



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0098384
(43) 공개일자 2022년07월12일

(51) 국제특허분류(Int. Cl.)
C12N 15/86 (2006.01) A61K 48/00 (2006.01)
C07K 14/475 (2006.01) C07K 14/62 (2006.01)
C07K 14/78 (2006.01) C12N 9/26 (2006.01)
(52) CPC특허분류
C12N 15/86 (2013.01)
A61K 48/0058 (2013.01)
(21) 출원번호 10-2022-7020169
(22) 출원일자(국제) 2022년11월19일
심사청구일자 없음
(85) 번역문제출일자 2022년06월14일
(86) 국제출원번호 PCT/US2020/061223
(87) 국제공개번호 WO 2021/102107
국제공개일자 2021년05월27일
(30) 우선권주장
62/937,583 2019년11월19일 미국(US)
(뒷면에 계속)

(71) 출원인
아스클레피오스 바이오파마슈티컬, 인크.
미국 노스 캐롤라이나주 27709 리서치 트라이앵글
파크 스위트 110 알렉산더 드라이브 티.더블유.
20
(72) 발명자
로버츠 마이클
미국 노스 캐롤라이나주 27709 리서치 트라이앵글
파크 스위트 110 알렉산더 드라이브 티.더블유.
20 아스클레피오스 바이오파마슈티컬, 인크. 내
이글레시아스 후안 마누엘
미국 노스 캐롤라이나주 27709 리서치 트라이앵글
파크 스위트 110 알렉산더 드라이브 티.더블유.
20 아스클레피오스 바이오파마슈티컬, 인크. 내
(뒷면에 계속)
(74) 대리인
문두현

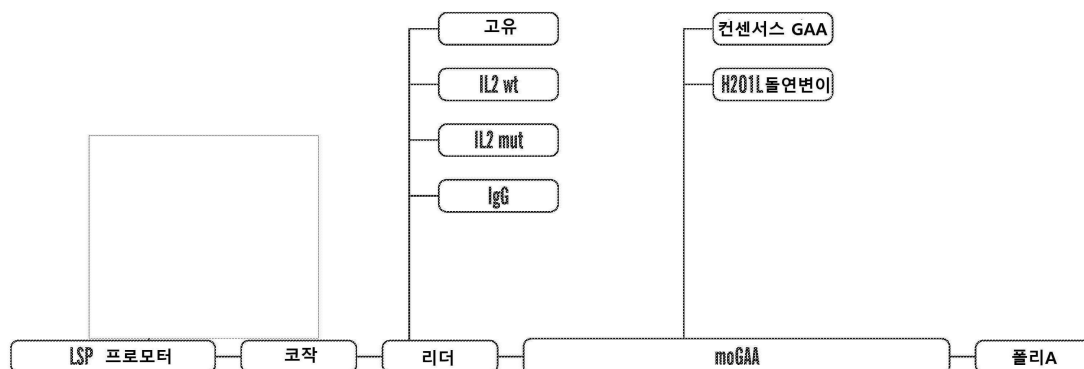
전체 청구항 수 : 총 73 항

(54) 발명의 명칭 **폼페병 및 리소좀 장애를 치료하기 위한 간-특이적 프로모터를 포함하는 치료적 아데노-관련 바이러스**

(57) 요약

리소좀 단백질, 예를 들어, 산 알파-글루코시다아제(GAA) 폴리펩타이드, 및 임의로 신호 펩타이드 및/또는 임의로 표적화 서열, 예를 들어, IGF2 표적화 펩타이드를 암호화하고, 간-특이적 프로모터(LSP)에 작동적으로 연결된 이중 핵산을 포함하고, GAA 폴리펩타이드가 간으로부터 분비되게 하고 리소좀을 표적으로 할 수 있는 rAAV 게놈을 포함하는 재조합 AAV(rAAV) 벡터. 특정 구현에는 간 분비 신호 펩타이드 및 인간 양이온-비의존적 만노오스-6-포스페이트 수용체(CI-MPR)에 결합하거나 IGF2 수용체에 대한 IGF2 표적화 펩타이드를 갖는, 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하며, 리소좀에 대한 GAA 폴리펩타이드의 적절한 세포화 국소화를 허용하는 재조합 AAV(rAAV) 벡터에 관한 것이다. 또한 rAAV 벡터로 리소좀 질환, 예를 들어, 글리코젠 축적 질환 II형(GSD II) 질환 및/또는 폼페병을 치료하는 세포, 및 방법이 포함된다.

대표도



(52) CPC특허분류

C07K 14/475 (2013.01)
C07K 14/62 (2013.01)
C07K 14/78 (2013.01)
C12N 9/2408 (2013.01)
C12N 2750/14143 (2013.01)
C12N 2810/851 (2013.01)
C12N 2830/008 (2013.01)
C12N 2830/42 (2013.01)
C12N 2830/50 (2013.01)

(72) 발명자

트레티아코바 안나

미국 노스 캐롤라이나주 27709 리서치 트라이앵글
 파크 스위트 110 알렉산더 드라이브 티.더블유. 20
 아스클레피오스 바이오파마슈티컬, 인크. 내

오캘러건 마이클 더블유.

미국 노스 캐롤라이나주 27709 리서치 트라이앵글
 파크 스위트 110 알렉산더 드라이브 티.더블유. 20
 아스클레피오스 바이오파마슈티컬, 인크. 내

프랑수아 아킬레

미국 노스 캐롤라이나주 27709 리서치 트라이앵글
 파크 스위트 110 알렉산더 드라이브 티.더블유. 20
 아스클레피오스 바이오파마슈티컬, 인크. 내

(30) 우선권주장

62/937,556 2019년11월19일 미국(US)
 63/023,570 2020년05월12일 미국(US)

명세서

청구범위

청구항 1

재조합 아데노바이러스 관련(AAV) 벡터로서, 그 게놈에

- a. 5' 및 3' AAV 역 말단 반복체(ITR) 서열 및
 - b. 5' ITR과 3' ITR 사이에 위치하는, 알파-글루코시다아제(GAA) 폴리펩타이드를 포함하는 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서, 상기 이중 핵산은
 - i. CRM_SP0412(서열 번호 86) 또는 SP0412(서열 번호 91), 또는 서열 번호 86 또는 서열 번호 91에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편,
 - ii. SP0422(서열 번호 92), 또는 서열 번호 92에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편,
 - iii. CRM_SP0239(서열 번호 87) 또는 SP0239(서열 번호 93) 또는 SP0238-UTR(서열 번호 147), 또는 서열 번호 87, 서열 번호 93 또는 서열 번호 147에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;
 - iv. CRM_SP0265(SP0131_A1)(서열 번호 88) 또는 SP0265(LVR_SP0131_A1) (서열 번호 94) 또는 SP0265-UTR(서열 번호 146), 또는 서열 번호 88, 서열 번호 94 또는 서열 번호 146에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;
 - v. CRM_SP0240(서열 번호 89) 또는 SP0240(서열 번호 95) 또는 SP0240-UTR(서열 번호 148), 또는 서열 번호 89, 서열 번호 95 또는 서열 번호 148에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;
 - vi. CRM_SP0246(서열 번호 90) 또는 SP0246(서열 번호 96) 또는 SP0246-UTR(서열 번호 149), 또는 서열 번호 90, 서열 번호 96 또는 서열 번호 149에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편
- 중 어느 하나로부터 선택된 간-특이적 프로모터에 작동적으로 연결된 재조합 아데노바이러스 관련(AAV) 벡터.

청구항 2

제1항에 있어서, 상기 이중 핵산 서열은 상기 GAA 폴리펩타이드에 융합된 분비 신호를 포함하는 융합 단백질을 암호화하거나, 상기 GAA 폴리펩타이드에 융합된 표적화 펩타이드를 포함하는 융합 폴리펩타이드를 암호화하거나, 상기 GAA 폴리펩타이드에 융합된 분비 신호 및 표적화 펩타이드를 포함하는 융합 단백질을 암호화하는, 재조합 AAV 벡터.

청구항 3

제2항에 있어서, 상기 AAV 게놈은 5'에서 3' 방향으로 다음을 포함하는, 재조합 AAV 벡터:

- a. 5' ITR,
- b. 간-특이적 프로모터 서열,
- c. 인트론 서열,
- d. 분비 신호 펩타이드를 암호화하는 핵산,
- e. IGF2 표적화 펩타이드를 암호화하는 핵산,
- f. 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산,
- g. 폴리A 서열, 및
- h. 3' ITR.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 분비 신호 펩타이드를 암호화하는 핵산은 AAT 신호 펩타이드, 피브로넥틴 신호 펩타이드(FN1), GAA 리더 서열, IL-2 wt 리더 서열, 변형된 IL-2 리더 서열, IL2(1-3) 리더 서열, IgG 리더 서열, AAT 리더 서열 중 어느 하나로부터 선택된 신호 서열, 또는 분비 신호 활성을 갖는 이의 활성 단편을 암호화하는, 제조함 AAV 벡터.

청구항 5

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 IGF2 표적화 펩타이드는 인간 양이온-비의존적 만노오스-6-포스페이트 수용체(CI-MPR) 또는 IGF2 수용체에 결합하는, 제조함 AAV 벡터.

청구항 6

제5항에 있어서, 상기 IGF2 표적화 펩타이드는 서열 번호 5를 포함하거나, IGF2 수용체에 결합하는 서열 번호 5에서 적어도 하나의 아미노 변형을 포함하는, 제조함 AAV 벡터.

청구항 7

제6항에 있어서, 서열 번호 5에서 상기 적어도 하나의 아미노 변형은 V43M 아미노산 변형(서열 번호 8 또는 서열 번호 9) 또는 Δ 2-7(서열 번호 6) 또는 Δ 1-7(서열 번호 7)인, 제조함 AAV 벡터.

청구항 8

제1항 또는 제2항에 있어서, 상기 핵산 서열은 야생형 GAA 폴리펩타이드 또는 변형된 GAA 폴리펩타이드를 암호화하는, 제조함 AAV 벡터.

청구항 9

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 핵산 서열은 인간 GAA 유전자 또는 인간 코돈 최적화된 GAA 유전자(coGAA) 또는 변형된 GAA 핵산 서열인, 제조함 AAV 벡터.

청구항 10

제1항 내지 제9항 중 어느 한 항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 핵산 서열은 (i) 생체내 향상된 발현을 위한 코돈 최적화, (ii) CpG 섬 감소, (iii) 정지 서열의 변형, (iv) 대안적인 리딩 프레임의 감소, 및 (v) 선천성 면역 반응 감소 중 어느 하나를 위해 서열 번호 11로부터 변형되는, 제조함 AAV 벡터.

청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 핵산 서열은 서열 번호 10의 H201L, H199R 또는 R233H로부터 선택된 적어도 1개, 적어도 2개 또는 적어도 3개 모두의 아미노산 변형을 포함하는 GAA 폴리펩타이드를 암호화하는, 제조함 AAV 벡터.

청구항 12

제1항 내지 제11항 중 어느 한 항에 있어서, 상기 암호화된 융합 폴리펩타이드는 GAA 폴리펩타이드에는 아미노 말단에 그리고 IGF2 표적화 펩타이드에는 C-말단에 위치한 적어도 1개의 아미노산에 대한 뉴클레오타이드 서열을 포함하는 스페이서를 추가로 포함하는, 제조함 AAV 벡터.

청구항 13

제12항에 있어서, 상기 IGF2 표적화 펩타이드를 암호화하는 핵산과 상기 GAA 폴리펩타이드를 암호화하는 핵산 사이에 위치하는 적어도 1개의 아미노산의 스페이서를 암호화하는 핵산을 추가로 포함하는, 제조함 AAV 벡터.

청구항 14

제1항 내지 제13항 중 어느 한 항에 있어서, 상기 GAA 유전자를 암호화하는 핵산의 3' 및 3' ITR 서열의 5'에 위치한 적어도 하나의 폴리A 서열을 추가로 포함하는, 제조함 AAV 벡터.

청구항 15

제1항 내지 제14항 중 어느 한 항에 있어서, 상기 이중 핵산 서열은 GAA 폴리펩타이드를 암호화하는 핵산의 3' 및 3' ITR 서열의 5'에 위치한, 콜라겐 안정성(CS) 서열 또는 3' UTR 서열, 또는 CS 및 3' UTR 서열을 추가로 포함하는, 재조합 AAV 벡터.

청구항 16

제1항 내지 제15항 중 어느 한 항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 핵산과 상기 폴리A 서열 사이에 위치한, 콜라겐 안정성(CS) 서열, 또는 3' UTR 서열, 또는 CS 및 3' UTR 서열을 암호화하는 핵산을 추가로 포함하는, 재조합 AAV 벡터.

청구항 17

제1항 내지 제16항 중 어느 한 항에 있어서, 상기 분비 신호 펩타이드를 암호화하는 서열의 5' 및 상기 프로모터의 3'에 위치한 인트론 서열을 추가로 포함하는, 재조합 AAV 벡터.

청구항 18

제17항에 있어서, 상기 인트론 서열은 MVM 서열 또는 HBB2 서열 또는 SV40 서열을 포함하는, 재조합 AAV 벡터.

청구항 19

제1항 내지 제18항 중 어느 한 항에 있어서, 상기 ITR은 삽입, 결실 또는 치환을 포함하는, 재조합 AAV 벡터.

청구항 20

제19항에 있어서, 상기 ITR에서 하나 이상의 CpG 섬이 제거되는, 재조합 AAV 벡터.

청구항 21

제1항 내지 제20항 중 어느 한 항에 있어서,

a. 상기 분비 신호 펩타이드를 암호화하는 핵산은

AAT 신호 펩타이드(예를 들어, 서열 번호 17) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 17-22와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산;

피브로넥틴 신호 펩타이드(FN1) (예를 들어, 서열 번호 18-21) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 18-21과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산;

동족 GAA 신호 펩타이드(서열 번호 175) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 175와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산;

hIGF2 신호 펩타이드(예를 들어, 서열 번호 22) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 22와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산;

IgG1 리더 펩타이드(서열 번호 177) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 177과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산;

wt IL2 리더 펩타이드(서열 번호 179) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 179와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산;

돌연변이 IL2 리더 펩타이드(서열 번호 181) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 181과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산

으로 이루어진 군 중 어느 하나로부터 선택되고,

b. 상기 GAA 폴리펩타이드를 암호화하는 핵산은 서열 번호 11, 서열 번호 72 또는 서열 번호 182, 또는 서열 번호 11, 서열 번호 72, 또는 서열 번호 182와 적어도 60%, 또는 70%, 또는 80%, 85% 또는 90% 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열로 이루어진 군 중 어느 하나로부터 선택되는, 재조합 AAV 벡터.

청구항 22

제1항 내지 제21항 중 어느 한 항에 있어서, 상기 IGF2 표적화 펩타이드는 서열 번호 5, 서열 번호 6, 서열 번호 7, 서열 번호 8 또는 서열 번호 9 중 어느 하나로부터 선택되는, 재조합 AAV 벡터.

청구항 23

제1항 내지 제22항 중 어느 한 항에 있어서, 상기 IGF2 표적화 펩타이드를 암호화하는 핵산은 상기 분비 신호 펩타이드를 암호화하는 핵산과 상기 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산 사이에 위치하는, 재조합 AAV 벡터.

청구항 24

제1항 내지 제23항 중 어느 한 항에 있어서, 상기 재조합 AAV 벡터는 키메라 AAV 벡터, 반수체 AAV 벡터, 하이브리드 AAV 벡터 또는 배수체(polyploid) AAV 벡터인, 재조합 AAV 벡터.

청구항 25

제1항 내지 제24항 중 어느 한 항에 있어서, 상기 재조합 AAV 벡터는 합리적인 반수체 벡터, 모자이크 AAV 벡터, 화학적으로 변형된 AAV 벡터, 또는 임의의 AAV 혈청형으로부터의 AAV 벡터인, 재조합 AAV 벡터.

청구항 26

제1항 내지 제25항 중 어느 한 항에 있어서, 상기 재조합 AAV 벡터는 AVXL32 벡터, AAVXL32.1 벡터, AAV8 벡터, 또는 적어도 하나의 AAV8 캡시드 단백질을 포함하는 반수체 AAV8 벡터로 이루어진 군으로부터 선택되는, 재조합 AAV 벡터.

청구항 27

제1항 내지 제26항 중 어느 한 항에 있어서, 상기 혈청형은 AAV3b인, 재조합 AAV 벡터.

청구항 28

제27항에 있어서, 상기 AAV3b 혈청형은 265D, 549A, Q263Y 중 어느 하나로부터 선택된 캡시드 단백질에 하나 이상의 돌연변이를 포함하는, 재조합 AAV 벡터.

청구항 29

제28항에 있어서, 상기 AAV3b 혈청형은 AAV3b265D, AAV3b265D549A, AAV3b549A 또는 AAV3bQ263Y, 또는 AAV3bSASTG 중 어느 하나로부터 선택되는, 재조합 AAV 벡터.

청구항 30

재조합 아데노바이러스 관련(AAV) 벡터로서, 그 게놈에

a. 5' 및 3' AAV 역 말단 반복체(ITR) 서열, 및

b. 5' ITR과 3' ITR 사이에 위치하는, 알파-글루코시다아제(GAA) 폴리펩타이드를 포함하는 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서, 상기 이중 핵산은 간-특이적 프로모터에 작동적으로 연결되고, 상기 간 특이적 프로모터는

i. CRM_SP0412(서열 번호 86) 또는 SP0412(서열 번호 91), 또는 서열 번호 86 또는 서열 번호 91에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편,

ii. SP0422(서열 번호 92), 또는 서열 번호 92에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능

적 단편,

iii. CRM_SP0239(서열 번호 87) 또는 SP0239(서열 번호 93) 또는 SP0238-UTR(서열 번호 147), 또는 서열 번호 87, 서열 번호 93 또는 서열 번호 147에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;

iv. CRM_SP0265(SP0131_A1) (서열 번호 88) 또는 SP0265(LVR_SP0131_A1) (서열 번호 94), 또는 SP0265-UTR(서열 번호 146) 또는 서열 번호 88, 서열 번호 94 또는 서열 번호 146에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;

v. CRM_SP0240(서열 번호 89) 또는 SP0240(서열 번호 95) 또는 SP0240-UTR(서열 번호 148), 또는 서열 번호 89, 서열 번호 95 또는 서열 번호 148에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;

vi. CRM_SP0246(서열 번호 90) 또는 SP0246(서열 번호 96) 또는 SP0246-UTR(서열 번호 149), 또는 서열 번호 90, 서열 번호 96 또는 서열 번호 149에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편

중 어느 하나로부터 선택되고,

상기 재조합 AAV 벡터는 혈청형 AAV3, AAV3b, AAV8로부터 선택된 캡시드 단백질을 포함하는, 재조합 AAV 벡터.

청구항 31

제30항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 이중 핵산 서열은 상기 GAA 폴리펩타이드를 암호화하는 핵산의 5'에 위치한 분비 신호 펩타이드를 암호화하는 핵산을 추가로 포함하는, 재조합 AAV 벡터.

청구항 32

제31항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 이중 핵산 서열은 상기 분비 신호 펩타이드를 암호화하는 핵산과 상기 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산 사이에 위치한 표적화 펩타이드를 암호화하는 핵산을 추가로 포함하는, 재조합 AAV 벡터.

청구항 33

제30항에 있어서, 상기 AAV 계놈은 5'에서 3' 방향으로 다음을 포함하는, 재조합 AAV 벡터:

- 5' ITR,
- 간 특이적 프로모터 서열,
- 인트론 서열,
- 분비 신호 펩타이드를 암호화하는 핵산,
- 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산,
- 폴리A 서열, 및
- 3' ITR.

청구항 34

제30항에 있어서, 상기 AAV 계놈은 5'에서 3' 방향으로 다음을 포함하는, 재조합 AAV 벡터:

- 5' ITR,
- 간 특이적 프로모터 서열,
- 인트론 서열,
- 표적화 펩타이드를 암호화하는 핵산,
- 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산,
- 폴리A 서열, 및
- 3' ITR.

청구항 35

제30항 내지 제34항 중 어느 한 항에 있어서, 상기 분비 신호 펩타이드는 AAT 신호 펩타이드, 피브로넥틴 신호 펩타이드(FN1), GAA 리더 서열, IL-2 wt 리더 서열, 변형된 IL-2 리더 서열, IL2(1-3) 리더 서열, IgG 리더 서열, AAT 리더 서열 중 어느 하나, 또는 분비 신호 활성을 갖는 이의 활성 단편, 또는 분비 신호 활성을 갖는 이의 활성 단편으로부터 선택되는, 재조합 AAV 벡터.

청구항 36

제30항 내지 제35항 중 어느 한 항에 있어서, 상기 표적화 펩타이드는 인간 양이온-비의존적 만노오스-6-포스페이트 수용체(CI-MPR) 또는 IGF2 수용체에 결합하는 IGF2 표적화 펩타이드 서열 중 어느 하나 또는 이의 기능적 변이체로부터 선택되는, 재조합 AAV 벡터.

청구항 37

제36항에 있어서, 상기 IGF2 표적화 펩타이드는 서열 번호 5를 포함하거나, CI-MPR 수용체에 대한 결합에 영향을 미치지 않는 서열 번호 5에서 적어도 하나의 아미노 변형을 포함하거나, 적어도 하나의 혈청 IGF-결합 단백질(IGFBP)에 대한 결합을 감소시키거나, 서열 번호 5와 적어도 85% 서열 동일성을 갖는 아미노산 서열을 포함하는, 재조합 AAV 벡터.

청구항 38

제30항 내지 제37항 중 어느 한 항에 있어서, 상기 핵산 서열은 서열 번호 10의 야생형 GAA 폴리펩타이드 또는 변형된 GAA 폴리펩타이드를 암호화하는, 재조합 AAV 벡터.

청구항 39

제30항 내지 제38항 중 어느 한 항에 있어서, 상기 핵산 서열은 서열 번호 10의 H201L, H199R 또는 R233H로부터 선택된 적어도 1개, 적어도 2개 또는 적어도 3개 모두의 아미노산 변형을 포함하는 GAA 폴리펩타이드를 암호화하는, 재조합 AAV 벡터.

청구항 40

제30항 내지 제39항 중 어느 한 항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 핵산 서열은 인간 GAA 유전자 또는 인간 코돈 최적화된 GAA 유전자(coGAA) 또는 변형된 GAA 핵산 서열인, 재조합 AAV 벡터.

청구항 41

제30항 내지 제40항 중 어느 한 항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 핵산 서열은 CpG 섬을 감소시키도록 코돈 최적화된, 재조합 AAV 벡터.

청구항 42

제30항 내지 제41항 중 어느 한 항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 핵산 서열은 선천성 면역 반응을 감소시키거나 CpG 섬을 감소시키도록 코돈 최적화되거나, 또는 선천성 면역 반응을 감소시키고 선천성 면역 반응을 감소시키도록 코돈 최적화된, 재조합 AAV 벡터.

청구항 43

제30항 내지 제42항 중 어느 한 항에 있어서, 상기 인트론 서열은 MVM 서열 또는 HBB2 서열을 포함하는, 재조합 AAV 벡터.

청구항 44

제30항 내지 제43항 중 어느 한 항에 있어서, 상기 ITR은 삽입, 결실 또는 치환을 포함하거나, 임의의 하나 이상의 CpG 섬이 제거되는, 재조합 AAV 벡터.

청구항 45

제44항에 있어서, 상기 재조합 AAV 벡터는 AAVXL32 또는 AAVXL32.1, 또는 AAV8 또는 적어도 하나의 AAV8 캡시

드 단백질을 포함하는 반수체 AAV8 벡터인, 재조합 AAV 벡터.

청구항 46

제30항 내지 제45항 중 어느 한 항에 있어서,

a. 상기 분비 신호 펩타이드를 암호화하는 핵산은

AAT 신호 펩타이드(예를 들어, 서열 번호 17) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 17-22와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산;

피브로넥틴 신호 펩타이드(FN1) (예를 들어, 서열 번호 18-21) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 18-21과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산;

동족 GAA 신호 펩타이드(서열 번호 175) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 175와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산;

hIGF2 신호 펩타이드(예를 들어, 서열 번호 22) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 22와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산;

IgG1 리더 펩타이드(서열 번호 177) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 177과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산;

wt IL2 리더 펩타이드(서열 번호 179) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 179와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산;

돌연변이 IL2 리더 펩타이드(서열 번호 181) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 181과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산

으로 이루어진 군 중 어느 하나로부터 선택되고,

b. 상기 GAA 폴리펩타이드를 암호화하는 핵산은 서열 번호 11, 서열 번호 72 또는 서열 번호 182, 또는 서열 번호 11, 서열 번호 72, 또는 서열 번호 182와 적어도 60%, 또는 70%, 또는 80%, 85% 또는 90% 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열로 이루어진 군 중 어느 하나로부터 선택되거나, 서열 번호 10의 H201L, H199R 또는 R233H로부터 선택된 적어도 1개, 적어도 2개 또는 적어도 3개 모두의 아미노산 변형을 갖는 GAA 폴리펩타이드를 암호화하는 핵산 서열인, 재조합 AAV 벡터.

청구항 47

제30항 내지 제46항 중 어느 한 항에 있어서, 상기 IGF2 표적화 펩타이드는 서열 번호 5, 서열 번호 6, 서열 번호 7, 서열 번호 8 또는 서열 번호 9 중 어느 하나로부터 선택되는, 재조합 AAV 벡터.

청구항 48

제30항 내지 제47항 중 어느 한 항에 있어서, 상기 IGF2 표적화 펩타이드는 서열 번호 8 또는 서열 번호 9 또는 이와 적어도 85% 서열 동일성을 갖는 기능적 변이체인, 재조합 AAV 벡터.

청구항 49

제1항 내지 제48항 중 어느 한 항의 재조합 AAV 벡터를 약제학적으로 허용되는 담체에 포함하는 약제학적 조성물.

청구항 50

핵산 서열로서,

GAA 폴리펩타이드를 암호화하는 핵산 서열에 작동적으로 연결된 간 특이적 프로모터를 포함하며, 여기서, 상기 간 특이적 프로모터는

- i. CRM_SP0412(서열 번호 86) 또는 SP0412(서열 번호 91), 또는 서열 번호 86 또는 서열 번호 91에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편, 또는
 - ii. SP0422(서열 번호 92), 또는 서열 번호 92에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편,
 - iii. CRM_SP0239(서열 번호 87) 또는 SP0239(서열 번호 93) 또는 SP0238-UTR(서열 번호 147), 또는 서열 번호 87, 서열 번호 93 또는 서열 번호 147에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;
 - iv. CRM_SP0265(SP0131_A1) (서열 번호 88) 또는 SP0265(LVR_SP0131_A1) (서열 번호 94) 또는 SP0265-UTR(서열 번호 146), 또는 서열 번호 88, 서열 번호 94 또는 서열 번호 146에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;
 - v. CRM_SP0240(서열 번호 89) 또는 SP0240(서열 번호 95) 또는 SP0240-UTR(서열 번호 148), 또는 서열 번호 89, 서열 번호 95 또는 서열 번호 148에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편; 또는
 - vi. CRM_SP0246(서열 번호 90) 또는 SP0246(서열 번호 96) 또는 SP0246-UTR(서열 번호 149), 또는 서열 번호 90, 서열 번호 96 또는 서열 번호 149에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편
- 중 어느 하나로부터 선택되는, 핵산 서열.

청구항 51

제조합 아데노바이러스 관련(rAAV) 벡터 계통을 위한 핵산 서열로서,

- a. 5' 및 3' AAV 역 말단 반복체(ITR) 핵산 서열 및
- b. 5' ITR 서열과 3' ITR 서열 사이에 위치하는, 분비 신호 펩타이드 및 알파-글루코시다아제(GAA) 폴리펩타이드를 포함하는 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서, 상기 이중 핵산 서열은 간-특이적 프로모터에 작동적으로 연결되고, 상기 간 특이적 프로모터는 다음의 간 특이적 프로모터 중 어느 하나로부터 선택된 핵산 서열:
- i. SP0422(서열 번호 92), 또는 서열 번호 92에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편,
- ii. CRM_SP0239(서열 번호 87) 또는 SP0239(서열 번호 93) 또는 SP0238-UTR(서열 번호 147), 또는 서열 번호 87, 서열 번호 93 또는 서열 번호 147에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;
- iii. CRM_SP0265(SP0131_A1) (서열 번호 88) 또는 SP0265(LVR_SP0131_A1) (서열 번호 94) 또는 SP0265-UTR(서열 번호 146), 또는 서열 번호 88, 서열 번호 94 또는 서열 번호 146에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;
- iv. CRM_SP0240(서열 번호 89) 또는 SP0240(서열 번호 95) 또는 SP0240-UTR(서열 번호 148), 또는 서열 번호 89, 서열 번호 95 또는 서열 번호 148에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편; 또는
- v. CRM_SP0246(서열 번호 90) 또는 SP0246(서열 번호 96) 또는 SP0246-UTR(서열 번호 149), 또는 서열 번호 90, 서열 번호 96 또는 서열 번호 149에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편.

청구항 52

제50항 또는 제51항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 이중 핵산 서열은 상기 분비 신호 펩타이드와 상기 알파-글루코시다아제(GAA) 폴리펩타이드 사이에 위치한 IGF2 표적화 펩타이드를 추가로 포함하는, 핵산 서열.

청구항 53

제50항 내지 제52항 중 어느 한 항에 있어서, 상기 분비 신호를 암호화하는 핵산은 서열 번호 17, 22-26, 177, 179, 181 중 어느 하나, 또는 이와 적어도 85% 서열 동일성을 갖는 핵산으로부터 선택되는, 핵산 서열.

청구항 54

제50항에 있어서, 상기 IGF2 표적화 펩타이드를 암호화하는 핵산은 서열 번호 2(IGF2- Δ 2-7), 서열 번호 3(IGF2- Δ 1-7), 또는 서열 번호 4(IGF2 V43M) 중 어느 하나, 또는 이와 적어도 85% 서열 동일성을 갖는 핵산으로부터 선택되는, 핵산 서열.

청구항 55

제50항 내지 제54항 중 어느 한 항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 핵산 서열은 인간 GAA 유전자 또는 인간 코돈 최적화된 GAA 유전자(coGAA) 또는 변형된 GAA 핵산 서열인, 핵산 서열.

