ggcagagcct acaaaccttt gcaccagagc	6360
cctccacata tcagtgtttg ttcatactca ctcaacagc aaatgtgact gctgagatta	6420
agattttaca caagatggtc tgtaatttca cagttagttt tatccatta ggtatgaaag	6480
aattagcata attcccccta 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
aaccacctg ttccagagtg acagacagtc cccaagacaa gccagcctga gccagagaga	6780
gaactgcaag agaaagtctc taatttaggt tctgttagat tcagacaagt gcaggtcac	6840
ctctctccac agctactcac ctctccagcc taacaaagcc tgcagtccac actccaacct	6900
tgggtgtctc cctcctagcc tctcccaaca tctgtctctc tgaccatctt ctgcactct	6960
catctacca tctccactg tctacagcct actcttgcaa ctaccatctc atttctgac	7020
atcctgtcta catcttctgc catactctgc catctacat accacctctt accatctacc	7080
acacatctt ttatctccat ccctctcaga agcctccaag ctgaatcctg ctttatgtgt	7140
tcatctcagc ccctgcatgg aaagctgacc ccagaggcag aactattccc agagagcttg	7200

gccaaagaaa acaaaactac cagcctggcc aggctcagga gtagtaagct gcagtgtctg 7260
 ttgtgttcta gcttcaacag ctgcaggagi tccactctca aatgtccac atttctcaca 7320
 tcctcctgat tctggctact acccatcttc aaagaacaga atatctcaca tcagcatact 7380
 gtgaaggact agtcatgggt gcagctgctc agagctgcaa agtcattctg gatggaggag 7440

 agcttacaaa catttcatga tgcctcccc gctctgatgg ctggagccca atccctacac 7500
 agactcctgc tgtatgtgtt ttccctttcac tctgagccac agccagaggg caggcattca 7560
 gtctcctctt caggctgggg ctggggcact gagaactcac ccaacacctt gctctcactc 7620
 cttctgcaaa acaagaaaga gctttgtgct gcagtagcca tgaagaatga aaggaaggct 7680
 ttaactaaaa aatgtcagag attatittca accccttact gtggatcacc agcaaggagg 7740
 aaacacaaca cagagacatt ttttccctc aaattatcaa aagaatcact gcatttgta 7800
 aagagagcaa ctgaatcagg aagcagagt ttgaacatat cagaagttag gaatctgcat 7860

 cagagacaaa tgcagtcag gtgtgttgct gcataccagc cctaattcatt agaagcctca 7920
 tggacttcaa acatcattcc ctctgacaag atgctctagc ctaactccat gagataaaat 7980
 aaatctgcct ttgagagcca aagaagagtc caccagcttc ttctcagtg gaacaagagc 8040
 tccagtcagg ttagtcagtc cagtcagta gaggagacca gtctgcatcc tctaattttc 8100
 aaaggcaaga agatttggtt accctggaca ccaggcacia gtgaggtcac agagctctta 8160
 gatatgcagt cctcatgagt gaggagacta aagcgcatgc catcaagact tcagtgtaga 8220
 gaaaacctcc aaaaaagcct cctcactact tctggaatag ctgagaggcc 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
 cacacctaa ctgacacaca ttccacagct gcattaatga atcgccaac gcgcggggag 8580
 aggcggtttg cgtattgggc gctcttcgc ttctcctc actgactcgc tgcgctcggt 8640
 cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacgtt tatccacaga 8700

 atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 8760
 taaaaaggcc gcgttgctgg cgtttttcca taggtccgc cccctgacg agcatcacia 8820
 aaatcgacgc tcaagtcaaa ggtggcgaaa cccgacagga ctataaagat accaggcgtt 8880
 tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgtta ccggatacct 8940
 gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgt gtaggtatct 9000
 cagttcgggt taggtcgctt gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 9060

cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 9120

atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 9180

tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag tatittggtat 9240

ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 9300

acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 9360

aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 9420

aaactcacgt taagggatit ttggtcatgag attatcaaaa aggatcttca cctagatcct 9480

tttaaattaa aaatgaagtt ttaaataaat ctaaagtata tatgagtaaa ctigtgtctga 9540

cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcac 9600

catagtggcc tgactcctgc aaaccacgtt gtgtctcaaa atctctgatg ttacattgca 9660

caagataaaa atatatcatc atgaacaata aaactgtctg cttacataaa cagtaataca 9720

aggggtgtta tgagccatat tcaacgggaa acgtcttgct cgaggccgag attaaattcc 9780

aacatggatg ctgatttata tgggtataaa tgggctcgcg ataatgtcgg gcaatcaggt 9840

gcgacaatct atcgattgta tgggaagccc gatgcgccag agttgtttct gaaacatggc 9900

aaaggtagcg ttgccaatga tgttacagat gagatgggtca gactaaactg gctgacggaa 9960

tttatgcctc ttccgacat caagcatttt atccgtactc ctgatgatgc atggttactc 10020

accactgcga tccccgggaa aacagcattc caggtattag aagaatatcc tgattcaggt 10080

gaaaatattg ttgatgcgt ggcagtggtc ctgcgccggt tgcattcgat tcctgtttgt 10140

aattgtcctt ttaacagcga tcgcgtattt cgtctcgtc aggcgcaatc acgaatgaat 10200

aacggtttgg ttgatgcgag tgattttgat gacgagcgta atggctggcc tgttgaacaa 10260

gtctggaaag aaatgcataa gcttttgcca ttctaccgg attcagtcgt cactcatggt 10320

gattttctac ttgataacct tatttttgac gaggggaaat taataggttg tattgatgtt 10380

ggacgagtcg gaatcgaga ccgataccag gatcttgcca tcctatggaa ctgcctcgg 10440

gagttttctc cttcattaca gaaacggctt tttcaaaaat atggtattga taatcctgat 10500

atgaataaat tgcagtttca tttgatgctc gatgagtttt tctaaggcg gcctgccacc 10560

ataccacgc cgaacaagc gctcatgagc ccgaagtggc gagcccgatc ttcccatcg 10620

gtgatgtcgg cgatatagcg gccagcaacc gcacctgtgg cgccggtgat gagggcgcg 10680

caagtcgacg tccggcagtc 10700

<210> 8

<211> 10700

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 8

ttggccactc ccctctctgcg cgctcgctcg ctactgagg cgggcgacc aaaggtcgcc	60
cgacgcccg gctttgccc ggccgctca gtgagcagc gagcgcgag agaggagtg	120
gccactatt agatctgat gccgcgctag ctctgggtat ttaagcccga gtgagcacgc	180
agggtctcca tttgaagcg ggaggttac cgttcgtcga ctactagtgg gtaccagagc	240
tcctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac	300
ctagttatta atagtaatca attacggggt cattagtcca tagcccatat atggagtcc	360
gcgttacata acttacggt aatggccgc ctggctgacc gccaacgac ccccgccat	420

tgacgtcaat aatgacgtat gttcccatag taacccaat agggactttc cattgacgtc	480
aatgggtgga gtatttacgg taaactgcc acttggcagt acatcaagt tatcatatgc	540
caagtacgcc ccctattgac gtcaatgac gtaaatggcc cgcctggcat tatgccagt	600
acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta	660
ccatggtcga ggtgagcccc acgttctgct tcaactctcc catctcccc cctccccac	720
ccccaatfff gtatttattt attttttaat tattttgtgc agcgatgggg gcgggggggg	780
ggggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga	840

gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa gtttcctttt atggcgaggc	900
ggcggcgggc gggccctat aaaaagcgaa gcgcgcggcg ggccggagtc gctgcgacgc	960
tgccctcgcc ccgtgcccc ctccgccgc gcctcgcgcc gcccgcgccg gctctgactg	1020
accggttac tcccacagt gagcgggcgg gacggccctt ctctccggg ctgtaattag	1080
cgcttggttt aatgacggct tgtttctttt ctgtggtgc gtgaaagcct tgaggggctc	1140
cgggagctag agcctctgct aacctgttc atgccttctt ctttttcta cagctcctgg	1200
gcaacgtgct ggttattgtg ctgtctcatc attttgcaa agaattctc gaagatccga	1260

agggaaagtc ttccacgact gtgggatccg ttcgaagata tcaccggtg agccaccatg	1320
gaattcagca gcccagcag agaggaatgc cccaagctc tgagccgggt gtcaatcatg	1380
gccgatctc tgacaggact gctgctgctt caggccgtgt cttgggcttc tggcgctaga	1440
ccttgcatcc ccaagagctt cggctacagc agcgtcgtgt gcgtgtgcaa tgccacctac	1500
tgcgacagct tcgacctcc tacctttcct gctctgggca ccttcagcag atacgagagc	1560
accagatccg gcagacggat ggaactgagc atgggacca tccaggcaa tcacacaggc	1620

actggcctgc tgctgacact gcagcctgag cagaaattcc agaaagtgaaggcttcggc 1680

ggagccatga cagatgccgc cgctctgaat atcctggctc tgtctccacc agctcagaac 1740

ctgctgctca agagctactt cagcgaggaa ggcatcggct acaacatcat cagagtgcc 1800

atggccagct gcgacttcag catcaggacc tacacctacg ccgacacacc cgacgatttc 1860

cagctgcaca acttcagcct gcctgaagag gacaccaagc tgaagatccc tctgatccac 1920

agagccctgc agctggcaca aagacccgtg tctctgctgg cctctccatg gacatctccc 1980

acctggctga aaacaaatgg cgccgtgaat ggcaaggcca gcctgaaagg ccaacctggc 2040

gacatctacc accagacctg ggccagatac ttcgtgaagt tctgggacgc ctatgccgag 2100

cacaagctgc agttttgggc cgtgacagcc gagaacgaac cttctgctgg actgctgagc 2160

ggctaccctt ttcagtgcct gggctttaca cccgagcacc agcgggactt tatcgccct 2220

gatctgggac ccacactggc caatagcacc caccataatg tgcggctgct gatgctggac 2280

gaccagagac tgcttctgcc ccactgggct aaagtgggtg tgacagatcc tgaggccgcc 2340

aaatacgtgc acggaatcgc cgtgactggg tatctggact ttctggcccc tgccaaggcc 2400

acactgggag agacacacag actgttcccc aacacatgc tgttcgccag cgaagcctgt 2460

gtgggcagca agttttggga acagagcgtg cggctcggca gctgggatag aggcattgag 2520

tacagccaca gcatcatcac caacctgctg taccacgtcg tcggctggac cgactggaat 2580

ctggccctga atcctgaagg cgccctaac tgggtccgaa acttcgtgga cagcccatc 2640

atcgtggaca tcaccaagga cacttctac aagcagccca tgttctacca cctgggacac 2700

ttcagcaagt tcatccccga gggctctcag cgcgttggac tgggtggcttc ccagaagaac 2760

gatctggacg ccgtggctct gatgcacct gatggatctg ctgtgggtgt ggtcctgaac 2820

cgcagcagca aagatgtgcc cctgaccatc aaggatcccg ccgtgggatt cctggaacaa 2880

atcagccctg gctactccat ccacacctac ctgtggcgta gacagtgaca attgttaatt 2940

aagtttaaac cctcgaggcc gcaagcttat cgataatcaa cctctggatt acaaaatttg 3000

tgaaagattg actggtatc ttaactatgt tgcctctttt acgctatgtg gatacgtgc 3060

tttaatgcct ttgtatcatg ctattgcttc ccgtatggct ttcatTTTT cctccttgta 3120

taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtcaggc aacgtggcgt 3180

ggtgtgcaact gtgtttgtcg acgcaacccc cactgggttg ggcatgcca ccacctgtca 3240

gtcctcttcc gggactttcg ctttccccct ccctattgcc acggcggaac tcatcgccgc 3300

ctgccttgcc cgctgctgga caggggctcg gctgttgggc actgacaatt ccgtggtgtt 3360

gtcggggaaa tcatcgtcct ttccttggct gctcgcctgt gttgccacct ggattctgcg 3420
cgggacgtcc ttctgtacg tcccttcggc cctcaatcca gcggaccttc cttcccgcgg 3480
cctgtctgcc gctctgcggc ctcttccgcg tcttcgcctt cgcctcaga cgagtcggat 3540
ctcccttttg gccgcctccc cgcctcgata ccgtcgacta gagctcgtg atcagcctcg 3600
actgtgcctt ctagttgcca gccatctgtt gtttgcctt ccccgctgcc ttccttgacc 3660
ctggaagggt ccactccac tgtctttcc taataaaatg aggaaattgc atcgattgt 3720
ctgagtaggt gtcattctat tctgggggtt ggggtggggc aggacagcaa gggggaggat 3780

tgggaagaca atagcaggca tgctggggag agatccacga taacaacag cttttttggg 3840
gtgaacatat tgactgaatt ccctgcaggt tggccactcc ctctctgcgc gctcgtcgc 3900
tactgaggc cggccgggca aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag 3960
tgagcgagcg agcgcgcaga gagggagtgg ccaactccat cactaggggt tctcgcggc 4020
gctcgtacgg tctcgaggaa ttcttcgagg ataacttgc aacctcattc taaaatgtat 4080
atagaagccc aaaagacaat acaaaaaata ttctttaga acaaatggg aaagaatgtt 4140
ccactaaata tcaagattta gagcaaagca tgagatgtgt ggggatagac agtgaggctg 4200

ataaaataga gtagagctca gaaacagacc cattgatata tgtaagtac ctatgaaaa 4260
aatatggcat ttacaatgg gaaaatgatg gtctttttct tttttagaaa aacagggaaa 4320
tatatttata tgtaaaaaat aaaagggaa ccatatgtca taccatacac acaaaaaaat 4380
tccagtgaat tataagtcta aatggagaag gcaaaacttt aaatctttta gaaaataata 4440
tagaagcatg cagaccagcc tggccaacat gatgaaaccc tctctactaa taataaaatc 4500
agtagaacta ctcaggacta ctttgagtgg gaagtcttt tctatgaaga cttctttggc 4560
caaaattagg ctctaaatgc aaggagatag tgcacatgc ctggctgcac ttactgataa 4620

atgatgttat caccatcttt aaccaaatgc acaggaacaa gttatggtac tgatgtgctg 4680
gattgagaag gagctctact tcttgacag gacacatttg tatcaactta aaaaagcaga 4740
tttttgccag cagaactatt cattcagagg taggaaactt agaatagatg atgtcactga 4800
ttagcatggc ttcccatct ccacagctgc ttcccacca ggttgccac agttgagttt 4860
gtccagtgt cagggtgcc cactctcag aagaagcccc acaccagccc ctctccaaat 4920
atgttggtg ttccttccat taaagtgacc ccactttaga gcagcaagt gatttctgtt 4980
tcttacagtt caggaaggag gagtcagctg tgagaacctg gagctgaga tgcttctaag 5040

tcccactgct actggggica gggaagccag actccagcat cagcagtcag gagcactaag 5100
cccttgcaa catcctgttt ctgagagaaa ctgcttccat tataatggtt gtcctttttt 5160
aagctatcaa gccaaacaac cagtgtctac cattattctc atcacctgaa gccaaagggtt 5220

ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct ccagcttctg tcttcagtca	5280
ctccactctt agcctgtctt gaatcaactc tgaccacagt tccctggagc cctgccacc	5340
tgctgcccc gccaccttct ccatctgcag tgctgtgcag ccttctgcac tcttgacagag	5400
ctaatagggt gagacttgaa ggaagaggag gaaagtttct cataatagcc ttgctgcaag	5460
ctcaaatggg aggtgggcac tgtgcccagg agccttggag caaaggctgt gcccacctc	5520
tgactgcac caggtttggg cttgacagag ataagaagcc ctggcttttg gagccaaaat	5580
ctaggtcaga cttaggcagg attctcaaag tttatcagca gaacatgagg cagaagaccc	5640
tttctgtcc agcttcttca ggctcaacct tcatcagaat agatagaaag agaggctgtg	5700
agggttctta aaacagaagc aaatctgact cagagaataa acaacctcct agtaaaactac	5760
agcttagaca gagcatctgg tggtagtgt gctcagtgct ctactcaact gtctggtatc	5820
agccctcatg aggacttctc tcttttccct catagacctc catctctgtt ttccttagcc	5880
tgacagaaatc tggatggcta ttcacagaat gcctgtgctt tcagagtgc attttttctc	5940
tggatattctg gttcaagcat ttgaaggtag gaaaggttct ccaagtgcaa gaaagccagc	6000
cctgagcctc aactgcctgg ctagtgtggt cagtaggatg caaaggctgt tgaatgccac	6060
aaggccaaac tttaacctgt gtaccacaag cctagcagca gaggcagctc tgctcactgg	6120
aactctctgt cttctttctc ctgagccttt tcttttctg agttttctag ctctctcaa	6180
ccttacctct gccctaccca ggacaaaccc aagagccact gtttctgtga tgtctctcc	6240
agccctaatt aggcacatg acttcagcct gaccttccat gctcagaagc agtgctaatc	6300
cacttcagat gagctgtctt atgcaacaca ggacagcctt acaaaccttt gcaccagagc	6360
cctccacata tcagtgtttg ttcatactca cttcaacagc aaatgtgact gctgagatta	6420
agattttaca caagatggct tgtaatttca cagttagttt tatccatta ggtatgaaag	6480
aattagcata attccctta aacatgaatg aatcttagat tttttaataa atagttttgg	6540
aagtaaagac agagacatca ggagcacaag gaatagcctg agaggacaaa cagaacaaga	6600
aagagtctgg aaatacacag gatgttcttg gcctctctaa agcaagtgca agcagatagt	6660
accagcagcc ccaggtatc agagcccagt gaagagaagt accatgaaag ccacagctct	6720
aaccaccctg ttccagagtg acagacagtc cccaagacaa gccagcctga gccagagaga	6780
gaactgcaag agaaagtctc taatttaggt tctgttagat tcagacaagt gcaggctcatc	6840
ctctctccac agctactcac ctctccagcc taacaaagcc tgcagtccac actccaaccc	6900
tgggtgtctc cctcctagcc tctcccaaca tctgtctctc tgaccatctt ctgcatctct	6960
catctcacca tctcccactg tctacagctt actcttgcaa ctaccatctc attttctgac	7020
atcctgtcta catcttctgc catactctgc catctacat accacctctt accatctacc	7080

acaccatctt ttatctccat ccctctcaga agcctccaag ctgaatcctg ctttatgtgt	7140
tcatctcagc ccttgcattg aaagctgacc ccagaggcag aactattccc agagagcttg	7200
gccaagaaaa acaaaactac cagcctggcc aggcctcagga gtagtaagct gcagtgtctg	7260
ttgtgttcta gcttcaacag ctgcaggagi tccactctca aatgctccac atttctcaca	7320
tcctctgat tctggctact acccatcttc aaagaacaga atatctcaca tcagcatact	7380
gtgaaggact agtcatgggt gcagctgctc agagctgcaa agtcattctg gatggaggag	7440
agcttacaaa catttcatga tgctccccc gctctgatgg ctggagccca atccctacac	7500
agactcctgc tgtatgtgtt ttcttttcac tctgagccac agccagaggg caggcattca	7560
gtctctcttt caggctgggg ctggggcact gagaactcac ccaacacctt gctctcactc	7620
cttctgcaaa acaagaaaga gctttgtgct gcagtagcca tgaagaatga aaggaaggct	7680
ttaactaaaa aatgtcagag attattttca accccttact gtggatcacc agcaaggagg	7740
aaacacaaca cagagacatt ttttccctc aaattatcaa aagaatcact gcatttgta	7800
aagagagcaa ctgaatcagg aagcagagtt ttgaacatat cagaagttag gaatctgcat	7860
cagagacaaa tgcagtcagt gttgtttgct gcataccagc cctaatactt agaagcctca	7920
tggacttcaa acatcatctc ctctgacaag atgctctagc ctaactccat gagataaaat	7980
aaatctgcct ttcagagcca aagaagagtc caccagcttc ttctcagtgt gaacaagagc	8040
tccagtcagg ttagtcagtc cagtcagta gaggagacca gtctgcatcc tctaattttc	8100
aaaggcaaga agatttgttt accctggaca ccaggcaciaa gtgaggtcac agagctctta	8160
gatatgcagt cctcatgagt gaggagacta aagcgcatgc catcaagact tcagtgtaga	8220
gaaaacctcc aaaaaagcct cctcactact tctggaatag ctgagaggcc gaggcggcct	8280
cggcctctgc ataaataaaa aaaattagtc agccatgggg cggagaatgg gcggaactgg	8340
gcggagttag gggcgggatg ggcggagtta ggggcgggac tatggttgct gactaatga	8400
gatgcatgct ttgcatactt ctgcctgctg gggagcctgg ggactttcca cacctggttg	8460
ctgactaatt gagatgcatg ctttgcatc ttctgcctgc tggggagcct ggggactttc	8520
cacaccctaa ctgacacaca ttccacagct gcattaatga atcgccaac gcgcggggag	8580
aggcggtttg cgtattgggc gctcttcgc ttctcgtc actgactcgc tgcgtcgggt	8640
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga	8700
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg	8760
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacia	8820

aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt	8880
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgtta ccgataacct	8940
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct	9000
cagttcgggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc	9060
cgaccgtgc gccttatccg gtaactatcg tcttgagtc aaccggtaa gacacgactt	9120
atgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtgc	9180
tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag tatttggtat	9240
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa	9300
acaaaccacc gctggtagcg gtgggttttt tgtttgcaag cagcagatta cgcgagaaa	9360
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga	9420
aaactcacgt taagggattt tggatcatg attatcaaaa aggatcttca cctagatcct	9480
tttaaatata aaatgaagtt ttaaatcaat cttaaagtata tatgagtaaa ctgggtctga	9540
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcac	9600
catagtggc tgactcctgc aaaccacgtt gtgtctcaaa atctctgatg ttacattgca	9660
caagataaaa atatatcatc atgaacaata aaactgtctg cttacataaa cagtaataca	9720
aggggtgtta tgagccatat tcaacgggaa acgtcttgct cgaggcccg attaaattcc	9780
aacatggatg ctgatttata tgggtataaa tgggctcgcg ataatgtcgg gcaatcaggt	9840
gcgacaatct atcgattgta tgggaagccc gatgcgccag agttgtttct gaaacatggc	9900
aaaggtagcg ttgccaatga tgttacagat gagatggtca gactaaactg gctgacggaa	9960
tttatgcctc ttccgacct caagcatttt atccgtactc ctgatgatgc atggttactc	10020
accactgcga tccccgggaa aacagcattc caggtattag aagaatatcc tgattcaggt	10080
gaaaatattg ttgatgcgtt ggcagtgttc ctgcgccggt tgcattcgat tctgtttgt	10140
aattgtcctt ttaacagcga tcgctatctt cgtctcgtc aggcgcaatc acgaatgaat	10200
aacggtttgg ttgatgcgag tgattttgat gacgagcgta atggctggcc tgttgaacaa	10260
gtctggaaag aaatgcataa gcttttgcca ttctaccgg attcagtcgt cactcatggt	10320
gattttctac ttgataacct tatttttgac gaggggaaat taataggttg tattgatgtt	10380
ggacgagtcg gaatgcaga ccgataccag gatcttgcca tctatggaa ctgcctcgg	10440
gagttttctc cticattaca gaaacggctt tttcaaaaat atggtattga taatcctgat	10500
atgaataaat tgcagtttca tttgatgtc gatgagtttt tctaaggcg gcctgccacc	10560
ataccacgc cgaacaagc gctcatgagc ccgaagtggc gagcccgatc ttccccatcg	10620
gtgatgtcgg cgatatagcg gccagcaacc gcacctgtgg cgccggtgat gagggcgcg	10680

caagtcgacg tccggcagtc 10700

<210> 9

<211> 10700

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 9

ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcgccctca gtgagcgagc gagcgcgag agagggagtg 120

gccaactcca tctactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggctctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300

ctagttatta atagtaatca attacggggg cattagtcca tagcccatat atggagtacc 360

gcgttacata acttacgcta aatggcccgc ctggctgacc gccaacgac ccccgcccat 420

tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc 480

aatgggtgga gtatttacgg taaactgcc acttggcagt acatcaagt tatcatatgc 540

caagtacgcc cctattgac gtcaatgacg gtaaattggc cgctggcat tatgeccagt 600

acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta 660

ccatggtcga ggtgagcccc acgttctgct tcaactctcc catctcccc cctctcccac 720

ccccaatfff gtatttatff attttttaat tattttgtgc agcgatgggg gcgggggggg 780

ggggggggcg cgcgccaggc ggggcggggc gggcgaggg gcggggcggg gcgagcgga 840

gaggtgcggc ggcagccaat cagagcgggc cgctccgaaa gtttcctttt atggcgaggc 900

ggcggcgggc gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgacgc 960

tgcttctgcc ccgtgccccg ctccgccgcc gcctcgcgcc gcccggccc gctctgactg 1020

accgcgttac tccacagggt gagcggggcg gacggccctt ctctccggg ctgtaattag 1080

cgcttggttt aatgacggct tgtttctttt ctgtggctgc gtgaaagcct tgaggggctc 1140

cgggagctag agcctctgct aaccatgttc atgcttctt ctttttcta cagctcctgg 1200

gcaactgtct ggtattgtg ctgtctcatc attttgcaa agaattctc gaagatccga 1260

agggaaagtc ttccacgact gtgggatccg ttcgaagata tcaccggttg agccaccatg 1320

gaattcagca gcccagcag agaggaatgc cccaagctc tgagccgggt gtcaatcatg 1380

gccggatctc tgacaggact gctgtgtctt caggccgtgt cttgggcttc tggcgctaga 1440

ccttgcatcc ccaagagctt cggtacagc agcgtcgtgt gcgtgtgcaa tgccacctac 1500
 tgcgacagct tgcacctcc tacctttcct gctctgggca ccttcagcag atacgagagc 1560
 accagatccg gcagacggat ggaactgagc atgggaccca tccaggccaa tcacacaggc 1620
 actggcctgc tgctgacact gcagcctgag cagaaattcc agaaagttaa aggccttcggc 1680
 ggagccatga cagatgccgc cgctctgaat atcctggctc tgtctccacc agctcagaac 1740
 ctgctgctca agagctactt cagcgaggaa ggcatcggct acaacatcat cagagtcccc 1800

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

 gatctgggac ccacactggc caatagcacc caccataatg tgcggctgct gatgctggac 2280
 gaccagagac tgcttctgcc ccactgggct aaagtgggtg tgacagatcc tgaggccgcc 2340
 aaatacgtgc acggaatcgc cgtgcactgg tatctggact ttctggcccc tgccaaggcc 2400
 acaactggag agacacacag actgttcccc aacacatgc tgttcgccag cgaagcctgt 2460
 gtgggcagca agttttggga acagagcgtg cggtcggca gctgggatag aggcattgag 2520
 tacagccaca gcatcatcac caacctgctg taccacgtcg tcggctggac cgactggaat 2580
 ctggccctga atcctgaagg cgccctaac tgggtccgaa acttcgtgga cagcccatc 2640

 atcgtggaca tcaccaagga cactttctac aagcagccca tgttctacca cctgggacac 2700
 ttcagcaagt tcatccccga gggctctcag cgcgttggac tgggtgcttc ccagaagaac 2760
 gatctggacg ccgtggctct gatgcacct gatggatctg ctgtggtggt ggtcctgaac 2820
 cgcagcagca aagatgtgcc cctgaccatc aaggatcccg ccgtgggatt cctggaacaa 2880
 atcagccctg gctactccat ccacacctac ctgtggcgta gacagtgaca attgttaatt 2940
 aagtttaaac cctcagggcc gcaagcttat cgataatcaa cctctggatt acaaaatttg 3000
 tgaaagattg actggtatc ttaactatgt tgctcctttt acgctatgtg gatacgtgc 3060

 tttaatgcct ttgtatcatg ctattgcttc ccgtatggct ttcattttct cctccttgta 3120
 taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtcaggc aacgtggcgt 3180
 ggtgtgcact gtgtttgtg acgcaacccc cactggttgg ggcatgcca ccacctgta 3240
 gctcctttcc gggactttcg ctttccccct ccctattgcc acggcggaac tcatcgccgc 3300