청구항 56

제55항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 핵산 서열은 (i) 생체내 향상된 발현을 위한 코돈 최적화, (ii) CpG 섬 감소, (iii) 정지 서열의 변형, (iv) 대안적인 리딩 프레임의 감소, 및 (v) 선천성 면역 반응 감소 중 어느 하나를 위해 서열 번호 11로부터 변형되는, 핵산 서열.

청구항 57

제55항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 핵산 서열은 CpG 섬을 감소시키거나 선천성 면역 반응을 감소시키거나, 또는, CpG 섬을 감소시키고 선천성 면역 반응을 감소시키거나, 및/또는 생체내 향상된 발현을 위해 코돈 최적화된, 핵산 서열.

청구항 58

제50항 내지 제57항 중 어느 한 항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 핵산 서열은 생체내 향상된 발현을 위해 코돈 최적화된, 핵산 서열.

청구항 59

제50항 내지 제58항 중 어느 한 항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 핵산은 서열 번호 11(전장 hGAA), 서열 번호 55(Dwight cDNA), 서열 번호 56(hGAA Δ 1-66), 또는 서열 번호 82(mod_hGAA) 또는 서열 번호 182 중 어느 하나, 이와 적어도 80%, 85%, 90%, 95% 또는 98% 동일성을 갖는 핵산 서열로부터 선택되는, 핵산 서열.

청구항 60

제50항 내지 제58항 중 어느 한 항에 있어서, 상기 GAA 폴리펩타이드를 암호화하는 핵산은 서열 번호 74(코돈 최적화된 1), 서열 번호 75(코돈 최적화된 2), 서열 번호 76(코돈 최적화된 3), 및 서열 번호 82(mod_hGAA) 중 어느 하나, 이와 적어도 80%, 85%, 90%, 95% 또는 98% 동일성을 갖는 핵산 서열로부터 선택되는, 핵산 서열.

청구항 61

제50항 내지 제58항 중 어느 한 항에 있어서, 상기 핵산은 서열 번호 10의 H201L, H199R 또는 R233H로부터 선택된 적어도 1개, 적어도 2개 또는 적어도 3개 모두의 아미노산 변형을 포함하는 GAA 폴리펩타이드를 암호화하는, 핵산 서열.

청구항 62

글리코젠 축적 질환 II형(GSD II, 폼페병, 산성 말타아제 결핍증)이 있거나, 알파-글루코시다아제(GAA) 폴리펩타이드가 결핍된 대상체를 치료하는 방법으로서, 상기 방법은 상기 재조합 AAV 벡터, 또는 상기 rAAV 계능 또는 전술한 청구항 중 어느 한 항의 핵산 서열 중 어느 하나를 상기 대상체에게 투여하는 단계를 포함하는 방법.

청구항 63

제62항에 있어서, GAA 폴리펩타이드는 대상체의 간으로부터 분비되고, 골격근 조직, 심장 근육 조직, 횡격막 근육 조직 또는 이들의 조합에 의해 분비된 GAA의 흡수가 있고, 여기서, 상기 분비된 GAA의 흡수는 조직(들)에서 리소좀 글리코겐 저장의 감소를 초래하는, 방법.

청구항 64

제62항에 있어서, 상기 대상체에 투여하는 단계가 근육내, 피하, 척수내, 수조내, 척수강내, 정맥내 투여 중 어느 하나로부터 선택되는, 방법.

청구항 65

제62항에 있어서, 상기 재조합 AAV 벡터가 키메라 AAV 벡터, 반수체 AAV 벡터, 하이브리드 AAV 벡터 또는 배수체 AAV 벡터인, 방법.

청구항 66

제62항에 있어서, 상기 재조합 AAV 벡터가 합리적인 반수체 벡터, 모자이크 AAV 벡터, 화학적으로 변형된 AAV 벡터, 또는 임의의 AAV 혈청형으로부터의 AAV 벡터인, 방법.

청구항 67

제62항에 있어서, 상기 재조합 AAV 벡터는 AAVXL32 벡터 또는 AAVXL32.1 벡터 또는 AAV8 벡터, 또는 적어도 하나의 AAV8 캡시드 단백질을 포함하는 반수체 AAV8 벡터인, 방법.

청구항 68

제62항에 있어서, 상기 재조합 AAV 벡터는 AAV8 벡터인, 방법.

청구항 69

리소좀 축적 질환(LSD)이 있는 대상체를 치료하는 방법으로서, 상기 방법은 재조합 AAV 벡터, 또는 rAAV 게놈 또는 전술한 청구항 중 어느 한 항의 핵산 서열 중 어느 하나를 상기 대상체에게 투여하는 단계를 포함하고, 여기서, 상기 AAV 벡터는 표 5B 또는 표 6B에서 임의의 폴리펩타이드로부터 선택된 폴리펩타이드를 발현하는, 방법.

청구항 70

제69항에 있어서, 상기 리소좀 축적 질환(LSD)은 표 5A 또는 표 6A에 나열된 것들 중 어느 하나로부터 선택되는, 방법.

청구항 71

제69항에 있어서, 상기 재조합 AAV 벡터는 키메라 AAV 벡터, 반수체 AAV 벡터, 하이브리드 AAV 벡터 또는 배수체 AAV 벡터인, 방법.

청구항 72

제69항에 있어서, 상기 재조합 AAV 벡터는 합리적인 반수체 벡터, 모자이크 AAV 벡터, 화학적으로 변형된 AAV 벡터, 또는 임의의 AAV 혈청형으로부터의 AAV 벡터인, 방법.

청구항 73

제69항에 있어서, 상기 재조합 AAV 벡터는 AAVXL32 벡터 또는 AAVXL32.1 벡터 또는 AAV8 벡터, 또는 적어도 하나의 AAV8 캡시드 단백질을 포함하는 반수체 AAV8 벡터인, 방법.

발명의 설명

기술 분야

서열 목록

[0001]

[0002] 본 발명은 35 U.S.C. § 119(e) 하에 2019년 11월 19일에 출원된 미국 가출원 62/937,556 및 2019년 11월 19일에 출원된 미국 가출원 62/937,583 및 2020년 5월 12일에 출원된 미국 가출원 63/023,570의 이익을 주장하고, 각각의 내용은 그 전체가 본원에 참조로 포함된다.

[0003] 서열 목록

[0004] 본 출원은 ASCII 형식으로 전자적으로 제출된 서열 목록을 포함하며 그 전체가 본원에 참조로 포함된다. 2020년 11월 17일에 생성된 상기 ASCII 사본의 명칭은 046192-096600WOPT SL.txt이며 크기는 840,179 바이트이다.

[0005] 발명의 분야

[0006] 본 발명은, 예를 들어, 알파-글루코시다아제(GAA) 폴리펩타이드와 같은 리소좀 효소의 표적화된 전위를 위한 아데노-관련 바이러스(AAV) 입자, 비리온 및 벡터, 및 예를 들어, 폼페병과 같은 리소좀 축적 질환 및 장애의 치료를 위한 사용 방법에 관한 것이다.

배경 기술

[0007] 40개 초과 리소좀 축적 질환(LSD)은 리소좀에 하나 이상의 리소좀 효소의 부재에 의해 직간접적으로 야기된다. LSD에 대한 효소 대체 요법이 활발히 추진되고 있다. 치료는 일반적으로 LSD 단백질이 흡수되어 M6P-의존적 방식으로 다양한 세포 유형의 리소좀으로 전달되어야 한다. 하나의 가능한 접근법은 LSD 단백질을 정제하고 탄수화물 모이어티를 M6P와 통합하도록 변형하는 것을 포함한다. 이 변형된 물질은 세포 표면 상의 M6P 수용체와의 상호작용으로 인해 변형되지 않은 LSD 단백질보다 더 효율적으로 세포에 흡수될 수 있다.

[0008] 효소 대체 요법에 대한 대안 또는 보조로서, GSD-II를 치료하기 위한 유전자 요법 접근법의 가능성이 조사되었다 (Amalfitano, A., et al., (1999) *Proc. Natl. Acad. Sci. USA* 96:8861-8866, Ding, E., et al. (2002) *Mol. Ther.* 5:436-446, Fraites, T. J., et al., (2002) *Mol. Ther.* 5:571-578, Tsujino, S., et al. (1998) *Hum. Gene Ther.* 9:1609-1616).

[0009] 그러나, 유전자, 특히 리소좀 축적 질환의 치료를 위한 리소좀 단백질 및 효소의 바이러스 또는 AAV 전달은 어려움이 있다. 일반적으로, 포유동물의 리소좀 효소는 세포질에서 합성되어 ER을 가로지르며 이들은 N-연결된 고 만노오스 유형 탄수화물로 글리코실화된다. 골지체에서, 고 만노오스 탄수화물은 리소좀에 이러한 단백질을 표적으로 하는 만노오스-6-포스페이트(M6P)를 추가함으로써 리소좀 단백질에 변형된다. M6P-변형된 단백질은 2개의 M6P 수용체 중 하나와의 상호작용을 통해 리소좀으로 전달된다. 그러나, 효소 대체 요법에 사용되는 재조합적으로 생성된 단백질은 종종 리소좀에 이들을 표적화하는 데 필요한 M6P의 추가가 부족하므로, 종종 환자에게 투여될 재조합적으로 생성된 효소의 고용량 및/또는 빈번한 주입을 필요로 한다.

[0010] 산 알파-글루코시다아제(GAA)는 글리코겐의 외부 가지를 포함하여 말토스 및 기타 선형 올리고사카라이드에서 알파 1-4 연결을 가수분해하여 리소좀에서 과잉 글리코겐을 분해하는 리소좀 효소이다(Hirschhorn et al. (2001) in *The Metabolic and Molecular Basis of Inherited Disease*, Scriver, et al., eds. (2001), McGraw-Hill: New York, p. 3389-3420). 다른 포유동물의 리소좀 효소와 마찬가지로, GAA는 세포질에서 합성되어 ER을 가로지르며 이는 N-연결된 고 만노오스 유형 탄수화물로 글리코실화된다. 골지체에서, 고 만노오스 탄수화물은 리소좀에 이러한 단백질을 표적으로 하는 만노오스-6-포스페이트(M6P)를 추가함으로써 리소좀 단백질에 변형된다. M6P-변형된 단백질은 2개의 M6P 수용체 중 하나와의 상호작용을 통해 리소좀으로 전달된다. 변형의 가장 유리한 형태는 2개의 M6P가 고 만노오스 탄수화물에 첨가되는 경우이다.

[0011] 리소좀에서 불충분한 GAA 활성은 산성 말타아제 결핍증(AMD), 글리코겐 축적 질환 II형(GSDII), 글리코겐증 II형, 또는 GAA 결핍증으로도 알려진 질환인 폼페병을 초래한다. 감소된 효소 활성은 GAA를 암호화하는 유전자에서 다양한 미스센스 및 넌센스 돌연변이로 인해 발생한다. 결과적으로, 글리코겐은 폼페병을 가진 환자의 모든 세포의 리소좀에 축적된다. 특히, 글리코겐 축적은 심장 및 골격근, 간, 및 기타 조직의 리소좀에서 가장 두드러진다. 축적된 글리코겐은 궁극적으로 근육 기능을 손상시킨다. 가장 심각한 형태의 폼페병에서는 심장 호흡 부전으로 인해 2세 이전에 사망한다.

[0012] 폼페병의 효과적인 치료가 필요하다. 폼페에 대한 효소 대체 치료제는 대상체의 근육 및 간 세포에 의해 투여되고 흡수될 재조합 GAA 단백질을 필요로 하며 이는 이후 M6P-의존적 방식으로 이들 세포에서 리소좀으로 수송된다. 그러나, 효소 대체 요법은 종종 영아 GSD II에 대해 합리적인 효능을 입증했지만, GAA 효소 대체 요법의 이점은 빈번한 주입의 필요성뿐만 아니라 재조합 hGAA 단백질에 대한 억제제 또는 중화 항체를 개발하는 대상체에 의해

제한된다(Amalfitano, A., et al. (2001) *Genet. In Med.* 3:132-138).

- [0013] 유전자 요법은 유전 장애를 치료할 뿐만 아니라 바이러스를 사용하여 후천성 및 퇴행성 질환의 장기간 비침습적 치료를 촉진할 수 있는 잠재력을 가지고 있다. 하나의 유전자 치료 벡터는 아데노-관련 바이러스(AAV)이다. AAV 자체는 효율적인 복제를 위해 헬퍼 바이러스가 필요한 비병원성 의존적 파르보바이러스이다. AAV는 이의 안전성과 단순성 때문에 유전자 요법을 위한 바이러스 벡터로 활용되어 왔다. AAV는 분열 세포 및 비분열 세포 둘 모두를 형질도입할 수 있는 광범위한 숙주 및 세포 유형 친화성(tropism)을 갖는다.
- [0014] 그러나, GAA 폴리펩타이드의 AAV 전달은 간에서의 충분한 발현 및/또는 혈당증을 경험했다고 보고된 환자의 리소좀으로의 전달과 관련하여 몇몇 어려움이 있다.
- [0015] 특히, 인간 대상체에서, GAA 폴리펩타이드를 암호화하는 rAAV 벡터의 투여는 저혈당증을 경험하거나 세포의 비특이적 업데이트로 인해 고혈당이 되는 다수의 환자를 초래하였다[예를 들어, 문헌(Byrne *et al.*, A study on the safety and efficacy of Reveglucosidase alfa in patients with late-onset Pompe disease; Orphanet J. of Rare diseases; 2017; 12: 144) 참조].
- [0016] 따라서, 예를 들어, GAA의 변형을 포함한 리소좀 폴리펩타이드 결핍을 치료하기 위해 시험관내 및 생체내에서 GAA와 같은 리소좀 폴리펩타이드를 생성하는 개선된 방법이 당업계에 필요하다. 더욱이, GAA 폴리펩타이드의 과발현으로부터 임의의 부작용을 감소시키고 저혈당증 위험을 감소시키는 데 도움이 되도록 간으로부터 분비를 개선할 뿐만 아니라 리소좀으로의 GAA의 표적화를 개선할 필요가 있다. 또한, 영향을 받는 조직 및 기관에 GAA 및 기타 리소좀 폴리펩타이드를 전신적으로 전달하는 방법이 필요하다. 특히, 임의의 잠재적인 부작용을 감소시키면서 대상체에게 GAA 단백질을 투여하고 환자 리소좀에 GAA 단백질을 표적화하기 위한 보다 효율적인 방법에 대한 필요성이 남아 있다.

발명의 내용

- [0017] 본원에 기재된 기술은 일반적으로 리소좀 축적 질환 및 장애, 예를 들어, 그러나 이에 제한되지 않는 폼페병을 치료하기 위한 유전자 요법 작제물, 방법 및 조성물에 관한 것이다. 보다 특히, 상기 기술은 리소좀 효소, 예를 들어, GAA 폴리펩타이드를 대상체에게 전달하도록, 보다 특히, 대상체의 간에 리소좀 효소, 예를 들어, 리소좀을 표적으로 하고 간 세포로부터 분비되는 GAA 폴리펩타이드를 전달하도록 구성된 아데노-연관(AAV) 비리온에 관한 것이다.
- [0018] 특히, 본원의 표 5A 또는 표 6A에 나열된 것들과 같은 역 말단 반복체(ITR)를 포함하는 뉴클레오타이드 서열, 프로모터, 이중 유전자, 폴리-A 테일 및 리소좀 축적 질환을 치료하는데 사용하기 위한 잠재적으로 다른 조절 요소를 포함하는, 예를 들어, 예시적인 예로서 rAAV 벡터를 사용하는 표적화된 바이러스 벡터가 본원에 기재되어 있으며, 여기서, 이중 유전자는, 예를 들어, GAA와 같은 리소좀 효소이고, 여기서, 벡터, 예를 들어, rAAV는 이중 리소좀 효소 유전자의 발현 및 질환, 예를 들어, 폼페병의 치료를 위해 적절한 조직 및/또는 기관에 전달되는 치료학적 유효량으로 환자에게 투여될 수 있다.
- [0019] 본 발명의 양태는 이하에 설명되는 예시적인 이점을 발생시키는 구성 및 사용에서의 특정 이점을 교시한다.
- [0020] 따라서, 본원에 기재된 특정 구현에는 역 말단 반복체(ITR)를 포함하고 ITR 사이에 위치하는 뉴클레오타이드 서열, 간 특이적 프로모터(LSP), 산 알파-글루코시다아제(GAA) 단백질을 암호화하는 이중 핵산 서열, 폴리-A 테일 및 폼페병을 치료하는데 사용하기 위한 잠재적으로 다른 조절 요소를 포함하는 rAAV 벡터이며, 여기서, GAA 단백질을 발현하는 rAAV는 폼페병을 가진 대상체의 치료를 위해 GAA 단백질을 암호화하는 이중 유전자의 발현을 위해 적절한 조직 및/또는 기관에 전달되는 치료학적 유효량으로 환자에게 투여될 수 있다.
- [0021] 보다 구체적으로, AAV 비리온 또는 게놈은 본원의 표 4에 나열된 임의의 프로모터, 또는 이의 기능적 변이체 또는 기능적 단편, 또는 리소좀 단백질, 예를 들어, GAA 단백질이 간에서 우선적으로 발현될 수 있게 하는 서열 번호 86, 91-96 또는 146-150으로부터 선택된 임의의 LSP, 또는 이의 기능적 변이체 또는 기능적 단편으로부터 선택된 LSP를 포함한다. 일부 구현예에서, 간-특이적 프로모터는 간에서 hGAA 단백질을 우선적으로 발현하는 반면, 또한 관심 있는 또 다른 조직, 예를 들어, 근육, 또는 CNS, 또는 근육 및 CNS 조직에서 hGAA를 어느 정도 발현할 수 있다. 일부 구현예에서, 발현된 리소좀 효소, 예를 들어, GAA 단백질은 GAA 단백질을 리소좀으로 표적화하고/하거나 신호 펩타이드(SP)와 융합하는 본원에 개시된 IGF2 표적화 펩타이드와 같은 표적화 서열을 갖는 GAA-융합 단백질로서 구성될 수 있고, GAA 단백질은 간에서 rAAV 게놈에 의해 발현되며, 이는 분비되어 포유동물 세포, 특히 근육 세포의 리소좀에 의해 흡수된다.

- [0022] 본원에 기재된 조성물 및 방법의 일부 구현예에서, 본원에 개시된 rAAV 벡터는 그 게놈에 5' 및 3' AAV 역 말단 반복체(ITR) 서열, 및 5' ITR과 3' ITR 사이에 위치하는, 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 이중 핵산 서열에 작동적으로 연결된 간 특이적 프로모터(LSP)를 포함하며, 여기서, 간-특이적 프로모터(LSP)는 서열 번호 86(CRM 0412), 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422), 서열 번호 93(SP0239), 서열 번호 94(SP0265, SP131_A1로도 지칭됨), 서열 번호 95(SP0240) 또는 서열 번호 96(SP0246), 또는 서열 번호 146(SP0265-UTR), 서열 번호 147(SP0239-UTR), 서열 번호 148(SP0240-UTR), 서열 번호 149(SP0246-UTR) 또는 서열 번호 150(SP0131-A1-UTR)으로부터 나열된 임의의 프로모터 또는 이의 기능적 단편 또는 변이체 또는 서열 번호 270-341 또는 342-430으로부터 선택된 임의의 LSP 또는 이의 기능적 단편 또는 변이체로부터 선택된 핵산 서열을 포함한다. 일부 구현예에서, GAA 폴리펩타이드는 IGF2 표적화 서열 또는 신호 서열에 융합되지 않는다. 일부 구현예에서, GAA 폴리펩타이드는 본원에 개시된 바와 같은 신호 서열 및/또는 본원에 개시된 바와 같은 IGF2 표적화 서열에 융합된다.
- [0023] 본원에 기재된 조성물 및 방법의 일부 구현예에서, 본원에 개시된 rAAV 벡터는 그 게놈에 5' 및 3' AAV 역 말단 반복체(ITR) 서열, 및 5' ITR과 3' ITR 사이에 위치하는, (i) 분비 신호 펩타이드 및/또는 IGF2 표적화 펩타이드; 및 (ii) 알파-글루코시다아제(GAA) 폴리펩타이드를 포함하는 융합 폴리펩타이드를 암호화하는 이중 핵산 서열에 작동적으로 연결된 간 특이적 프로모터(LSP)를 포함하고, 여기서, 간-특이적 프로모터(LSP)는 본원의 표 4에 나열된 임의의 프로모터 또는 이의 기능적 변이체 또는 기능적 단편, 또는 서열 번호 86, 91-96, 또는 146-150으로부터 선택된 임의의 LSP 또는 이의 기능적 변이체 또는 기능적 단편으로부터 선택된다.
- [0024] 일부 구현예에서, 본원에 개시된 rAAV 벡터는 그 게놈에 5' 및 3' AAV 역 말단 반복체(ITR) 서열, 및 5' ITR과 3' ITR 사이에 위치하는, (i) 분비 신호 펩타이드(리더 펩타이드로도 지칭됨); 및 (ii) 알파-글루코시다아제(GAA) 폴리펩타이드를 포함하는 융합 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서, 이중 핵산은 본원의 표 4에 나열된 임의의 프로모터 또는 이의 기능적 변이체 또는 기능적 단편, 또는 서열 번호 86, 91-96, 또는 146-150으로부터 선택된 임의의 LSP 또는 이의 기능적 변이체 또는 기능적 단편 또는 본원의 표 4로부터 선택된 임의의 LSP 또는 이의 기능적 변이체 또는 기능적 단편으로부터 선택된 간-특이적 프로모터(LSP)에 작동적으로 연결된다. 예시적인 리더 서열은 본원에 기재된 바와 같은 고유 GAA 리더 서열, AAT 서열, IL2(1-3), IL2 리더 서열(IL2 wt), 변형된 IL2 리더 서열(IL2 mut), 피브로넥틴(FN1) 신호 서열, 또는 IgG 리더 서열 또는 이의 기능적 변이체를 포함하지만 이에 제한되지 않는다. 일부 구현예에서, AAV 벡터는 LSP와 리더 서열 사이에 위치한 코작(Kozak) 서열을 포함한다.
- [0025] 일부 구현예에서, 본원에 개시된 rAAV 벡터는 그 게놈에 5' 및 3' AAV 역 말단 반복체(ITR) 서열, 및 5' ITR과 3' ITR 사이에 위치하는, (i) IGF2 표적화 펩타이드, 및 (ii) 알파-글루코시다아제(GAA) 폴리펩타이드를 포함하는 융합 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서, 이중 핵산은 본원의 표 4에 나열된 임의의 프로모터 또는 이의 기능적 변이체 또는 기능적 단편, 또는 서열 번호 86, 91-96, 또는 146-150으로부터 선택된 임의의 LSP 또는 이의 기능적 변이체 또는 기능적 단편으로부터 선택된 간-특이적 프로모터(LSP)에 작동적으로 연결된다.
- [0026] 추가의 구현예에서, 본원에 개시된 rAAV 벡터는 그 게놈에 5' 및 3' AAV 역 말단 반복체(ITR) 서열, 및 5' ITR과 3' ITR 사이에 위치하는, 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하고 (즉, GAA 폴리펩타이드가 이중 신호 펩타이드 (또는 리더 서열)에 융합되지 않거나, 또는 본원에 기재된 바와 같은 IGF2 표적화 서열에 융합되지 않은 경우), 여기서, 이중 핵산은 본원의 표 4에 나열된 임의의 프로모터 또는 이의 기능적 변이체 또는 기능적 단편, 또는 서열 번호 86, 91-96, 또는 146-150으로부터 선택된 임의의 LSP 또는 이의 기능적 변이체 또는 기능적 단편으로부터 선택된 간-특이적 프로모터(LSP)에 작동적으로 연결된다.
- [0027] 일부 구현예에서, rAAV 벡터는 간 특이적 캡시드, 예를 들어, 그 전체가 본원에 참조로 포함된 W02019/241324에 개시된 바와 같은 XL32 및 XL32.1로부터 선택되는 간 특이적 캡시드를 포함한다. 일부 구현예에서, rAAV 벡터는 그 전체가 본원에 참조로 포함된 W02019/241324에 개시된 바와 같은 AAVXL32 또는 AAVXL32.1, 또는 AAV8 벡터, 또는 적어도 하나의 AAV8 캡시드 단백질(예를 들어, VP1, VP2, 또는 VP3 중 적어도 하나는 AAV8 혈청형으로부터 유래됨)을 포함하는 반수체 AAV 벡터이고, 일부 구현예에서, AAV 벡터는 적어도 2개의 AAV8 캡시드 단백질을 포함하는 반수체 AAV 벡터이다. 일부 구현예에서, AAV 벡터는 그 전체가 본원에 참조로 포함된 W02019241324A1 또는 국제 특허 출원 PCT/US2019/036676에 기재된 캡시드를 포함한다. 일부 구현예에서, AAV 벡터는 W02019241324A1에 개시된 바와 같은 서열 번호 1-3 중 어느 하나의 뉴클레오타이드 서열; 또는 (b) W02019241324A1에 개시된 바와 같은 서열 번호 4-6 중 어느 하나를 암호화하는 뉴클레오타이드 서열과 적어도 90% 동일한 핵산 AAV 캡시드 암호화 서열에 의해 암호화된 캡시드를 포함한다. 일부 구현예에서, AAV 캡시드는

AAV 벡터 게놈 및 본 발명의 AAV 캡시드를 포함하는 AAV 입자와 함께 W02019241324A1에 개시된 바와 같은 서열 번호 4-6 중 어느 하나와 적어도 90% 동일한 아미노산 서열을 포함한다.

[0028] 일부 구현예에서, rAAV 벡터는 AAV 벡터가 간 세포를 형질도입하도록 캡시드 단백질을 포함하고, 일부 구현예에서, AAV 벡터가 근육 및 간 세포를 형질도입하도록 캡시드 단백질을 포함한다.

[0029] 방법 및 조성물에 사용하기 위한 예시적인 LSP는 SP0412(서열 번호 91) 또는 이의 기능적 변이체이다. 대안적인 구현예에서, LSP는 서열 번호 86(CRM 0412), 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422), 서열 번호 93(SP0239), 서열 번호 94(SP0265, SP131_A1로도 지칭됨), 서열 번호 95(SP0240) 또는 서열 번호 96(SP0246), 또는 서열 번호 146(SP0265-UTR), 서열 번호 147(SP0239-UTR), 서열 번호 148(SP0240-UTR), 서열 번호 149(SP0246-UTR) 또는 서열 번호 150(SP0131-A1-UTR) 중 어느 하나 또는 이의 기능적 단편 또는 변이체로부터 선택될 수 있다.

[0030] 본원에 기재된 조성물 및 방법의 일부 구현예에서, 분비 신호 펩타이드는 AAT 신호 펩타이드, 피브로넥틴 신호 펩타이드(FN1), GAA 신호 펩타이드, 고유 GAA 리더 서열, AAT 서열, IL2(1-3), IL2 리더 서열(IL2 wt), 변형된 IL2 리더 서열(IL2 mut), 또는 IgG 리더 서열 중 어느 하나 또는 분비 신호 활성을 갖는 이의 기능적 변이체로부터 선택된다.

[0031] 본원에 기재된 조성물 및 방법의 일부 구현예에서, 알파-글루코시다아제(GAA) 폴리펩타이드는 GAA 폴리펩타이드의 N-말단에서 IGF2 표적화 펩타이드에 연결된다. 일부 구현예에서, IGF2 표적화 펩타이드는 인간 산 알파-글루코시다아제(GAA) 폴리펩타이드(서열 번호 10) (즉, 인간 산 알파-글루코시다아제(GAA) 폴리펩타이드의 잔기 70-952의 N-말단에 연결됨) 또는 서열 번호 10의 아미노산 70-952와 적어도 85% 서열 동일성을 갖는 GAA 폴리펩타이드의 아미노산 70에서 N-말단에 연결된다. 대안적인 구현예에서, IGF2 표적화 펩타이드는 인간 산 알파-글루코시다아제(GAA) 폴리펩타이드(서열 번호 10) (즉, 인간 산 알파-글루코시다아제(GAA) 폴리펩타이드의 잔기 40-952의 N-말단에 연결됨) 또는 서열 번호 10의 아미노산 40-952와 적어도 85% 서열 동일성을 갖는 GAA 폴리펩타이드의 아미노산 40에서 N-말단에 연결된다. 본원에 기재된 조성물 및 방법의 일부 구현예에서, GAA 폴리펩타이드는 야생형 GAA 핵산 서열(예를 들어, 서열 번호 11 또는 서열 번호 72)에 의해 암호화되거나, 예를 들어, 생체내 발현 증가, CpG 섬 감소 및/또는 대상체에서 선천성 면역 반응 감소 중 어느 하나를 위한 코돈 최적화된 GAA 핵산 서열일 수 있다. 예시적인 코돈 최적화된 GAA 핵산 서열은 서열 번호 73, 서열 번호 74, 서열 번호 75, 서열 번호 76 및 서열 번호 182를 포함하지만 이에 제한되지 않는다.

[0032] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 제조용 AAV 벡터는 간-특이적 프로모터(LSP), 예를 들어, 본원의 표 4의 어느 하나, 또는 이의 기능적 변이체 또는 이의 기능적 변이체로부터 선택된 간 특이적 프로모터를 포함하지만 이에 제한되지 않는다. 방법 및 조성물에 사용하기 위해 포함된 예시적인 LSP는 SP0412 및 이의 기능적 변이체를 포함한다. 대안적인 구현예에서, LSP는 본원에 개시된 바와 같은 SP0422, SP0131A1, SP0239, SP0240 또는 SP0246 중 어느 하나, 또는 이의 기능적 변이체로부터 선택된 핵산 서열을 포함할 수 있다. 예를 들어, 간 특이적 프로모터는 서열 번호 86(CRM 0412), 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422) 중 어느 하나, 또는 이의 기능적 변이체 또는 기능적 단편으로부터 선택된 핵산 서열을 포함할 수 있다. 대안적인 구현예에서, 간 특이적 프로모터는 서열 번호 93(SP0239), 서열 번호 94(SP0265, SP131_A1로도 칭함), 서열 번호 95(SP0240) 또는 서열 번호 96(SP0246), 또는 서열 번호 146(SP0265-UTR), 서열 번호 147(SP0239-UTR), 서열 번호 148(SP0240-UTR), 서열 번호 149(SP0246-UTR), 또는 서열 번호 150(SP0131-A1-UTR) 중 어느 하나로부터 선택된 핵산 서열을 포함할 수 있다. 본원에 개시된 조성물 및 방법의 일부 구현예에서, 간-특이적 프로모터는, 그 전체가 본원에 참조로 포함된 가출원 62,937,556의 표 4A 또는 4B에 이전에 개시된 바와 같은 간-특이적 시스-조절 요소(CRE), 합성 간-특이적 시스-조절 모듈(CRM) 또는 서열 번호 270-341(CRM을 포함할 수 있는 최소 LSP) 또는 서열 번호 342-430(예시적인 합성 LSP) 중 어느 하나, 또는 이의 기능적 단편 또는 기능적 변이체로부터 선택된 프로모터 서열을 포함하는 합성 간-특이적 프로모터를 포함한다. 이러한 간-특이적 프로모터 요소는 최소 간-특이적 프로모터(예를 들어, 서열 번호 86, 270-341 참조) 또는 간-특이적 근위 프로모터(예를 들어, 서열 번호 91-96, 146-150 및 342-430 참조)를 포함할 수 있다. 예를 들어, 서열 번호 86(CRM 0412), 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422), 또는 이의 기능적 변이체 또는 기능적 단편이다.

[0033] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 제조용 AAV 벡터는 간-특이적 프로모터(LSP), 예를 들어, 서열 번호 86, 91-96, 146-150, 370-430 중 어느 하나, 또는 이의 기능적 변이체 또는 기능적 단편으로부터 선택된 간 특이적 프로모터를 포함하지만 이에 제한되지 않는다.

[0034] 예를 들어, 본원의 표 4에 개시된 간-특이적 프로모터의 기능적 변이체 또는 기능적 단편, 또는 서열 번호 86,

91-96, 또는 146-150, 또는 370-430 또는 이의 기능적 변이체 또는 기능적 단편으로부터 선택된 임의의 LSP는 원래의 비변형된 참조 서열과 적어도 약 75% 서열 동일성, 또는 적어도 약 80% 서열 동일성, 적어도 약 90% 서열 동일성, 적어도 약 95% 서열 동일성, 적어도 약 98% 서열 동일성을 갖고, 또한 상응하는 비변형된 프로모터 서열의 프로모터 활성의 적어도 35%, 또는 프로모터 활성의 적어도 약 45%, 또는 프로모터 활성의 적어도 약 50%, 또는 프로모터 활성의 적어도 약 60%, 또는 프로모터 활성의 적어도 약 75%, 또는 프로모터 활성의 적어도 약 80%, 또는 프로모터 활성의 적어도 약 85%, 또는 프로모터 활성의 적어도 약 90%, 또는 프로모터 활성의 적어도 약 95%를 갖는다.

[0035] 예를 들어, 서열 번호 92(SP0422) 또는 서열 번호 91(SP0412)의 기능적 변이체 또는 기능적 단편은 서열 번호 92 또는 서열 번호 91과 적어도 약 75% 서열 동일성, 또는 서열 번호 92 또는 서열 번호 91과 적어도 약 80% 서열 동일성, 서열 번호 92 또는 서열 번호 91과 적어도 약 90% 서열 동일성, 서열 번호 92 또는 서열 번호 91과 적어도 약 95% 서열 동일성, 서열 번호 92 또는 서열 번호 91, 또는 원래의 비변형된 서열과 적어도 약 98% 서열 동일성을 갖고, 또한 각각 서열 번호 92 또는 서열 번호 91의 상응하는 비변형된 프로모터 서열의 프로모터 활성의 적어도 35%, 또는 프로모터 활성의 적어도 약 45%, 또는 프로모터 활성의 적어도 약 50%, 또는 프로모터 활성의 적어도 약 60%, 또는 프로모터 활성의 적어도 약 75%, 또는 프로모터 활성의 적어도 약 80%, 또는 프로모터 활성의 적어도 약 85%, 또는 프로모터 활성의 적어도 약 90%, 또는 프로모터 활성의 적어도 약 95%를 갖는다.

[0036] 기능적 단편은 적어도 35%, 또는 적어도 약 45%, 또는 적어도 약 50%, 또는 적어도 약 75%, 또는 적어도 약 80%, 또는 적어도 약 85%, 또는 적어도 약 90%의 절단되지 않은(untrunkated) 프로모터를 갖는 프로모터의 일부이다. 일부 구현예에서, 기능적 단편은 비변형된 프로모터 서열의 인접 부분을 포함한다. TTR(서열 번호 431)이 예시적인 LSP로서 본원의 실시예에 개시되어 있지만, 당업자는 TTR 프로모터(서열 번호 431)를 본원의 표 4에 나열된 간-특이적 프로모터 중 임의의 하나 이상, 예를 들어, 적어도 서열 번호 92(SP0422) 또는 서열 번호 91(SP0412), 또는 서열 번호 92(SP0422) 또는 서열 번호 91(SP0412)의 기능적 변이체 또는 단편을 포함하는 핵산 서열, 또는 서열 번호 93(SP0239), 서열 번호 94(SP131_A1), 서열 번호 95(SP0240), 서열 번호 96(SP0246), 또는 서열 번호 146(SP0265-UTR), 서열 번호 147(SP0239-UTR), 서열 번호 148(SP0240-UTR), 서열 번호 149(SP0246-UTR) 또는 서열 번호 150(SP0131-A1-UTR) 중 어느 하나를 포함하는 핵산 서열, 또는 서열 번호 270-341 또는 342-430, 또는 이의 기능적 변이체 또는 기능적 단편으로 대체할 수 있다. 일부 구현예에서, LSP는 간에서 hGAA 단백질을 우선적으로 발현하지만, 또한 관심 있는 또 다른 조직, 예를 들어 근육, 또는 CNS, 또는 근육 및 CNS 조직에서 hGAA를 어느 정도 발현할 수 있다.

[0037] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 재조합 AAV 벡터는 본원에 개시된 바와 같은 야생형 GAA 폴리펩타이드(wtGAA) 또는 변형된 GAA 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서, GAA 폴리펩타이드의 하나 이상의 아미노산은 변형되는데, 예를 들어 H199R, R223H, H201L 변형이다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, 재조합 AAV 벡터는 인간 GAA 유전자 또는 인간 코돈 최적화된 GAA 유전자(coGAA)인 GAA 폴리펩타이드를 암호화하는 이중 핵산 서열, 또는 H199R, R223H, H201L로부터 선택된 하나 이상의 변형을 포함하는 변형된 GAA 폴리펩타이드를 암호화하는 코돈 최적화된 변형된 GAA 핵산 서열을 포함한다. 본원에 개시된 방법 및 방법의 모든 양태에서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 생체내 향상된 발현, CpG 섬 감소, 또는 선천성 면역 반응 감소 중 어느 하나 이상을 위해 코돈 최적화된다. 본원에 개시된 방법 및 방법의 모든 양태에서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 CpG 섬을 감소시키고 선천성 면역 반응을 감소시키도록 코돈 최적화된다. 일부 구현예에서, 야생형 GAA 폴리펩타이드를 암호화하는 핵산 서열은 서열 번호 182에 개시되고 본원에 기재된 바와 같은 변형을 포함한다.

[0038] 본원 기술의 또 다른 양태는 본원에 개시된 임의의 재조합 AAV 벡터 조성물 및 약제학적으로 허용되는 담체를 포함하는 약제학적 조성물에 관한 것이다.

[0039] 본원 기술의 또 다른 양태는 다음 순서로 다음을 포함하는 핵산 서열을 포함하는 조성물에 관한 것이다: 5' ITR, H199R, R223H, H201L, 및 3' ITR로부터 선택된 하나 이상의 변형을 포함하는 변형된 GAA 폴리펩타이드를 암호화하는 핵산을 포함하는 핵산 서열에 작동적으로 연결된 간 특이적 프로모터(LSP). 하나의 양태에서, 핵산 서열은 임의로, LSP와 GAA 폴리펩타이드를 암호화하는 핵산 사이에 위치한 리더 서열 (또는 신호 서열)을 암호화하는 핵산 서열을 추가로 포함하고, 여기서, 리더 서열은 본원에 개시된 바와 같은 고유 GAA 리더 서열, AAT 서열, IL2(1-3), IL2 리더 서열(IL2 wt), 변형된 IL2 리더 서열(IL2 mut), 피브로넥틴(FN1), 또는 IgG 리더 서열 중 어느 하나 또는 이의 기능적 변이체로부터 선택된다. 일부 구현예에서, 핵산 서열은 임의로, LSP와 리더 서열 사이에 위치한 코작 서열을 추가로 포함한다. 일부 구현예에서, 핵산 서열은 임의로, 리더 서열과 GAA 폴

리펩타이드를 암호화하는 핵산 사이에 위치하는 IGF2 표적화 펩타이드를 추가로 포함한다. 일부 구현예에서, 핵산 서열은 임의로, GAA 폴리펩타이드를 암호화하는 핵산 및 폴리A 서열의 3'에 위치한 3' UTR을 추가로 포함한다. 일부 구현예에서, 핵산 서열은 임의로, 바람직하게는 LSP와 코작 서열 사이에 LSP의 3'와 GAA 폴리펩타이드를 암호화하는 핵산의 5'에 위치하는 인트론 서열을 포함한다. rAAV 벡터 또는 rAAV 게놈에 대한 예시적인 작제물은 도 5a-5g에 나타낸다.

[0040] 본원 기술의 또 다른 양태는 5' ITR, H199R, R223H, H201L로부터 선택된 하나 이상의 변형을 포함하는 변형된 GAA 폴리펩타이드를 암호화하는 핵산 서열에 작동적으로 연결된 간 특이적 프로모터(LSP), 폴리A 서열 및 3' ITR 서열을 포함하는 핵산 서열을 포함하는 조성물에 관한 것이며, 여기서, 폴리A 서열은 전장 또는 절단된 폴리A 신호 서열일 수 있다. 본원 기술의 또 다른 양태는 5' ITR, H199R, R223H, H201L로부터 선택된 하나 이상의 변형을 포함하는 변형된 GAA 폴리펩타이드를 암호화하는 핵산 서열에 작동적으로 연결된 간 특이적 프로모터(LSP), 전장 폴리A 서열, 말단 반복 서열 및 3' ITR 서열을 포함하는 핵산 서열을 포함하는 조성물에 관한 것이며, 여기서, 핵산은 AAV P5 프로모터 서열이 결여되어 있다.

[0041] 본원 기술의 또 다른 양태는 (a) 분비 신호 펩타이드를 암호화하는 핵산, (b) IGF2 표적화 펩타이드를 암호화하는 핵산, 및 (c) GAA 폴리펩타이드를 암호화하는 핵산의 순서로, 핵산 서열에 작동적으로 연결된 간 특이적 프로모터(LSP)를 포함하는 핵산 서열을 포함하는 조성물에 관한 것이다.

[0042] 본원 기술의 또 다른 양태는 재조합 아데노바이러스 관련(rAAV) 벡터 게놈에 대한 핵산 서열을 포함하는 조성물에 관한 것이며, 상기 핵산 서열은 (a) 5' 및 3' AAV 역 말단 반복체(ITR) 핵산 서열, 및 (b) 5' ITR 서열과 3' ITR 서열 사이에 위치하는, 분비 신호 펩타이드 및 알파-글루코시다아제(GAA) 폴리펩타이드를 포함하는 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서, 이중 핵산은 상기 기재된 바와 같은 간-특이적 프로모터에 작동적으로 연결된다. 예시적인 간-특이적 프로모터는 SP0412 또는 SP0422 또는 이의 기능적 변이체이다. 일부 구현예에서, 본원에 개시된 바와 같은 방법 및 조성물에 사용하기 위한 간-특이적 프로모터는 본원의 표 4에 개시된 바와 같은 간-특이적 시스-조절 요소(CRE), 합성 간-특이적 시스-조절 모듈(CRM) 또는 합성 간-특이적 프로모터를 포함한다.

[0043] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 핵산 서열은 GAA 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서, 핵산 서열은 인간 GAA 유전자 또는 인간 코돈 최적화된 GAA 유전자(coGAA) 또는 변형된 GAA 핵산 서열이다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, 핵산 서열은 생체내 향상된 발현, CpG 섬 감소 또는 선천성 면역 반응 감소 중 임의의 하나 이상을 위해 코돈 최적화된(coGAA) GAA 유전자인 이중 핵산 서열을 포함한다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, 핵산 서열은 CpG 섬을 감소시키고 선천성 면역 반응을 감소시키기 위해 코돈 최적화된 (coGAA) GAA 유전자인 이중 핵산 서열을 포함한다.

[0044] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 핵산 서열은 서열 번호 11(전장 hGAA), 서열 번호 55(Dwight cDNA), 서열 번호 56(hGAA Δ 1-66) 또는 서열 번호 182(modGAA, H199R, R223H) 중 어느 하나로부터 선택된 GAA 폴리펩타이드를 암호화하는 이중 핵산 서열, 또는 서열 번호 170(modGAA; H199R, R223H), 서열 번호 171(modGAA; H199R, R223H, H201L)의 아미노산 서열을 갖는 GAA 폴리펩타이드를 암호화하는 핵산 서열, 또는 서열 번호 11, 55, 56 또는 182 중 어느 하나와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 GAA 폴리펩타이드를 암호화하는 핵산 서열을 포함한다.

[0045] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 핵산 서열은 GAA 폴리펩타이드를 암호화하는 이중 핵산 서열 (여기서, GAA 폴리펩타이드를 암호화하는 핵산은 서열 번호 74(코돈 최적화된 1), 서열 번호 75(코돈 최적화된 2), 및 서열 번호 76(코돈 최적화된 3), 또는 서열 번호 182(modGAA, H199R, R223H) 중 어느 하나로부터 선택된다), 또는 서열 번호 74, 75, 76 또는 182 중 어느 하나와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열을 포함한다.

[0046] 본원 기술의 또 다른 양태는 질환을 치료하기 위한 방법에서 본원에 개시된 rAAV 및 핵산 조성물의 용도에 관한 것이다. 특히, 본원의 기술의 하나의 양태는 글리코겐 축적 질환 II형(GSD II, 폰페병, 산성 말타아제 결핍증)이 있거나 알파-글루코시다아제(GAA) 폴리펩타이드가 결핍된 대상체를 치료하는 방법에서, 본원에 개시된 rAAV 벡터 조성물 및 핵산 조성물의 용도에 관한 것이며, 상기 방법은 재조합 AAV 벡터, 또는 rAAV 게놈 또는 본원에 개시된 핵산 서열을 대상체에게 투여하는 단계를 포함한다. 본원에 개시된 방법의 일부 구현예에서, 발현된 GAA 폴리펩타이드는 대상체의 간으로부터 분비되고 골격근 조직, 심장 근육 조직, 횡격막 근육 조직 또는 이들의 조합에 의해 분비된 GAA의 흡수가 있고, 여기서, 분비된 GAA의 흡수는 조직(들)에서 리소좀 글리코겐 저장의 감소를 초래한다. 개시된 방법의 일부 구현예에서, 재조합 AAV 벡터, 또는 rAAV 게놈 또는 핵산 서열은

임의의 적합한 투여 방법, 예를 들어, 그러나 이에 제한되지 않는 근육내, 피하, 척수내, 수조내, 척수강내, 정맥내 투여 중 어느 하나로부터 선택된 투여 방법에 의해 대상체에게 투여된다. 일부 구현예에서, 본원에 개시된 약제학적 조성물은 본원에 개시된 방법에 사용될 수 있다.

[0047] 본원 기술의 또 다른 양태는 본원에 개시된 바와 같은 rAAV 조성물, rAAV 게놈 조성물, 또는 핵산 조성물 중 임의의 하나 이상을 포함하는 세포에 관한 것이다. 일부 구현예에서, 세포는 인간 세포, 또는 비인간 세포 포유동물 세포, 또는 곤충 세포이다.

[0048] 본원 기술의 또 다른 양태는 본원에 개시된 바와 같은 rAAV 조성물, rAAV 게놈 조성물, 또는 핵산 조성물 중 임의의 하나 이상을 포함하는 숙주 동물에 관한 것이다. 일부 구현예에서, 숙주 동물은 포유동물, 비인간 포유동물 또는 인간이다.

[0049] 본원 기술의 또 다른 양태는 본원에 개시된 바와 같은 rAAV 조성물, rAAV 게놈 조성물, 또는 핵산 조성물 중 임의의 하나 이상을 포함하는 적어도 하나의 세포를 포함하는 숙주 동물에 관한 것이다. 일부 구현예에서, 이러한 변형된 세포를 포함하는 숙주 동물은 포유동물, 비인간 포유동물 또는 인간이다.

[0050] 일부 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터, rAAV 게놈을 암호화하는 핵산, 및 약제학적으로 허용되는 담체를 포함하는 약제학적 제형이 본원에 개시되어 있다.

[0051] 본 발명의 양태는 아래에 설명된 예시적인 이점을 발생시키는 구성 및 사용의 특정 이점을 교시한다. 본 발명의 양태의 다른 특징 및 이점은 예로서 본 발명의 양태의 원리를 예시하는 첨부 도면과 함께 취해진 다음의 보다 상세한 설명으로부터 명백해질 것이다.

도면의 간단한 설명

[0052] 본 출원은 파일은 컬러로 실행된 적어도 하나의 도면을 포함한다. 컬러 도면을 갖는 이 특허 출원 간행물의 사본은 요청 및 필요한 수수료 지불 시 특허청에 의해 제공될 것이다. 첨부 도면은 본 발명의 양태를 예시한다. 이러한 도면에서:

도 1은 적어도 하나의 구현예에 따라 전혈에서 측정된 바와 같이, 이배체 게놈당 벡터 게놈의 y축 및 상이한 AAV 혈청형 AAV3b, AAV3ST, AAV8 및 AAV9의 x축을 도시하는 그래프이다.

도 2는 적어도 하나의 구현예에 따라, 좌측, 중앙 및 우측 간 엽에서 측정된 바와 같이, 이배체 게놈당 벡터 게놈의 y축 및 상이한 AAV 혈청형 AAV3b, AAV3ST, AAV8, 및 AAV9의 x축을 도시하는 그래프이다.

도 3a-3b는 본원에 개시된 바와 같은 방법 및 조성물에 유용한 rAAV 벡터의 생성을 위한 예시적인 플라스미드이다. **도 3a**는 적어도 하나의 구현예에 따라 생산자 세포주, 예를 들어, pro-10 세포주에서 rAAV 벡터의 생성을 위한 pAAV-LSPHGA 플라스미드의 플라스미드 맵의 예시이고, 여기서, 플라스미드는 5' ITR, LSP, hGAA 핵산 서열, 3' UTR, 폴리A 서열 및 3' ITR을 포함하며, 여기서, ITR은 AAV2로부터 유래한다. **도 3b**는 도 3a의 플라스미드 맵의 예시의 보다 상세한 맵을 나타낸다.

도 4a-4g는 발현되는 예시적인 리소좀 단백질로서 hGAA를 사용하는, 표적화 펩타이드를 갖는 본원에 개시된 바와 같은 rAAV 게놈에 대한 예시적인 핵산 작제물의 예시이다. **도 4a**는 5' ITR, 분비 신호 펩타이드(SS), 표적화 펩타이드 (TP) 및 인간 GAA (hGAA) 폴리펩타이드를 암호화하는 이중 핵산에 작동적으로 연결된 간 특이적 프로모터(LSP), 및 3' ITR을 포함하는, rAAV 게놈에 대한 핵산 작제물을 나타낸다. **도 4b**는 도 4a와 동일한 요소를 포함하고 hGAA 폴리펩타이드의 적어도 하나의 폴리A 신호 3' 및 3'-ITR의 5'를 추가로 포함하는, 본 명세서에 개시된 바와 같은 rAAV 게놈에 대한 예시적인 핵산 작제물을 나타낸다. **도 4c**는 프로모터의 인트론 서열 3'를 포함하는 것을 제외하고는 도 4b와 동일한 요소를 포함하는, 본원에 개시된 바와 같은 rAAV 게놈에 대한 예시적인 핵산 작제물을 나타낸다. **도 4d**는 콜라겐 안정성(CS) 서열 및/또는 hGAA 폴리펩타이드 핵산 서열의 3'에 위치하고 폴리A 서열 앞에 위치하는 3' UTR 서열을 포함하는 것을 제외하고는 도 4c와 동일한 요소를 포함하는, 본원에 개시된 바와 같은 rAAV 게놈에 대한 예시적인 핵산 작제물을 나타낸다. **도 4e**는 hGAA 폴리펩타이드를 암호화하는 핵산과 표적화 펩타이드(TP), 예를 들어, IGF2 표적화 펩타이드를 암호화하는 핵산 사이에 위치하는 적어도 1개의 아미노산의 스페이서를 암호화하는 핵산을 또한 포함하는 것을 제외하고는 도 4d와 동일한 요소를 포함하는, 본원에 개시된 rAAV 게놈에 대한 예시적인 핵산 작제물을 나타낸다. **도 4f**는 도 4e와 동일한 요소를 포함하는, 본원에 개시된 rAAV 게놈에 대한 예시적인 핵산 작제물을 나타내고, 여기서, 프로모터는 간 프로모터이고, 인트론 서열은 MVM 또는 HBB2 인트론 서열로부터 선택되고, 분비 신호 펩타이드는 FN1 신호 펩타이드(예를 들어, hFN1, ratFN1), AAT 신호 펩타이드 또는 hGAA 신호 펩타이드 중 어느 하나로부터 선택되고; 표

적화 펩타이드는 본원에 개시된 바와 같은 IGF2 표적화 펩타이드이고, 적어도 폴리A 서열은 hGHPA 또는 synPA 폴리A 서열로부터 선택된다. 도 4g는 IGF2 표적화 펩타이드가 서열 번호 2(IGF2 Δ 2-7), 서열 번호 3(IGF2 Δ 1-7), 또는 서열 번호 4(IGF2 V43M)로부터 선택된 핵산 서열인 경우를 제외하고는 도 4f와 동일한 요소를 포함하는, 본원에 개시된 바와 같은 rAAV 게놈에 대한 예시적인 핵산 작제물을 나타낸다.

도 5a-5g는 rAAV 게놈에 대한 예시적인 핵산 작제물을 나타낸다. 도 5a는 5' ITR, hGAA 폴리펩타이드를 암호화하는 핵산에 작동적으로 연결된 간 특이적 프로모터, 폴리A 서열(예를 들어, hGHPA, synPA, RBG 또는 SV40 폴리A 서열 중 임의의 하나 이상) 및 3' ITR을 포함하는, 예시적인 rAAV 게놈의 개략도이다. 도 5b는 5' ITR, 신호 분비 펩타이드(예를 들어, FN1, AAT 또는 동족 GAA 신호 펩타이드, IL2, mutIL2, IgG 중 어느 하나로부터 선택됨)를 암호화하는 핵산에 작동적으로 연결된 간 특이적 프로모터, 인간 GAA 폴리펩타이드 및 폴리A 서열을 암호화하는 핵산 및 3' ITR을 포함하는, 예시적인 rAAV 게놈의 개략도이다. 도 5c는 5' ITR, 인트론 서열(예를 들어, MVM, SV40 또는 HBB2 인트론 서열)에 작동적으로 연결된 간 특이적 프로모터, 신호 분비 펩타이드(예를 들어, FN1, AAT 또는 동족 GAA 신호 펩타이드, IL2, mutIL2, IgG 중 어느 하나로부터 선택됨)를 암호화하는 핵산, 인간 GAA 폴리펩타이드를 암호화하는 핵산 및 폴리A 서열 및 3' ITR을 포함하는, 예시적인 rAAV 게놈의 개략도이다. 도 5d는 GAA를 암호화하는 핵산의 3'과 적어도 하나의 폴리A 서열(예를 들어, hGHPA 및/또는 synPA 폴리A 서열) 사이에 위치한 콜라겐 안정성(CS) 서열 또는 3' UTR을 포함하는, 도 5c와 유사한 작제물의 개략도이다. 일부 구현예에서, 작제물은 본원에 개시된 바와 같은 CS 서열 및 3' UTR 서열 둘 모두를 포함한다. 일부 구현예에서, CS 서열은 본원에 개시된 바와 같은 3' UTR 서열로 대체될 수 있다. 도 5a-5d에서, 예시적인 간 특이적 프로모터는 본원의 표 4에 개시된 것들 중 어느 하나로부터 선택될 수 있고, 서열 번호 86, 91-96, 또는 146-150, 또는 서열 번호 86, 91-96, 또는 146-150과 적어도 85% 서열 동일성을 갖는 서열을 포함하지만 이에 제한되지 않는다. 도 5e는 5' ITR과 3' ITR 서열 사이의 측면에 위치한 핵산을 포함하는 품폐병을 치료하기 위해 본원에 개시된 바와 같은 방법 및 조성물에 유용한 AAV 벡터의 하나의 구현예의 개략도로, 상기 핵산은 5'에서 3' 방향으로: LSP 프로모터, 코작 서열, 신호 서열(도 5e에서 리더 서열로 지칭됨), hGAA를 암호화하는 핵산 및 폴리A 서열을 포함한다. 일부 구현예에서, 리더 서열은 고유 GAA 리더 서열, IL2 리더 서열(IL2 wt), 변형된 IL2 리더 서열(IL2 mut) 또는 IgG 리더 서열 중 어느 하나 또는 이의 기능적 변이체로부터 선택될 수 있으며; hGAA 서열은 컨센서스 hGAA 핵산 서열 또는 적어도 H201L 돌연변이 또는 본원에 개시된 기타 변형(예를 들어, H199R, R223H)을 갖는 hGAA 핵산으로부터 선택될 수 있다. 도 5f는 5'에서 3' 방향으로: 간 특이적 프로모터, 인트론 서열, 코작 서열, 신호 서열(리더 서열로도 칭함), IGF2 표적화 펩타이드 서열(도 5f에서 "GILT"로 칭함), hGAA를 암호화하는 핵산, 임의로 3' UTR 서열, 및 폴리A 서열을 포함하는, 품폐병을 치료하기 위해 본원에 개시된 바와 같은 방법 및 조성물에 유용한 AAV 벡터의 또 다른 구현예의 개략도로서, 이는 상이한 구현예를 나타내는데, 예를 들어, 프로모터는 본원에 개시된 바와 같은 임의의 LSP, 예를 들어, 상이한 수준의 발현을 갖는 LSP, 예컨대 고발현 수준 LSP(LSP-H), 중간 발현 수준 LSP(LSP-M) 또는 저발현 수준 LSP(LSP-L)로부터 선택될 수 있고, 인트론 서열은 HBB2, MVM, SV40 및 기타 인트론 서열로부터 선택될 수 있고, 리더 서열은 고유 GAA 리더 서열, AAT 서열(도 5f에서 A1AT로 칭함), IL2(1-3), IL2 리더 서열(IL2 wt), 변형된 IL2 리더 서열(IL2 mut), 피브로넥틴(FN1, 도 5f에서 FBN으로 칭함), 또는 IgG 리더 서열 중 어느 하나 또는 이의 기능적 변이체로부터 선택될 수 있고; 본원에 기재된 IGF2 표적화 펩타이드, 예를 들어, WT IGF2(서열 번호 1), Δ 2-7, V43M(서열 번호 9), Δ 2-7V43M 중 어느 하나 또는 이의 기능적 변이체로부터 선택된 IGF2 표적화 펩타이드 서열, 및 본원에 개시된 바와 같은 코돈 최적화된 hGAA 핵산 서열, 예를 들어, 임의로 적어도 H201L 돌연변이 및/또는 본원에 개시된 바와 같은 기타 변형(예를 들어, H199R, R223H)도 포함할 수 있는 C1-10, 및 예를 들어, RBG 또는 SV40 폴리A로부터 선택된 폴리A 서열을 포함한다. LSP-H, M-LSP 및 LSP-L로 지정된 LSP는 간에서 hGAA를 우세하고 우선적으로 발현하지만, 예를 들어, 근육에서 하나 이상의 다른 조직에서 hGAA를 발현할 수 있는 간 특이적 프로모터를 나타낸다. 이러한 LSP는 근육 조직에서의 일부 발현뿐만 아니라 근육 세포에 의한 전신 분비 및 흡수를 위한 간에서의 발현을 허용한다. 도 5g는 품폐병을 치료하기 위해 본원에 개시된 바와 같은 방법 및 조성물에 유용한 AAV 벡터 작제물의 상이한 구현예의 개략도를 나타내고, 여기서, 작제물 1(상단 패널)은 5'에서 3' 방향으로, 5' ITR, AAV P5 프로모터, 간-특이적 프로모터(LSP), hGAA 핵산 서열, 절단된 폴리A 서열(t-pA), 및 3' ITR을 포함하는 rAAV 벡터 작제물을 나타내고; 작제물 2(하단 패널)은 예시적인 rAAV 벡터 작제물을 나타내고, 여기서, P5 AAV 프로모터 단편이 제거되고, 상기 작제물은 5'에서 3' 방향으로, 5' ITR, 간-특이적 프로모터(LSP), hGAA 핵산 서열, 전장 폴리A 서열(f1-pA), 안티센스 방향의 종격자 서열 및 3' ITR(센스 방향으로)을 포함한다.

도 6은 본원에 개시된 바와 같은 rAAV 게놈을 생성하기 위한 Gibson 클로닝(Gibson cloning) 기술의 예시를 나타낸다. 특히, 삼중 결찰은 핵산 서열의 3개 블록을 함께 결찰하기 위해 수행되며, 이어서, 프로모터, 예를

들어, 간 특이적 프로모터, 및 5' 및 3' ITR을 갖는 벡터로 클로닝되어 rAAV 게놈을 생성할 수 있다. 깃승 클로닝 방법론을 사용하여 다음 rAAV 게놈을 사용하였다: 서열 번호 57(AAT-V43M-wtGAA(델타1-69aa)); 서열 번호 58(ratFN1-IGF2V43M-wtGAA(델타1-69aa)); 서열 번호 59(hFN1-IGF2V43M-wtGAA(델타1-69aa)); 서열 번호 60(AAT-IGF2 Δ 2-7-wtGAA(델타 1-69)); 서열 번호 61(FN1rat-IGF Δ 2-7-wtGAA(델타 1-69)); 서열 번호 62(hFN1-IGF Δ 2-7-wtGAA(델타 1-69)).