ctgccttgcc cgctgctgga caggggctcg gctgttgggc actgacaatt ccgtggtgtt	3360
gtcggggaaa tcatcgtcct ttccttggct gctcgctgt gttgccacct ggattctgcg	3420
cgggacgtcc ttctgctacg tcccttcggc cctcaatcca gcggaccttc cttcccgcgg	3480
cctgctgccg gctctgcgcg ctcttcgcg tcttcgcctt cgccctcaga cgagtccgat	3540
ctcccttttg gccgcctccc cgcctcgata ccgtcgacta gagctcgctg atcagcctcg	3600
actgtgcctt ctagttgcc a gccatctgtt gtttgccctt ccccgtgcc ttccttgacc	3660
ctggaaggtg ccactccac tgtcctttcc taataaaatg aggaaattgc atcgcatgtt	3720
ctgagtaggt gtcatcttat tctggggggt ggggtggggc aggacagcaa gggggaggat	3780
tggaagaca atagcaggca tgctggggag agatccacga taacaaacag cttttttggg	3840
gtgaacatat tgactgaatt ccctgcaggt tggccactcc ctctctgcgc gctcgctcgc	3900
tactgaggc cgcccgggca aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag	3960
tgagcgagcg agcgcgcaga gagggagtgg ccaactccat cactaggggt tcctgcggcc	4020
gctcgtacgg tctcgaggaa ttcctgcagg ataacttgc aacctcattc taaaatgtat	4080
atagaagccc aaaagacaat aacaaaaata tttttgtaga acaaaatggg aaagaatgtt	4140
ccactaaata tcaagattta gagcaaagca tgagatgtgt ggggatagac agtgaggctg	4200
ataaaataga gtagagctca gaaacagacc cattgatata tgtaagtgc ctatgaaaaa	4260
aatatggcat ttacaatgg gaaaatgatg gtctttttct ttttagaaa aacagggaaa	4320
tatatattata tgtaaaaaat aaaaggaac 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 ctggtgcac ttactgataa	4620
atgatgttat caccatcttt aaccaaagc acaggaacaa gttatggtac tgatgtgctg	4680
gattgagaag gagctctact tcttgacag gacacattg tatcaactta aaaaagcaga	4740
tttttgccag cagaactatt cattcagagg taggaaactt agaatagatg atgtcactga	4800
ttagcatggc ttcccactt ccacagctgc tcccaccca ggttgccac agttgagttt	4860
gtccagtgtc cagggtgcc cactctcagt aagaagcccc acaccagccc ctctccaaat	4920
atgttggtg ttccttccat taaagtgacc ccactttaga gcagcaagtg gatttctgtt	4980
tcttacagtt caggaaggag gagtcagctg tgagaacctg gagcctgaga tgcttctaag	5040
tcccactgct actggggta gggaagccag actccagcat cagcagtcag gagcactaag	5100

cccttgccaa catcctgttt ctacagaaaa ctgcttccat tataatgggt gtcctttttt 5160

aagctatcaa gccaaacaac cagtgtctac cattattctc atcacctgaa gccaaagggtt 5220

ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct ccagcttctg tcttcagtea 5280

ctccactctt agcctgtctt gaatcaactc tgaccacagt tccctggagc ccttgccacc 5340

tgctgcccct gccaccttct ccatctgcag tgctgtgcag ccttctgcac tcttgagag 5400

ctaatagggt gagacttgaa ggaagaggag gaaagtttct cataatagcc ttgctgcaag 5460

ctcaaatggg aggtgggcac tgtgcccagg agccttggag caaaggctgt gcccaacctc 5520

tgactgcac caggtttgggt cttgacagag ataagaagcc ctggcttttg gagccaaaat 5580

ctaggtcaga ctiaggcagg atttcaaaag ttatcagca gaacatgagg cagaagaccc 5640

tttctgtcc agcttcttca ggctcaacct tcatcagaat agatagaaag agaggctgtg 5700

agggttctta aaacagaagc aaatctgact cagagaataa acaacctcct agtaaacac 5760

agcttagaca gagcatctgg tggtagtgt gctcagtgct ctactcaact gcttggtatc 5820

agccctcatg aggacttctc ttctttccct catagacctc catctctgtt ttccttagcc 5880

tgcaaaaac tggatggcta ttcacagaat gctgtgctt tcagagttgc atttttctc 5940

tggatattctg gttaagcat ttgaaggtag gaaaggttct ccaagtgcaa gaaagccagc 6000

cctgagcctc aactgcctgg ctagtgtggt cagtaggatg caaaggctgt tgaatgccac 6060

aaggccaaac tttaacctgt gtaccacaag cctagcagca gaggcagctc tgctcactgg 6120

aactctctgt cttctttctc ctgagccttt tcttttctg agttttctag ctctcctcaa 6180

ccttacctct gccctaccca ggacaaacc aagagccact gtttctgtga tgctcctcc 6240

agccctaatt aggcacatg acttcagcct gaccttccat gctcagaagc agtgctaate 6300

catttcagat gagctgtct atgcaacaca ggcagagcct acaaaccttt gcaccagagc 6360

cctccacata tcagtgtttg ttcatactca cttcaacagc aaatgtgact gctgagatta 6420

agattttaca caagatggtc tgtaatttca cagttagttt tatccatta 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

aaccacctg ttccagagtg acagacagtc cccaagacaa gccagcctga gccagagaga 6780

gaactgcaag agaaagtctc taatttaggt tctgttagat tcagacaagt gcaggctac 6840

ctctctccac agctactcac ctctccagcc taacaaagcc tgcagtccac actccaaccc	6900
tgggtgtctca cctcctagcc tctcccaaca tctgtctctc tgaccatctt ctgcatctct	6960
catctcacca tctcccactg tctacagcct actcttgcaa ctaccatctc attttctgac	7020
atcctgtcta catcttctgc catactctgc catctacat accacctctt accatctacc	7080
acaccatctt ttatctccat cctctcaga agcctccaag ctgaatcctg ctttatgtgt	7140
tcatctcagc cctgcatgg aaagctgacc ccagaggcag aactattccc agagagcttg	7200
gccaagaaaa aaaaaactac cagcctggcc aggctcagga gtagtaagct gcagtgtctg	7260
ttgtgttcta gcttcaacag ctgcaggagt tccactctca aatgctccac atttctcaca	7320
tctctctgat tctggctact acccatcttc aaagaacaga atatctcaca tcagcatact	7380
gtgaaggact agtcatgggt gcagctgctc agagctgcaa agtcattctg gatggaggag	7440
agcttacaaa catttcatga tgcctcccc gctctgatgg ctggagccca atccctacac	7500
agactcctgc tgiatgtgtt ttctttcac tctgagccac agccagaggg caggcattca	7560
gtctcctctt caggctgggg ctggggcact gagaactcac ccaacacctt gctctcactc	7620
cttctgcaaa acaagaaaga gctttgtgct gcagtagcca tgaagaatga aaggaaggct	7680
ttaactaaaa aatgtcagag attattttca accccttact gtggatcacc agcaaggagg	7740
aaacacaaca cagagacatt ttttccctc aaattatcaa aagaatcact gcatttgta	7800
aagagagcaa ctgaatcagg aagcagagt ttgaacatat cagaagttag gaatctgcat	7860
cagagacaaa tgcagtcag gtgtgttgct gcataccagc cctaactatt agaagcctca	7920
tggacttcaa acatcattcc ctctgacaag atgctctagc ctaactccat gagataaaat	7980
aaatctgcct ttcagagcca aagaagagtc caccagcttc ttctcagtg gaacaagagc	8040
tccagtcagg ttagtcagtc cagtgcagta gaggagacca gtctgcatcc tctaattttc	8100
aaaggcaaga agatttgttt accctggaca ccaggcacia gtgaggtcac agagctctta	8160
gatatgcagt cctcatgagt gaggagacta aagcgcatgc catcaagact tcagtgtaga	8220
gaaaacctcc aaaaaagcct cctcactact tctggaatag ctgagaggcc 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 ctttcatac ttctgcctgc tggggagcct ggggactttc	8520
cacaccctaa ctgacacaca ttccacagct gcattaatga atcgccaac gcgcggggag	8580
aggcggtttg cgtattgggc gctcttcgc ttctcgcctc actgactcgc tgcgctcggt	8640
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatcgggt tatccacaga	8700

atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg	8760
taaaaaggcc gcgttgctgg cgtttttcca taggtccgc cccctgacg agcatcacia	8820
aaatcgacgc tcaagtcaaa ggtggcgaaa cccgacagga ctataaagat accaggcgtt	8880
tccccctgga agtccctcg tgcgtctcc tgttccgacc ctgccgtta ccgataacct	8940
gtccgccttt ctcccttcgg gaagcgtggc gcttttctcat agtcacgct gtaggtatct	9000
cagttcggtg taggtcgctt gctccaagct gggctgtgtg cacgaacccc ccgttcagcc	9060
cgaccctgc gccttatccg gtaactatcg tcttgagtcc aaccggtaa gacacgactt	9120
atgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc	9180
tacagagtgc ttgaagtgtt ggcctaacta cggctacact agaagaacag tatttggtat	9240
ctgcgtctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa	9300
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa	9360
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga	9420
aaactcacgt taagggtatt tgggtcatgag attatcaaaa aggatcttca cctagatcct	9480
tttaaatata aaatgaagtt ttaaatcaat cttaaagtata tatgagtaaa cttggtctga	9540
cagttaccaaa tgcttaata gtaggcacc tatctcagcg atctgtctat ttcgttcac	9600
catagtggc tgactctgc aaaccacgtt gtgtctcaaa atctctgatg ttacattgca	9660
caagataaaa atatatcatc atgaacaata aaactgtctg cttacataaa cagtaataca	9720
aggggtgtta tgagccatat tcaacgggaa acgtcttgct cgaggccgcg attaaattcc	9780
aacatggatg ctgatttata tgggtataaa tgggctcgcg ataagtgcgg gcaatcaggt	9840
gcgacaatct atcgattgta tgggaagccc gatgcgccag agttgtttct gaaacatggc	9900
aaaggtagcg ttgccaatga tgttacagat gagatggtca gactaaactg gctgacggaa	9960
tttatgcctc ttccgacct caagcatttt atccgtactc ctgatgatgc atggttactc	10020
accactgcga tccccgggaa aacagcattc caggtattag aagaatatcc tgattcaggt	10080
gaaaatattg ttgatgcgtt ggcagtgttc ctgcgccggt tgcattcgat tcctgtttgt	10140
aattgtcctt ttaacagcga tcgcgtattt cgtctcgctc aggcgcaatc acgaatgaat	10200
aacggttttg ttgatgcgag tgattttgat gacgagcgta atggctggcc tgttgaacaa	10260
gtctggaaag aaatgcaata gcttttgcca ttctaccgg attcagtcgt cactcatggt	10320
gatttctcac ttgataacct tatttttgac gaggggaaat taataggttg tattgatgtt	10380
ggacgagtcg gaatcgaga ccgataccag gatcttgcca tcctatggaa ctgcctcggt	10440
gagttttctc cttcattaca gaaacggctt ttcaaaaaat atggtattga taatcctgat	10500
atgaataaat tgcagtttca ttgatgctc gatgagtttt tctaaggcg gcctgccacc	10560

ataccacgc cgaacaagc gctcatgagc ccgaagtggc gagcccgatc ttcccatcg 10620

gtgatgtcgg cgatatagc gccagcaacc gcacctgtgg cgccggtgat gagggcgcg 10680

caagtgcagc tccggcagtc 10700

<210> 10

<211> 10700

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 10

ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgcccgggc aaagcccggg 60

cgtcgggcga cctttggtcg cccggcctca gtgagcgagc gagcgcgag agagggagtg 120

gccaaactcca tctactagggg ttcttgctag ctctgggtat ttaagcccga gtgagcacgc 180

agggctctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240

tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac 300

ctagttatta atagtaatca attacggggt cattagtcca tagcccatat atggagtacc 360

gcgttacata acttacggta aatggcccgc ctggctgacc gccaacgac ccccgcccat 420

tgacgtcaat aatgacgtat gtcccatag 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 tcaactctcc catctcccc cctccccac 720

ccccaatttt gtatttatat attttttaat tattttgtgc agcgatgggg gcgggggggg 780

ggggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgagcgga 840

gaggtgcggc ggcagccaat cagagcggcg cgtccgaaa gtttcctttt atggcgaggc 900

ggcggcgggc gcgccctat aaaaagcgaa gcgcgcggcg ggccggagtc gctgcgacgc 960

tgccctcgcc ccgtgccccg ctccgccgcc gcctcgccg ccccgcccc gctctgactg 1020

accgcgttac tcccacaggt gagcgggcgg gacggccctt ctctccggg ctgtaattag 1080

cgcttgggtt aatgacggct tgtttctttt ctgtggctgc gtgaaagcct tgaggggctc 1140

cgggagctag agcctctgct aaccatgttc atgccttctt ctttttcta cagctcctgg 1200

gcaacgtgct ggttattgtg ctgtctcacc attttgcaa agaattctc gaagatccga 1260

agggaaagtc ttccacgact gtgggatccg ttccaagata 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 tcgacctcc tacctttcct gctctgggca ccttcagcag atacgagagc	1560
accagatccg gcagacggat ggaactgagc atgggaccca tccaggccaa tcacacaggc	1620
actggcctgc tgctgacact gcagcctgag cagaaattcc agaaagtga aggcttcggc	1680
ggagccatga cagatgccgc cgctctgaat atcctggctc tgtctccacc agctcagaac	1740
ctgctgctca agagctactt cagcgaggaa ggcatcggt acaacatcat cagagtcccc	1800
atggccagct gcgacttcag catcaggacc tacacctacg ccgacacacc cgacgatttc	1860
cagctgcaca acitcagcct gcctgaagag gacaccaagc tgaagatccc tctgatccac	1920
agagccctgc agctggcaca aagaccctgt tcaactgctgg cctctccatg gacatctccc	1980
acctggctga aaacaaatgg cgccgtgaat ggcaaggga gcctgaaagg ccaacctggc	2040
gacatctacc accagacctg ggccagatac ttctgtaagt tcctggacgc ctatgccgag	2100
cacaagctgc agttttgggc cgtgacagcc gagaacgaac cttctgctgg actgctgagc	2160
ggctaccctt ttcagtgcct gggctttaca cccgagcacc agcgggactt tatcgccgt	2220
gatctgggac ccacactggc caatagcacc caccataatg tgcggctgct gatgctggac	2280
gaccagagac tgcttctgcc ccaactgggt aaagtgggtc tgacagatcc tgaggccgcc	2340
aaatacgtgc acggaatgc cgtgcactgg tatctggact ttctggcccc tgccaaggcc	2400
acactgggag agacacacag actgttcccc aacacatgc tgttcgccag cgaagcctgt	2460
gtgggcagca agttttggga acagagcgtg cggtctggca gctgggatag aggcatgcag	2520
tacagccaca gcatcatcac caacctgctg taccacgtcg tcggctggac cgactggaat	2580
ctggccctga atcctgaagg cgccctaac tgggtccgaa acttcgtgga cagcccatc	2640
atcgtggaca tcaccaagga cactttctac aagcagccca tgtttacca cctgggacac	2700
ttcagcaagt tcattccccg gggctctcag cgcgttggac tgggtggcttc ccagaagaac	2760
gatctggacg ccgtggctct gatgcaccct gatggatctg ctgtggtggt ggtcctgaac	2820
cgcagcagca aagatgtgcc cctgaccatc aaggatcccg ccgtgggatt cctggaaca	2880
atcagccctg gctactccat ccacacctac ctgtggcgta gacagtgaca attgttaatt	2940
aagtttaaac cctcgaggcc gcaagcttat cgataatcaa cctctggatt acaaaatttg	3000
tgaagattg actggtattc ttaactatgt tgctcctttt acgctatgtg gatacgtgc	3060
tttaatgcct ttgtatcatg ctattgcttc ccgtatggct ttcattttct cctccttgta	3120