도 7은 핵산 서열 블록(1, 2 및 3)의 깃승 클로닝을 사용하여 AAT-V43M-wtGAA(델타1-69aa)를 포함하는, 서열 번호 57의 예시적인 rAAV 게놈의 생성을 나타낸다. 당업자는 TTR 간 프로모터를 서열 번호 86, 91-96, 146-150 중 어느 하나, 또는 이의 기능적 변이체 또는 기능적 단편으로부터 선택된 프로모터를 포함하지만 이에 제한되지 않는 본원의 표 4에 개시된 간 특이적 프로모터 중 어느 하나로 쉽게 대체할 수 있다. 또한 AAT-V43M-wtGAA(델타1-69aa) 벡터에는 IGF2(V43M) 표적화 펩타이드를 암호화하는 핵산 서열의 3' 및 wtGAA(Δ 1-69) 효소를 암호화하는 핵산의 5'에 위치하는 3 아미노산(3aa) 스페이서 핵산 서열(예시적인 3aa 서열 "G-A-P"를 서열 번호 31로서 나타냄) 및 폴리A 서열의 3' 및 3'ITR 서열의 5'에 위치하는 핵산 서열(도 8에서 "스페이서" 서열로 지칭됨)의 위치를 나타낸다.

도 8은 핵산 서열 블록(8, 2 및 3)의 깃승 클로닝을 사용하여 hFN1-IGF Δ 2-7-wtGAA(델타 1-69)를 포함하는 서열 번호 62의 rAAV 게놈의 생성을 나타낸다. 당업자는 TTR 간 프로모터를 서열 번호 86, 91-96, 146-150 중 어느 하나, 또는 이의 기능적 변이체 또는 기능적 단편으로부터 선택된 프로모터를 포함하지만 이에 제한되지 않는 본원의 표 4에 개시된 간 특이적 프로모터로 쉽게 대체할 수 있다. 또한 hFN1-IGF Δ 2-7-wtGAA(델타 1-69) 벡터에는 IGF Δ 2-7표적화 펩타이드를 암호화하는 핵산 서열의 3' 및 wtGAA(Δ 1-69) 효소를 암호화하는 핵산의 5'에 위치하는 3 아미노산 (3aa) 스페이서 핵산 서열(예시적인 3aa 서열 "G-A-P"를 서열 번호 31로서 나타냄) 및 폴리A 서열의 3' 및 3'ITR 서열의 5'에 위치하는 스테퍼 핵산 서열(도 13에서 "스페이서" 서열로 지칭됨)의 위치를 나타낸다.

도 9a-9f는 야생형 GAA를 발현하는 rAAV 게놈의 예시적인 작제물의 개략도를 나타낸다. **도 9a**는 후보 1_AAT_hIGF2-V43M_wtGAA_del1-69_스테퍼.V02(서열 번호 79)의 예시적인 rAAV 게놈 작제물의 개략도를 나타낸다. **도 9b**는 후보 2_FIBrat_hIGF2-V43M_wtGAA_del1-69_스테퍼.V02(서열 번호 80)의 예시적인 rAAV 게놈 작제물의 개략도를 나타낸다. **도 9c**는 후보 3_FIBhum_hIGF2-V43M_wtGAA_del1-69_스테퍼.V02(서열 번호 81)의 예시적인 rAAV 게놈 작제물의 개략도를 나타낸다. **도 9d**는 후보 4_AAT_GILT_wtGAA_del1-69_스테퍼.V02(서열 번호 82)의 예시적인 rAAV 게놈 작제물의 개략도를 나타낸다. **도 9e**는 후보 5_FIBrat_GILT_wtGAA_del1-69_스테퍼.V02(서열 번호 83)의 예시적인 rAAV 게놈 작제물의 개략도를 나타낸다. **도 9f**는 후보 6_FIBhum_GILT_wtGAA_del1-69_스테퍼.V02(서열 번호 84)의 예시적인 rAAV 게놈 작제물의 개략도를 나타낸다. 당업자는 도 9a-9f에 나타낸 TTR 간 프로모터를 서열 번호 86, 91-96, 또는 146-150 중 어느 하나로부터 선택된 프로모터를 포함하지만 이에 제한되지 않는 임의의 LPS, 예를 들어, 본원의 표 4에 개시된 임의의 간 특이적 프로모터로 쉽게 대체할 수 있다. 더욱이, TTR 프로모터는 간 및 또한 적어도 하나의 다른 관심 있는 조직, 예를 들어 근육 또는 CNS에서 우선적으로 hGAA 폴리펩타이드를 발현할 수 있는 LSP로 대체될 수 있으며, 일부 구현예에서, TTR 프로모터는 간 및 근육 및 CNS 조직에서 우선적으로 hGAA 폴리펩타이드를 발현할 수 있는 LSP로 대체될 수 있다. 일부 구현예에서, 발현된 리소좀 효소, 예를 들어, GAA 단백질을 리소좀으로 표적화 하고/하거나 신호 펩타이드(SP)와 융합하는 본원에 개시된 바와 같은 IGF2 표적화 펩타이드와 같은 표적화 서열을 갖는 GAA-융합 단백질로서 구성될 수 있으며, 상기 GAA 단백질은 간에서 rAAV 게놈에 의해 발현되며, 이는 분비되어 포유동물 세포, 특히 근육 세포의 리소좀에 의해 흡수된다.

이들은 단지 예시를 위한 예시적인 작제물이므로, 또한 wtGAA 서열을 본원에 개시된 바와 같은 코돈 최적화된 서열, 또는 본원에 개시된 바와 같이 CpG 섬을 감소시키고/시키거나 선천 면역을 감소시키도록 변형된 GAA 서열로 쉽게 대체할 수 있다(도 11b 참조).

도 10은 예시적인 간-특이적 프로모터 SP0244 및 SP0239에 의해 구동되는 마우스에서의 평균 생체내 루시퍼라제 발현을 나타낸다. 발현 수준은 평균 생물발광 강도 총 플러스(초당 광자)로서 나타낸다. 오차 막대는 평균의 표준 오차이다. 동물에 식염수(n=10)만 주입하는 경우, 루시퍼라제 생물발광이 검출되지 않는다. 동물에 LP1 프로모터(n=9)에 작동적으로 연결된 루시퍼라제를 포함하는 작제물을 주사하는 경우, 루시퍼라제 생물발광이 검출된다. 예시적인 간-특이적 프로모터의 활성을 시험하기 위해, 동물에 SP0244 프로모터(n=8) 및 SP0239 프로모터(n=10)에 작동적으로 연결된 루시퍼라제를 포함하는 등가 작제물을 주사한다. 프로모터 SP0244 및 SP0239는 대조군 LP1보다 생체내에서 더 높은 루시퍼라제 발현을 나타내었다.

도 11a-11d는 생체내에서 AAV에 의한 GAA 단백질 발현을 최적화하기 위한 GAA 폴리펩타이드 및 핵산 작제물을 암호화하는 핵산 서열에 대한 예시적인 변형을 나타낸다. **도 11a**는 화살표로 나타난 대안적인 리딩(reading) 프레임 및 3개의 CpG 섬과 함께, 본 명세서에 개시된 바와 같은 간 특이적 프로모터, 예를 들어, 표 4의 LSP에 작동적으로 연결된 야생형 GAA(wtGAA) 뉴클레오타이드 서열의 개략도를 나타낸다. **도 11b**는 CpG 섬을 제거하기 위해 GAA를 암호화하는 핵산 서열에 대한 변형을 보다 상세히 나타내는 도 11a와 유사한 개략도를 나타낸다. **도 11c**는 (i) 대안적인 리딩 프레임을 감소시키고, (ii) CpG 섬의 수를 감소시키고, (iii) 최적의 코작 서열에 대한 변형을 포함하도록 변형된 서열 번호 182의 wtGAA 핵산 서열에 대한 변형을 나타낸다. **도 11d**는 대안적인 리딩 프레임, CpG 섬의 수 및 최적 코작 서열에 대한 변형을 감소시키기 위한 GAA 폴리펩타이드를 암호화하는 핵산 서열의 변형을 나타내는 또 다른 개략도이다.

도 12는 간 특이적 프로모터 하에 GAA를 발현하기 위한 LSP를 포함하는 예시적인 rAAV 작제물의 개략도를 나타낸다. LSP는 스테퍼 서열이 있거나 없는 본원의 표 4에 개시된 임의의 간 특이적 프로모터로부터 선택될 수 있다.

도 13a-13b는 Huh 7 세포 및 HEPG2 세포에서 간 특이적 프로모터 SP0412 및 SP0422를 포함하는 작제물로부터의 GAA 발현을 나타낸다. **도 13a**는 Huh 7 세포에서 간 특이적 프로모터 SP0412(서열 번호 91) 및 SP0422(서열 번호 92)를 포함하는 작제물로부터의 GAA 발현의 웨스턴 블롯을 나타낸다. 도 13a는 프로모터 412(서열 번호 91) 및 422(서열 번호 92)를 사용한 hGAA의 발현이 "LSP SS"로 지칭되는 LP1 프로모터(서열 번호 432)를 사용한 발현과 비교하여 Huh7 세포에서 hGAA의 상당히 더 높은 발현을 유도한다는 것을 나타낸다. **도 13b**는 HEPG2 세포에서 간 특이적 프로모터 SP0412(서열 번호 91) 및 SP0422(서열 번호 92)를 포함하는 작제물로부터의 GAA 발현의 웨스턴 블롯을 나타낸다. GAA 폴리펩타이드는 다음 플라스미드를 사용하여 생성된 rAAV로부터 발현되었다: LSP NEW(서열 번호 160), 412 NEW(서열 번호 159), TTR NEW(서열 번호 155), LSP ss(LP-1을 갖는 AAV), 412 TTR, 422 스테퍼(서열 번호 158), 422 TTR, 412 스테퍼(서열 번호 156). 도 13b는 프로모터 412(서열 번호 91) 및 422(서열 번호 92)를 사용한 hGAA의 발현이 "LSP SS"로 지칭되는 LP1 프로모터(서열 번호 432)를 사용한 발현과 비교하여 HepG2 세포에서 hGAA의 상당히 더 높은 발현을 유도한다는 것을 나타낸다.

상기 기재된 도면은 예시적인 구현에 중 적어도 하나에서 본 발명의 양태를 도시하며, 이는 다음 설명에서 더욱 상세하게 정의된다. 상이한 도면에서 동일한 번호로 언급되는 본 발명의 특징, 요소 및 양태는 하나 이상의 구현에 따라 동일하거나 동등하거나 유사한 특징, 요소 또는 양태를 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0053] 본원에 기재된 개시내용은 일반적으로 GAA 폴리펩타이드와 같은 리소솜 단백질을 대상체에게 전달하기 위한 유전자 요법을 위한 rAAV 계통을 위한 재조합 AAV(rAAV) 벡터 및 작제물에 관한 것이다. 특히, 본원에 기재된 기술은 일반적으로 간에서 발현되고 포유동물 세포, 예를 들어, 인간 심장 및 골격근 세포의 리소솜을 효과적으로 표적화하는 리소솜 단백질, 예를 들어, GAA 폴리펩타이드를 생성하기 위한 rAAV 벡터 또는 rAAV 계통에 관한 것이다. 예를 들어, 상기 기술은 형질도입된 간 세포가 GAA 폴리펩타이드를 분비하고 분비된 GAA 폴리펩타이드가 골격근 조직, 심장 근육 조직, 횡격막 근육 조직 또는 이들의 조합에서 리소솜을 표적으로 하는 간 세포를 형질도입하기 위한 rAAV 벡터에 관한 것이다.
- [0054] 따라서, 본원에 기재된 기술의 하나의 양태는 세포, 예를 들어, 간세포로부터 보다 효과적으로 분비된 다음, 포유동물 세포, 예를 들어, 인간 심장 및 골격근 세포의 리소솜을 표적으로 하는 리소솜 단백질, 예를 들어, GAA 또는 변형된 GAA를 생성하는 데 사용될 수 있는 rAAV 계통을 포함하는 rAAV 벡터를 제공한다.
- [0055] 특히, 일부 구현예에서, 리소솜 단백질, 예를 들어, GAA 폴리펩타이드는 그 자체로 발현된다. 일부 구현예에서, 리소솜 단백질은 적어도 리소솜 단백질, 예를 들어, 간으로부터의 GAA 폴리펩타이드의 분비를 촉진하는 신호 펩타이드를 포함하는 융합 단백질로서 발현된다. 일부 구현예에서, GAA 폴리펩타이드 또는 변형된 GAA는 적어도 간으로부터 GAA 폴리펩타이드의 분비를 촉진하는 신호 펩타이드, 및 또한 포유동물 세포, 예를 들어, 근육 세포, 예를 들어, 인간 심장 및 골격근 세포에서 리소솜에 대한 효과적인 표적화를 허용하는 표적화 서열을 포함하는 융합 단백질로서 발현된다. 일부 구현예에서, 표적화 펩타이드는 본원에 기재된 IGF2 표적화 펩타이드이다.
- [0056] 본원에 기재된 기술의 하나의 양태는 폼페병과 같은 질환을 치료하기 위해, 그리고 추가로 폼페병 치료를 위해 사용하기 위한 역 말단 반복체(ITR), 간 특이적 프로모터, 이중 유전자, 폴리-A 테일 및 잠재적으로 다른 조절 요소를 포함하는 뉴클레오타이드 서열을 포함하는 rAAV 벡터에 관한 것이며, 여기서, 이중 유전자는 GAA이고,

여기서 rAAV GAA는 이중 유전자의 발현 및 질환의 치료를 위해 적절한 조직 및/또는 기관에 전달되는 치료학적 유효량으로 환자에게 투여될 수 있다.

[0057] 본원에 기재된 기술의 하나의 양태는 5'에서 3' 방향으로 다음을 그 게놈에 포함하는 rAAV 벡터에 관한 것이다: 5'- 및 3'-AAV 역 말단 반복체(ITR) 서열, 및 5' ITR과 3' ITR 사이에 위치하는, **알파-글루코시다아제(GAA) 폴리펩타이드**를 암호화하는 이중 핵산 서열로서, 여기서, 이중 핵산은 간 특이적 프로모터, 예를 들어, 본원의 표 4에 개시된 간 특이적 프로모터, 또는 이의 기능적 변이체에 작동적으로 연결된다. 본원에 기재된 기술의 또 다른 양태는 5'에서 3' 방향으로 다음을 그 게놈에 포함하는 rAAV 벡터에 관한 것이다: 5'- 및 3'-AAV 역 말단 반복체(ITR) 서열, 및 5' ITR과 3' ITR 사이에 위치하는, **분비 신호 펩타이드(SS)**를 암호화하는 이중 핵산 서열, **알파-글루코시다아제(GAA) 폴리펩타이드**를 암호화하는 핵산 서열로서, 여기서, 이중 핵산은 간 특이적 프로모터, 예를 들어, 본원의 표 4에 개시된 간 특이적 프로모터 또는 이의 기능적 변이체에 작동적으로 연결된다.

[0058] 본원에 기재된 기술의 하나의 양태는 5'에서 3' 방향으로 다음을 그 게놈에 포함하는 rAAV 벡터에 관한 것이다: 5'- 및 3'-AAV 역 말단 반복체(ITR) 서열, 및 5' ITR과 3' ITR 사이에 위치하는, (i) **분비 신호 펩타이드(SS)**, (ii) **IGF2 표적화 펩타이드**; 및 (iii) **알파-글루코시다아제(GAA) 폴리펩타이드**를 포함하는 융합 폴리펩타이드를 암호화하는 이중 핵산 서열로서, 여기서, 이중 핵산은 간 특이적 프로모터, 예를 들어, 본원의 표 4에 개시된 간 특이적 프로모터 또는 이의 기능적 변이체에 작동적으로 연결된다.

[0059] 본원에 기재된 기술의 모든 구현예의 모든 양태에서, 간 특이적 프로모터는 리소솜 단백질, 예를 들어, 간에서 우선적으로 hGAA 폴리펩타이드를 발현한다. 본원에 기재된 기술의 모든 구현예의 모든 양태에서, 간 특이적 프로모터는 간 및 적어도 하나의 다른 관심 있는 조직, 예를 들어, 근육 또는 CNS에서 우선적으로 리소솜 단백질, 예를 들어, hGAA 폴리펩타이드를 발현하고, 일부 구현예에서, LSP는 간 및 근육 및 CNS 조직에서 우선적으로 hGAA 폴리펩타이드를 발현할 수 있는 LSP로 대체될 수 있다. 본원에 기재된 기술의 모든 구현예의 모든 양태에서, AAV 벡터가 근육을 표적화하는 적어도 하나의 캡시드 단백질을 포함하는 일부 구현예에서, 간 특이적 프로모터는 또 다른 프로모터, 예를 들어, 근육 프로모터로 대체될 수 있다.

[0060] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 분비 신호 펩타이드는 AAT 신호 펩타이드, 피브로넥틴 신호 펩타이드(FN1), GAA 신호 펩타이드, 또는 분비 신호 활성을 갖는 이의 활성 단편 중 어느 하나로부터 선택된다.

[0061] 일부 구현예에서, 본원에 기재된 rAAV 벡터는 임의의 혈청형 형태로부터 유래된다. 일부 구현예에서, rAAV 벡터는 AAV3b265D 비리온, AAV3b265D549A 비리온, AAV3b549A 비리온, AAV3bQ263Y 비리온, 또는 AAV3bSASTG 비리온(즉, Q263A/T265 돌연변이를 포함하는 AAV3b 캡시드를 포함하는 비리온)을 포함하지만 이에 제한되지 않는 AAV3b 혈청형이다. 일부 구현예에서, rAAV 벡터는 간 특이적 캡시드, 예를 들어, 그 전체가 본원에 참조로 포함된 WO2019/241324에 개시된 바와 같은 XL32 및 XL32.1로부터 선택된 간 특이적 캡시드를 포함한다. 일부 구현예에서, rAAV 벡터는 그 전체가 본원에 참조로 포함된 WO2019/241324에 개시된 바와 같은 AAVXL32 또는 AAVXL32.1이다. 일부 구현예에서, rAAV 벡터는 rAAV8 벡터, 또는 AAV8로부터의 적어도 하나의 캡시드 단백질(즉, VP1, VP2 또는 VP3 중 임의의 하나 이상은 AAV8 또는 이의 키메라 단백질로부터 유래한다)을 포함하는 반수체 rAAV 벡터이다. 일부 구현예에서, AAV 벡터는 그 전체가 본원에 참조로 포함된 WO2019241324A1 또는 국제 특허 출원 PCT/US2019/036676에 개시된 캡시드를 포함한다. 일부 구현예에서, AAV 벡터는 WO2019241324A1에 개시된 바와 같은 서열 번호 1-3 중 어느 하나의 뉴클레오타이드 서열 또는 (b) WO2019241324A1에 개시된 바와 같은 서열 번호 4-6 중 어느 하나를 암호화하는 뉴클레오타이드 서열과 적어도 90% 동일한 핵산 AAV 캡시드 암호화 서열에 의해 암호화된 캡시드를 포함한다. 일부 구현예에서, AAV 캡시드는 본 발명의 AAV 벡터 게놈 및 AAV 캡시드를 포함하는 AAV 입자와 함께 WO2019241324A1에 개시된 바와 같은 서열 번호 4-6 중 어느 하나와 적어도 90% 동일한 아미노산 서열을 포함한다. 일부 구현예에서, rAAV 벡터는 AAV 벡터가 간 세포를 형질도입하도록 캡시드 단백질을 포함하고, 일부 구현예에서, rAAV 벡터는 AAV 벡터가 근육 및 간 세포를 형질도입하도록 캡시드 단백질을 포함한다. rAAV가 근육 세포의 형질도입을 가능하게 하는 캡시드 단백질을 포함하는 이러한 구현예에서, LSP는 또 다른 프로모터, 예를 들어, 근육 프로모터, 또는 간 세포 및 근육 세포에서 단백질을 발현하는 프로모터로 대체될 수 있다.

[0062] I. 정의

[0063] 다음 용어는 본원의 설명 및 첨부된 청구범위에서 사용된다:

[0064] 용어 "a," "an," "the" 및 본 발명을 설명하는 맥락에서(특히 다음 청구범위의 맥락에서) 사용되는 유사한 참조

는 본원에 달리 나타내거나 문맥상 명백히 모순되지 않는 한 단수와 복수를 둘 모두 포함하는 것으로 해석되어야 한다. 또한, 확인된 요소에 대한 서수 지표 - 예컨대, "제1", "제2", "제3" 등 - 는 요소 사이를 구별하는데 사용되며, 이러한 요소의 필수 또는 제한된 수를 나타내거나 암시하지 않으며, 달리 구체적으로 언급되지 않는 한 이러한 요소의 특정 위치 또는 순서를 나타내지 않는다. 본원에 기재된 모든 방법은 본원에 달리 표시되지 않거나 문맥상 명백히 모순되지 않는 한 임의의 적절한 순서로 수행될 수 있다. 본원에 제공된 임의의 및 모든 예 또는 예시적인 언어(예를 들어, "~와 같은")의 사용은 본 발명을 더 잘 설명하기 위한 것일 뿐 달리 청구된 본 발명의 범위에 제한을 두지 않는다. 본 명세서에서는 어떤 언어도 본 발명의 실행에 필수적인 임의의 청구되지 않은 요소를 나타내는 것으로 해석되어서는 안된다.

[0065] 또한, 폴리뉴클레오타이드 또는 폴리펩타이드 서열의 길이, 용량, 시간, 온도 등의 양과 같은 측정 가능한 값을 언급할 때 본원에 사용된 용어 "약"은 명시된 양의 $\pm 20\%$, $\pm 10\%$, $\pm 5\%$, $\pm 1\%$, $\pm 0.5\%$, 또는 심지어 $\pm 0.1\%$ 의 변동을 포함하는 것을 의미한다.

[0066] 본원에 사용된 바와 같이, "및/또는"은 대안("또는")으로 해석될 때 조합의 부족뿐만 아니라 하나 이상의 연관된 나열된 항목의 임의의 및 모든 가능한 조합을 지칭하고 포함한다.

[0067] 본원에 사용된 바와 같이, "~로 본질적으로 이루어진"이라는 전환 문구는 청구항의 범위가 청구범위에 인용된 특정 재료 또는 단계를 포함하는 것으로 해석되어야 하며 청구된 발명의 "기본적이고 신규한 특성(들)에 실질적으로 영향을 미치지 않는 것들"을 의미한다. 문헌[*In re Herz*, 537 F.2d 549, 551-52, 190 USPQ 461,463 (CCPA 1976) (원문 강조)]을 참조하고; MPEP § 2111.03도 참조한다. 따라서, 본 발명의 청구범위에서 사용될 때 "~로 본질적으로 이루어진"이라는 용어는 "포함하는"과 동등한 것으로 해석되도록 의도되지 않는다. 문맥이 달리 나타내지 않는 한, 본원에 기재된 본 발명의 다양한 특징이 임의의 조합으로 사용될 수 있다는 것이 구체적으로 의도된다.

[0068] 더욱이, 본 발명은 또한 본 발명의 일부 구현예에서, 본원에 기재된 임의의 특징 또는 특징의 조합이 배제되거나 생략될 수 있음을 고려한다.

[0069] 추가로 설명하기 위해, 예를 들어, 특정 아미노산이 A, G, I, L 및/또는 V로부터 선택될 수 있음을 명세서가 나타내는 경우, 이 언어는 또한 각각의 이러한 하위조합이 본원에 명시적으로 기재된 것처럼 아미노산이 이러한 아미노산(들)의 임의의 하위세트, 예를 들어, A, G, I 또는 L; A, G, I 또는 V; A 또는 G; 단지 L로부터 선택될 수 있음을 나타낸다. 더욱이, 이러한 언어는 또한 명시된 아미노산 중 하나 이상이 거부될 수 있음을 나타낸다(예를 들어, 부정 단서에 의해). 예를 들어, 특정 구현예에서, 각각의 이러한 가능한 거부가 본원에 명시적으로 명시된 것처럼, 아미노산은 A, G 또는 I가 아니고; A가 아니고; G 또는 V 등이 아니다.

[0070] 본원에 사용된 용어 "파르보바이러스"는 자발적으로 복제되는 파르보바이러스 및 디펜도바이러스를 포함한 파르보비리다에 계열을 포함한다. 자발적 파르보바이러스는 파르보바이러스, 에리트로바이러스, 텐소바이러스, 이테라바이러스 및 콘트라바이러스 속의 구성원을 포함한다. 예시적인 자발적 파르보바이러스는 마우스의 미세바이러스, 소 파르보바이러스, 개 파르보바이러스, 닭 파르보바이러스, 고양이 범백혈구감소증 바이러스, 고양이 파르보바이러스, 거위 파르보바이러스, H1 파르보바이러스, 사향오리 파르보바이러스, B19 바이러스, 및 현재 알려져 있거나 나중에 발견되는 임의의 다른 자발적 파르보바이러스를 포함하지만 이에 제한되지 않는다. 다른 자발적 파르보바이러스는 당업자에게 알려져 있다. 예를 들어, 문헌(BERNARD N. FIELDS et al., VIROLOGY, volume 2, chapter 69 (4th ed., Lippincott-Raven Publishers))을 참조한다.

[0071] 본원에 사용된 바와 같이, 용어 "아데노-관련 바이러스"(AAV)는 AAV 1형, AAV 2형, AAV 3형(3A형 및 3B형 포함), AAV 4형, AAV 5형, AAV 6형, AAV 7형, AAV 8형, AAV 9형, AAV 10형, AAV 11형, 조류 AAV, 소 AAV, 개 AAV, 말 AAV, 양 AAV, 및 현재 알려져 있거나 나중에 발견되는 임의의 다른 AAV를 포함하지만 이에 제한되지 않는다. 예를 들어, 문헌(BERNARD N. FIELDS et al., VIROLOGY, volume 2, chapter 69 (4th ed., Lippincott-Raven Publishers))을 참조한다. 다수의 비교적 새로운 AAV 혈청형 및 분기군(clade)이 확인되었다[예를 들어, 문헌(Gao et al., (2004) *J. Virology* 78:6381-6388; Moris et al., (2004) *Virology* 33:375-383); 및 또한 2019년 11월 19일에 출원된 미국 가출원 62,937,556에 개시된 표 1 및 국제 출원 W02020/102645 및 W02020/102667의 표 1 참조, 이들 각각은 그 전체가 본원에 참조로 포함됨].

[0072] AAV 및 자발적 파르보바이러스의 다양한 혈청형의 게놈 서열뿐만 아니라 천연 역 말단 반복체(ITR), Rep 단백질 및 캡시드 서브유닛의 서열은 당업계에 알려져 있다. 이러한 서열은 문헌에서 또는 GenBank와 같은 공공 데이터베이스에서 찾을 수 있다. 예를 들어, GenBank 승인 번호 NC_002077, NC_001401, NC_001729, NC_001863,

NC_001829, NC_001862, NC_000883, NC_001701, NC_001510, NC_006152, NC_006261, AF063497, U89790, AF043303, AF028705, AF028704, J02275, J01901, J02275, X01457, AF288061, AH009962, AY028226, AY028223, NC_001358, NC_001540, AF513851, AF513852, AY530579를 참조하고; 그 개시내용은 파르보바이러스 및 AAV 핵산 및 아미노산 서열을 교시하기 위해 본원에 참조로 포함된다. 또한, 예를 들어, 문헌(Srivistava et al., (1983) *J. Virology* 45:555; Chiarini et al., (1998) *J. Virology* 71:6823; Chiarini et al., (1999) *J. Virology* 73:1309; Bantel-Schaal et al., (1999) *J. Virology* 73:939; Xiao et al., (1999) *J. Virology* 73:3994; Muramatsu et al., (1996) *Virology* 221:208; Shade et al., (1986) *J. Viral.* 58:921; Gao et al., (2002) *Proc. Nat. Acad. Sci. USA* 99:11854; Morris et al., (2004) *Virology* 33:375-383; 국제 특허 공개 WO 00/28061, WO 99/61601, WO 98/11244; 미국 특허 번호 6,156,303)을 참조하고; 그 개시내용은 파르보바이러스 및 AAV 핵산 및 아미노산 서열을 교시하기 위해 본원에 참조로 포함된다. 또한, 2019년 11월 19일에 출원된 62,937,556에 개시된 표 1 및 표 5 또는 국제 출원 WO2020/102645 및 WO2020/102667에 개시된 표 1을 참조하고, 이들 각각은 그 전체가 본원에 참조로 포함된다. 자발적 파르보바이러스 및 AAV의 캡시드 구조는 문헌(BERNARD N. FIELDS et al., *VIROLOGY*, volume 2, chapters 69 & 70 (4th ed., Lippincott-Raven Publishers))에 보다 자세히 기재되어 있다. 또한, AAV2(Xie et al., (2002) *Proc. Nat. Acad. Sci.* 99:10405-10), AAV4(Padron et al., (2005) *J. Viral.* 79: 5047-58), AAV5(Walters et al., (2004) *J. Viral.* 78: 3361-71) 및 CPV(Xie et al., (1996) *J. Mal. Biol.* 6:497-520 and Tsao et al., (1991) *Science* 251: 1456-64)의 결정 구조에 대한 설명을 참조한다.

[0073] 본원에 사용된 용어 "친화성"은 특정 세포 또는 조직으로 바이러스가 우선적으로 진입한 후 임의로 세포에서 바이러스 계놈에 의해 운반되는 서열(들)의 발현(예를 들어, 전사 및 임의로 번역), 예를 들어 재조합 바이러스의 경우 관심 있는 이중 핵산(들)의 발현을 지칭한다.

[0074] 본원에 사용된 바와 같이, "전신 친화성" 및 "전신 형질도입" (및 등가 용어)은 본 발명의 바이러스 캡시드 또는 바이러스 벡터가 신체(예를 들어, 뇌, 폐, 골격근, 심장, 간, 신장 및/또는 췌장) 전체에 걸쳐 조직에 대해 친화성을 나타내고/ 나타내거나 형질도입한다는 것을 나타낸다. 본 발명의 구현예에서, 중추 신경계(예를 들어, 뇌, 신경 세포 등)의 전신 형질도입이 관찰된다. 다른 구현예에서, 심장 근육 조직의 전신 형질도입이 달성된다.

[0075] 본원에 사용된 바와 같이, "선택적 친화성" 또는 "특이적 친화성"은 특정 표적 세포 및/또는 특정 조직으로의 바이러스 벡터의 전달 및/또는 특정 조직의 특이적 형질도입을 의미한다.

[0076] 달리 표시되지 않는 한, "효율적인 형질도입" 또는 "효율적인 친화성," 또는 유사한 용어는 적합한 대조군(예를 들어, 대조군의 형질도입 또는 친화성 각각의 적어도 약 50%, 60%, 70%, 80%, 85%, 90%, 95%, 100%, 125%, 150%, 175%, 200%, 250%, 300%, 350%, 400%, 500% 이상)을 참조하여 결정될 수 있다. 특정 구현예에서, 바이러스 벡터는 간 세포 및 근육 세포에 대해 효율적으로 형질도입하거나 효율적인 친화성을 갖는다. 적합한 대조군은 원하는 친화성 및/또는 형질도입 프로파일을 포함한 다양한 인자에 따라 달라질 것이다.

[0077] 유사하게는, 바이러스가 표적 조직에 대해 "효율적으로 형질도입하지 않는지" 또는 "효율적인 친화성을 갖지 않는지" 또는 유사한 용어를 적합한 대조군을 참조하여 결정될 수 있다. 특정 구현예에서, 바이러스 벡터는 콩팥, 생식선 및/또는 생식 세포에 대해 효율적으로 형질도입하지 않는다(즉, 효율적인 친화성을 갖지 않음). 특정 구현예에서, 조직(들) (예를 들어, 콩팥)의 형질도입(예를 들어, 바람직하지 않은 형질도입)은 원하는 표적 조직(들) (예를 들어, 간, 골격근, 횡격막 근육, 심장 근육 및/또는 중추 신경계의 세포)의 형질도입 수준의 20% 이하, 10% 이하, 5% 이하, 1% 이하, 0.1% 이하이다.

[0078] 본 발명의 일부 구현예에서, 본 발명의 캡시드를 포함하는 AAV 입자는 30개의 특정 조직/세포의 효율적인 형질도입의 다중 표현형 및 형질도입이 바람직하지 않은 특정 조직/세포의 경우 매우 낮은 수준의 형질도입(예를 들어, 감소된 형질도입)을 입증할 수 있다.

[0079] 본원에 사용된 바와 같이, 용어 "폴리펩타이드"는 달리 표시되지 않는 한 펩타이드와 단백질 둘 모두를 포함한다.

[0080] "폴리뉴클레오타이드"는 뉴클레오타이드 염기의 서열이고, RNA, DNA 또는 DNA-RNA 하이브리드 서열(자연적으로 발생하는 그리고 비자연적으로 발생하는 뉴클레오타이드 둘 모두를 포함)일 수 있지만, 대표적인 구현예에서 단일 또는 이중 가닥 DNA 서열이다.

[0081] 용어 "이중 뉴클레오타이드 서열" 및 "이중 핵산 분자"는 본원에서 상호교환적으로 사용되며 바이러스에서 자연

적으로 발생하지 않는 핵산 서열을 지칭한다. 일반적으로, 이중 핵산 분자 또는 이중 뉴클레오타이드 서열은 (예를 들어, 세포 및/또는 대상체로의 전달을 위해) 관심 있는 폴리펩타이드 및/또는 비번역 RNA를 암호화하는 오픈 리딩 프레임에 포함한다.

- [0082] "키메라 핵산"은 융합 폴리펩타이드를 암호화하기 위해 함께 공유적으로 연결된 2개 이상의 핵산 서열을 포함한다. 핵산은 DNA, RNA, 또는 이들의 하이브리드일 수 있다.
- [0083] 용어 "융합 폴리펩타이드"는 전형적으로 펩타이드 결합에 의해 함께 공유적으로 연결된 2개 이상의 폴리펩타이드를 포함한다.
- [0084] 본원에 사용된 바와 같이, "단리된" 폴리뉴클레오타이드(예를 들어, "단리된 DNA" 또는 "단리된 RNA")는 자연적으로 발생하는 유기체 또는 바이러스의 다른 성분의 적어도 일부, 예를 들어, 세포 또는 바이러스 구조적 성분 또는 폴리뉴클레오타이드와 관련하여 일반적으로 발견되는 기타 폴리펩타이드 또는 핵산으로부터 적어도 부분적으로 분리된 폴리뉴클레오타이드를 의미한다. 대표적인 구현예에서, "단리된" 뉴클레오타이드는 출발 물질과 비교하여 적어도 약 10배, 100배, 1000배, 10,000배 이상 농축된다.
- [0085] 마찬가지로, "단리된" 폴리펩타이드는 자연적으로 발생하는 유기체 또는 바이러스의 다른 성분의 적어도 일부, 예를 들어, 세포 또는 바이러스 구조적 성분 또는 폴리펩타이드와 관련하여 일반적으로 발견되는 기타 폴리펩타이드 또는 핵산으로부터 적어도 부분적으로 분리된 폴리펩타이드를 의미한다. 대표적인 구현예에서, "단리된" 폴리펩타이드는 출발 물질과 비교하여 적어도 약 10배, 100배, 1000배, 10,000배 이상 농축된다.
- [0086] "단리된 세포"는 자연 상태에서 일반적으로 결합되는 다른 성분으로부터 분리된 세포를 지칭한다. 예를 들어, 단리된 세포는 배양 배지 중의 세포 및/또는 본 발명의 약제학적으로 허용되는 담체 중의 세포일 수 있다. 따라서, 단리된 세포는 대상체에 전달 및/또는 도입될 수 있다. 일부 구현예에서, 단리된 세포는 대상체로부터 제거되고 생체외에서 본원에 기재된 바와 같이 조작된 다음, 대상체로 복귀되는 세포일 수 있다.
- [0087] 비리온의 집단은 본원에 기재된 임의의 방법에 의해 생성될 수 있다. 하나의 구현예에서, 집단은 적어도 10^1 개 비리온이다. 하나의 구현예에서, 집단은 적어도 10^2 개 비리온, 적어도 10^3 개 비리온, 적어도 10^4 개 비리온, 적어도 10^5 개 비리온, 적어도 10^6 개 비리온, 적어도 10^7 개 비리온, 적어도 10^8 개 비리온, 적어도 10^9 개 비리온, 적어도 10^{10} 개 비리온, 적어도 10^{11} 개 비리온, 적어도 10^{12} 개 비리온, 적어도 10^{13} 개 비리온, 적어도 10^{14} 개 비리온, 적어도 10^{15} 개 비리온, 적어도 10^{16} 개 비리온, 또는 적어도 10^{17} 개 비리온이다. 비리온의 집단은 이질적일 수 있거나 균질할 수 있다(예를 들어, 실질적으로 균질하거나 완전히 균질함).
- [0088] 본원에 사용된 용어로서 "실질적으로 균질한 집단"은 그 안에 오염된 비리온(동일하지 않은 것들)이 거의 내지 전혀 없는 대부분이 동일한 비리온의 집단을 지칭한다. 실질적으로 균질한 집단은 동일한 비리온(예를 들어, 원하는 비리온)의 적어도 90%이고, 동일한 비리온의 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 적어도 99.9%일 수 있다.
- [0089] 완전히 균질한 비리온 집단은 동일한 비리온만을 함유한다.
- [0090] 본원에 사용된 바와 같이, 바이러스 벡터 또는 바이러스 입자 또는 바이러스 입자 집단을 "단리하다" 또는 "정제하다" (또는 문법적 등가물)는 바이러스 벡터 또는 바이러스 입자 또는 바이러스 입자 집단이 출발 물질의 다른 성분 중 적어도 일부로부터 적어도 부분적으로 분리된다는 것을 의미한다. 대표적인 구현예에서, "단리된" 또는 "정제된" 바이러스 벡터 또는 바이러스 입자 또는 바이러스 입자 집단은 출발 물질과 비교하여 적어도 약 10배, 100배, 1000배, 10,000배 이상 농축된다.
- [0091] 달리 표시되지 않는 한, "효율적인 형질도입" 또는 "효율적인 친화성" 또는 유사한 용어는 적합한 대조군(예를 들어, 대조군의 형질도입 또는 친화성 각각의 적어도 약 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, 125%, 150%, 175%, 200%, 250%, 300%, 350%, 400%, 500% 이상)을 참조하여 결정될 수 있다. 특정 구현예에서, 바이러스 벡터는 신경 세포 및 심근세포에 대해 효율적으로 형질도입하거나 효율적인 친화성을 갖는다. 적합한 대조군은 원하는 친화성 및/또는 형질도입 프로파일을 포함한 다양한 인자에 따라 달라질 것이다.
- [0092] "치료용 폴리펩타이드"는 세포 또는 대상체에서 단백질의 부재 또는 결함으로부터 초래되는 증상을 완화, 감소, 예방, 지연 및/또는 안정화할 수 있는 폴리펩타이드 및/또는 그렇지 않으면 대상체에게 이점, 예를 들어, 질환

의 증상을 감소 또는 제거하기 위한 효소 대체, 또는 이식 생존가능성의 개선 또는 면역 반응의 유도를 부여하는 폴리펩타이드이다.

- [0093] 용어 "이종 뉴클레오타이드 서열" 및 "이종 핵산 분자"는 본원에서 상호교환적으로 사용되며 바이러스에서 자연적으로 발생하지 않는 핵산 서열을 지칭한다. 일반적으로, 이종 핵산 분자 또는 이종 뉴클레오타이드 서열은 (예를 들어, 세포 및/또는 대상체로의 전달을 위해) 관심 있는 폴리펩타이드 및/또는 비번역 RNA, 예를 들어, GAA 폴리펩타이드를 암호화하는 오픈 리딩 프레임에 포함한다.
- [0094] 본원에 사용된 바와 같이, 용어 "바이러스 벡터," "벡터" 또는 "유전자 전달 벡터"는 핵산 전달 비히클로 기능하고 비리온 내에 패키징된 벡터 게놈(예를 들어, 바이러스 DNA[vDNA])을 포함하는 바이러스(예를 들어, AAV) 입자를 지칭한다. 대안적으로, 일부 맥락에서, 용어 "벡터"는 벡터 게놈/vDNA만을 지칭하기 위해 사용될 수 있다.
- [0095] "rAAV 벡터 게놈" 또는 "rAAV 게놈"은 하나 이상의 이종 핵산 서열을 포함하는 AAV 게놈(즉, vDNA)이다. rAAV 벡터는 일반적으로 바이러스를 생성하기 위해 시스템에서 역 말단 반복체(들) (TR(들))만 필요하다. 다른 모든 바이러스 서열은 필요 없고 트랜스로 공급될 수 있다(Muzyczka, (1992) *Curr. Topics Microbial. Immunol.* 158:97). 전형적으로, rAAV 벡터 게놈은 벡터에 의해 효율적으로 패키징될 수 있는 전이유전자의 크기를 최대화하기 위해 하나 이상의 TR 서열만을 보유할 것이다. 구조적 및 비구조적 단백질 암호화 서열은 트랜스로(예를 들어, 플라스미드와 같은 벡터로부터, 또는 서열을 패키징 세포에 안정적으로 통합함으로써) 제공될 수 있다. 본 발명의 구현예에서, rAAV 벡터 게놈은 적어도 하나의 ITR 서열(예를 들어, AAV TR 서열), 임의로 2개의 ITR(예를 들어, 2개의 AAV TR)을 포함하고, 이는 전형적으로 벡터 게놈의 5' 및 3' 말단에 있고 이중성 핵산의 측면에 위치할 것이지만 이에 인접할 필요는 없다. TR은 서로 동일하거나 상이할 수 있다.
- [0096] 용어 "말단 반복체" 또는 "TR"은 헤어핀 구조를 형성하고 역 말단 반복체(즉, 복제, 바이러스 패키징, 통합 및/또는 프로바이러스 구제 등과 같은 원하는 기능을 매개하는 ITR)로서 기능하는 임의의 바이러스 말단 반복체 또는 합성 서열을 포함한다. TR은 AAV TR 또는 비-AAV TR일 수 있다. 예를 들어, 다른 파르보바이러스(예를 들어, 개 파르보바이러스(CPV), 마우스 파르보바이러스(MVM), 인간 파르보바이러스 B-19)의 것과 같은 비-AAV TR 서열 또는 임의의 다른 적합한 바이러스 서열(예를 들어, SV40 복제의 기원으로서 작용하는 SV40 헤어핀)은 절단, 치환, 결실, 삽입 및/또는 추가에 의해 추가로 변형될 수 있는 TR로서 사용될 수 있다. 또한, TR은 Samulski 등의 미국 특허 번호 5,478,745에 기재된 "이중-D 서열"과 같이 부분적으로 또는 완전히 합성될 수 있다.
- [0097] "AAV 역 말단 반복체" 또는 "AAV ITR"을 포함한 "AAV 말단 반복체" 또는 "AAV TR"은 혈청형 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 또는 12 또는 현재 알려져 있거나 나중에 발견되는 임의의 다른 AAV를 포함하지만 이에 제한되지 않는 임의의 AAV로부터 유래할 수 있다. AAV 말단 반복체는 말단 반복체가 원하는 기능, 예를 들어, 복제, 바이러스 패키징, 통합 및/또는 프로바이러스 구제 등을 매개하는 한, 천연 말단 반복 서열을 가질 필요가 없다(예를 들어, 천연 AAV TR 또는 AAV ITR 서열은 삽입, 결실, 절단 및/또는 미스센스 돌연변이에 의해 변경될 수 있다).
- [0098] AAV 단백질 VP1, VP2 및 VP3은 서로 상호작용하여 20면체 대칭의 AAV 캡시드를 형성하는 캡시드 단백질이다. VP1.5는 미국 공개 번호 2014/0037585에 기재된 AAV 캡시드 단백질이다.
- [0099] 본 발명의 바이러스 벡터는 추가로 국제 특허 공개 WO 00/28004 및 문헌(Chao et al., (2000) *Molecular Therapy* 2:619)에 기재된 바와 같은 "표적화단" 바이러스 벡터(예를 들어, 유도 친화성을 가짐) 및/또는 "하이브리드" 파르보바이러스(즉, 바이러스 TR 및 바이러스 캡시드가 상이한 파르보바이러스에서 유래함)일 수 있다.
- [0100] 본 발명의 바이러스 벡터는 추가로 국제 특허 공개 WO 01/92551(그 개시내용은 그 전체가 본원에 참조로 포함됨)에 기재된 바와 같은 듀플렉싱된 파르보바이러스 입자일 수 있다. 따라서, 일부 구현예에서, 이종 가닥(듀플렉스) 게놈은 본 발명의 바이러스 캡시드 내로 패키징될 수 있다.
- [0101] 또한, 바이러스 캡시드 또는 게놈 요소는 삽입, 결실 및/또는 치환을 포함한 다른 변형을 포함할 수 있다.
- [0102] 본원에 사용된 "키메라" 캡시드 단백질은 야생형에 비해 아미노산 서열에서 하나 이상의 (예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10개 등) 아미노산 잔기의 삽입 및/또는 결실뿐만 아니라, 야생형에 비해 캡시드 단백질의 아미노산 서열에서 하나 이상의 (예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10개 등) 아미노산 잔기의 치환에 의해 변형된 AAV 캡시드 단백질(예를 들어, VP1, VP2 또는 VP3 중 임의의 하나 이상)을 의미한다. 일부 구현예에서, 하나의 AAV 혈청형으로부터의 완전 또는 부분 도메인, 기능적 영역, 에피토프 등은 본 발명의 키메라 캡시드 단백

질을 생성하기 위해 임의의 조합으로 상이한 AAV 혈청형의 상응하는 야생형 도메인, 기능적 영역, 에피토프 등을 대체할 수 있다. 키메라 캡시드 단백질의 생성은 당업계에 잘 알려진 프로토콜에 따라 수행될 수 있고 상당한 수의 키메라 캡시드 단백질이 본 발명의 캡시드에 포함될 수 있는 문헌 및 본원에 기재되어 있다.

[0103] 본원에 사용된 바와 같이, 용어 "반수체 AAV"는 그 전체가 본원에 참조로 포함된 국제 출원 W02018/170310 또는 미국 출원 US2018/037149에 기재된 바와 같은 AAV를 의미할 것이다. 일부 구현예에서, 비리온의 집단은 비리온 입자를 작제할 수 있는 반수체 AAV 집단이고, 여기서, AAV 캡시드 단백질, VP1, VP2 및 VP3으로 이루어진 군으로부터 적어도 하나의 바이러스 단백질은 AAV 게놈을 캡슐화할 수 있는 비리온 입자를 형성하는 데 필요한 다른 바이러스 단백질 중 적어도 하나와 상이하다. 존재하는 각 바이러스 단백질(VP1, VP2 및/또는 VP3)의 경우, 해당 단백질은 동일한 유형이다(예를 들어, 모두 AAV2 VP1). 하나의 예에서, 바이러스 단백질 중 적어도 하나는 키메라 바이러스 단백질이고 다른 2개의 바이러스 단백질 중 적어도 하나는 키메라가 아니다. 하나의 구현예에서, VP1 및 VP2는 키메라이고, VP3만이 키메라가 아니다. 예를 들어, 단지 바이러스 입자는 AAV2로부터 단지 VP3과 쌍을 이루는 키메라 AAV2/8(AAV2의 N-말단 및 AAV8의 C-말단); 또는 단지 AAV2로부터 단지 VP3과 쌍을 이루는 키메라 VP1/VP2 28m-2P3(VP3 개시 코돈의 돌연변이가 없는 AAV8로부터의 N-말단 및 AAV2로부터의 C-말단)로부터 VP1/VP2로 구성된다. 또 다른 구현예에서, VP3만이 키메라이고, VP1 및 VP2는 키메라가 아니다. 또 다른 구현예에서, 바이러스 단백질 중 적어도 하나는 완전히 상이한 혈청형으로부터 유래한다. 예를 들어, 키메라 VP1/VP2 28m-2P3만 단지 AAV3으로부터 VP3과 쌍을 이룬다. 또 다른 예에서, 키메라가 존재하지 않는다.

[0104] 용어 "하이브리드" AAV 벡터 또는 파르보바이러스는 바이러스 TR 또는 ITR 및 바이러스 캡시드가 상이한 파르보 바이러스로부터 유래된 rAAV 벡터를 지칭한다. 하이브리드 벡터는 국제 특허 공개 WO 00/28004 및 문헌(Chao et al., (2000) Molecular Therapy 2:619)에 기재되어 있다. 예를 들어, 하이브리드 AAV 벡터는 전형적으로 아데노바이러스 복제 및 패키징에 충분한 아데노바이러스 5' 및 3' 시스 ITR 서열(즉, 아데노바이러스 말단 반 복제 및 PAC 서열)을 포함한다.

[0105] 용어 "배수체(polyploid) AAV"는, 예를 들어, 2개 이상의 AAV 혈청형으로부터의 캡시드로 구성되고 더 높은 형질도입을 위해 개별 혈청형으로부터 이점을 취할 수 있지만 특정 구현예에서 모체로부터 친화성을 제거하지 않는 AAV 벡터를 지칭한다.

[0106] 본원에 사용된 용어 "GAA" 또는 "GAA 폴리펩타이드"는 성숙(~76 또는 ~67 kDa) 및 전구체(예를 들어, ~110 kDa) GAA 뿐만 아니라 생물학적 기능(즉, 천연 GAA 단백질의 적어도 하나의 생물학적 활성을 가지며, 예를 들어, 상기 정의된 바와 같이 글리코젠을 가수분해할 수 있음) 및 GAA 변이체(예를 들어, 각각은 그 전체가 본원에 참조로 포함된 문헌(Kunita et al., (1997) Biochemica et Biophysica Acta 1362:269; GAA polymorphisms and SNPs are described by Hirschhorn, R. and Reuser, A. J. (2001) in The Metabolic and Molecular Basis for Inherited Disease (Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D. Eds.), pp. 3389-3419, McGraw-Hill, New York, see pages 3403-3405)에 기재된 바와 같은 GAA II)를 보유하는 변형된(예를 들어, 삽입(들), 결실(들) 및/또는 치환(들)에 의해) GAA 단백질 또는 이의 단편을 포함한다. 당업계에 알려진 임의의 GAA 암호화 서열이 사용될 수 있으며, 예를 들어, 도 8 및 9의 암호화 서열; GenBank Accession number NM_00152 및 문헌(Hoefsloot et al., (1988) EMBO J. 7:1697 and Van Hove et al., (1996) Proc. Natl. Acad. Sci. USA 93:65 (human), GenBank Accession number NM_008064 (mouse), 및 Kunita et al., (1997) Biochemica et Biophysics Acta 1362:269 (quail))을 참조하고; 그 개시내용은 GAA 암호화 및 비암호화 서열의 교시를 위해 본원에 참조로 포함된다.

[0107] 용어 "양이온-비의존적 만노오스-6-포스페이트 수용체(CI-MPR)," "M6P/IGF-II 수용체," "CI-MPR/IGF-II 수용체," "IGF-II 수용체" 또는 "IGF2 수용체" 또는 이의 약어는 M6P 및 IGF-II 둘 모두에 결합하는 세포 수용체를 지칭하는 본원에서 상호교환적으로 사용된다.

[0108] 용어 "표적화 펩타이드"는 또한 본원에서 사용된 "표적화 서열"로도 지칭되며, 특정 세포내(subcellular) 구획, 예를 들어 포유동물의 리소좀을 표적화하는 펩타이드를 지칭하는 것으로 의도된다. 본원에서 사용하기 위해 포함된 표적화 펩타이드는 만노오스-6-포스페이트-비의존적인 리소좀 표적화 펩타이드이다. 예시적인 표적화 서열은 본원에 개시된 바와 같은 IGF2 표적화 펩타이드이다.

[0109] 용어 "IGF2 서열"은 "IGF2 표적화 서열" 또는 "IGF2 리더 서열" 및 "IGF2 표적화 펩타이드"와 함께 사용되며 본원에서 상호교환적으로 사용되며 세포 표면 상의 CI-MBR에 결합하는 IGF2 폴리펩타이드의 서열을 지칭한다. 특히, IGF2 서열은 서열 번호 5의 IGF2 흡수 서열의 일부를 포함하거나 서열 번호 5의 아미노산의 변형을 포함하는 펩타이드이다. IGF2 표적화 펩타이드는 인간 양이온-비의존적 만노오스-6-포스페이트 수용체(CI-MPR 또는

CA-M6P 수용체)의 반복체 11-12, 반복체 11 또는 아미노산 1508-1566으로 본질적으로 이루어진 수용체 도메인에 결합하는 펩타이드 서열을 지칭한다.

[0110] 용어 "리더 서열"은 본원에서 "분비 신호 서열" 또는 "신호 서열" 또는 "신호 펩타이드" 또는 이의 변형이라는 용어와 상호교환적으로 사용되고, 천연 폴리펩타이드로 관찰되는 분비 수준과 비교하여 세포로부터 작동 가능하게 연결된 폴리펩타이드(예를 들어, GAA 펩타이드 또는 IGF2-GAA 융합 단백질)의 분비(상기 정의된 바와 같음)를 향상시키는 기능을 하는 아미노산 서열을 지칭하도록 의도된다. 상기 정의된 바와 같이, "향상된" 분비란, 세포로부터 분비되는 세포에 의해 합성된 리소좀 폴리펩타이드의 상대적 비율이 증가된다는 것을 의미하고; 분비된 단백질의 절대량도 증가할 필요는 없다. 본 발명의 특정 구현예에서, GAA-폴리펩타이드의 본질적으로 전부(즉, 적어도 95%, 97%, 98%, 99% 이상)가 분비된다. 그러나, 분비 수준이 천연 GAA 폴리펩타이드와 비교하여 향상되는 한, GAA 폴리펩타이드의 본질적으로 전부 또는 심지어 대부분이 분비될 필요는 없다. 예시적인 리더 서열은 본원에 개시된 바와 같은 고유 GAA 리더 서열(동족 GAA 리더 서열로도 지칭됨), AAT 서열, IL2(1-3), IL2 리더 서열(IL2 wt), 변형된 IL2 리더 서열(IL2 mut), 피브로넥틴(FN1, FBN으로도 지칭됨), 또는 IgG 리더 서열 또는 이의 기능적 변이체를 포함하지만 이에 제한되지 않는다.

[0111] 본원에 사용된 바와 같이, 용어 "아미노산"은 임의의 자연적으로 발생하는 아미노산, 이의 변형된 형태, 및 합성 아미노산을 포함한다. 자연적으로 발생하는 좌선성(L-) 아미노산은 그 전체가 본원에 포함된 미국 공개 2018/0371496의 표 2에 개시되어 있다. 대안적으로, 아미노산은 변형된 아미노산 잔기일 수 있고/있거나(비제한적인 예는 미국 공개 2018/0371496의 표 4에 나타냄) 번역 후 변형(예를 들어, 아세틸화, 아미드화, 포르밀화, 하이드록실화, 메틸화, 인산화 또는 황산화)에 의해 변형된 아미노산일 수 있다. 또한, 비자연 발생 아미노산은 문헌(Wang et al., Annu Rev Biophys Biomol Struct. 35:225-49 (2006))에 기재된 바와 같이 "비천연" 아미노산일 수 있다. 이들 비천연 아미노산은 AAV 캡시드 단백질에 관심 있는 분자를 화학적으로 연결하는데 유리하게 사용될 수 있다.

[0112] 추가로 설명하기 위해, 예를 들어, 특정 아미노산이 A, G, I, L 및/또는 V로부터 선택될 수 있음을 명세서가 나타내는 경우, 이 언어는 또한 각각의 이러한 하위조합이 본원에 명시적으로 명시된 것처럼 아미노산이 이러한 아미노산(들)의 임의의 하위세트, 예를 들어, A, G, I 또는 L; A, G, I 또는 V; A 또는 G; 단지 L로부터 선택될 수 있음을 나타낸다. 더욱이, 이러한 언어는 또한 명시된 아미노산 중 하나 이상이 거부될 수 있음을 나타낸다(예를 들어, 부정 단서에 의해). 예를 들어, 특정 구현예에서, 이러한 각 가능한 거부가 본원에 명시적으로 개시된 것처럼, 아미노산은 A, G 또는 I가 아니고; A가 아니고; G 또는 V 등이 아니다.

[0113] 용어 "시스-조절 요소" 또는 "CRE"는 당업자에게 잘 알려진 용어이며, 인접 유전자(즉, 시스로)의 전사를 조절(regulate)하거나 조절(modulate)할 수 있는 인핸서, 프로모터, 인슐레이터 또는 사이린서와 같은 핵산 서열을 의미한다. CRE는 이들이 조절하는 유전자 근처에서 발견된다. CRE는 전형적으로 전사 인자(TF)에 결합하여 유전자 전사를 조절하는데, 즉, 이들은 TF 결합 부위(TFBS)를 포함한다. 단일 TF는 많은 CRE에 결합할 수 있으며 따라서 많은 유전자의 발현(다면발현(pleiotropy))을 제어할 수 있다. CRE는 항상 그런 것은 아니지만 일반적으로 이들이 조절하는 유전자의 전사 개시 부위(TSS)의 업스트림에 위치한다. "인핸서"는 작동적으로 회합된 유전자의 전사를 향상(즉, 상향조절)하는 CRE이며, 이들이 조절하는 유전자의 업스트림, 다운스트림, 및 심지어 인트론 내에서도 발견할 수 있다. 다중 인핸서는 하나의 유전자의 전사를 조절하기 위해 조정된 방식으로 작용할 수 있다. 이 맥락에서 "사이린서"는 유전자의 전사를 방지하거나 하향조절하는 역할을 하는 리프레서(repressor)로 불리는 TF에 결합하는 CRE와 관련이 있다. 용어 "사이린서"는 또한 mRNA 분자의 번역을 억제하는 단백질에 결합하는 메신저 RNA의 3' 비번역 영역에서의 영역을 지칭할 수 있지만, 이 사용법은 CRE를 설명할 때 사용하는 것과는 구별된다. 일반적으로, 본 발명의 CRE는 간-특이적 인핸서(중종 간-특이적 CRE 또는 간-특이적 CRE 인핸서 등으로 지칭됨)이다. 본 맥락에서, CRE는 전사 개시 부위(TSS)로부터 1500개 이하의 뉴클레오타이드, 보다 바람직하게는 TSS로부터 1000개 이하의 뉴클레오타이드, 더욱 바람직하게는 TSS로부터 500개 이하의 뉴클레오타이드, 적합하게는 TSS로부터 250개, 200개, 150개 또는 100개 이하의 뉴클레오타이드에 위치하는 것이 바람직하다. 본 발명의 CRE는 바람직하게는 길이가 비교적 짧고, 바람직하게는 100개 이하의 뉴클레오타이드 길이이고, 예를 들어, 이들은 90, 80, 70, 60개 이하의 뉴클레오타이드 길이일 수 있다.

[0114] 용어 "시스-조절 모듈" 또는 "CRM"은 2개 이상의 CRE로 구성된 기능 모듈을 의미하고; 본 발명에서 CRE는 전형적으로 간-특이적 인핸서이다. 따라서, 본 출원에서, CRM은 전형적으로 복수의 간-특이적 인핸서 CRE를 포함한다. 전형적으로, CRM 내의 다중 CRE는 함께 작용하여(예를 들어, 부가적으로 또는 상승적으로) CRM이 작동적으로 회합된 유전자의 전사를 향상시킨다. CRM 내에서 CRE의 서플(즉, 재정렬), 반전(즉, 역방향) 및 간격 변경에 대한 보존 가능한 범위가 있다. 따라서, 본 발명의 CRM의 기능적 변이체는 참조된 CRM의 변이체를

포함하며, 여기서, CRE는 서플 및/또는 반전되고/되거나 CRE 사이의 간격이 변경되었다.