taaatacctgg ttgctgtctc tttatgagga gttgtggccc gttgtcaggc aacgtggcgt	3180
ggtgtgcaact gtgtttgctg acgcaacccc cactggttgg ggcatgtcca ccacctgtca	3240
gtctctttcc gggactttcg ctttccccc cctattgcc acggcggaac tcatcgccgc	3300
ctgccttgcc cgctgctgga caggggctcg gctgttgggc actgacaatt ccgtggtgtt	3360
gtcggggaaa tcatcgctct ttccttggct gctgcctgt gttgccacct ggattctgcg	3420
cgggacgtcc ttctgtacg tcccttcggc cctcaatcca gcggaccttc cttcccgagg	3480
cctgtctccg gctctgcggc ctcttccgcg tcttcgctt cgccctcaga cgagtcggat	3540
ctccctttgg gccgcctccc cgcacgata ccgtcgacta gagctcgctg atcagcctcg	3600
actgtgcctt ctagttgcca gccatctgtt gtttgccct ccccgctgcc ttccttgacc	3660
ctggaaggtg ccaactccac tgtcctttcc taataaaatg aggaattgc atcgcatgtt	3720
ctgagtaggt gtcattctat tctggggggt ggggtggggc aggacagcaa gggggaggat	3780
tgggaagaca atagcaggca tgctggggag agatccacga taacaaacag cttttttggg	3840
gtgaacatat tgactgaatt ccctgcagga ggaacccta gtgatggagt tggccactcc	3900
ctctctgcgc gctcgtcgc tcaactgagg cgcccgggca aagcccgggc gtcgggcgac	3960
ctttggtcgc ccggcctcag tgagcgagcg agcgcgcaga gagggagtgg ccaagcggcc	4020
gctcgtacgg tctcgaggaa ttcctgcagg ataacttgc aacctcattc taaaatgtat	4080
atagaagccc aaaagacaat aacaaaaata ttcctgtaga acaaatggg aaagaatgtt	4140
ccactaaata tcaagattta gagcaaagca tgagatgtgt ggggatagac agtgaggctg	4200
ataaaataga gtagagctca gaaacagacc catgatata tgtaagtgc ctatgaaaaa	4260
aatatggcat ttiacaatgg gaaaatgatg gtctttttct ttttagaaa aacagggaaa	4320
tatatattata tgtaaaaaat aaaagggaac ccatatgtca taccatacac aaaaaaaat	4380
tccagtgaat tataagtcta aatggagaag gcaaaacttt aaatctttta gaaaataata	4440
tagaagcatg cagaccagcc tggccaacat gatgaaacct tctctactaa taataaaatc	4500
agtagaacta ctcaggacta ctttgagtgg gaagtcttt tctatgaaga cttctttggc	4560
caaaattagg ctctaaatgc aaggagatag tgcacatgc ctggctgcac ttactgataa	4620
atgatgttat caccatcttt aaccaaatgc acaggaacaa gttatggtac tgatgtgctg	4680
gattgagaag gagctctact tcttgacag gacacattg tatcaactta aaaaagcaga	4740
tttttgccag cagaactatt cattcagagg taggaaactt agaatagatg atgtcactga	4800
ttagcatggc ttccccatct ccacagctgc ttcccacca ggttgccac agttgagttt	4860

gtccagtgtc cagggtgcc cactctcagt aagaagcccc acaccagccc ctctccaaat 4920
atgttggttg ttctttccat taaagtgacc ccactttaga gcagcaagtg gattttctgtt 4980
tcttacagtt caggaaggag gagtgcagtg tgagaacctg gagcctgaga tgcttctaag 5040
tcccactgct acitggggta gggaagccag actccagcat cagcagtcag gagcactaag 5100
cccttgccaa catcctgttt ctgagagaaa ctgcttccat tataatggtt gtcttttttt 5160
aagctatcaa gccaaacaac cagtgtctac cattattctc atcacctgaa gccaaagggtt 5220
ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct ccagcttctg tcttcagtca 5280

ctccactctt agcctgtctt gaatcaactc tgaccacagt tccctggagc ccttgccacc 5340
tgctgcccci gccaccttct ccatctgcag tgcgtgcag cttctgcac tcttgagag 5400
ctaataagtg gagacttgaa ggaagaggag gaaagtttct cataatagcc ttgctgcaag 5460
ctcaaatggg aggtggggcac tgtgcccagg agccttggag caaaggctgt gcccacctc 5520
tgactgcac caggttttgt cttgacagag ataagaagcc ctggcttttg gagccaaaat 5580
ctaggtcaga cttaggcagg attctcaaag ttatcagca gaacatgagg cagaagaccc 5640
ttctgtctcc agcttcttca ggctcaacct tcatcagaat agatagaaag agaggctgtg 5700

agggttctta aaacagaagc aaatctgact cagagaataa acaacctcct agtaactac 5760
agcttagaca gagcatctgg tggtagtgt gctcagtgct ctactcaact gtctggtatc 5820
agccctcatg aggacttctc ttctttccct catagacctc catctctgtt ttccttagcc 5880
tgcagaaatc tggatggcta ttcacagaat gcctgtgctt tcagagttgc atttttctc 5940
tggtagttctg gttcaagcat ttgaaggtag gaaaggttct ccaagtgcaa gaaagccagc 6000
cctgagcctc aactgcctgg ctagtgtgtt cagtaggatg caaaggctgt tgaatgccac 6060
aaggccaaac tttaacctgt gtaccacaag cctagcagca gaggcagctc tgctcactgg 6120

aactctctgt cttctttctc ctgagccttt tcttttctg agttttctag ctctctcaa 6180
ccttacctct gccctacca ggacaaacc aagagccact gtttctgtga tgcctctcc 6240
agccctaatt aggcacatg acttcagcct gaccttccat gctcagaagc agtgctaate 6300
cacttcagat gagctgtct atgcaacaca ggcagagcct acaaaccttt gcaccagagc 6360
cctccacata tcagtgtttg ttcatactca cttcaacagc aaatgtgact gctgagatta 6420
agattttaca caagatggtc tgtaatttca cagttagttt tatccatta ggtatgaaag 6480
aattagcata attcccctta aacatgaatg aatcttagat ttttaataa atagttttgg 6540

aagtaaagac agagacatca ggagcacaag gaatagcctg agaggacaaa cagaacaaga 6600
aagagtctgg aaatacacag gatgttcttg gcctcctcaa agcaagtgca agcagatagt 6660
accagcagcc ccaggctatc agagcccagt gaagagaagt accatgaaag ccacagctct 6720

aaccacctg ttccagagtg acagacagtc cccaagacaa gccagcctga gccagagaga 6780
gaactgcaag agaaagtttc taatttaggt tctgttagat tcagacaagt gcaggtcac 6840
ctctctccac agctactcac ctctccagcc taacaaagcc tgcagtccac actccaaccc 6900
tgggtgtctca cctcctagcc tctcccaaca tctgtctctc tgaccatctt ctgcatctct 6960

catctcacca tctcccactg tctacagcct actcttgcaa ctaccatctc attttctgac 7020
atcctgtcta catctttctgc catactctgc catctacat accacctctt accatctacc 7080
acaccatctt ttatctccat ccctctcaga agcctccaag ctgaatcctg ctttatgtgt 7140
tcctctcagc cctgcatgg aaagctgacc ccagaggcag aactattccc agagagcttg 7200
gccaagaaaa acaaaactac cagcctggcc aggctcagga gtagtaagct gcagtgtctg 7260
ttgtgttcta gcttcaacag ctgcaggagt tccactctca aatgctccac atttctcaca 7320
tcctcctgat tctggctact acccatcttc aaagaacaga atatctcaca tcagcatact 7380

gtgaaggact agtcatgggt gcagctgctc agagctgcaa agtcattctg gatggaggag 7440
agcttacaaa catttcatga tgctccccc gctctgatgg ctggagccca atccctacac 7500
agactcctgc tgatgtgtt ttcttttcac tctgagccac agccagaggg caggcattea 7560
gtctctctct caggctgggg ctggggcact gagaactcac ccaacacctt gctctcactc 7620
cttctgcaaa acaagaaaga gctttgtgct gcagtagcca tgaagaatga aaggaaggct 7680
ttaactaaaa aatgtcagag attattttca accccttact gtggatcacc agcaaggagg 7740
aaacacaaca cagagacatt ttttccctc aaattatcaa aagaatcact gcatttgta 7800

aagagagcaa ctgaatcagg aagcagagtt ttgaacatat cagaagttag gaatctgcat 7860
cagagacaaa tgcagtcatg gttgtttgct gcataccagc cctaatactt agaagcctca 7920
tggacttcaa acatatttc ctctgacaag atgctctagc ctaactccat gagataaaat 7980
aaatctgcct ttccagagcca aagaagagtc caccagcttc ttctcagtgt gaacaagagc 8040
tccagtcagg ttagtcagtc cagtcagta gaggagacca gtctgcatcc tctaattttc 8100
aaaggcaaga agatttggtt accctggaca ccaggcaciaa gtgaggtcac agagctctta 8160
gatatgcagt cctcatgagt gaggagacta aagcgcatgc catcaagact tcagtgtaga 8220

gaaaacctcc aaaaaagcct cctcactact tctggaatag ctccagaggcc 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 atcgccaac gcgcggggag 8580

aggcggtttg cgtattgggc gctcttccgc ttctctgctc actgactcgc tgcgctcggt	8640
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga	8700
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg	8760
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacia	8820
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcggt	8880
tccccctgga agctccctcg tgcgtctctc tgttccgacc ctgccgctta ccggatacct	8940
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct	9000
cagttcgggt taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc	9060
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt	9120
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtgc	9180
tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag tatittggtat	9240
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa	9300
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa	9360
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga	9420
aaactcacgt taagggattt tggatcatgag attatcaaaa aggatcttca cctagatcct	9480
tttaaatata aaatgaagtt ttaaatcaat ctaaaagtata tatgagtaaa cttggtctga	9540
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcac	9600
catagttgcc tgactctcgc aaaccacgtt gtgtctcaaa atctctgatg ttacattgca	9660
caagataaaa atatatcatc atgaacaata aaactgtctg cttacataaa cagtaataca	9720
aggggtgtta tgagccatat tcaacgggaa acgtcttgct cgaggccgag attaaattcc	9780
aacatggatg ctgatttata tgggtataaa tgggctcgcg ataatgtcgg gcaatcaggt	9840
gcgacaatct atcgattgta tgggaagccc gatgcgccag agttgtttct gaaacatggc	9900
aaaggtagcg ttgccaatga tgttacagat gagatgggtca gactaaactg gctgacggaa	9960
tttatgcctc ttccgaccat caagcatttt atccgtactc ctgatgatgc atggttactc	10020
accactgcga tccccgggaa aacagcattc caggtattag aagaatatcc tgattcaggt	10080
gaaaatattg ttgatgcgct ggcagtgttc ctgcgccggt tgcatcagat tcctgtttgt	10140
aattgtcctt ttaacacgca tcgcgtatct cgtctcgtc aggcgcaatc acgaatgaat	10200
aacggtttgg ttgatgcgag tgattttgat gacgagcgta atggctggcc tgttgaacaa	10260
gtctggaaaag aaatgcaata gcttttgcca ttctaccgg attcagtcgt cactcatggt	10320

gattttctcac ttgataacct tatttttgac gaggggaaat taataggttg tattgatgtt 10380
 ggacgagtcg gaatcgaga ccgataccag gatcttgcca tcctatggaa ctgcctcgg 10440
 gagttttctc cttcattaca gaaacggctt tttcaaaaat atggtattga taatcctgat 10500
 atgaataaat tgcagtttca tttgatgctc gatgagtttt tctaaggcg gcctgccacc 10560
 ataccacgc cgaacaagc gctcatgagc ccgaagtggc gagcccgatc ttcccatcg 10620
 gtgatgtcgg cgatatagc gccagcaacc gcacctgtgg cgccggtgat gagggcgcg 10680
 caagtcgacg tccggcagtc 10700