- [0115] 본원에 사용된 바와 같이, 문구 "프로모터"는 일반적으로 전사가 발생하기 위해 필요한, 즉 전사를 개시하는 데 필요한 전사될 핵산 서열의 업스트림에 위치하는 DNA의 영역을 지칭한다. 프로모터는 이들의 제어 하에 암호화 서열의 전사의 적절한 활성화 또는 억제를 허용한다. 프로모터는 전형적으로 복수의 TF에 의해 인식되고 결합되는 특정 서열을 포함한다. TF는 프로모터 서열에 결합하여 유전자의 암호화 영역으로부터 RNA를 합성하는 효소인 RNA 폴리머라제를 동원한다. 매우 많은 프로모터가 당업계에 알려져 있다.
- [0116] 본원에 사용된 용어 "합성 프로모터"는 자연에서 발생하지 않는 프로모터에 관한 것이다. 본 문맥에서, 이는 전형적으로 최소 (또는 코어) 프로모터 또는 간-특이적 근위 프로모터에 작동적으로 연결된 본 발명의 합성 CRE 및/또는 CRM을 포함한다. 본 발명의 CRE 및/또는 CRM은 프로모터에 작동적으로 연결된 유전자의 간-특이적 전사를 향상시키는 역할을 한다. 합성 프로모터의 일부는 자연적으로 발생할 수 있지만(예를 들어, 최소 프로모터 또는 프로모터에서 하나 이상의 CRE), 완전한 실체로서의 합성 프로모터는 자연적으로 발생하지 않는다.
- [0117] 본원에 사용된 바와 같이, "최소 프로모터"(코어 프로모터"로도 알려져 있음)는 그 자체로 불활성이거나 대체로 불활성이지만 다른 전사 조절 요소와 조합될 때 전사를 매개할 수 있는 짧은 DNA 분절을 지칭한다. 최소 프로모터 서열은 원핵생물 및 진핵생물 유전자를 포함한 상이한 공급원으로부터 유래될 수 있다. 최소 프로모터의 예는 상기 논의되었으며, 도파민 베타-하이드록실라제 유전자 최소 프로모터, 거대세포바이러스(CMV) 즉시 초기 유전자 최소 프로모터(CMV-MP) 및 헤르페스 티미딘 키나제 최소 프로모터(MinTK)를 포함한다. 최소 프로모터는 전형적으로 전사 개시 부위(TSS) 및 업스트림 바로 다음에 있는 요소, RNA 폴리머라제 II에 대한 결합 부위 및 일반적인 전사 인자 결합 부위(종종 TATA 박스)를 포함한다.
- [0118] 본원에 사용된 바와 같이, "근위 프로모터"는 최소 프로모터와 일차 조절 요소를 포함하는 경향이 있는 유전자의 업스트림의 근위 서열에 관한 것이다. 종종 TSS의 업스트림에서 대략 250개의 염기 쌍을 확장하며 특정 TFBS를 포함한다. 본 경우에, 근위 프로모터는 적합하게는 본 발명의 하나 이상의 CRE 또는 CRM과 조합될 수 있는 자연적으로 발생하는 간-특이적 근위 프로모터이다. 그러나, 근위 프로모터는 합성될 수 있다.
- [0119] 본 발명의 맥락에서 시스-조절 요소(CRE), 시스-조절 모듈(CRM), 프로모터 또는 기타 핵산 서열의 "기능적 변이체"는 참조 서열과 동일한 방식으로, 예를 들어, 간-특이적 시스-조절 인핸서 요소, 간-특이적 시스-조절 모듈 또는 간-특이적 프로모터로서 기능하는 능력을 보유하는 참조 서열의 변이체이다. 이러한 기능적 변이체에 대한 대체 용어는 "생물학적 등가물" 또는 "등가물"을 포함한다.
- [0120] 간-특이적 인핸서로서 기능하는 주어진 시스-조절 요소의 능력은 주로 참조 서열에 결합하는 동일한 간-특이적 전사 인자(TF)에 결합하는 서열의 능력에 의해 결정되는 것으로 이해될 것이다. 따라서, 대부분의 경우에, 시스-조절 요소의 기능적 변이체는 참조 시스-조절 요소와 동일한 TF에 대한 TFBS를 포함할 것이다. 기능적 변이체의 전사 인자 결합 부위(TFBS)가 참조 시스-조절 요소와 동일한 상대 위치(즉, 순서)에 있는 것이 바람직하지만 필수적인 것은 아니다. 기능적 변이체의 TFBS가 참조 서열과 동일한 방향에 있는 것이 바람직하지만 필수적인 것은 아니다(TFBS는, 일부 경우에, 예를 들어, 참조 서열에서 서열에 관한 역보체로서, 역방향으로 존재할 수 있음에 주목할 것이다). 또한, 기능적 변이체의 TFBS가 참조 서열과 동일한 가닥에 있는 것이 바람직하지만 필수적인 것은 아니다. 따라서, 바람직한 구현예에서, 기능적 변이체는 참조 서열과 동일한 순서로 동일한 배향으로 그리고 동일한 가닥 상에, 동일한 TF에 대해 TFBS를 포함한다. 또한 TFBS 사이에 있는 서열(일부 경우에 스페이서 서열 등으로 지칭됨)은 시스-조절 요소의 기능에 덜 중요하다는 것이 이해될 것이다. 이러한 서열은 전형적으로 상당히 다양할 수 있으며 이들의 길이가 변경될 수 있다. 그러나, 바람직한 구현예에서, 간격(즉, 인접한 TFBS 사이의 거리)은 참조 서열에서와 같이 기능적 변이체에서 실질적으로 동일하다(예를 들어, 이는 20% 초과, 바람직하게는 10% 이하로 변하지 않으며, 보다 바람직하게는 동일함). 일부 경우에 시스-조절 인핸서 요소의 기능적 변이체가 역방향으로 존재할 수 있음이 명백할 것이며, 예를 들어, 이는 상기 기재된 바와 같은 시스-조절 인핸서 요소의 역보체 또는 이의 변이체일 수 있다.
- [0121] 기능적 변이체와 참조 서열 사이의 서열 동일성의 수준은 또한 지표 또는 보유된 기능일 수 있다. 시스-조절 요소의 TFBS에서 높은 수준의 서열 동일성은 일반적으로 스페이서 서열(서열의 임의의 보존에 대한 요구 사항이 거의 또는 전혀 없는 경우)에서 서열 동일성보다 더 중요하다. 그러나, 기능적 TFBS의 서열이 컨센서스 서열과 정확히 일치할 필요가 없다면 TFBS 내에서도 상당한 정도의 서열 변이가 수용될 수 있음을 이해할 것이다.
- [0122] 주어진 기능적 변이체에서 TFBS에 결합하는 하나 이상의 TF의 능력은 전기이동성 이동 검정(EMSA), 결합 검정, 염색질 면역침전(ChIP), 및 ChIP-시퀀싱(ChIP-seq)을 포함하지만 이에 제한되지 않는 당업계에 알려진 임의의

관련 수단에 의해 결정될 수 있다. 바람직한 구현예에서, 주어진 기능적 변이체에 결합하는 하나 이상의 TF의 능력은 EMSA에 의해 결정된다. EMSA를 수행하는 방법은 당업계에 잘 알려져 있다. 적합한 접근법은 상기 인용한 Sambrook 등에 설명되어 있다. 이 절차를 설명하는 많은 관련 문서를 사용할 수 있다(예를 들어, Hellman and Fried, Nat Protoc. 2007; 2(8): 1849-1861).

[0123] 프로모터와 관련하여 용어 "간-특이적" 또는 "간-특이적 발현"은 다른 조직(예를 들어, 비장, 근육, 심장, 폐, 및 뇌)과 비교하여 우선적 또는 우세한 방식으로 간에서 (또는 간 유래 세포에서) 유전자의 발현을 향상시키거나 유도하는 시스-조절 요소, 시스-조절 모듈 또는 프로모터의 능력을 지칭한다. 유전자의 발현은 mRNA 또는 단백질의 형태일 수 있다. 일부 구현예에서, 간-특이적 발현은 다른(즉, 비-간) 조직 또는 세포에서 무시할 수 있는 발현이 있도록 하는 것으로, 즉, 발현은 매우 간 특이적이다. 일부 구현예에서, 간-특이적 프로모터는 간에서 우선적으로 발현을 유도하지만, 낮은 수준에서 관심 있는 또 다른 조직, 예를 들어, 근육에서도 유전자의 발현을 유도할 수 있다.

[0124] 간-특이적 시스-조절 인핸서 요소로서 기능하는 시스-조절 요소의 능력은 당업자에 의해 쉽게 평가될 수 있다. 따라서, 당업자는 상기 인용된 특정 시스-조절 요소의 임의의 변이체가 기능적으로 남아 있는지 여부를 쉽게 결정할 수 있다(즉, 상기 정의된 바와 같은 기능적 변이체임). 예를 들어, 평가될 임의의 주어진 시스-조절 요소는 최소 프로모터(예를 들어, CMV-MP의 업스트림에 위치함)에 작동적으로 연결될 수 있고 유전자(전형적으로 리포터 유전자)의 간-특이적 발현을 유도하는 시스-조절 요소의 능력이 측정된다. 대안적으로, 시스-조절 인핸서 요소의 변이체는 참조 시스-조절 인핸서 요소 대신 합성 간-특이적 프로모터로 치환될 수 있고, 상기 변형된 프로모터에 의해 구동되는 간-특이적 발현에 대한 효과는 결정되고 비변형된 형태와 비교할 수 있다. 유사하게는, 간-특이적 발현을 유도하는 시스-조절 모듈 또는 프로모터의 능력은 (예를 들어, 하기 실시예에 기재된 바와 같이) 당업자에 의해 쉽게 평가될 수 있다. 참조 프로모터의 변이체에 의해 구동되는 유전자의 발현 수준은 참조 서열에 의해 구동되는 발현 수준과 비교될 수 있다. 일부 구현예에서, 변이체 프로모터에 의해 구동되는 간-특이적 발현 수준이 참조 프로모터에 의해 구동되는 발현 수준의 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 90%, 또는 적어도 100%인 경우, 변이체는 기능적으로 남아 있다고 말할 수 있다. 간-특이적 발현 향상을 평가하기 위한 적합한 핵산 작제물 및 리포터 검정은 쉽게 작제될 수 있으며 아래에 제시된 예는 적합한 방법론을 제공한다.

[0125] 간-특이성은 유전자(예를 들어, 치료제 또는 리포터 유전자)의 발현이 간 유래 세포에서 우선적으로 또는 우세하게 발생하는 경우에 확인될 수 있다. 예를 들어, 발현 수준이 다른 유형의 세포(즉, 비-간 유래 세포)보다 간 유래 세포에서 상당히 더 큰 경우 우선적 또는 우세한 발현이 정의될 수 있다. 예를 들어, 간 유래 세포에서의 발현은 비-간 세포보다 적합하게는 적어도 5배 초과, 바람직하게는 비-간 세포보다 적어도 10배 초과, 일부 경우에는 50배 이상일 수 있다. 편의상, 간-특이적 발현은 콩팥 유래 세포주(예를 들어, HEK-293), 자궁경부 조직 유래 세포주(예를 들어, HeLa) 및/또는 폐 유래 세포주(예를 들어, A549)에서의 발현 수준과 비교하여 간 세포주(예를 들어, Huh7 및/또는 HepG2 세포와 같은 간 유래 세포주) 또는 간 1차 세포에서의 발현 수준의 비교를 통해 적합하게 입증될 수 있다.

[0126] 본 발명의 합성 간-특이적 프로모터는 바람직하게는 대상체의 간에서의 발현 촉진, 예를 들어, 전이유전자, 바람직하게는 치료적 전이유전자의 간-특이적 발현을 구동하는 데 적합하다. 일부 구현예에서, 본 발명의 간-특이적 프로모터는 서열 번호 432의 LP1 프로모터보다 적어도 1.5배 초과, 바람직하게는 LP1 프로모터보다 2배 초과, 보다 바람직하게는 LP1 프로모터보다 3배 초과, 및 보다 더 바람직하게는 LP1 프로모터(서열 번호 432)보다 5배 초과 수준으로 간-특이적 전이유전자 발현을 촉진하는데 적합하다. 이러한 발현은 간 유래 세포, 예를 들어, Huh7, 및/또는 HepG2 세포 또는 1차 간 세포(적합하게는 1차 인간 간세포)에서 결정된다. 일부 구현예에서, 본 발명의 합성 간-특이적 프로모터는 비-간 유래 세포(예를 들어, HEK-293, HeLa, 및/또는 A549 세포)에서 LP1 프로모터(서열 번호 432)보다 적어도 1.5배 미만인 수준으로 유전자 발현을 촉진하는데 적합하다.

[0127] 본 발명의 바람직한 합성 간-특이적 프로모터는 간-특이적 전이유전자 발현을 촉진하는 데 적합하고 간 세포에서 TBG 프로모터(서열 번호 435)의 활성의 적어도 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 125%, 150%, 175%, 200%, 250%, 300%, 350% 또는 400%의 활성을 갖는다.

[0128] 본 발명의 합성 간-특이적 프로모터는 바람직하게는 간-유래 세포에서 서열 번호 433의 CMV-IE 프로모터보다 적어도 1.5배 초과, 바람직하게는 간-유래 세포(예를 들어, HEK-293, HeLa, 및/또는 A549 세포)에서 CMV 프로모터보다 적어도 2배 초과 수준에서 간-특이적 발현을 촉진하는 데 적합하다. 본원에 개시된 합성 간 특이적 프

로모터는 LSP-H, LSP-M 및 LSP-L 프로모터일 수 있으며, 이는 간에서의 고발현, 중발현 및 저발현을 지칭하고, 일부 구현예에서, LSP-H, LSP-M 및 LSP-L은 간에서 단백질을 우선적으로 또는 우세하게 발현할 수 있지만, 또한 하나 이상의 다른 조직, 예를 들어, 근육 및/또는 뇌에서 단백질을 발현할 수 있다. 본원에 개시된 이러한 LSP-H, LSP-M 및 LSP-L 프로모터는 간에서 단백질의 적어도 90%, 또는 적어도 80%, 또는 적어도 70% 또는 적어도 60%, 또는 적어도 50%를 우선적으로 발현할 수 있으며, 또한 또 다른 조직, 예를 들어, 근육 조직에서 적어도 10%, 또는 적어도 20%, 또는 적어도 30%, 또는 적어도 40% 또는 적어도 50%를 발현한다. 일부 구현예에서, 예를 들어, 폼페 또는 리소좀 질환의 치료를 위해 본원에 개시된 방법 및 조성물에 유용한 LSP-H, LSP-M 및 LSP-L 프로모터는 간에서 우선적 또는 우세한 방식으로 유전자 발현을 유도하거나 향상시키지만 근육 조직에서 단백질의 적어도 일부를 발현할 수도 있다.

[0129] 용어 "동일성" 및 "동일한" 등은 2개의 중합체 분자 사이, 예를 들어, 2개의 DNA 분자 사이와 같은 2개의 핵산 분자 사이의 서열 유사성을 지칭한다. 서열 정렬 및 서열 동일성의 결정은, 예를 들어, 문헌(Tatusova and Madden 1999 (FEMS Microbiol Lett 174: 247-250))에 의해 기술된 "Blast 2 서열" 알고리즘과 같은 문헌(Altschul et al. 1990 (J Mol Biol 215: 403-10))에 의해 원래 기술된 기본 로컬 정렬 검색 도구(BLAST: Basic Local Alignment Search Tool)를 사용하여 수행할 수 있다.

[0130] 본원에 사용된 용어 "합성"은 자연에서 발생하지 않는 핵산 분자를 의미한다. 본 발명의 합성 핵산 발현 작제물은 전형적으로 재조합 기술에 의해 인공적으로 생성된다. 이러한 합성 핵산은 자연적으로 발생하는 서열(예를 들어, 프로모터, 인핸서, 인트론 및 기타 이러한 조절 서열)을 포함할 수 있지만, 이들은 비-자연적으로 발생하는 맥락에서 존재한다. 예를 들어, 합성 유전자 (또는 유전자의 일부)는 전형적으로 자연에서 인접하지 않은 하나 이상의 핵산 서열(키메라 서열)을 포함하고/하거나, 치환, 삽입, 결실 및 이들의 조합을 포함할 수 있다.

[0131] 본원에 사용된 "스페이서 서열" 또는 "스페이서"는 2개의 기능적 핵산 서열을 분리하는 핵산 서열이다. 기능적 핵산 서열(예를 들어, 시스-조절 요소)이 원하는 대로 기능하는 것을 방해하지 않는다면 본질적으로 임의의 서열을 가질 수 있다(예를 들어, 이것은 사일렌서 서열을 포함하거나 원하는 전사 인자의 결합을 방지하는 등의 경우에 발생할 수 있음). 전형적으로, 이것은 인접한 기능적 핵산 서열을 서로 이격시키기 위해서만 존재하기 때문에 비기능적이다.

[0132] 본원에 사용된 용어 "약제학적으로 허용되는"은 당업계와 일치하며, 약제학적 조성물의 다른 성분과 상용성이며 이의 수용자에게 유해하지 않음을 의미한다.

[0133] 용어 "치료하다", "치료하는" 또는 "~의 치료" (및 이의 문법적 변형)란, 대상체의 병태의 중증도가 감소, 적어도 부분적으로 개선 또는 안정화되고/되거나 적어도 하나의 임상 증상에서 약간의 완화, 경감, 감소 또는 안정화가 달성되고/되거나 질환 또는 장애의 진행이 지연됨을 의미한다.

[0134] 용어 "예방하다", "예방하는" 및 "예방" (및 이의 문법적 변형)은 대상체에서 질환, 장애 및/또는 임상 증상(들)의 발병의 예방 및/또는 지연하고/하거나, 본 발명의 방법의 부재시 발생할 수 있는 것과 비교하여 질환, 장애 및/또는 임상 증상(들)의 발병 중증도의 감소를 지칭한다. 예방은, 예를 들어, 질환, 장애 및/또는 임상 증상(들)의 완전한 부재와 같이 완전할 수 있다. 예방은 또한 부분적일 수 있어, 대상체에서 질환, 장애 및/또는 임상 증상(들)의 발생 및/또는 발병의 중증도가 본 발명이 없을 때 발생하는 것보다 실질적으로 적다.

[0135] 본원에 사용된 "치료 유효" 양은 대상체에게 일부 개선 또는 이점을 제공하기에 충분한 양이다. 대안적으로 언급하면, "치료 유효" 양은 대상체에서 적어도 하나의 임상 증상에서 약간의 완화, 경감, 감소 또는 안정화를 제공할 양이다. 당업자는 일부 이점이 대상체에게 제공되는 한 치료 효과가 완전하거나 치유적일 필요는 없음을 인식할 것이다.

[0136] 본원에 사용된 "예방 유효" 양은 대상체에서 질환, 장애 및/또는 임상 증상의 발병을 예방 및/또는 지연하고/하거나, 본 발명의 방법의 부재시 발생할 수 있는 것과 비교하여 대상체에서 질환, 장애 및/또는 임상 증상의 발병의 중증도를 감소 및/또는 지연시키기에 충분한 양이다. 당업자는 일부 예방 이점이 대상체에게 제공되는 한 예방 수준이 완전할 필요는 없음을 인식할 것이다.

[0137] 문구 "치료학적 유효량" 및 유사한 문구는, 예를 들어, 간에서 치료 유전자를 발현하고 혈장으로 분비하기 위해 원하는 특정 약리학적 효과를 제공하는 대상체에서의 용량 또는 혈장 농도를 의미한다. 비록 이러한 투여량이 당업자에 의해 치료학적 유효량인 것으로 간주되더라도, 치료학적 유효량은 본원에 기재된 병태를 치료하는데 항상 효과적인 것은 아닐 수 있다는 점을 강조한다. 치료학적 유효량은 투여 경로 및 투여 형태, 대상체의 연

령 및 체중 및/또는 치료되는 질환 또는 병태에 기반하여 달라질 수 있다.

[0138] 용어 "개체", "대상체" 및 "환자"는 상호교환적으로 사용되며 치료가 필요한 질환 또는 병태를 가진 임의의 모든 개별 대상체를 지칭한다. 본 개시내용의 목적을 위해, 대상체는 영장류, 바람직하게는 인간 또는 또 다른 포유동물, 예컨대 개, 고양이, 말, 돼지, 염소 또는 소 등일 수 있다.

[0139] 발현될 관심 있는 유전자를 포함하는 DNA 벡터를 포함하여 AAV 또는 AAV의 양태와 관련되거나 이를 개시하거나 이를 설명하는 본원에 참조용으로 포함된 추가 특허는 다음과 같다: 미국 특허 번호 6,491,907; 7,229,823; 7,790,154; 7,201,898; 7,071,172; 7,892,809; 7,867,484; 8,889,641; 9,169,494; 9,169,492; 9,441,206; 9,409,953; 및 9,447,433; 9,592,247; 및 9,737,618.

[0140] II. rAAV 게놈 요소

[0141] 본원에 개시된 바와 같이, 기술의 하나의 양태는 캡시드 및 이의 캡시드 내에 "rAAV 벡터 게놈"으로 지칭되는 뉴클레오타이드 서열을 포함하는 rAAV 벡터에 관한 것이다. rAAV 벡터 게놈("rAAV 게놈"으로도 지칭됨)은 2개의 역 말단 반복체(ITR, 예를 들어, 5'-ITR 및 3'-ITR)를 포함하지만 이에 제한되지 않는 다중 요소를 포함하고, ITR 사이에 위치하는 추가 요소는 프로모터, 이중 유전자 및 폴리-A 테일을 포함한다.

[0142] 일부 구현예에서, 본원에 개시된 rAAV 게놈은 5' ITR 및 3' ITR 서열, 및 5' ITR과 3' ITR 사이에 위치하는, 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산을 암호화하는 이중 핵산에 작동적으로 연결된 프로모터, 예를 들어, 본원에 개시된 간 특이적 프로모터 서열을 포함하며, 여기서, 이중 핵산 서열은 임의로 하기 요소: 인트론 서열, 분비 신호 펩타이드를 암호화하는 핵산, IGF2 표적화 펩타이드를 암호화하는 핵산, 및 폴리 A 서열 중 하나 이상을 추가로 포함할 수 있다.

[0143] 일부 구현예에서, 본원에 개시된 rAAV 게놈은 5' ITR 및 3' ITR 서열, 및 5' ITR과 3' ITR 사이에 위치하는, 분비 펩타이드를 암호화하는 이중 핵산 및 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산에 작동적으로 연결된 프로모터(즉, 이중 핵산은 신호 펩타이드-GAA 폴리펩타이드를 포함하는 GAA 융합 폴리펩타이드를 암호화함)를 포함하고, 여기서, rAAV 게놈은 임의로, 인트론 서열, 콜라겐 안정성(CS) 서열, 폴리A 테일 및 적어도 1개의 아미노산의 스페이서를 암호화하는 핵산 중 하나 이상을 추가로 포함한다. 일부 구현예에서, 본원에 개시된 rAAV 게놈은 5' ITR 및 3' ITR 서열, 및 5' ITR과 3' ITR 사이에 위치하는, 분비 펩타이드(예를 들어, FN1, AAT 또는 GAA 신호 펩타이드)를 암호화하는 이중 핵산 및 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산에 작동적으로 연결된 본원에 개시된 바와 같은 간 특이적 프로모터를 포함하고, 여기서, rAAV 게놈은 임의로, 인트론 서열(예를 들어, MVM 또는 HBB2 인트론 서열), 콜라겐 안정성(CS) 서열, 폴리A 테일 및 적어도 1개의 아미노산의 스페이서를 암호화하는 핵산 중 하나 이상을 추가로 포함한다.

[0144] 일부 구현예에서, 본원에 개시된 rAAV 게놈은 5' ITR 및 3' ITR 서열, 및 5' ITR과 3' ITR 사이에 위치하는, 분비 펩타이드, 표적화 펩타이드 및 GAA 폴리펩타이드를 암호화하는 이중 핵산에 작동적으로 연결된 프로모터(즉, 이중 핵산은 신호 펩타이드-표적화 서열-GAA 폴리펩타이드를 포함하는 GAA 융합 폴리펩타이드를 암호화함)를 포함하고, 여기서, 표적화 펩타이드는 본원에 기재된 바와 같은 IGF2 표적화 펩타이드이고, rAAV 게놈은 임의로 인트론 서열, 콜라겐 안정성(CS) 서열, 폴리A 테일 및 적어도 1개의 아미노산의 스페이서를 암호화하는 핵산 중 하나 이상을 추가로 포함할 수 있다.

[0145] rAAV 게놈에서 각 요소는 본원에 논의된다.

[0146] A. 알파-글루코시다아제(GAA) 폴리펩타이드

[0147] 알파-글루코시다아제(GAA) 폴리펩타이드는 글리코사이드 가수분해효소의 계열 31의 구성원이다. 인간 GAA는 110 kDa 전구체로서 합성된다(Wisselaar et al. (1993) J. Biol. Chem. 268(3): 2223-31). 효소의 성숙한 형태는 70 및 76 kDa의 단량체의 혼합물이다(Wisselaar et al. (1993) J. Biol. Chem. 268(3): 2223-31). 전구체 효소는 7개의 잠재적인 글리코실화 부위를 가지며 이들 중 4개는 성숙한 효소에 유지된다(Wisselaar et al. (1993) J. Biol. Chem. 268(3): 2223-31). 성숙한 효소를 생성하는 단백질분해 절단 사건은 후기 엔도솜 또는 리소솜에서 발생한다(Wisselaar et al. (1993) J. Biol. Chem. 268(3): 2223-31).

[0148] rAAV 벡터 게놈은 GAA 폴리펩타이드를 암호화할 수 있으며, 예를 들어, 인간 GAA의 아미노산 잔기 40-952 또는 70-952, 또는 아미노산 잔기 40-790 또는 70-790과 같은 더 작은 부분을 포함할 수 있다.

[0149] 하나의 구현예에서, GAA 폴리펩타이드는 IGF2 표적화 서열에 융합될 수 있다. 일부 구현예에서, IGF2 표적화 서열은 아미노산 40, 또는 아미노산 70, 또는 인간 GAA 폴리펩타이드의 아미노산 40 또는 70의 1개 또는 2개 부

분 내의 아미노산에 융합된다. 일부 구현예에서, 본원에 개시된 바와 같은 IGF2 표적화 펩타이드는 세포의 수용체에 대한 리간드이며, 예를 들어, IGF2 표적화 펩타이드는 인간 양이온-비의존적 만노오스-6-포스페이트 수용체(CI-MPR) 또는 IGF2 수용체에 결합한다.

- [0150] C-말단 160개 아미노산은 성숙한 70 및 76 kDaI GAA 폴리펩타이드 종에는 부재이다. 그러나, 특정 폼페 대립유전자는 이 영역에 대한 GAA 활성 탭의 완전한 손실을 초래한다(Becker et al. (1998) J. Hum. Genet. 62:991). 이 돌연변이의 표현형은 70 또는 76 kDaI 종의 일부는 아니지만 단백질의 C-말단 부분이 단백질의 기능에서 중요한 역할을 한다는 것을 나타낸다. 또한 단백질의 C-말단 부분은 처리 동안 단백질의 나머지 부분으로부터 절단되지만 주요 종과 회합되어 남아 있는 것으로 보고되었다(Moreland et al. (Nov. 1, 2004) J. Biol. Chem., Manuscript 404008200). 따라서, C-말단 잔기는 단백질의 촉매 활성화에 직접적인 역할을 할 수 있고/있거나 단백질의 N-말단 부분의 적절한 폴딩을 촉진하는 데 관여할 수 있다.
- [0151] 천연 GAA 유전자는 신호 서열 및 인접한 추정 막관통 도메인, 3개의 이황화 연결을 함유하는 약 45개 아미노산의 시스테인이 풍부한 도메인인 트레포일 도메인(PFAM PF00088) (Thim (1989) FEBS Lett. 250:85), 성숙한 70/76 kDaI 폴리펩타이드에 의해 정의된 도메인, 및 C-말단 도메인을 보유하는 전구체 폴리펩타이드를 암호화한다. 기능적 GAA의 생성에는 트레포일 도메인과 C-말단 도메인이 둘 모두 필요하고 C-말단 도메인이 단백질 폴딩 동안 트레포일 도메인과 상호작용하여 아마도 트레포일 도메인에서 적절한 이황화 결합 형성을 촉진하는 것이 가능하다고 보고되었다.
- [0152] GAA 폴리펩타이드는 그 전체가 본원에 참조로 포함된 미국 특허 5,962,313 및 6,537,785에 기재되어 있다. 당업자는 분비 신호 펩타이드(SS) 또는 대안적으로 표적화 펩타이드(예를 들어, IGF2 표적화 펩타이드)가 융합될 수 있는 GAA의 특정 위치를 인식할 수 있다. 따라서, 하나의 양태에서, 본 발명은 GAA 융합 단백질에 관한 것으로, 여기서, SP 또는 IGF2 표적화 펩타이드는 서열 번호 10의 인간 GAA의 아미노산 40, 68, 69, 70, 71, 72, 779, 787, 789, 790, 791, 792, 793, 또는 796, 또는 서열 번호 170-174의 변형된 GAA 단백질, 또는 이의 일부에 융합된다.
- [0153] 본원에 개시된 방법 및 조성물의 일부 구현예에서, AAV에 의해 발현된 인간 GAA 단백질은 서열 번호 10의 아미노산, 또는 이의 단편 또는 변이체, 예를 들어, 서열 번호 10의 잔기 40, 68, 69, 70, 71, 72, 779, 787, 789, 790, 791, 792, 793, 또는 796에서 시작하는 인간 GAA 단백질을 포함한다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, AAV에 의해 발현된 인간 GAA 단백질은 서열 번호 10의 아미노산, 또는 서열 번호 10과 적어도 60%, 또는 70%, 또는 80%, 85% 또는 90% 또는 95%, 또는 98%, 또는 99% 동일한 단백질을 포함한다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, AAV에 의해 발현된 인간 GAA 단백질은 서열 번호 10의 잔기 40, 68, 69, 70, 71, 72, 779, 787, 789, 790, 791, 792, 793, 또는 796에서 시작하는 인간 GAA 단백질, 또는 이와 적어도 60%, 또는 70%, 또는 80%, 85% 또는 90% 또는 95%, 또는 98%, 또는 99% 동일한 단백질을 포함한다. 일부 구현예에서, AAV에 의해 발현된 인간 GAA 단백질은 서열 번호 170(modGAA; H199R, R223H) 또는 서열 번호 171(modGAA; H199R, R223H, H201L) 중 어느 하나의 잔기 40, 68, 69, 70, 71, 72, 779, 787, 789, 790, 791, 792, 793, 또는 796에서 시작하는 아미노산 또는 이와 적어도 60%, 또는 70%, 또는 80%, 85% 또는 90% 또는 95%, 또는 98%, 또는 99% 동일한 단백질을 포함한다.
- [0154] 일부 구현예에서, 당업자는 분비 신호 펩타이드(SS) 또는 대안적으로 표적화 펩타이드(예를 들어, IGF2 표적화 펩타이드)가 융합될 수 있는 GAA의 특정 위치를 인식할 수 있다. 예를 들어, 그 전체가 본원에 포함된 국제 특허 출원 WO2018046774A1은 분비 신호 펩타이드(SS) 또는 대안적으로 표적화 펩타이드(예를 들어, IGF2 표적화 펩타이드)가 부착될 수 있는 절단된 GAA 폴리펩타이드를 개시한다. 신호 펩타이드 또는 IGF2 표적화 펩타이드는 그 전체가 본원에 참조로 포함된 2019년 11월 19일에 출원된 미국 가출원 62,937,556 및 2019년 11월 15일에 출원된 국제 출원 WO 2020/102667에 개시된 바와 같이 GAA 절단된 단백질의 출발 아미노산인 임의의 절단된 GAA 폴리펩타이드 또는 절단된 변형된 GAA 폴리펩타이드에 부착될 수 있다.
- [0155] 일부 구현예에서, 본원에 기재된 바와 같은 rAAV 게놈에 의해 암호화된 GAA-융합 폴리펩타이드는, 예를 들어, 인간 GAA의 아미노산 잔기 40-952 또는 잔기 70-952, 또는 더 작은 부분, 예컨대 아미노산 잔기 40-790 또는 70-790을 포함할 수 있다. 하나의 구현예에서, 분비 신호 펩타이드(SS) 또는 표적화 펩타이드, 예를 들어, IGF2 표적화 펩타이드는 아미노산 40, 또는 아미노산 70, 또는 아미노산 40 또는 70의 1개 또는 2개 위치 내의 아미노산에 융합된다.
- [0156] 일부 구현예에서, 분비 신호 펩타이드(SS) 및 GAA 폴리펩타이드 및 임의로 IGF2 표적화 펩타이드(즉, SS-GAA 융합 폴리펩타이드 또는 SS-IGF2-GAA 융합 단백질)를 포함하는 융합 단백질은 인간 산 알파-글루코시다아제(GAA)

의 아미노산 잔기 40-952 또는 잔기 70-952를 포함한다(서열 번호 10). 일부 구현예에서, GAA 폴리펩타이드의 N-말단은 SS의 C-말단에 부착되고, 일부 구현예에서, GAA 폴리펩타이드의 N-말단은 IGF2 표적화 펩타이드의 C-말단에 부착되고, IGF2 표적화 펩타이드의 N-말단은 분비 신호 펩타이드의 C-말단에 부착된다.