<210> 11

<211> 11188

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 11

ttggccaactc cctctctgcg cgctcgctcg ctcaactgagg ccgggcgacc aaaggtcgcc 60
 cgacgcccgg gctttgcccg ggcgccctca gtgagcgagc gagcgcgag agagggagt 120
 gccaaactatt agatctgatg gccgcgctag ctctgggtat ttaagcccga gtgagcacgc 180
 agggctctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240
 gtggtgactg agatgttttc taggaaacac aaaagataca aaaaagaaca cgtggaagga 300
 tagccaaaaa ggggggctgc cccatttcc tgcacccgc tgcgatggct ggcaccattt 360

ggaagacttc gagatacact gttgagcgca gtaagacaac agtgtatctc gaagtcttcc 420
 agatggggcc agccggtcca ctctgtatcc aggccagttc tgcaaggcgt tcgaggacca 480
 cccccctccc ctgcccacca ggggtggtctc atacagaact tataagattc ccaaattcaa 540
 agacatttca cgtttatggt gatttcccag aacacatagc gacatgcaa tattgcaggg 600
 cgccactccc ctgtccctca cagecatctt cctgccaggg cgacgcgcg ctgggtgttc 660
 ccgcctagtg aactgggcc cgcatctct tggagcgggt tgatgacgtc agcgtttccc 720
 atggtgaatc ctiaggttct agaaccggtg acgtctccca tggagaagct tggatctgaa 780

ttcggtacct agttattaat agtaatcaat tacggggtca ttagttcata gccatata 840
 ggagttccgc gttacataac ttacggtaaa tggccgcct ggctgaccgc ccaacgaccc 900
 ccgcccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca 960
 ttgacgtcaa tgggtggagt atttacgta aactgcccac ttggcagtac atcaagtgt 1020
 tcatatgcca agtacgccc ctattgacgt caatgacggt aaatggccc cctggcatta 1080
 tgcccagta atgaccttat gggactttcc tacttggcag tacatctacg tattagtc 1140

cgctattacc atggtcgagg tgagcccccac gttctgcttc actctcccca tctccccccc	1200
ctcccccccc ccaattttgt atttatttat tttttaatta ttttgtgcag cgatgggggc	1260
gggggggggg ggggggcgcg cgccaggcgg ggccggggcgg ggcgaggggc ggggcggggc	1320
gaggcggaga ggtgcggcgg cagccaatca gagcggcgcg ctccgaaagt ttccttttat	1380
ggcgaggcgg cgccggcggc ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc	1440
tgcgacgtg ctttcgcccc gtgccccgt cgcgcgcgc ctcgcgcgc cgcgccggc	1500
tctgactgac cgcgttactc ccacaggtga gcggggcggga cgcccttct cctccgggt	1560
gtaattagcg cttgggttaa tgacggcttg tttcttttct gtggctgctg gaaagccttg	1620
aggggctccg ggagctagag cctctgctaa ccatgttcat gccttcttct ttttctaca	1680
gctcctgggc aacgtgtcgg ttattgtgct gtctcatcat tttggcaaag aattcctcga	1740
agatccgaag ggaaagtctt ccacgactgt gggatccgtt cgaagatac accggttgag	1800
ccacatgga attcagcagc ccacgagag aggaatgccc caagcctctg agccgggtgt	1860
caatcatggc cggatctctg acaggactgc tgctgcttca ggccgtgtct tgggcttctg	1920
gcgctagacc ttgcatcccc aagagcttcg gctacagcag cgtcgtgtgc gtgtgcaatg	1980
ccacctactg cgacagcttc gaccctctta cctttctctg tctgggcacc ttcagcagat	2040
acgagagcac cagatccggc agacggatgg aactgagcat gggacccatc caggccaatc	2100
acacaggcac tggcctgctg ctgacactgc agcctgagca gaaattccag aaagtgaag	2160
gcttcggcgg agccatgaca gatgccgcg ctctgaatat cctggctctg tctccaccag	2220
ctcagaacct gctgctcaag agctacttca gcgaggaagg catcggtac aacatcatca	2280
gagtgcccat ggccagctgc gacttcagca tcaggaccta cacctacgc gacacaccg	2340
acgatttcca gctgcacaac ttcagcctgc ctgaagagga caccaagctg aagatccctc	2400
tgatccacag agccctgcag ctggcacaaa gaccctgtc actgctggcc tctccatgga	2460
catctccac 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
tcgccgtga tctgggaccc acactggcca atagcaccca ccataatgtg cggctgctga	2760
tgctggacga ccagagactg cttctgcccc actgggctaa agtgggtgctg acagatcctg	2820
aggccgcaa atacgtgcac ggaatcgccg tgcaactggta tctggacttt ctggccctg	2880

ccaaggccac actgggagag acacacagac tgttcccaa caccatgctg ttcgccagcg 2940
aagcctgtgt gggcagcaag ttttgggaac agagcgtgcg gctcggcagc tgggatagag 3000
gcatgcagta cagccacagc atcatcacca acctgctgta ccacgtcgtc ggctggaccg 3060
actggaatct ggccctgaat cctgaaggcg gcctaactg ggtccgaaac ttcgtggaca 3120
gccccatcat cgtggacatc accaaggaca cttctacaa gcagcccatg ttctaccacc 3180
tgggacactt cagcaagtgc atccccgagg gctctcagcg cgttggactg gtggcttccc 3240
agaagaacga tctggacgcc gtggctctga tgcaccctga tggatctgct gtggtggtgg 3300

tcctgaaccg cagcagcaaa gatgtgcccc tgaccatcaa ggatcccgcc gtgggattcc 3360
tggaacaat cagccctggc tactccatcc acacctacct gtggcgtaga cagtgacaat 3420
tgttaattaa gttaaaccg tcgaggccgc aagcttatcg ataataacc tctggattac 3480
aaaatttgtg aaagattgac tggattctt aactatgttg ctctttttac gctatgtgga 3540
tacgtgctt taatgcctt gtatcatgct attgcttccc gtatggctt cattttctcc 3600
tccttgata aatcctggtt gctgtctct tatgaggagt tgtggcccg tgcaggcaa 3660
cgtggcgtag gtgtcactgt gtttctgac gcaacccca ctggttgggg cattgccacc 3720

acctgtcagc tcctttccgg gactttcgt tccccctcc ctattgccac ggcggaactc 3780
atcgccgct gccttggcg ctgctggaca ggggctcggc tgttgggcac tgacaattcc 3840
gtggtgtgt cggggaaatc atgctcttt cttggctgc tcgctgtgt tgccacctgg 3900
attctgcgcg ggacgtcct ctgctacgtc cttcggccc tcaatccagc ggaccttct 3960
tcccgcggc tctgcccgc tctgggct cttccgcgtc ttcgcttcg cctcagacg 4020
agtcggatct cctttgggc cgcctcccc catcgatacc gtcgactaga gctcgtgat 4080
cagcctcgac tgtccttct agttgccagc catctgtgt ttgccctcc cccgtgcctt 4140

ccttgacct ggaaggtgcc actccactg tcctttccta ataaatgag gaaattgcat 4200
cgcatgtct gagtaggtgt cattctattc tggggggtgg ggtggggcag gacagcaagg 4260
gggaggattg ggaagacaat agcaggcatg ctggggagag atccacgata acaaacagct 4320
tttttgggt gaacatatg actgaattcc ctgcaggtg gccactcct ctctgcgcgc 4380
tcgctcgtc actgaggccg cccgggcaa gcccggcgt cgggcgacct ttggtcgccc 4440
ggcctcagt agcgagcag cgcgcagaga gggagtggc aactccatca ctagggttc 4500
ctgcggcgc tcgtacggtc tcgaggaatt cctgcaggat aactgcca cctattcta 4560

aaatgtatat agaagccaa aagacaataa caaaaatatt cttgtagaac aaaatgggaa 4620
agaatgttc actaaatac aagatttga gcaaagcatg agatgtgtg ggatagacag 4680
tgaggctgat aaaatagagt agagctcaga aacagacca ttgatatat taagtacct 4740

atgaaaaaaa tatggcattt tacaatggga aaatgatggt ctttttcttt tttagaaaaa	4800
cagggaataa tatttatatg taaaaaataa aagggaaccc atatgtcata ccatacacac	4860
aaaaaaattc cagtgaatta taagtctaaa tggagaaggc aaaactttaa atcttttaga	4920
aaataatata gaagcatgca gaccagcctg gccaacatga tgaaaccttc tctactaata	4980
ataaaatcag tagaactact caggactact ttgagtggga agtccttttc tatgaagact	5040
tctttggcca aaattaggct ctaaagcaaa ggagatagtg catcatgcct ggctgcactt	5100
actgataaat gatgttatca ccatctttaa ccaaatgcac aggaacaagt tatggtactg	5160
atgtgctgga ttgagaagga gctctacttc ctgacagga cacatttgta tcaacttaaa	5220
aaagcagatt ttgccagca gaactattca ttcagaggta ggaaacttag aatagatgat	5280
gtcactgatt agcatggctt ccccatctcc acagctgctt cccaccagg ttgccacag	5340
ttgagtttgt ccagtgtca gggctgcca ctctcagtaa gaagccccac accagccct	5400
ctccaaatat gtggctgtt ccttcatta aagtgacccc acttagagc agcaagtgga	5460
tttctgttc ttacagtca ggaaggagga gtcagctgtg agaacctgga gcctgagatg	5520
cttctaagtc ccaactgtac tggggtcagg gaagccagac tccagcatca gcagttagga	5580
gcactaagcc ctgccaaaca tctgtttct cagagaaact gcttcatta taatggttgt	5640
ccttttttaa gctatcaagc caaacaacca gtgtctacca ttattctcat cacctgaagc	5700
caagggttct agcaaaagtc aagctgtctt gtaatggttg atgtgcctcc agcttctgtc	5760
ttcagtcact ccaactctag cctgctctga atcaactctg accacagttc cctggagccc	5820
ctgccacctg ctgcccctgc caccttctcc atctgcagtg ctgtgcagcc ttctgcactc	5880
ttgcagagct aatagggtga gacttgaagg aagaggagga aagtttctca taatagcctt	5940
gctgcaagct caaatgggag gtgggcactg tgcccaggag ccttgagca aaggctgtgc	6000
ccaacctctg actgcatcca ggtttggtct tgacagagat aagaagccct ggcttttgga	6060
gccaaaatct aggtcagact taggcaggat tctcaaagtt tatcagcaga acatgaggca	6120
gaagaccctt tctgtccag cttcttcagg ctcaaccttc atcagaatag atagaaagag	6180
aggctgtgag ggttcttaaa acagaagcaa atctgactca gagaataaac aacctcctag	6240
taaactacag cttagacaga gcatctggtg gtgagtgtgc tcagtgtcct actcaactgt	6300
ctggtatcag cctcatgag gacttctctt ctttccctca tagacctcca tctctgtttt	6360
ccttagcctg cagaaatctg gatggctatt cacagaatgc ctgtgctttc agagtgtcat	6420
tttttctctg gtattctggt tcaagcattt gaaggtagga aaggttctcc aagtgaaga	6480
aagccagccc tgagcctcaa ctgctggct agtgtgtca gtaggatgca aaggctgttg	6540
aatgccacaa ggccaaactt taacctgtgt accacaagcc tagcagcaga ggcagctctg	6600

ctcactggaa ctctctgtct tctttctcct gagccttttc ttttctgag ttttctagct 6660

ctcctcaacc ttacctctgc cctacccagg acaaaccxaa gagccactgt ttctgtgatg 6720

tcctctccag ccctaattag gcatcatgac ttcagcctga ccttccatgc tcagaagcag 6780

tgctaatacca cttcagatga gctgctctat gcaacacagg cagagcctac aaacctttgc 6840

accagagccc tccacataat agtggtttgt catactcact tcaacagcaa atgtgactgc 6900

tgagattaag attttacaca agatggctctg taatttcaca gttagtttta tcccattagg 6960

tatgaaagaa ttagcataat tccccttaa catgaatgaa tcttagattt tttataaat 7020

agttttggaa gtaaagacag agacatcagg agcacaagga atagcctgag aggacaaaca 7080

gaacaagaaa gagtctggaa atacacagga tgttcttggc ctctcaaag caagtgaag 7140

cagatagtag cagcagcccc aggctatcag agcccagtga agagaagtag catgaaagcc 7200

acagctctaa ccacctgtt ccagagtgc agacagtcac caagacaagc cagcctgagc 7260

cagagagaga actgcaagag aaagtcttca atttaggttc tgttagattc agacaagtgc 7320

aggctacact ctctccacag ctactcacct ctccagccta acaaagcctg cagtcacac 7380

tccaaccctg gtgtctcacc tcttagctc tccaacatc ctgctctctg accatcttct 7440

gcatctctca tctcaccatc tccactgtc tacagcctac tcttgcaact accatctcat 7500

tttctgacat cctgtctaca tttctgcca tactctgcca tctaccatac cactctttac 7560

catctaccac accatctttt atctccatcc ctctcagaag cctccaagct gaatcctgt 7620

ttatgtgttc atctcagccc ctgcatggaa agctgacccc agaggcagaa ctattcccag 7680

agagcttggc caagaaaaac aaaactacca gcctggccag gctcaggagt agtaagctgc 7740

agtgtctgtt gtgtctttag ttcaacagct gcaggagttc cactctcaa tgctccacat 7800

ttctcacatc ctctgattc tggctactac ccactttcaa agaacagaat atctcacatc 7860

agcatactgt gaaggactag tcatgggtgc agctgctcag agctgcaaag tcattctgga 7920

tgggtggagag cttacaaaca ttcatgatg ctccccccgc tctgatggct ggagcccaat 7980

ccctacacag actcctgtctg tatgtgtttt ctttctactc tgagccacag ccagagggca 8040

ggcattcagt ctctcttca ggctggggct ggggcactga gaactcacc aacaccttgc 8100

tctcactcct tctgcaaaac aagaaagagc tttgtgctgc agtagccatg aagaatgaaa 8160

ggaaggcttt aactaaaaaa tgtcagagat tattttcaac cccttactgt ggatcaccag 8220

caaggaggaa acacaacaca gagacatttt tccccctcaa attatcaaaa gaatcactgc 8280

atttgtaaa gagagcaact gaatcaggaa gcagagtttt gaacatatca gaagttagga 8340

atctgcatca gagacaaatg cagtcattgt tgtttgctgc ataccagccc taatcattag	8400
aagcctcatg gacttcaaac atcattccct ctgacaagat gctctagcct aactccatga	8460
gataaaataa atctgccttt cagagccaaa gaagagtcca ccagcttctt ctcagtgtga	8520
acaagagctc cagtcagggt agtcagtcca gtgcagttaga ggagaccagt ctgcatcctc	8580
taattttcaa aggcaagaag atttgtttac cctggacacc aggcacaagt gaggtcacag	8640
agctcttaga tatgcagtcc tcatgagtga ggagactaaa gcgcatgcca tcaagacttc	8700
agtgtagaga aaacctccaa aaaagcctcc tcactacttc tggatagct cagaggccga	8760
ggcggcctcg gcctctgcat aaataaaaaa aattagttag ccatggggcg gagaatgggc	8820
ggaactgggc ggagttaggg gcgggatggg cggagttagg ggcgggacta tggttgctga	8880
ctaattgaga tgcatgcttt gcatacttct gcctgctggg gagcctgggg actttccaca	8940
cctggttctg gactaattga gatgcagtct ttgcatactt ctgcctgctg gggagcctgg	9000
ggactttcca caccctaact gacacacatt ccacagctgc attaatgaat cggccaacgc	9060
gcggggagag gcggtttcgc tattgggcgc tcttcgctt cctcgtcac tgactcgtg	9120
cgctcggctg ttcggctcgc gcgagcggta tcagctcact caaaggcggg aatacggtta	9180
tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gaaaaggcc	9240
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgcc cctgacgag	9300
catcacaaaa atcgacgtc aagtcagagg tggcgaaacc cgacaggact ataaagatac	9360
caggcggttc cccctggaag ctccctcgtg cgctctcctg ttccgacct gccgcttacc	9420
ggatacctgt ccgcctttct ccttcggga agcgtggcg tttctcatag ctcacgtgt	9480
aggtatctca gttcgggtga ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc	9540
gttcagcccc accgctgcgc cttatccggt aactatcgtc ttgagtcaa cccggtaga	9600
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta	9660
ggcggtgcta cagagtcttt gaagtgggtg ctaactacg gctacactag aagaacagta	9720
tttggtatct gcgtctgct gaagccagt accttcggaa aaagagttag tagctctga	9780
tccggcaaac aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg	9840
cgcagaaaaa aaggatctca agaagatcct ttgatcttt ctacggggtc tgacgtcag	9900
tggaaacgaa actcacgta agggattttg gtcatgagat tatcaaaaag gatcttcacc	9960
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact	10020
tggcttgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt	10080
cgttcatcca tagttgctg actcctgcaa accacgttgt gtctcaaat ctctgatgtt	10140
acattgcaca agataaaaa ataatcatcat gaacaataaa actgtctgct tacataaaca	10200

gtaatacaag ggggtgttatg agccatattc aacgggaaac gtcttgctcg aggccgcgat 10260
 taaattccaa catggatgct gatttataatg ggtataaatg ggctcgcgat aatgtcgggc 10320
 aatcaggtgc gacaatctat cgattgtatg ggaagcccgga tgcgccagag ttgtttctga 10380
 aacatggcaa aggtagcgtt gccaatgatg ttacagatga gatggtcaga ctaaactggc 10440

tgacggaatt tatgcctctt ccgaccatca agcattttat ccgtactcct gatgatgcat 10500
 ggttactcac cactgcgac cccgggaaaa cagcattcca ggtattagaa gaatatcctg 10560
 attcaggtga aaatattgtt gatgcgctgg cagtgttctt gcgccggttg cattcgattc 10620
 ctgtttgtaa ttgtcctttt aacagcgatc gcgtatttcg tctcgtcag 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
 gcctcgggtga gttttctcct tcattacaga aacggctttt tcaaaaatat ggtattgata 10980
 atcctgatat gaataaatg cagtttcatt tgatgctcga tgagttttct taaggcgggc 11040
 ctgccacat acccagcgc aaacaagcgc tcatgagccc gaagtggcga gcccgatctt 11100
 ccccatcggg 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 aattcggtag ctagttataa 60
 tagtaatcaa ttacggggtc attagtcat agcccatata tggagtccg cgttacataa 120
 cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata 180
 atgacgtatg ttcccatagt aacccaata gggactttcc attgacgtca atgggtggag 240
 tatttacggt aaactgccc cttggcagta catcaagtgt atcatatgcc aagtacgccc 300
 cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta catgacctta 360

tgggactttc ctacttgcca gtacatctac gtattagtca tcgtattac catggtcgag 420
 gtgagcccca cgttctgctt cactctcccc atctcccccc cctccccacc cccaattttg 480

tatttatatta ttttttaatt attttgtgca gcgatggggg cggggggggg gggggggcgc	540
gcgccaggcg gggcggggcg gggcgaggcg cggggcgggg cgaggcggag aggtgcggcg	600
gcagccaatc agagcggcgc gctccgaaag tttctttta tggcgaggcg gcggcggcgg	660
cggccctata aaaagcgaag cgcgcgggcg gcgggagtcg ctgcgacgt gccttcgccc	720
cgtgccccgc tccgcccgcg cctcgcgcg cccgccccgg ctctgactga ccgcgttact	780
cccacaggtg agcgggcggg acggcccttc tcctccgggc tgtaattagc gcttggttta	840
atgacggctt gtttcttttc tgtggctgcg tgaaagcctt gaggggctcc gggagctaga	900
gcctctgcta accatgttca tgctttcttc ttttctctac agctcctggg caacgtgctg	960
gttattgtgc tgctcatca ttttggcaaa gaattcctcg aagatccgaa gggaaagtct	1020
tccacgactg tgggatccgt tcgaagatat caccggttga gccacatgg aattcagcag	1080
ccccagcaga gaggaatgcc ccaagcctct gagccgggtg tcaatcatgg ccggtatctt	1140
gacaggactg ctgctgcttc aggcctgttc ttgggcttct ggcgctagac ctgcatccc	1200
caagagcttc ggctacagca gcgtcgtgtg cgtgtgcaat gccacctact gcgacagctt	1260
cgacctctct acctttctg ctctgggcac cttcagcaga tacgagagca ccagatccgg	1320
cagacggatg gaactgagca tgggacccat ccaggccaat cacacaggca ctggcctgct	1380
gctgacactg cagcctgagc agaaattcca gaaagtgaaa ggcttcggcg gagccatgac	1440
agatgccgcg gctctgaata tcctggctct gtctccacca gctcagaacc tgctgctcaa	1500
gagctacttc agcgaggaag gcatcggcta caacatcatc agagtgccca tggccagctg	1560
cgacttcagc atcaggacct acacctacgc cgacacaccc gacgatttcc agctgcacaa	1620
cttcagcctg ccigaagagg acaccaagct gaagatccct ctgatccaca gagccctgca	1680
gctggcacia 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 gctaccctt	1920
tcagtgcctg ggctttacac ccgagcacca gcgggacttt atgccccgtg atctgggacc	1980
cacactggcc aatagacccc accataatgt gcggctgctg atgctggacg accagagact	2040
gcttctgccc cactgggcta aagtgggtgt gacagatcct gaggcgcgca aatactgca	2100
cggaatcgcc gtgcactggt atctggactt tctggcccct gccaggcca cactgggaga	2160
gacacacaga ctgttcccca acaccatgtt gttcgcagc gaagcctgtg tgggcagcaa	2220
gttttgggaa cagagcgtgc ggctcggcag ctgggataga ggcatgcagt acagccacag	2280
catcatcacc aacctgtgt accacgtcgt cggtgggacc gactggaatc tggccctgaa	2340

tcctgaaggc ggcctaact ggggtccgaaa cttcgtggac agcccatca tcgtggacat	2400
caccaaggac accttctaca agcagcccat gtctaccac ctgggacact tcagcaagtt	2460
catccccgag ggcctcagc gcgttggact ggtggcttcc cagaagaacg atctggacgc	2520
cgtggctctg atgcacctg atggatctgc tgtgggtgtg gtcctgaacc gcagcagcaa	2580
agatgtgcc ctgacctca aggatccgc cgtgggattc ctggaacaa tcagccctgg	2640
ctactccatc cacacctacc tgtggcgtag acagtgacaa ttgttaatta agtttaaac	2700
ctcgaggccg caagcttacc gataatcaac ctcctggatta caaaatttgt gaaagattga	2760
ctggtattct taactatgtt gctcctttta cgtatgttg atacgtgct ttaatgcctt	2820
tgtatcatgc tattgcttcc cgtatggctt tcattttctc ctccttgat aaatcctgtt	2880
tgtgtctctt ttatgaggag ttgtggcccg ttgtcaggca acgtggcgtg gtgtgcactg	2940
tgtttgtga cgcaaccccc actggttggg gcattgccac cacctgtcag ctcctttccg	3000
ggactttgc tttccccctc cctattgcc cggcgggaact catcgccgc tgccttgccc	3060
gctgctggac aggggctcgg ctgttgggca ctgacaattc cgtggtgttg tcggggaaat	3120
catgctctt tcttggctg ctcgcctgtg ttgccactg gattctgcgc gggacgtcct	3180
tctgtactgt ccttcggcc ctcaatccag cggaccttc ttcgcgggc ctgctgccgg	3240
ctctgcggcc tcttcggct cttgccttc gccctcagac gactcggatc tcccttggg	3300
ccgcctccc gcatcgatac cgtcgactag agctcgtga tcagcctcga ctgtgccttc	3360
tagttgccag ccatctgtt tttgccctc ccccgctcct tcttgacct tggagggtgc	3420
cactccact gtcttttct aataaaatga ggaaattgca tcgcattgtc tgagtaggtg	3480
tcattctatt ctggggggtg ggggtgggca ggacagcaag ggggaggatt gggaagacaa	3540
tagcaggcat gctggggaga gatccacat aacaaacagc tttttgggg tgaacatatt	3600
gactgaattc cctgcaggtt ggccactccc tctctgcgc ctcgtcgtc cactgaggcc	3660
gcccgggcaa agcccgggcg tcgggcgacc ttggtcgcc cggcctcagt gagcgagcga	3720
gcgcgcagag agggagtggc caactccatc actaggggtt cctgcggccg ctcgtacggt	3780
ctcgaggaat tctgcagga taacttgcca acctattct aaaatgtata tagaagccca	3840
aaagacaata acaaaaatat tctttagaa caaatggga aagaatgttc cactaaatat	3900
caagatttag agcaaagcat gagatgtgtg gggatagaca gtgaggctga taaaatagag	3960
tagagctcag aaacagacc attgatatat gtaagtgacc tatgaaaaa atatggcatt	4020
ttacaatggg aaaatgatgg tcttttctt ttttagaaaa acagggaat atatttatat	4080

gtaaaaata aaaggaacc catatgtcat accatacaca caaaaaaatt ccagtgaatt 4140

ataagtctaa atggagaagg caaaacttta aatcttttag aaaataatat agaagcatgc 4200

agaccagcct ggccaacatg atgaaacct ctctactaat aataaaatca gtagaactac 4260

tcaggactac tttagtgagg aagtcctttt ctatgaagac ttccttgacc aaaattaggc 4320

tctaaatgca aggagatagt gcatcatgcc tggctgcact tactgataaa tgatgttacc 4380

accatcttta accaaatgca caggaacaag ttatggtact gatgtgctgg attgagaagg 4440

agctctactt ccttgacagg acacatttgi atcaactta aaaagcagat ttttgccagc 4500

agaactattc attcagaggt aggaacttta gaatagatga tgtcactgat tagcatggct 4560

tccccatctc cacagctgct tcccaccag gttgccaca gttgagtttg tccagtgcctc 4620

agggtgccc actctcagta agaagcccca caccagcccc tctccaaata tgttggtgt 4680

tccttcatt aaagtgacc cacttttagag cagcaagtgg atttctgttt cttacagttc 4740

aggaaggagg agtcagctgt gagaacctgg agcctgagat gcttctaagt cccactgcta 4800

ctggggtcag ggaagccaga ctccagcatc agcagtcagg agcactaagc ccttgccaac 4860

atctgtttc tcagagaaac tgcttcatt ataatggttg tcctttttta agctatcaag 4920

ccaaacaacc agtgtctacc attattctca tcacctgaag ccaagggttc tagcaaaagt 4980

caagctgtct tgtaatggtt gatgtgcctc cagcttctgt cttcagtcac tccactctta 5040

gcctgctctg aatcaactct gaccacagtt ccttgagacc cctgccacct gctgccctg 5100

ccacttctc catctgcagt gctgtgcagc ctctgcact cttgcagagc taataggtgg 5160

agacttgaag gaagaggagg aaagtttctc ataatagcct tgctgcaagc tcaaatggga 5220

ggtgggcact gtgcccagga gccttgagc aaaggctgtg cccaacctct gactgcatcc 5280

aggtttggtc ttgacagaga taagaagccc tggcttttgg agccaaaatc taggtcagac 5340

ttaggcagga ttctcaaagt ttatcagcag aacatgaggc agaagacct ttctgtcca 5400

gcttcttcag gctcaacctt catcagaata gatagaaaga gaggtgtga gggttctta 5460

aacagaagca aatctgactc agagaataaa caacctccta gtaactaca gcttagacag 5520

agcatctggt ggtagtgtg ctgagtgcc tactcaactg tctggtatca gccctcatga 5580

ggacttctct tcttccctc atagacctcc atctctgttt tcttagcct gcagaaatct 5640

ggatggctat tcacagaatg cctgtgcttt cagagttgca tttttctct ggtattctgg 5700

ttcaagcatt tgaaggtagg aaagtttctc caagtgaag aaagccagcc ctgagcctca 5760

actgcctggc tagtgtggtc agtaggatgc aaaggctgtt gaatgccaca aggccaaact 5820

ttaacctgtg taccacaagc ctagcagcag aggcagctct gctcactgga actctctgtc	5880
ttctttctcc tgagcccttt cttttcctga gttttctagc tctcctcaac cttacctctg	5940
ccctaccagc gacaaaccca agagccactg tttctgtgat gtcctctcca gccctaatta	6000
ggcatcatga cttcagcctg accttccatg ctcagaagca gtgctaatacc acttcagatg	6060
agctgtctta tgcaacacag gcagagccta caaacctttg caccagagcc ctccacatat	6120
cagtgtttgt tcatactcac ttcaacagca aatgtgactg ctgagattaa gatttttacac	6180
aagatggctc gtaatttcac agtttagttt atcccattag gtaigaaaga 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 cagacagctc ccaagacaag ccagcctgag ccagagagag aactgcaaga	6540
gaaagtttct aatttaggtt ctgttagatt cagacaagtg caggctatcc tctctccaca	6600
gctactcacc tctccagcct aacaaagcct gcagtcacac ctccaaccct ggtgtctcac	6660
ctcctagcct ctccaacat cctgtctctc gaccatcttc tgcatctctc atctcaccat	6720
ctcccactgt ctacagccta ctcttgcaac taccatctca tttctgaca tctgtctac	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 ccatctcaa atgctccaca tttctcacat cctctgatt	7080
ctggtcacta cccatcttca aagaacagaa tatctcacat cagcatactg tgaaggacta	7140
gtcatgggtg cagctgctca gagctgcaaa gtcattctgg atggctggaga gcttacaac	7200
atttcatgat gctcccccg ctctgatggc tggagcccaa tccctacaca gactcctgct	7260
gtatgtgttt tcttttact ctgagccaca gccagagggc aggcattcag tctcctcttc	7320
aggctggggc tggggcactg agaactcacc caacaccttg ctctcactcc ttctgcaaaa	7380
caagaaagag ctttgtgtg cagtagccat gaagaatgaa aggaaggctt taactaaaaa	7440
atgtcagaga ttattttcaa ccccttactg tggatcacca gcaaggagga aacacaacac	7500
agagacattt tttcccctca aattatcaaa agaactactg catttggtta agagagcaac	7560
tgaatcagga agcagagttt tgaacatac agaagttagg aatctgcac agagacaaat	7620
gcagtcattg ttgtttgctg cataccagcc ctaatcatta gaagcctcat ggacttcaaa	7680