[0157] (i) 변형된 GAA (modGAA)

[0158] 일부 구현예에서, GAA 단백질은 둘 모두 그 전체가 본원에 참조로 포함되는 US2014/0186326, 및 문헌(Moreland et al., Gene, 2012; 491(25-30))에 개시된 바와 같은 H201L 변이체를 포함한다. 특히, 아미노산 위치 201에서 히스티딘(His)은 류신(L) 잔기로 변경되어 76 kD GAA 전단백질이 성숙한 70 kD GAA 단백질로 신속하게 처리될 수 있도록 한다.

[0159] 특히, 일부 구현예에서, 본원에 개시된 바와 같은 융합 단백질은 N-말단 70-kDa 처리 부위에서 또는 그 근처에서 증가된 소수성을 초래하는 아미노산의 변형과 함께 서열 번호 10의 GAA 폴리펩타이드를 포함한다. 일부 경우에, GAA 펩타이드는 서열 번호 10의 위치 190-209에 상응하는 하나 이상의 아미노산에서 변형된다. 추가 구현예에서, 폴리펩타이드는 서열 번호 10의 아미노산 위치 195-209에 상응하는 하나 이상의 아미노산에서 변형된다. 추가 구현예에서, 변형은 서열 번호 10의 아미노산 위치 200-204에 상응하는 하나 이상의 아미노산에 있다. 특정 구현예에서, 변형은 서열 번호 10의 위치 201에 상응하는 아미노산에 있다. 추가 구현예에서, 변형은 하나 이상의 아미노산을 더 소수성인 아미노산으로 치환하는 것이다. 다른 구현예에서, 변형은 하나 이상의 소수성 아미노산의 삽입이다. 더 추가의 구현예에서, 소수성 아미노산은 류신 및 티로신, 또는 류신 또는 티로신의 보존성 아미노산으로부터 선택된다.

[0160] 특정 구현예에서, GAA는 적어도 하나의 아미노산을 더 소수성인 아미노산으로 치환함으로써 N-말단 70-kDa 처리 부위에서 또는 그 근처에서 소수성을 증가시키도록 변형된다. 일부 구현예에서, 치환은 N-말단 70-kDa 처리 부위의 업스트림 또는 다운스트림에 5개의 아미노산 내에서 이루어질 수 있다. 특정 예에서, 아미노산 치환은 서열 번호 10의 위치 195 내지 209에 상응하는 아미노산에서 이루어질 수 있다. 다른 경우에, 아미노산 치환은 서열 번호 10의 위치 200 내지 204에 상응하는 아미노산에서 이루어질 수 있다. 추가 구현예에서, 변형된 인간 GAA는 서열 번호 10의 아미노산 위치 201에 상응하는 위치에 소수성 아미노산을 함유한다. 일부 구현예에서, GAA는 N-말단 70-kDa 처리 부위에서 또는 그 근처에서 하나 이상의 소수성 아미노산을 삽입함으로써 변형된다. 추가 변형은 N-말단 70-kDa 처리 부위에서 또는 그 근처에서 하나 이상의 아미노산의 결실을 포함한다.

[0161] 특정 구현예에서, 변형된 인간 GAA는 N-말단 70-kDa 처리 부위의 하나 초과 위치에서 또는 N-말단 70-kDa 처리 부위의 5개 아미노산 내에서 소수성 아미노산(천연 또는 합성)을 함유하도록 제공된다. 하나의 구현예에서, 변형된 아미노산 중 하나는 서열 번호 10의 아미노산 201에 상응하는 위치에 있다.

[0162] 다양한 구현예에서, 소수성 아미노산은 발린, 류신, 이소류신, 메티오닌, 페닐알라닌, 트립토판, 티로신, 시스테인 또는 알라닌으로부터 선택된다. 추가 구현예에서, 소수성 아미노산은 류신 또는 티로신이다. 일부 구현예에서, 변형된 인간 GAA는 소수성 성질을 나타내는 합성 또는 비천연 아미노산을 함유한다. 일반적으로, 치환된 아미노산은 야생형 아미노산보다 더 소수성이므로 N-말단 70 kDa 처리 부위에서 또는 그 근처에서 소수성을 증가시킨다.

[0163] 하나의 예시적인 구현예에서, 변형된 GAA는 서열 번호 10의 아미노산 201에 상응하는 위치에 류신을 갖는다. 또 다른 구현예에서, 변형된 GAA는 서열 번호 10의 아미노산 201에 상응하는 위치에 티로신을 갖는다.

[0164] 일부 구현예에서, 변형된 인간 GAA 단백질은 서열 번호 10(GAA(H199R))의 아미노산 위치 199에서 His(H)에서 아르기닌(R) (H199R)으로 변형된 폴리펩타이드, 또는 서열 번호 10(GAA(R223H))의 아미노산 위치 223에서 아르기닌(R)에서 히스티딘(H) (R223H)으로 변형된 폴리펩타이드를 포함한다. 일부 구현예에서, 변형된 인간 GAA 단백질은 서열 번호 10의 아미노산 위치 199에서 His(H)에서 아르기닌(R) (H199R)으로 변형된 폴리펩타이드 및 서열 번호 10(GAA(H199R-R223H))의 아미노산 위치 223에서 아르기닌(R)에서 히스티딘(H) (R223H)으로 변형된 폴리펩타이드를 포함한다. 일부 구현예에서, 변형된 인간 GAA 단백질은 서열 번호 170 또는 H199R 또는 R223H, 또는 둘 모두의 적어도 하나의 변형을 갖는 서열 번호 170의 적어도 500, 550, 600, 650, 700, 750, 800, 850, 또는 900개 아미노산에 대해 적어도 80%, 90%, 95%, 또는 99% 상동성의 변이체를 포함한다. 일부 구현예에서, GAA의 동족 리더 서열(즉, 서열 번호 175 또는 서열 번호 170의 아미노산 1-27)은 본원에 개시된 바와 같은 IGF2 표적화 펩타이드, 또는 서열 번호 176의 리더 서열, 또는 IL2 야생형 리더 펩타이드(서열 번호 178), 변형된 IL2 리더 펩타이드(서열 번호 180) 또는 서열 번호 176, 178 또는 180과 적어도 90% 서열 동일성의 리더 펩타이드로 대체된다.

- [0165] 일부 구현예에서, 변형된 인간 GAA 단백질은 서열 번호 171 또는 적어도 H210L의 변형을 포함하는 서열 번호 171의 적어도 500, 550, 600, 650, 700, 750, 800, 850, 또는 900개 아미노산과 적어도 80%, 90%, 95%, 또는 99% 상동성의 변이체를 포함한다. 일부 구현예에서, GAA의 동족 리더 서열(즉, 서열 번호 175 또는 서열 번호 171의 아미노산 1-27)은 본원에 개시된 바와 같은 IGF2 표적화 펩타이드, 또는 서열 번호 176의 리더 서열, 또는 IL2 야생형 리더 펩타이드(서열 번호 178), 변형된 IL2 리더 펩타이드(서열 번호 180) 또는 서열 번호 176, 178 또는 180과 적어도 90% 서열 동일성의 리더 펩타이드로 대체된다.
- [0166] 일부 구현예에서, 변형된 인간 GAA 단백질은 서열 번호 10의 H199R, R223H, 또는 H201L로부터 선택된 적어도 하나의 변형, 또는 이들 변형 중 적어도 1개를 갖는 서열 번호 10의 적어도 500, 550, 600, 650, 700, 750, 800, 850, 또는 900개 아미노산과 적어도 80%, 90%, 95%, 또는 99% 상동성의 변이체를 갖는 폴리펩타이드를 포함한다. 일부 구현예에서, 변형된 인간 GAA 단백질은 서열 번호 10의 H199R, R223H, 또는 H201L로부터 선택된 적어도 2개의 변형, 또는 이들 변형 중 적어도 2개를 갖는 서열 번호 10의 적어도 500, 550, 600, 650, 700, 750, 800, 850, 또는 900개 아미노산과 적어도 80%, 90%, 95%, 또는 99% 상동성의 변이체를 갖는 폴리펩타이드를 포함한다. 일부 구현예에서, 변형된 인간 GAA 단백질은 서열 번호 10(GAA-H199R-H201L-R223H)의 3개의 변형 H199R, R223H, 및 H201L, 또는 이들 3개의 변형을 갖는 서열 번호 10의 적어도 500, 550, 600, 650, 700, 750, 800, 850, 또는 900개 아미노산과 적어도 80%, 90%, 95%, 또는 99% 상동성의 변이체를 갖는 폴리펩타이드를 포함한다.
- [0167] 특정 구현예에서, 서열 번호 10의 적어도 500, 550, 600, 650, 700, 750, 800, 850, 또는 900개 아미노산과 적어도 80%, 90%, 95%, 또는 99% 상동성을 갖는 변형된 인간 GAA가 제공되고, 여기서, 변형된 인간 GAA는 N-말단 70-kDa 처리 부위에서 더 소수성인 아미노산으로 치환된 적어도 하나의 아미노산을 갖는다.
- [0168] 일부 구현예에서, 변형된 인간 GAA의 적어도 50%는 20, 30, 또는 40시간 내에 리소솜에서 70-kDa 형태로 처리된다. 더 추가의 구현예에서, 실질적으로 모든 변형된 인간 GAA는 55, 65, 또는 75시간 내에 리소솜에서 70-kDa 형태로 처리된다.
- [0169] 특정 구현예에서, 본 발명의 변형된 인간 GAA는 성숙한 70-kDa 형태 또는 이의 상응하는 변이체로의 보다 신속한 단백질분해 처리에 의해 확인될 수 있다. 다른 구현예에서, 본원에 기재된 바와 같은 변형된 인간 GAA는 천연 인간 GAA의 단백질분해 처리 동안 생성되지 않는 82-kDa 중간체 폴리펩타이드의 생성에 의해 확인될 수 있다. 추가 구현예에서, 변형된 인간 GAA는 비변형된 인간 GAA의 단백질분해 처리 동안 생성되는 76-kDa 중간체 폴리펩타이드의 부재에 의해 확인될 수 있다.
- [0170] 특정 구현예에서, 폴리펩타이드는 서열 번호 10 또는 서열 번호 170-171의 적어도 500개 아미노산과 적어도 80% 동일성을 갖는다. 일부 경우에, 폴리펩타이드는 서열 번호 10 또는 서열 번호 170-171의 적어도 500개 아미노산과 적어도 90% 동일성을 갖는다. 다른 경우에, 폴리펩타이드는 서열 번호 10 또는 서열 번호 170-171의 적어도 500개 아미노산과 적어도 95% 동일성을 갖는다.
- [0171] 특정 구현예에서, 아미노산 201에서 소수성 잔기로 변형된 GAA 폴리펩타이드, 예를 들어, 예를 들어, H201L 변형은 비변형된 인간 산 알파-글루코시다아제 단백질과 비교할 때보다 신속한 리소솜 프로테아제 처리를 나타낸다. 일부 구현예에서, GAA 예비-폴리펩타이드의 적어도 50%는 발현 20시간 이내에 70-kDa 성숙한 GAA 형태로 단백질분해 처리된다. 다른 구현예에서, 실질적으로 모든 GAA 예비-폴리펩타이드는 발현 55시간 이내에 70-kDa 성숙한 GAA 형태로 단백질분해 처리된다.
- [0172] 일부 구현예에서, 서열 번호 10의 아미노산 1-27의 동족 GAA 리더 펩타이드(즉, MGVRHPPCSHRLAVCALVSLATAALL, 서열 번호 175)는 상이한 신호 펩타이드(리더 펩타이드)로 대체된다. 예를 들어, GAA의 동족 리더 펩타이드(서열 번호 175)는 다음 중 어느 하나로 대체될 수 있다: (i) 서열 번호 177의 핵산 서열에 의해 암호화된 MEFGLSWVFLVALLKGVQCE(서열 번호 176)의 아미노산 서열을 갖는 IgG1 리더 펩타이드(본원에서 "201 리더 펩타이드" 또는 "2011p"로 지칭됨), (ii) 서열 번호 179의 핵산 서열에 의해 암호화된 wt IL2 lp: MYRMQLLSICALSLALVTNS(서열 번호 178), 또는 (iii) 서열 번호 181의 핵산 서열에 의해 암호화된 mut IL2 lp: MYRMQLLLIALSLALVTNS(서열 번호 180). 일부 구현예에서, 동족 GAA 리더 펩타이드(서열 번호 175)가 존재하는 채로 남아 있고 추가적인 신호 펩타이드가 추가되는데, 예를 들어, 신호 펩타이드 AAT, FN1 중 임의의 하나 이상, 서열 번호 177의 핵산 서열에 의해 암호화된 MEFGLSWVFLVALLKGVQCE(서열 번호 176)의 아미노산 서열을 갖는 IgG1 리더 펩타이드("201 리더 펩타이드" 또는 "2011p"로도 지칭됨), (ii) 서열 번호 179의 핵산 서열에 의해 암호화된 wt IL2 lp: MYRMQLLSICALSLALVTNS(서열 번호 178), 또는 (iii) 서열 번호 181의 핵산 서열에 의해 암호화된

호화된 mutIL2 lp: MYRMQLLL~~LL~~I~~LL~~ALS~~LL~~ALVTNS(서열 번호 180)이다.

[0173] 일부 구현예에서, GAA는 N-연결된 글리코실화 부위, O-연결된 글리코실화 부위 또는 둘 모두와 같은 글리코실화 부위를 추가하거나 제거하도록 변형된다. 특정 구현예에서, 글리코실화 부위의 추가 또는 제거는 N-말단 결실, C-말단 결실, 내부 결실, 무작위 점 돌연변이유발, 또는 부위 지정 돌연변이유발에 의해 달성된다. 일부 구현예에서, 예시적인 GAA 변형은 하나 이상의 아스파라긴(Asn) 잔기(들)의 추가, 또는 아스파라긴(Asn) 잔기(들)를 생성하기 위한 하나 이상의 돌연변이, 또는 하나 이상의 아스파라긴(Asn) 잔기(들)의 결실을 포함한다. 특정 구현예에서, GAA에 존재하는 N-연결된 및/또는 O-연결된 글리코실화 부위의 전부 또는 일부가 돌연변이된다. 일부 구현예에서, GAA 변형은 GAA의 생물학적 활성, 물리적 구조 및/또는 기질 결합 가능성에 관한 정보를 산출할 것이다.

[0174] (ii) GAA를 암호화하는 핵산

[0175] 일부 구현예에서, rAAV 게놈은 이중 신호 서열 또는 표적화 펩타이드에 융합되지 않은 전체 GAA 폴리펩타이드(예를 들어, N-말단/축매 및 C-말단 도메인)를 암호화하는 이중 핵산 서열을 포함한다.

[0176] 일부 구현예에서, rAAV 게놈은 전체 GAA 폴리펩타이드(예를 들어, N-말단/축매 및 C-말단 도메인)를 암호화하는 GAA 핵산 서열의 3' 말단에 프레임 내 융합된 분비 신호 펩타이드 또는 IGF2 표적화 펩타이드를 암호화하는 이중 핵산 서열을 포함한다. 예를 들어, 분비 신호 펩타이드, 또는 IgF2 표적화 펩타이드를 암호화하는 이중 핵산 서열은 70 kDa 및 76 kDa GAA 폴리펩타이드를 암호화하는 GAA 핵산 서열의 3' 말단에 프레임 내 융합되며, 이러한 두 폴리펩타이드 모두는 rAAV 벡터가 포유동물 세포를 형질도입할 때 rAAV 게놈으로부터 발현된다. 일부 구현예에서, GAA 핵산의 발현은 rAAV 게놈 내의 2개의 프로모터 또는 바이시스트론 작제물의 발현을 구동하는 1개의 프로모터에 의해 구동될 수 있다.

[0177] 본원에 개시된 방법 및 조성물의 일부 구현예에서, rAAV 벡터는 야생형 GAA 핵산 서열, 예를 들어, 서열 번호 11 또는 서열 번호 72 또는 서열 번호 182인 GAA 단백질을 암호화하는 핵산 서열을 포함한다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, rAAV 벡터는 (i) 생체내 향상된 발현, (ii) CpG 섬 감소, (iii) 선천성 면역 반응 감소 중 하나 이상에 대해 코돈 최적화된 GAA 핵산 서열인 GAA 단백질을 암호화하는 핵산 서열을 포함한다. 본원에 개시된 바와 같은 방법 및 rAAV 조성물에 사용하기 위해 포함되는 예시적인 코돈 최적화된 GAA 핵산 서열은 서열 번호 73, 서열 번호 74, 서열 번호 75, 서열 번호 76 또는 서열 번호 182, 또는 서열 번호 73, 서열 번호 74, 서열 번호 75, 서열 번호 76 또는 서열 번호 182와 적어도 60%, 또는 70%, 또는 80%, 85% 또는 90% 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열 중 어느 하나로부터 선택될 수 있다.

[0178] 또한, 일부 구현예에서, 본원에 개시된 바와 같은 방법 및 rAAV 조성물에 사용하기 위해 포함되는 GAA 핵산 서열은 다음 변형 중 적어도 하나 이상으로 추가로 변형된다: (i) 적어도 하나, 또는 둘 또는 일부 구현예에서, 모든 대안적인 리딩 프레임의 제거, (ii) 하나 이상의 CpG 섬의 제거, (iii) 코작 서열의 변형, (iv) 번역 종결자 서열의 변형 및 (v) 프로모터와 코작 서열 사이의 스페이스 제거.

[0179] 예를 들어, 일부 구현예에서, rAAV 조성물은 서열 번호 182의 hGAA 뉴클레오타이드 서열, 또는 서열 번호 182와 적어도 60%, 또는 70%, 또는 80%, 85% 또는 90% 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하며, 여기서, 서열 번호 182는 GAA에 대한 야생형 핵산 서열과 비교하여 표 1A에 나타난 다음 요소를 포함한다;

[0180] [표 1A] modGAA 핵산 서열의 요소.

변형	염기쌍 (서열 번호 440 에 기반한 뉴클레오타이드 번호)
동족 리더 펩타이드	954-1034 (81bp)
Asp (T154C)에 대해 GAT 에서 GAC 로	1152-1154
Asn (T1901C)에 대해 AST 로 AAC 로	1899-1901
Asp (T1910C)에 대해 GAT 에서 GAC 로	1909-1910
Asp (T1967C)에 대해 GAT 에서 GAC 로	1965-1967
CpG (T2024G)를 제거하기 위해 GTT 에서 GTG 로	2022-2024
gln (A2156G)에 대해 CAA 에서 CAG 로	2154-2156
CpG (C2165T)를 제거하기 위해 CAC 에서 GAT 로	2163-2165
Val (A2393G)에 대해 GTA 에서 GTG 로	2391-2393
Asn (T2561C)에 대해 AAT 에서 AAC 로	2259-2561
CpG (C3401G)를 제거하기 위해 GTC 에서 GTG 로	3399-3401
Asp (T3536C)에 대해 GAT 에서 GAC 로	3534-3536
비최적 정지의 대체	3810-3821 (12bp)

[0181]

[0182] 일부 구현예에서, rAAV 조성물은 서열 번호 182의 hGAA 뉴클레오타이드 서열, 또는 서열 번호 182와 적어도 60%, 또는 70%, 또는 80%, 85% 또는 90% 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하며, 여기서, hGAA 뉴클레오타이드 서열은 3개의 잠재적인 전염증성 CpG 모티프와 다수의 대안적인 리딩 프레임(ARF)을 제거하는 일련의 점 돌연변이로 변형되었으며, 여기서, 서열 번호 182는 GAA에 대한 야생형 핵산 서열과 비교하여 표 1B에 나타난 바와 같은 다음 점 돌연변이를 포함하며, 여기서, 표 1B의 넘버링은 GAA 개시 코돈 ATG에서 "A"가 제1 뉴클레오타이드인 것으로 가정한다.

[0183] [표 1B]

NT#	목적	원래 NT	새로운 NT	NT#	목적	원래 NT	새로운 NT
201	ORF 제거	T	C	1212	CpG 파괴	C	T
949	ORF 제거	T	C	1440	ORF 제거	A	G
957	ORF 제거	T	C	1608	ORF 제거	T	C
1014	ORF 제거	T	C	2448	CpG 파괴	C	G
1071	CpG 파괴	T	G	2583	ORF 제거	T	C
1203	ORF 제거	A	G				

[0184]

[0185] 일부 구현예에서, 서열 번호 182(예를 들어, 서열 번호 182의 뉴클레오타이드 1-81)에서 동족 리더 펩타이드를 암호화하는 핵산 서열은 2011p, wtIL2 lp 또는 mutIL2 lp 중 어느 하나를 암호화하는 핵산 서열로 대체될 수 있다. 따라서, 일부 구현예에서, 서열 번호 182의 핵산 잔기 1-81(GAA의 동족 리더 펩타이드를 암호화함)은 서열 번호 177(2011p), 서열 번호 179(wtIL2 lp) 또는 서열 번호 181(mutIL2 lp)의 핵산 서열, 또는 서열 번호 177, 179 또는 181과 적어도 60%, 또는 70%, 또는 80%, 85% 또는 90% 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열로 대체될 수 있다.

[0186] 일부 구현예에서, rAAV 벡터 또는 rAAV 게놈은 서열 번호 170(동족 GAA신호 서열 및 H199R, R223H 변형을 갖는 GAA 폴리펩타이드), 또는 서열 번호 171(동족 GAA신호 서열 및 H199R, H201L, R223H 변형을 갖는 GAA 폴리펩타이드)을 포함하는 GAA 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함한다. 서열 번호 170의 GAA 폴리펩타이드는 서열 번호 182의 핵산 서열에 의해 암호화된다. 따라서, 일부 구현예에서, rAAV 벡터는 H199R, R223H 변형을 포함하는 변형된 GAA 폴리펩타이드를 암호화하는 서열 번호 182의 핵산을 포함한다. 서열 번호 171의 GAA 폴리펩타이드는 서열 번호 182의 핵산 서열에 의해 암호화되며, 여기서, 서열 번호 182의 염기쌍(bp) 667-669는 CAC에서 UUA, UUG, CUU, CUC CUA, CUG 중 어느 하나로 변경(히스타딘(H)에서 류신(L) 아미노산으로의 변경을 초래함)되거나; 여기서, 서열 번호 182의 bp 668은 A에서 U로 변경된다. 따라서, 일부 구현예에서, rAAV

벡터는 서열 번호 182의 핵산을 포함하며, 여기서, 서열 번호 182의 bp 667-669는 CAC에서 UUA, UUG, CUU, CUC CUA, CUG 중 어느 하나로 변경(아미노산을 히스타딘(H)에서 류신(L)으로 변경함)되거나; 여기서, 서열 번호 182의 bp 668은 H199R, H201L 및 R223H 변형을 포함하는 변형된 GAA 폴리펩타이드를 암호화하는 A에서 U로 변경된다.

[0187] 일부 구현예에서, rAAV 벡터 또는 rAAV 게놈은 서열 번호 172(동족 신호 펩타이드가 IgG신호 서열 및 H199R, R223H 변형으로 대체된 GAA 폴리펩타이드), 또는 서열 번호 173(동족 신호 펩타이드가 wtIL2 신호 서열 및 H199R, R223H 변형으로 대체된 동족 신호 펩타이드), 서열 번호 174(동족 신호 펩타이드가 mutIL3 신호 서열 및 H199R, R223H 변형으로 대체된 GAA 폴리펩타이드) 중 어느 하나로부터 선택된 GAA 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함한다

[0188] 일부 구현예에서, rAAV 벡터 또는 rAAV 게놈은 서열 번호 182를 포함하는 이중 핵산 서열을 포함하며, 여기서, 서열 번호 182의 bp 1-81은 서열 번호 172(H199R, R223H 변형을 갖는 IgG 리더-GAA)의 GAA 폴리펩타이드를 암호화하는 서열 번호 177(IgG신호 서열)의 핵산으로 대체된다. 일부 구현예에서, rAAV 벡터는 서열 번호 182를 포함하는 이중 핵산 서열을 포함하며, 여기서, 서열 번호 182의 bp 1-81은 A에서 U로 변경되며, 여기서, 서열 번호 182의 bp 1-81은 서열 번호 172(H199R, H201L 및 R223H 변형을 갖는 IgG 리더-GAA)의 GAA 폴리펩타이드를 암호화하는 서열 번호 177(IgG 신호 펩타이드)의 핵산으로 대체된다.

[0189] 일부 구현예에서, rAAV 벡터 또는 rAAV 게놈은 서열 번호 182를 포함하는 이중 핵산 서열을 포함하며, 여기서, 서열 번호 182의 bp 1-81은 서열 번호 173(H199R, R223H 변형을 갖는 wtIL2 신호 펩타이드-GAA)의 GAA 폴리펩타이드를 암호화하는 서열 번호 179(wtIL2 신호 펩타이드)의 핵산으로 대체된다. 일부 구현예에서, rAAV 벡터는 서열 번호 182를 포함하는 이중 핵산 서열을 포함하며, 여기서, 서열 번호 182의 bp 668은 A에서 U로 변경되며, 여기서, 서열 번호 182의 bp 1-81은 서열 번호 173(H199R, H201L 및 R223H 변형을 갖는 wtIL2 신호 펩타이드-GAA)의 GAA 폴리펩타이드를 암호화하는 서열 번호 179(wtIL2 신호 펩타이드)의 핵산으로 대체된다.

[0190] 일부 구현예에서, rAAV 벡터는 서열 번호 182를 포함하는 이중 핵산 서열을 포함하며, 여기서, 서열 번호 182의 bp 1-81은 서열 번호 174(H199R, R223H 변형을 갖는 mutIL2 신호 펩타이드-GAA)의 GAA 폴리펩타이드를 암호화하는 서열 번호 181(mutIL2 신호 펩타이드)의 핵산으로 대체된다. 일부 구현예에서, rAAV 벡터는 서열 번호 182를 포함하는 이중 핵산 서열을 포함하며, 여기서, 서열 번호 182의 bp 668은 A에서 U로 변경되며, 여기서, 서열 번호 182의 bp 1-81은 서열 번호 174(H199R, H201L, 및 R223H 변형을 갖는 mutIL2 신호 펩타이드-GAA)의 GAA 폴리펩타이드를 암호화하는 서열 번호 181(mutIL2 신호 펩타이드)의 핵산으로 대체된다.

[0191] GAA의 C-말단 도메인은 70/76 kDa 종과 함께 트랜스로 기능하여 활성 GAA를 생성한다. 촉매 도메인과 C-말단 도메인 사이의 경계는 계열 31 가수분해효소의 대부분의 구성원에는 없고 GAA에 4개의 연속적인 프롤린 잔기를 포함하는 18개 미만의 아미노산의 짧은 영역에서의 존재에 기반하여 대략 아미노산 잔기 791에 있는 것으로 보인다. 성숙한 종과 관련된 C-말단 도메인은 아미노산 잔기 792에서 시작하는 것으로 보고되었다(Moreland et al. (Nov. 1, 2004) *J. Biol. Chem.*, Manuscript 404008200). 따라서, 일부 구현예에서, C-말단 도메인을 제외하고 GAA 핵산 서열은 전체 GAA 폴리펩타이드를 암호화한다. 따라서, 이러한 구현예에서, rAAV 벡터는 별도의 폴리펩타이드로서 GAA의 C-말단 도메인을 발현하는 포유동물 세포를 형질도입하는데 사용될 수 있다.

[0192] B. 분비 신호 펩타이드

[0193] 천연 GAA 신호 펩타이드는 ER에서 절단되지 않아 천연 GAA 폴리펩타이드가 ER에서 막 결합되도록 한다(Tsuji et al. (1987) *Biochem. Int.* 15(5):945-952). GAA의 막 회합의 분열은 내인성 GAA 신호 펩타이드 (및 임의로 인접한 서열)를 GAA에 대한 대안적인 신호 펩타이드로 대체함으로써 달성될 수 있다.