catcattccc tctgacaaga tgctctagcc taactccatg agataaaaata aatctgcctt 7740
 tcagagccaa agaagagtc 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 ctactactt ctggaatagc tcagaggccg aggcggcctc ggcctctgca 8040
 taaataaaaa aaattagtca gccatggggc ggagaatggg cggaactggg cggagttagg 8100
 ggcgggatgg gcggagttag gggcgggact atggttgctg actaattgag atgcatgctt 8160
 tgcatacttc tgctgtctgg ggagcctggg gactttccac acctggttgc tgactaattg 8220
 agatgcatgc ttgcatact tctgcctgct ggggagcctg gggactttcc acaccctaac 8280
 tgacacacat tccacagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc 8340

glattgggcg ctcttccgct tectcgtca ctgactcgt gcgctcggtc gttcggctgc 8400
 ggcgagcggg atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata 8460
 acgcaggaag gaacatgtga gcaaaaaggc agcaaaaaggc caggaaccgt aaaaaggcgg 8520
 cgttgtctggc gtttttccat aggcctcgcc cccctgacga gcatcacaaa aatcgacgt 8580
 caagtccagag gtggcgaac ccgacaggac tataaagata ccaggcgttt cccctggaa 8640
 gtcctctgt gcgtctcct gttecgacct tgccgcttac cggatactg tccgctttc 8700
 tcccttcggg aagcgtggcg ttttctcata gctcacgtg taggtatctc agttcgggtg 8760

aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgtgcg 8820
 ccttatccgg taactatcgt cttgagtcca acccggttaag acacgactta tcgccactgg 8880
 cagcagccac tggtaacagg attagcagag cgaggtatgt aggcgggtgct acagagttct 8940
 tgaagtgggt gcctaactac ggctacacta gaagaacagt atttgggtatc tgcgtctgc 9000
 tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 9060
 ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 9120
 aagaagatcc ttgatcttt tctacggggt ctgacgtca gtggaacgaa aactcacgtt 9180

aagggatattt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa 9240
 aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 9300
 gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcaccc atagtgcct 9360
 gactcctgca aaccacgttg tgtctcaaaa tctctgatgt tacattgcac aagataaaaa 9420
 tatatcatca tgaacaataa aactgtctgc ttacataaac agtaatacaa ggggtgttat 9480
 gagccatatt caacgggaaa cgtcttgctc gaggccgca ttaaattcca acatggatgc 9540

tgatttatat gggatataat gggctcgcga taatgtcggg caatcagggtg cgacaatcta 9600

tcgattgtat gggaagcccg atgcgccaga gttgtttctg aaacatggca aaggtagcgt 9660

tgccaatgat gttacagatg agatggtcag actaaactgg ctgacggaat ttatgcctct 9720

tccgaccatc aagcatttta tccgtactcc tgatgatgca tggttactca ccaactgcgat 9780

ccccgggaaa acagcattcc aggtattaga agaatatcct gattcagggtg aaaatatgtt 9840

tgatgcgctg gcagtgttcc tgcgccggtt gcattcgatt cctgtttgta attgtccttt 9900

taacagcgat cgctatttc gtctcgctca ggcgcaatca cgaatgaata acggtttggg 9960

tgatgcgagt gattttgatg acgagcgtaa tggctggcct gttgaacaag tctggaaaga 10020

aatgcataag cttttgccat tctcaccgga ttcagtcgtc actcatgggtg atttctcact 10080

tgataacctt atttttgacg aggggaaatt aataggttgt attgatgttg gacgagtcgg 10140

aatcgcgac cgataaccagg atcttgccat cctatggaac tgccctcggtg agttttctcc 10200

ttcattacag aaacggcttt ttcaaaaata tggatattgat aatcctgata tgaataaatt 10260

gcagtttcat ttgatgctcg atgagttttt ctaaggcgcg cctgccacca taccacgccc 10320

gaaacaagcg ctcatgagcc cgaagtggcg agcccgatct tcccctcgg tgatgtcggc 10380

gatataggcg ccagcaaccg cacctgtggc gccggtgatg agggcgcgcc aagtcgacgt 10440

ccggcagttc tggccactcc ctctctgcgc gctcgtcgc tcaactgaggc cgggcgacca 10500

aaggtcgccc gacgcccggg ctttgcccgg gcggcctcag tgagcgagcg agcgcgcaga 10560

gaggagtggt ccaactccat cactaggggt tctgtctagc tctgggtatt taagcccag 10620

tgagcacgca gggctctcat tttgaagcgg gaggttacgc gttcgtcgac tactagtggg 10680

taccagagcg tggtagctga gatgttttct aggaaacaca aaagatacaa aaaagaacac 10740

gtggaaggat agccaaaaag gggggtgcc ccatttct gcaccccgct gcgatggctg 10800

gcacatttg gaagacttcg agatacactg ttgagcgag taagacaaca gtgtatctcg 10860

aagtcttcca gatggggcca gccgtccac tctgtatcca ggccagttct gcaaggcgtt 10920

cgaggaccac cccctcccc tgcaccacag ggtggtctca tacagaactt ataagattcc 10980

caaatccaaa gacatttcac gtttatgggtg atttcccaga acacatagcg acatgcaaat 11040

attgcagggc gccactcccc tgtccctcac agccatcttc ctgccagggc gcacgcgcgc 11100

tgggtgttcc cgctagtga cactgggccc gcgattcctt ggagcgggtt gatgacgtca 11160

gcgtttccca tggatgaatcc ctaggtt 11187

<210> 13

<211> 10960

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 13

ttggccactc ccctctctgcg cgctcgctcg ctactgagg cgggcgacc aaaggtcgcc	60
cgacgcccg gctttgccg ggcgccctca gtgagcgagc gagcgcgag agagggagt	120
gccaaactcca tcaactagggg ttctgctag ctctgggtat ttaagcccga gtgagcacgc	180
agggtctcca tttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc	240
tcctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac	300
cctagttatt aatagtaatc aattacgggg tcattagttc atagccata tatggagttc	360
cgcgttacat aacttacggt aaatggcccg cctggctgac cgccaacga ccccgccca	420
ttgacgtcaa taatgacgta tgttccata gtaacgcaa tagggacttt ccattgacgt	480
caatgggtgg actatttacg gtaactgcc cacttggcag tacatcaagt gtatcatatg	540
ccaagtacgc ccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgccag	600
tacatgacct tatgggactt tctacttgg cagtacatct acgtattagt catcgtatt	660
accatggtcg aggtgagccc cacgttctgc ttactctcc ccatctccc cccctccca	720
cccccaattt tgtatttatt tatttttta ttattttgtg cagcgatggg ggcggggggg	780
gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg	840
agaggtgcgg cggcagccaa tcagagcggc gcgtccgaa agtttcttt tatggcgagg	900
cggcgcgggc ggcgcccta taaaaagcga agcgcgcggc gggcgggagt cgtgcgacg	960
ctgccttcgc cccgtgccc gctccgccgc cgctcgcgc cgccgcccc ggctctgact	1020
gaccgcgtta ctccacagg tgagcgggcg ggacggccct tctctccgg gctgtaatta	1080
gcgcttggtt taatgacggc ttgtcctggt ggcgagggga ggggggtggt cctcgaacgc	1140
cttcgagaac tggcctggat acagagtga ccggctggcc ccatctgga gacttcgaga	1200
tacactgttg tcttactgcg ctcaacagtg tatctgaag tcttccaat ggtgccagcc	1260
atcgagcgg ggtgcaggaa atgggggcag ccccccttt tggctatcct tccacgtgtt	1320
ctttttgta tcttttgtt ttctagaaa acatctcagt caccaccttt ctgtggtgc	1380
gtgaaagcct tgaggggctc ggggagctag agcctctgct aacctgttc atgccttctt	1440
cttttcccta cagctcctgg gcaacgtgct ggttattgtg ctgtctcatc attttgcaa	1500
agaattctc gaagatccga agggaaagtc ttccagact gtgggatccg ttcgaagata	1560

tcaccggttg agccaccatg gaattcagca gccccagcag agaggaatgc cccaagcctc	1620
tgagccgggt gtcaatcatg gccgcatctc tgacaggact gctgctgctt caggccgtgt	1680
cttgggcttc tggcgctaga ccttgcaccc ccaagagctt cggtacagc agcgtcgtgt	1740
gcgtgtgcaa tgccacctac tgcgacagct tcgaccctcc tacctttcct gctctgggca	1800
ccttcagcag atacgagagc accagatccg gcagacggat ggaactgagc atgggaccca	1860
tccaggccaa tcacacaggc actggcctgc tgcctgacact gcagcctgag cagaaattcc	1920
agaaagttaa aggccttcggc ggagccatga cagatgccgc cgctctgaat atcctggctc	1980
tgtctccacc agctcagaac ctgctgctca agagctactt cagcgaggaa ggcatcggct	2040
acaacatcat cagagtgcgc atggccagct gcgacttcag catcaggacc tacacctacg	2100
ccgacacacc cgacgatttc cagctgcaca acttcagcct gcctgaagag gacaccaagc	2160
tgaagatccc tctgatccac agagccctgc agctggcaca aagaccctgt tactgtctgg	2220
cctctccatg gacatctccc acctggctga aaacaaatgg cgccgtgaat ggcaagggca	2280
gcctgaaagg ccaacctggc gacatctacc accagacctg ggccagatac ttcgtgaagt	2340
tcctggagcg ctatgccgag cacaagctgc agttttgggc cgtgacagcc gagaacgaac	2400
cttctgctgg actgctgagc ggctaccctt ttcagtgcct gggctttaca cccgagcacc	2460
agcgggactt tatcgccctg gatctgggac ccacactggc caatagcacc caccataatg	2520
tgcggctgct gatgctggac gaccagagac tgcttctgcc ccactgggct aaagtgggtg	2580
tgacagatcc tgaggccgcc aaatacgtgc acggaatcgc cgtgactgg tatctggact	2640
ttctggcccc tgccaaggcc aactgggag agacacacag actgttcccc aacaccatgc	2700
tgttcgccag cgaagcctgt gtgggcagca agttttggga acagagcgtg cggctcggca	2760
gctgggatag aggcctgagc tacagccaca gcatcatcac caacctgctg taccagctcg	2820
tcggctggac cgactggaat ctggccctga atcctgaagg cggccctaac tgggtccgaa	2880
acttcgtgga cagccccatc atcgtggaca tcaccaagga caccctctac aagcagccca	2940
tgttctacca cctgggacac ttcagcaagt tcacccccga gggctctcag cgcgttggac	3000
tgggtggcttc ccagaagaac gatctggacg ccgtggctct gatgcacct gatggatctg	3060
ctgtggtggt ggtcctgaac cgcagcagca aagatgtgcc cctgaccatc aaggatccc	3120
ccgtgggatt cctggaaaca atcagccctg gctactccat ccacacctac ctgtggcgta	3180
gacagtgaca attgttaatt aagtttaaac cctcgaggcc gcaagcttat cgataatcaa	3240
cctctggatt acaaaatttg tgaaagattg actggtattc ttaactatgt tgctcctttt	3300

acgctatgtg gatacgtcgc tttaatgcct ttgtatcatg ctattgcttc cegtatggct 3360

ttcattttct ctcctttgta taaatcctgg ttgctgtctc tttatgagga gttgtggccc 3420

gttgtcaggc aacgtggcgt ggtgtgcaact gtgtttgctg acgcaacccc cactggttgg 3480

ggcattgccca ccacctgtca gctcctttcc gggactttcg ctttccccct ccctattgcc 3540

acggcggaac tcatcgccgc ctgccttgcc cgtgctgga caggggctcg gctgttgggc 3600

actgacaatt ccgtggtgtt gtccgggaaa tcatcgtcct ttctttggct gctcgccgtg 3660

gttgccacct ggattctgcg cgggacgtcc ttctgctacg tcccttcggc cctcaatcca 3720

gcggaccttc cttcccgcg cctgtgcgcg gctctgcggc ctcttcgcgc tcttcgcctt 3780

cgccctcaga cgagtcggat ctccctttgg gccgcctccc cgcatcgata ccgtcgacta 3840

gagctcgctg atcagcctcg actgtgcctt ctagtgtcca gccatctgtt gtttgccctt 3900

ccccgtgcc ttctttgacc ctggaagggt ccaactccac tgtcctttcc taataaaatg 3960

aggaaattgc atcgattgt ctgagtaggt gtcattctat tctggggggt ggggtggggc 4020

aggacagcaa gggggaggat tgggaagaca atagcaggca tgctggggag agatccacga 4080

taacaaacag cttttttggg gtgaacatat tgactgaatt ccctgcaggt tggccactcc 4140

ctctctgcgc gctcgtcgc tcactgaggc cgcccgggca aagcccgggc gtcgggcgac 4200

ctttggtcgc ccggcctcag tgagcgagcg agcgcgcaga gagggagtgg ccaactccat 4260

cactaggggt tctcgggcc gctcgtacgg tctcgaggaa ttctgcagg ataacttgc 4320

aacctcattc taaaatgtat atagaagccc aaaagacaat acaaaaaata ttctttaga 4380

acaaaatggg aaagaatgtt ccaactaaata tcaagattta gagcaaagca tgagatgtgt 4440

ggggatagac agtgaggctg ataaaataga gtagagctca gaaacagacc cattgatata 4500

tgtaatgtac ctatgaaaaa aatatggcat ttacaatgg 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 tgcacatgc 4860

ctggctgcac ttactgataa atgatgttat caccatcttt aaccaaagtc acaggaacaa 4920

gttatggtac tgatgtgctg gattgagaag gagctctact tccttgacag gacacatttg 4980

tatcaactta aaaaagcaga tttttgccag cagaactatt cattcagagg taggaaactt 5040

agaatagatg atgtcactga ttagcatggc ttcccatctt ccacagctgc ttccaccca 5100

ggttgccac agttgagttt gtccagtgtc cagggtgcc cactctcagt aagaagcccc 5160

acaccagccc ctctccaaat atgttggctg ttccttccat taaagtgacc ccactttaga 5220
gcagcaagtg gatttctgtt tcttacagtt caggaaggag gagtacgtg tgagaacctg 5280
gagcctgaga tgcttctaag tcccactgct actggggtca gggaagccag actccagcat 5340
cagcagtcag gagcactaag ccttggccaa catcctgttt ctgagagaaa ctgcttccat 5400

tataatggtt gtcctttttt aagctatcaa gccaaacaac cagtgtctac cattattctc 5460
atcacctgaa gccaaagggt ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct 5520
ccagcttctg tcttcagtca ctccactctt agcctgctct gaatcaactc tgaccacagt 5580
tccttgagc ccttgccacc tctgcccct gccaccttct ccatctgcag tctgtgcag 5640
ccttctgcac tcttgacag ctaatagggt gagacttgaa ggaagaggag gaaagtctt 5700
cataatagcc ttgctgcaag ctcaaatggg aggtgggcac tgtgccagg agccttgag 5760
caaaggctgt gcccaacctc tgactgcac caggtttggt cttgacagag ataagaagcc 5820

ctggcttttg gagccaaaat ctaggtcaga cttaggcagg attctcaaag tttatcagca 5880
gaacatgagg cagaagacct tttctgtcc agcttcttca ggctcaacct tcatcagaat 5940
agatagaaag agaggtgtg agggttctta aaacagaagc aaatctgact cagagaataa 6000
acaacctct agtaaacac agcttagaca gagcatctgg tggtagtgt gctcagtgct 6060
ctactcaact gtctggtatc agccctcatg aggacttctc tcttttcct catagacctc 6120
catctctgtt ttcttagcc tgcagaaac tggatggcta ttcacagaat gcctgtgctt 6180
tcagattgc atttttctc tggattctg gttcaagcat ttgaagtag gaaagttct 6240

ccaagtcaa gaaagccagc cctgagcctc aactgcctgg ctagtgtgt cagtaggatg 6300
caaaggctgt tgaatgccac aaggccaaac tttaacctgt gtaccacaag ctagcagca 6360
gaggcagctc tctcactgg aactctctgt cttctttctc ctgagccttt tcttttctg 6420
agttttctag ctctctcaa ccttacctct gccctacca ggacaaacct aagagccact 6480
gtttctgtga tctctctcc agccctaatt aggcacatg acttcagcct gaccttccat 6540
gctcagaagc agtgctaac cacttcagat gagctgctc atgcaacaca ggcagagcct 6600
acaaaccttt gcaccagagc cctccacata tcagtgtttg ttcatactca cttcaacagc 6660

aaatgtgact gctgagatta agattttaca caagatggc tgtaatttca cagttagttt 6720
tatccatta ggtatgaaag aattagcata attccctta aacatgaatg aatcttagat 6780
ttttaataa atagttttgg aagtaaagac agagacatca ggagcacaag gaatagcctg 6840
agaggacaaa cagaacaaga aagagtctgg aaatacacag gatgttcttg gcctctcaa 6900
agcaagtgca agcagatagt accagcagcc ccaggctatc agagcccagt gaagagaagt 6960
accatgaaag ccacagctct aaccacctg ttccagagtg acagacagtc cccaagacaa 7020

gccagcctga gccagagaga gaactgcaag agaaagtttc taatttaggt tctgttagat 7080

ttagacaagt gcaggtcatc ctctctccac agctactcac ctctccagcc taacaaagcc 7140

tgcagtcac actccaaccc tgggtgtctca cctcctagcc tctcccaaca tectgtctc 7200

tgaccatctt ctgcatctct catctcacca tctccactg tctacagcct actcttgcaa 7260

ctaccatctc attttctgac atctgtctca catcttctgc catactctgc catctacat 7320

accacatctt accatctacc acaccatctt ttatctccat cctctcaga agcctccaag 7380

ctgaatcctg ctttatgtgt tcatctcagc ccttgcattg aaagctgacc ccagaggcag 7440

aactattccc agagagcttg gccaaagaaa acaaaactac cagcctggcc aggtcagga 7500

gtagtaagct gcagtgtctg ttgtgttcta gcttcaacag ctgcaggagt tccactctca 7560

aatgctccac atttctcaca tctcctgat tctggtcact acccatcttc aaagaacaga 7620

atatctcaca tcagcactgt gtgaaggact agtcatgggt gcagctgtctc agagctgcaa 7680

agtcatctct gatgggtggag agcttcaaaa 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 accccttact 7980

gtggatcacc agcaaggagg aaacacaaca cagagacatt ttttccctc aaattatcaa 8040

aagaatcact gcatttgtta aagagagcaa ctgaatcagg aagcagagtt ttgaacatat 8100

cagaagttag gaatctgcat cagagacaaa tgcagtcatt gttgtttgct gcataccagc 8160

cctaattcatt agaagcctca tggacttcaa acatattcc ctctgacaag atgctctagc 8220

ctaactccat gagataaaa aaatctgcct ttcagagcca aagaagagtc caccagcttc 8280

ttctcagtggt gaacaagagc tccagtcagg ttagtcagtc cagtcagta gaggagacca 8340

gtctgcatcc tctaattttc aaaggcaaga agatttgttt accctggaca ccaggcacia 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

tatgggtgct gactaattga gatgcatgct ttgcatactt ctgcctgctg gggagcctgg 8700

ggactttcca cacctggttg ctgactaatt gagatgcatg ctttgcatac ttctgcctgc 8760

tggggagcct ggggactttc cacaccctaa ctgacacaca ttccacagct gcattaatga	8820
atcgccaac gcgcggggag aggcgggttg cgtattgggc gctcttccgc ttctcgcctc	8880
actgactcgc tgcgctcggc cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg	8940
gtaatacggc tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc	9000
cagcaaaagg ccaggaaccg taaaaggcc gcgttgctgg cgtttttcca taggtccgc	9060
ccccctgacg agcatcaca aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga	9120
ctataaagat accaggcggt tccccctgga agctccctcg tgcgctctcc tgttccgacc	9180
ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat	9240
agctcacgct gtaggtatct cagttcgggt taggtcgttc gctccaagct gggctgtgtg	9300
cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc	9360
aaccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga	9420
gcgaggtatg taggcgggtc tacagagttc ttgaagtggg ggcctaacta cggctacact	9480
agaagaacag tatttggat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt	9540
ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtgggttttt tgtttgcaag	9600
cagcagatta cgcgcagaaa aaaaggatct caagaagatc cttgatctt ttctacgggg	9660
tctgacgctc agtggaaacga aaactcacgt taagggtttt tggatcatgag attatcaaaa	9720
aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata	9780
tatgagtaaa ctgggtctga cagttacaa tgcttaatca gtgaggcacc tatctcagcg	9840
atctgtctat ttctgtctc catagttgcc tgactcctgc aaaccacgtt gtgtctcaaa	9900
atctctgatg ttacattgca caagataaaa atatatcatc atgaacaata aaactgtctg	9960
cttacataaa cagtaatata 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 ttccgacat caagcatttt atccgtactc	10260
ctgatgatgc atggttactc accactgcga tccccgggaa aacagcattc caggtattag	10320
aagaatatcc tgattcaggt gaaaatattg ttgatgcgct ggcagtgttc ctgcgccgt	10380
tgcattcgat tctgtttgt aattgtcctt ttaacagcga tcgctattt cgtctcgcctc	10440
aggcgcaatc acgaatgaat aacggttttg ttgatgcgag tgattttgat gacgagcgta	10500
atggctggcc tgttgaacaa gtctggaaaag 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 ataccacgc cgaacaagc gctcatgagc ccgaagtggc 10860

gagcccgatc ttcccatcg gtgatgtcgg cgatataggc gccagcaacc gcacctgtgg 10920
 cgccggtgat gaggcgcgcg 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 tctaccttt cctgctctgg gcaccttcag cagatacgag 240

agcaccagat cgggcagacg gatggaactg agcatgggac ccatccaggc caatcacaca 300

ggcactggcc tgetgctgac actgcagcct gagcagaaat tccagaaagt gaaaggcttc 360

ggcggagcca tgacagatgc cgccgctctg aatatactgg ctctgtctcc accagctcag 420

aacctgctgc tcaagagcta cttcagcgag gaaggcatcg gctacaacat catcagagtg 480

cccatggcca gctgcgactt cagcatcagg acctacacct acgccgacac acccgacgat 540

ttccagctgc acaacttcag cctgcctgaa gaggacacca agctgaagat ccctctgatac 600
cacagagccc tgcagctggc acaaagaccc gtgtcactgc tggcctctcc atggacatct 660
cccacctggc tgaatacaaa tggcgccgtg aatggcaagg gcagcctgaa aggccaacct 720
ggcgacatct accaccagac ctgggccaga tacttcgtga agttcctgga cgcctatgcc 780
gagcacaagc tgcagttttg ggccgtgaca gccgagaacg aaccttctgc tggactgctg 840
agcggctacc ctttcagtg cctgggcttt acaccgagc accagcggga ctttatcgcc 900

cgtgatctgg gaccacact ggccaatagc acccaccata atgtgcggct gctgatgctg 960
gacgaccaga gactgcttct gcccactgg gctaaagtgg tgctgacaga tcttgaggcc 1020
gccaaatacg tgcacggaat cgccgtgcac tggatatctgg 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
catttcagca agttcatccc cgagggtctc cagcgcgttg gactggtggc ttccagaag 1440
aacgatctgg acgccgtggc tctgatgcac cctgatggat ctgctgtggt ggtggtcctg 1500
aaccgcagca gcaaagatgt gccctgacc 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

atgtacgcc ttttctgct ggccagcctg ctgggcgccg ccctggccgg ccccgctgtg	60
ggcctgaagg agtgcacccg cggcagcgcc gtgtgtgtgcc agaactgtgaa gaccgccagc	120
gactgcggcg ccgtgaagca ctgcctgcag accgtgtgga acaagccac cgtgaagagc	180
ctgccttgcg acatctgcaa ggacgtgtgt accgccgccg gcgacatgct gaaggacaac	240
gccaccgagg aggagatcct ggtgtacctg gagaagacct gcgactggct gcccagccc	300
aacatgagcg ccagctgcaa ggagatctgt gacagctacc tgcccgtgat cctggacatc	360
atcaagggcg agatgagccg ccccgccgag gtgtgcagcg ccctgaacct gtgcgagagc	420
ctgcagaagc acctggccga gctgaaccac cagaagcagc tggagagcaa caagatcccc	480
gagctggaca tgaccgaggt ggtggccccc ttcatggcca acatccccct gctgctgtac	540
ccccaggacg gccccgcag caagccccag cccaaggaca acggcgacgt gtgccaggac	600
tgcatccaga tggtagccga catccagacc gccgtgcgca ccaacagcac ctctgtgcag	660
gccctggtgg agcacgtgaa ggaggagtgc gaccgcctgg gccccgcat ggccgacatc	720
tgcaagaact acatcagcca gtacagcgag atcgccatcc agatgatgat gcacatgcag	780
cccaaggaga tctgcgcctt ggtgggcttc tgcgacgagg tgaaggagat gcccattgag	840
accttggtgc ccgccaaggt ggccagcaag aacgtgatcc ccgccttga 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 ccccagctg	1140
gtgtgcagca tgctgcacct gtgcagcggc accgcctgc ccgcctgac cgtgcacgtg	1200
accagccca aggacggcgg ctcttgcgag gtgtgcaaga agctggtggg ctacctggac	1260
cgcaacctgg agaagaacag caccaagcag gagatcctgg ccgccttga gaagggtgc	1320
agcttctgc ccgacccta ccagaagcag tgcgaccagt tcgtggccga gtacagccc	1380
gtgtgatcg agatcctggt ggaggtgatg gacccagct tcgtgtgcct gaagatcggc	1440
gcctgcccc gcgcccacaa gcccctgtg ggacccgaga agtgcatctg gggccccagc	1500
tactgtgtcc agaacaccga gaccgccgc cagtgaacg 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

atgggcccgt gctgcttcta caccgccggc accctgagcc tgctgctgct ggtgaccagc	60
gtgaccctgc tggtagcccg cgtgttccag aaggccgtgg accagagcat cgagaagaag	120
atcgtgctgc gcaacggcac cgaggccttc gacagctggg agaagccccc cctgcccgtg	180
tacaccagtg tctacttctt caacgtgacc aaccccagg agatcctgcg cggcgagacc	240
ccccgcgtgg aggagggtggg cccctacacc taccgcgagc tgcgcaacaa ggccaacatc	300
cagttcggcg acaacggcac caccatcagc gccgtgagca acaaggccta cgtgttcgag	360
cgcgaccaga gcgtgggcga cccaagatc gacctgatcc gcaccctgaa catccccgtg	420
ctgaccgtga tcgagtggag ccagggtgcac ttcttgcgcg agatcatcga ggccatgctg	480
aaggcctacc agcagaagct gtctgtgacc cacaccgtgg acgagctgct gtggggctac	540
aaggacgaga tctgagcct gatccacgtg ttccgccccg acatcagccc ctacttcggc	600
ctgtttctacg agaagaacgg caccaacgac ggcgactacg tgttctgac cggcgaggac	660
agctacctga acttcacaa gatcgtggag tggaacggca agaccagcct ggactggtgg	720
atcacgcaga agtgcaacat gatcaacggc accgacggcg acagcttcca cccctgac	780
accaaggacg aggtgctgta cgtgttcccc agcgacttct gccgcagcgt gtacatcacc	840
ttcagcgact acgagagcgt gcagggcctg cccgccttcc gctacaaggt gcccgcgag	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 tgggtgttccc cgtgatgtac ctgaacgaga gcgtgcacat cgacaaggag	1260
accgccagcc gcctgaagag catgatcaac accaccctga tcatcacaa catcccctac	1320
atcatcatgg cctggggcgt gttcttcggc ctggtgttca cctggctggc ctgcaagggc	1380
cagggcagca tggacgaggg caccgccgac gagcgcgccc cctgatccg 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

ttagaaata 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 gaggtagcac gcagggtctc cattttgaag cgggagggtta	60
<210> 29	
<211> 145	
<212> DNA	
<213> Unknown	
<220><223> AAV2 ITR	
<400> 29	
aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcaactgagg	60

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc 120
gagcgcgag agagggagtg gccaa 145

<210> 30

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 30

tattagatct gatggccgcg 20

<210> 31

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 31

tccatcacta ggggttcctg 20