[0194] 따라서, 대표적인 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터 및 rAAV 게놈은 내인성 GAA 신호 펩타이드 대신 분비 신호 펩타이드를 암호화하는 이중 핵산 서열에 부착된, 표적 세포로 전달될 GAA 폴리펩타이드를 암호화하는 이중 핵산을 추가로 포함한다. 이중 핵산은 분비 신호 펩타이드를 암호화하는 분절과 작동적으로 회합되어 전사 및 번역 시 GAA 폴리펩타이드와 작동적으로 회합된(예를 들어, 분비를 지시하는) 분비 신호 서열을 포함하는 융합 폴리펩타이드가 생성된다.

[0195] 일부 구현예에서, AAV 벡터는 내인성 GAA 신호 펩타이드(예를 들어, 서열 번호 10의 아미노산 1-27)를 포함하는 GAA 폴리펩타이드("고유 GAA" 또는 "동족 GAA" 신호 펩타이드로도 지칭함)를 암호화한다. 일부 구현예에서, AAV 벡터는 내인성 GAA 신호 펩타이드(예를 들어, 서열 번호 10의 아미노산 1-27("고유 GAA" 또는 "동족 GAA" 신호 펩타이드로도 지칭됨) 및 추가적인 이중 (비천연) 신호 서열을 포함하는 GAA 폴리펩타이드를 암호화한다.

일부 구현예에서, GAA의 아미노산 1-27의 내인성 신호 펩타이드가 결합된 GAA 폴리펩타이드는 분비 신호에 융합된다. 본원에 기재된 조성물 및 방법의 일부 구현예에서, 분비 신호는 GAA 폴리펩타이드, 또는 융합 폴리펩타이드, 예를 들어 간 세포로부터 혈액으로의 IGF2 표적화 펩타이드-GAA 융합 폴리펩타이드의 분비를 보조하는 일반적인 목적을 제공하며, 이는 이동하여 본원에 기재된 바와 같이 포유동물 세포, 예를 들어, 인간 심장 및 골격근 세포의 리소솜으로 표적화될 수 있다. 일부 구현예에서, 이중 분비 신호는 AAT 신호 펩타이드, 피브로넥틴 신호 펩타이드(FN1), GAA 신호 펩타이드, 또는 분비 신호 활성을 갖는 AAT, FN1 또는 GAA 신호 펩타이드의 활성 단편으로부터 선택된다.

[0196] 일부 구현예에서, 분비 신호 펩타이드는 관심 있는 폴리펩타이드에 이중성(즉, 외래성 또는 외인성)이다. 예를 들어, 이중 분비 신호 펩타이드는 피브로넥틴 분비 신호 펩타이드이고, 관심 있는 폴리펩타이드는 피브로넥틴이 아니다. 일부 구현예에서, 분비 신호 펩타이드는 AAT 신호 펩타이드, 피브로넥틴 신호 펩타이드(FN1), 또는 분비 신호 활성을 갖는 AAT, FN1 또는 GAA 신호 펩타이드의 활성 단편 중 어느 하나로부터 선택된다. 대안적인 구현예에서, 분비 신호 펩타이드는 GAA에 대해 이중성이 아니며, 즉 신호 펩타이드는 GAA 신호 펩타이드(즉, 천연 GAA 폴리펩타이드의 잔기 1-27)이다.

[0197] 일부 구현예에서, 서열 번호 10의 아미노산 1-27의 동족 GAA 신호 펩타이드(즉, MGVRRPPCSHRLAVCALVSLATAALL, 서열 번호 175)는 상이한 또는 이중 리더 펩타이드로 대체된다. 예를 들어, GAA의 동족 리더 펩타이드(서열 번호 175)는 다음으로부터 선택된 이중 신호 펩타이드 중 어느 하나로 대체될 수 있다: (i) 서열 번호 177의 핵산 서열에 의해 암호화된 MEFGLSWVFLVALLKGVQCE(서열 번호 176)의 아미노산 서열을 갖는 IgG1 리더 펩타이드(본원에서 "201 리더 펩타이드" 또는 "201p"로 지칭됨), (ii) 서열 번호 179의 핵산 서열에 의해 암호화된 wtIL2lp: MYRMQLLSIALSLALVTNS(서열 번호 178), 또는 (iii) 서열 번호 181의 핵산 서열에 의해 암호화된 mutIL2lp: MYRMQLLLIALSLALVTNS(서열 번호 180), 또는 서열 번호 176, 178 또는 180 중 어느 하나와 적어도 90% 서열 동일성을 갖는 리더 펩타이드.

[0198] 일반적으로, 분비 신호 펩타이드는 융합 폴리펩타이드의 아미노-말단(N-말단)에 있을 것이다(즉, 분비 신호 펩타이드를 암호화하는 핵산 분절은 본원에 개시된 바와 같은 rAAV 벡터 또는 rAAV 게놈에서 GAA 펩타이드 또는 GAA-융합 펩타이드를 암호화하는 이중 핵산에 대해 5'이다). 대안적으로, 분비 신호는, 분비 신호가 이와 작동적으로 회합되어 있고 세포로부터 관심 있는 GAA 폴리펩타이드 또는 GAA 융합 폴리펩타이드의 분비를 지지하는 한 (GAA 폴리펩타이드로부터 신호 펩타이드의 절단과 함께 또는 절단 없이), 카복실-말단에 있거나 GAA 폴리펩타이드 또는 GAA 융합 폴리펩타이드(예를 들어, IGF2-GAA 융합 폴리펩타이드) 내에 매립될 수 있다.

[0199] 분비 신호는 GAA 폴리펩타이드와 작동적으로 회합되거나 GAA 융합 폴리펩타이드는 분비 경로를 표적으로 한다. 대안적으로 언급하면, 분비 신호는 GAA-폴리펩타이드 또는 GAA 융합 폴리펩타이드가 분비 신호 펩타이드가 부재일 때보다 더 높은 수준(즉, 더 많은 양)으로 세포로부터 분비되도록 GAA 폴리펩타이드와 작동적으로 회합되어 있다. 일반적으로, 전형적으로 적어도 약 20%, 30%, 40%, 50%, 70%, 80%, 85%, 90%, 95% 이상의 GAA-폴리펩타이드 또는 IGF2-GAA 융합 폴리펩타이드(단독 및/또는 신호 펩타이드와 융합)는 분비 신호 펩타이드가 부착되지 않은 경우와 비교하여 신호 펩타이드가 부착될 때 세포로부터 분비된다. 다른 구현예에서, 본질적으로 모든 검출 가능한 폴리펩타이드(단독으로 및/또는 융합 폴리펩타이드의 형태로)는 세포로부터 분비된다.

[0200] "세포로부터 분비된"이라는 문구란, 폴리펩타이드는 다음을 포함하지만 이에 제한되지 않는 세포 외부의 임의의 구획(예를 들어, 체액 또는 공간)으로 분비될 수 있다: 간질 공간(interstitial space), 혈액, 림프, 뇌척수액, 콩팥 세뇨관, 기도 통로(예를 들어, 폐포, 세기관지, 기관지, 비강 등), 위장관(예를 들어, 식도, 위, 소장, 결장 등), 눈의 유리체액, 및 달팽이관 내림프 등.

[0201] 하나의 구현예에서, rAAV 게놈은 GAA-융합 폴리펩타이드에 융합된 분비 신호 펩타이드(SP)를 암호화하는 이중 핵산을 포함하고, 여기서, GAA-융합 폴리펩타이드는 GAA 폴리펩타이드에 융합된 표적화 펩타이드(예를 들어, IGF2 표적화 펩타이드)를 포함한다. 본원에 사용된 바와 같이, GAA는 또한 상기 기재된 변형된 GAA를 지칭한다. 따라서, 본원에 개시된 신호 펩타이드는 본원에 기재된 바와 같은 rAAV 벡터로 형질도입되거나 rAAV 게놈을 포함하는 세포로부터의 GAA 폴리펩타이드 또는 IGF2-GAA 융합 폴리펩타이드의 분비 효능을 증가시킨다.

[0202] 따라서, 일부 구현예에서, 본원에 개시된 rAAV 게놈은 5' ITR 및 3' ITR 서열, 및 5' ITR과 3' ITR 사이에 위치하는, 분비 펩타이드를 암호화하는 이중 핵산 및 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산에 작동적으로 연결된 프로모터(즉, 이중 핵산은 신호 펩타이드-GAA 폴리펩타이드를 포함하는 GAA 융합 폴리펩타이드를 암호화함)를 포함한다.

- [0203] 대안적인 구현예에서, 본원에 개시된 rAAV 게놈은 5' ITR 및 3' ITR 서열, 및 5' ITR과 3' ITR 사이에 위치하는, 분비 펩타이드를 암호화하는 이중 핵산 및 알파-글루코시다아제(GAA) 융합 폴리펩타이드를 암호화하는 핵산에 작동적으로 연결된 프로모터를 포함하며, 여기서, 융합 단백질은 IGF2 표적화 펩타이드 및 GAA 폴리펩타이드를 포함한다(즉, 이중 핵산은 신호 펩타이드-IGF2-GAA 폴리펩타이드를 포함하는 GAA 융합 폴리펩타이드를 암호화함).
- [0204] 일반적으로, 분비 신호 펩타이드는 소포체 내에서 절단되고, 일부 구현예에서, 분비 신호 펩타이드는 분비 전에 GAA 폴리펩타이드로부터 절단된다. 그러나, 세포로부터의 GAA 폴리펩타이드 또는 IGF2-GAA 융합 폴리펩타이드의 분비가 강화되고 GAA 폴리펩타이드가 기능적인 한, 분비 신호 펩타이드가 절단될 필요는 없다. 따라서, 일부 구현예에서, 분비 신호 펩타이드는 부분적으로 또는 전체적으로 보유된다.
- [0205] 일부 구현예에서, 본원에 개시된 바와 같은 rAAV 게놈 또는 단리된 핵산은 분비 신호 펩타이드에 작동적으로 연결된 GAA 폴리펩타이드를 포함하는 키메라 폴리펩타이드를 암호화하는 핵산을 포함하고, 키메라 폴리펩타이드는 rAAV 벡터로 형질도입된 세포로부터 발현 및 생성되고 GAA 폴리펩타이드는 세포로부터 분비된다. GAA 폴리펩타이드 또는 GAA 융합 폴리펩타이드(예를 들어, IGF2-GAA 융합 폴리펩타이드)는 분비 신호 펩타이드의 전부 또는 일부의 절단 후에 분비될 수 있다. 대안적으로, GAA 폴리펩타이드 또는 GAA 융합 폴리펩타이드(예를 들어, IGF2-GAA 융합 폴리펩타이드)는 분비 신호 펩타이드를 보유할 수 있다(즉, 분비 신호는 분할되지 않음). 따라서, 이러한 맥락에서, "GAA 폴리펩타이드 또는 GAA 융합 폴리펩타이드"는 분비 펩타이드를 포함하는 키메라 폴리펩타이드일 수 있다.
- [0206] 본 발명의 분비 신호 서열은 관심 있는 폴리펩타이드를 분비 경로로 유도하는 한, 임의의 특정 길이로 제한되지 않는다. 대체적인 구현예에서, 신호 펩타이드는 적어도 약 6, 8, 10, 12, 15, 20, 25, 30 또는 35개 아미노산 길이이고 최대 약 40, 50, 60, 75, 또는 100개 이상의 아미노산 길이이다.
- [0207] 본원에 개시된 바와 같은 rAAV 게놈 및 rAAV 벡터에 암호화된 분비 신호 펩타이드는 자연적으로 발생하는 분비 신호 서열 또는 이의 변형을 포함하거나, 이로 본질적으로 이루어지거나, 이로 이루어질 수 있다. 세포로부터의 분비를 지시하는 다수의 분비된 단백질 및 서열이 당업계에 알려져 있으며, 그 전체가 본원에 참고로 포함된 미국 특허 제9,873,868호에 개시되어 있다. 예시적인 분비된 단백질 (및 이들의 분비 신호)는 에리트로포이에틴, 응고 인자 IX, 시스타틴, 락토프라노스페린, 혈장 프로테아제 C1 억제제, 아포지단백질(예를 들어, APO A, C, E), MCP-1, α -2-HS-당단백질, α -1-마이크로글로불린, 보체(예를 들어, C1q, C3), 비트로넥틴, 림포톡신- α , 아주로시딘, VIP, 메탈로프로테이나제 억제제 2, 글리피칸-1, 췌장 호르몬, 클루스테린, 간세포 성장 인자, 인슐린, α -1-안티키모트립신, 성장 호르몬, IV형 콜라게나제, 구아닐린, 프로페딘, 프로엔케팔린 A, 인히빈 β (예를 들어, A 사슬), 프리알부민, 안지오테닌, 루트로핀(예를 들어, β 사슬), 인슐린-유사 성장 인자 결합 단백질 1 및 2, 전활성체 폴리펩타이드, 피브리노겐(예를 들어, β 사슬), 위 트리아실글리세롤 리파제, 미드카인, 호중구 디펜신 1, 2, 및 3, α -1-항트립신, 기질 gla-단백질, α -트립타제, 담즙염 활성화 리파제, 키모트립시노젠 B, 엘라스틴, IG 람다 사슬 V 영역, 혈소판 인자 4 변이체, 크로모그라닌 A, WNT-1 원암 유전자 단백질. 온코스타틴 M, β -네오엔돌핀-다이노르핀, 폰 빌레브란트 인자(von Willebrand factor), 혈장 세린 프로테아제 억제제, 혈청 아밀로이드 A 단백질, 니도젠, 피브로넥틴, 레닌, 오스테오넥틴, 히스타틴 3, 포스포리파제 A2, 연골 기질 단백질, GM-CSF, 마트릴리신, 신경내분비 단백질 7B2, 태반 단백질 11, 젤라틴, M-CSF, 트랜스코발라민 I, 락타제-플로리진 가수분해효소, 엘라스타제 2B, 펩시노젠 A, MIP 1- β , 프로락틴, 트립시노젠 II, 가스트린-방출 펩타이드 II, 심방 나트륨이뇨 인자, 분비된 알칼리성 포스파타제, 췌장 α -아밀라제, 세크레토그라닌 I, β -카제인, 세로트랜스페린, 조직 인자 경로 억제제, 폴리트로핀 β -사슬, 응고 인자 XII, 성장 호르몬-방출 인자, 전립선 정액 혈장 단백질, 인터루킨(예를 들어, 2, 3, 4, 5, 9, 11), 인히빈(예를 들어, 알파 사슬), 안지오테닌노젠, 티로글로불린, IG 중쇄 또는 경쇄, 플라스미노젠 활성화제 억제제-1, 리소자임 C, 플라스미노젠 활성화제, 항류코프로테이나제 1, 스타테린, 피브린-1, 이소형 B, 유로모듈린, 티록신-결합 글로불린, 엑소닌-1, 자궁내막 α -2 글로불린, 인터페론(예를 들어, 알파, 베타, 감마), β -2-마이크로글로불린, 프로콜시스토키닌, 프로가스트리신, 전립선 산 포스파타제, 뼈 시알로단백질 II, 콜리파제, 알츠하이머 아밀로이드 A4 단백질, PDGF(예를 들어, A 또는 B 사슬), 응고 인자 V, 트리아실글리세롤 리파제, 합도글로부린-2, 코르티코스테로이드-결합 글로불린, 트리아실글리세롤 리파제, 프로릴락신 H2, 폴리스타틴 1 및 2, 혈소판 당단백질 IX, GCSF, VEGF, 해파린 보조인자 II, 안티트롬빈-III, 백혈병 억제 인자, 간질 콜라게나제, 플레오토로핀, 소형 유도성 사이토카인 A1, 멜라닌-농축 호르몬, 안지오테닌-전환 효소, 췌장 트립신 억제제, 응고 인자 VIII, α -태아단백질, α -락트알부민, 세노겔린 II, 카파 카제인, 글루카곤, 티로트로핀 베타 사슬, 트랜스코발라민 II, 트롬보스폰딘 1, 부갑상선 호르몬, 바소프레신 코펩타인, 조직 인자, 모틸린, MP1F-1, 키니노젠, 신경내분비

전환효소 2, 줄기 세포 인자 프로콜라겐 α1 사슬, 혈장 칼리크레인 케라티노사이트 성장 인자 뿐만 아니라 임의의 다른 분비된 호르몬, 성장 인자, 사이토카인, 효소, 응고 인자, 우유 단백질, 면역글로불린 사슬 등을 포함하지만 이에 제한되지 않는다.

[0208] 일부 구현예에서, 본원에 개시된 바와 같은 rAAV 게놈 및 rAAV 벡터에 암호화된 기타 분비 신호 펩타이드는 프리프로-카텝신 L(예를 들어, GenBank 승인 번호 KHRTL, NP_037288; NP_034114, AAB81616, AAA39984, P07154, CAA68691; 이들의 개시내용은 그 전체가 본원에 참조로 포함됨) 및 프리프로-알파 2 유형 콜라겐(예를 들어, GenBank 승인 번호 CAA98969, CAA26320, CGHU2S, NP_000080, BAA25383, P08123; 이들의 개시내용은 그 전체가 본원에 참조로 포함됨)뿐만 아니라 대립형질 변이, 변형 및 이의 기능적 단편(피브로넥틴 분비 신호 서열과 관련하여 상기 논의된 바와 같음)으로부터의 분비 신호 서열로부터 선택될 수 있지만 이에 제한되지 않는다. 예시적인 분비 신호 서열은 프리프로카텝신 L(시궁쥐(*Rattus norvegicus*), MTPLLLLAVLCLGTALA [서열 번호 27]; 승인 번호 CAA68691) 및 프로프로-알파 2형 콜라겐(호모 사피엔스, MLSFVDTRITLLLLAVTLCLATC [서열 번호 28]; 승인 번호 CAA98969)을 포함한다. 또한, 프리프로카텝신 L 및 프리프로-알파 2형 콜라겐 또는 이의 기능적 단편으로부터의 전장 분비 신호 서열을 포함하는 더 긴 아미노산 서열이 포함된다(피브로넥틴 분비 신호 서열과 관련하여 상기 논의된 바와 같음).

[0209] 일부 구현예에서, 분비 신호 펩타이드는 부분적으로 또는 전체적으로 간 세포에 의해 생성되는 분비된 폴리펩타이드로부터 유래된다. 일부 구현예에서, 분비 신호 펩타이드는 추가로, 전체적으로 또는 부분적으로 합성 또는 인공적일 수 있다. 합성 또는 인공 분비 신호 펩타이드는 당업계에 알려져 있으며, 예를 들어, 그 개시내용은 그 전체가 본원에 참조로 포함된 문헌(Barash et al., "Human secretory signal peptide description by hidden Markov model and generation of a strong artificial signal peptide for secreted protein expression," Biochem. Biophys. Res. Comm. 294:835-42 (2002))을 참조한다. 특정 구현예에서, 분비 신호 펩타이드는 다음의 인공 분비 신호를 포함하거나 이로 본질적으로 이루어지거나 이로 이루어진다: MWWRLWWLLLLLLLLWPMVWA(서열 번호 29) 또는 1, 2, 3, 4, 또는 5개 아미노산 치환을 갖는 이의 변형(임의로, 보존성 아미노산 치환, 보존성 아미노산 치환은 당업계에 알려져 있음).

[0210] 본원에 개시된 바와 같은 방법 및 조성물에 사용하기 위한 예시적인 신호 펩타이드는 표 2에 개시된 임의의 신호 펩타이드, 또는 이의 기능적 변이체로부터 선택될 수 있다. 예시적인 신호 펩타이드는 피브로넥틴(FN1) 또는 AAT이다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, rAAV 벡터 조성물은 분비 신호 펩타이드를 암호화하는, 예를 들어, AAT 신호 펩타이드(예를 들어, 서열 번호 17), 피브로넥틴 신호 펩타이드(FN1) (예를 들어, 서열 번호 18-21), GAA 신호 펩타이드, hIGF2 신호 펩타이드(예를 들어, 서열 번호 22) 또는 분비 신호 활성을 갖는 이의 활성 단편으로부터 선택된 분비 신호 펩타이드를 암호화하는 핵산, 예를 들어, 서열 번호 17-22와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산을 포함한다.

[0211] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 분비 신호를 암호화하는 핵산은 서열 번호 17, 81-21, 22-26 중 어느 하나, 또는 서열 번호 17 또는 22-26 중 어느 하나와 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열로부터 선택된다.

[0212] 일부 구현예에서, FN1 또는 AAT 신호 펩타이드를 2019년 11월 19일에 출원된 62,937,556 또는 2019년 11월 15일에 출원된 PCT/US19/61653에 개시된 간에서 발현된 단백질에 대한 신호 펩타이드를 포함한 임의의 신호 펩타이드로 쉽게 대체할 수 있다.

[0213] 피브로넥틴 분비 신호 펩타이드:

[0214] 일부 구현예에서, 분비 신호 펩타이드는 피브로넥틴 분비 신호 펩타이드이며, 이 용어는 자연적으로 발생하는 서열의 변형을 포함한다(아래에서 더 자세히 설명되어 있음).

[0215] 일부 구현예에서, 분비 신호 펩타이드는 피브로넥틴 신호 펩타이드, 예를 들어, 인간 피브로넥틴의 신호 서열 또는 래트 피브로넥틴의 신호 서열이다. 본원에 기재된 rAAV 게놈 및 rAAV 벡터에 사용하기 위해 포함된 피브로넥틴(FN1) 신호 서열 및 변형된 FN1 신호 펩타이드는 그 전체가 본원에 참조로 포함된 미국 특허 7,071,172 및 2019년 11월 19일에 출원된 가출원 62/937,556의 표 3에 개시되어 있다. 예시적인 피브로넥틴 분비 신호 서열의 예는 그 전체가 본원에 참조로 포함된 미국 특허 7,071,172의 표 1에 나열된 것들을 포함하지만 이에 제한되지 않는다.

[표 2] 예시적인 피브로넥틴(FN1) 분비 신호 펩타이드

종	분비 신호 서열	핵산 서열
에이지. 사피엔스(<i>H. Sapiens</i>)	MLRGPGPGLLLLAVQCLGTAV PSTGA (서열 번호 20)	ATG CTT AGG GGT CCG GGG CCC GGG CTG CTG CTG CTG GCC GTC CAG TGC CTG GGG ACA GCG GTG CCC TCC ACG GGA GCC (서열 번호 25)
알, 노르베기쿠스 (<i>R. Norvegicus</i>)	MLRGPGPGRLLLLAVLCLGTSV RCTETGKSKR (서열 번호 18)	5'- ATGCTCAGGGGTCCGGGACCCGGGCGGCT GCTGCTGCTAGCAGTCCTGTGCCTGGGGAC ATCGGTGCGCTGCACCGAAACCGGAAGA GCAAGAGG-3 (서열 번호 23) (뉴클레오타이드 208-303)
알, 노르베기쿠스 (<i>R. Norvegicus</i>)	MLRGPGPGRLLLLAVLCLGTSV RCTETGKSKR ↑ LALQIV (서열 번호 19)	5'-ATG CTC AGG GGT CCG GGA CCC GGG CGG CTG CTG CTG CTA GCA GTC CTG TGC CTG GGG ACA TCG GTG CGC TGC ACC GAA ACC GGG AAG AGC AAG AGG ↑ CAG GCT CAG CAA ATC GTG-3'. (서열 번호 24) (↑는절단 분위를 나타냄)
엑스. 라비스(<i>X. laevis</i>)	MRRGALTGLLLVLCLSVVLRA APSATSKKRR (서열 번호 21)	ATG CGC CGG GGG GCC CTG ACC GGG CTG CTC CTG GTC CTG TGC CTG AGT GTT GTG CTA CGT GCA GCC CCC TCT GCA ACA AGC AAG AAG CGC AGG (서열 번호 26)

[0217]

[0218] 시궁쥐의 피브로넥틴 분비 신호 서열을 암호화하는 예시적인 뉴클레오타이드 서열은 GenBank 승인 번호 X15906 (이의 개시내용은 본원에 참조로 포함됨)에서 발견된다. 또 다른 추가 예시적인 서열로서, 인간 피브로넥틴 1, 전사체 변이체 1의 분비 신호 펩타이드를 암호화하는 뉴클레오타이드 서열(승인 번호 NM_002026, 뉴클레오타이드 268-345; 승인 번호 NM_002026의 개시내용은 그 전체가 본원에 참조로 포함됨). 또 다른 예시적인 분비 신호 서열은 아프리카발톱개구리(*Xenopus laevis*) 피브로넥틴 단백질의 분비 신호 펩타이드를 암호화하는 뉴클레오타이드 서열에 의해 암호화된(승인 번호 M77820, 뉴클레오타이드 98-190; 승인 번호 M77820의 개시내용은 그 전체가 본원에 참조로 포함됨).

[0219] 또 다른 구현예에서, 피브로넥틴 신호 서열(FN1, 뉴클레오타이드 208-303, 5'-CTC AGG GGT CCG GGA CCC GGG CGG CTG CTG CTG CTA GCA GTC CTG TGC CTG GGG ACA TCG GTG CGC TGC ACC GAA ACC GGG AAG AGC AAG AGG-3', 서열 번호 23)은 래트 피브로넥틴 mRNA 서열(Genbank 승인 #X15906) 및 다음 펩타이드 신호 서열에 대한 코드로부터 유래되었다: Met Leu Arg Gly Pro Gly Pro Gly Arg Leu Leu Leu Leu Ala Val Leu Cys Leu Gly Thr Ser Val Arg Cys Thr Glu Thr Gly Lys Ser Lys Arg (서열 번호 18). 본원에 개시된 방법 및 조성물의 일부 구현예에서, 재조합 AAV 벡터는 분비 신호 활성을 갖는 피브로넥틴 신호 펩타이드(FN1) 또는 이의 활성 단편인 분비 신호 펩타이드를 암호화하는 이중 핵산 서열을 포함하고(예를 들어, FN1 신호 펩타이드는 서열 번호 18-21 중 어느 하나, 또는 서열 번호 18-21과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 가짐), 상기 이중 핵산 서열은 서열 번호 5, 서열 번호 6, 서열 번호 7, 서열 번호 8 또는 서열 번호 9 중 어느 하나, 또는 서열 번호 5-9와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 IGF2 펩타이드로부터 선택된 IGF2 표적화 펩타이드를 암호화한다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, 재조합 AAV 벡터는 AAT 신호 펩타이드 또는 분비 신호 활성을 갖는 이의 활성 단편인 분비 신호 펩타이드를 암호화하는 이중 핵산 서열을 포함하고(예를 들어, AAT 신호 펩타이드는 서열 번호 17의 서열, 또는 서열 번호 17과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산을 가짐) 이중 핵산 서열은 서열 번호 5, 서열 번호 6, 서열 번호 7, 서열 번호 8 또는 서열 번호 9 중 어느 하나, 또는 서열 번호 5-9와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 IGF2 펩타이드로부터 선택된 IGF2 표적화 펩타이드를 암호화한다.

[0220] 당업자는 분비 신호 서열이 펩티다제 절단 부위(↑로 식별됨)의 C-말단 측에서 1개, 2개, 3개, 4개, 5개 또는 6개 이상 모두의 아미노산을 암호화할 수 있다는 것을 인식할 것이다(예를 들어, 표 2의 서열 번호 19 및 24 참조). 당업자는 절단 부위의 카복시-말단 측 상의 추가 아미노산(예를 들어, 1, 2, 3, 4, 5, 6개 이상의 아미노산)이 분비 신호 서열에 포함될 수 있다는 것을 인식할 것이다.

[0221] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 재조합 AAV 벡터는 5' ITR과 3' ITR 사이에 위치하는, 분비 신호 펩타이드를 암호화하는 이중 핵산 서열 및 hGAA 폴리펩타이드를 암호화하는 핵산을 포함하거나 이로 이루어

어지며, 여기서, 신호 서열을 암호화하는 핵산 서열은 AAT 신호 펩타이드(예를 들어, 서열 번호 17), 피브로넥틴 신호 펩타이드(FN1) (예를 들어, 서열 번호 18-21), 동족 GAA 신호 펩타이드(서열 번호 175), hIGF2 신호 펩타이드(예를 들어, 서열 번호 22), IgG1 리더 펩타이드(서열 번호 177), wt IL2 리더 펩타이드(서열 번호 179), 돌연변이 IL2 리더 펩타이드(서열 번호 181) 중 어느 하나 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 17-22, 175, 177, 179 또는 181과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산으로부터 선택되며, 여기서, 신호 펩타이드를 암호화하는 핵산은 본원에 개시된 바와 같은 hGAA 폴리펩타이드를 암호화하는 핵산의 5'에 위치하며, 여기서, 신호 서열을 암호화하는 핵산 및 hGAA 폴리펩타이드는 본원의 표 4에 개시된 임의의 LSP, 또는 이의 기능적 변이체에 작동적으로 연결된다.

[0222] 본 발명의 구현예에서, 기능적 단편은 본원에 구체적으로 개시된 서열과 비교하여 적어도 약 50%, 70%, 80%, 90% 이상의 분비 신호 활성을 갖거나 심지어 더 높은 수준의 분비 신호 활성을 갖는다.

[0223] 펩티다제 절단 부위

[0224] 일부 구현예에서, 하나 이상의 외인성 펩티다제 절단 부위는 분비 신호 펩타이드 - GAA 융합 폴리펩타이드 내로, 예를 들어, 분비 신호 펩타이드와 GAA 폴리펩타이드 사이에 삽입될 수 있다. 특정 구현예에서, 자가프로테아제(예를 들어, 구제역 바이러스 2A 자가프로테아제)는 분비 신호 펩타이드와 GAA 폴리펩타이드 또는 IGF2-GAA 융합 폴리펩타이드 사이에 삽입된다. 다른 구현예에서, 외인성 프로테아제의 추가에 의해 조절될 수 있는 프로테아제 인식 부위가 사용된다(예를 들어, 트립신에 대한 Lys-Arg 인식 부위, 아스페르길루스(*Aspergillus*) KEX2-유사 프로테아제의 Lys-Arg 인식 부위, 메탈로프로테아제에 대한 인식 부위, 세린 프로테아제에 대한 인식 부위 등). 천연 프로테아제 부위를 결실하거나 불활성화시키기 위한 GAA 폴리펩타이드의 변형은 본원에 포함되며 2019년 11월 19일에 출원된 미국 가출원 62,937,556 및 2019년 11월 15일에 출원된 국제 출원 PCT/US19/61653에 개시되어 있다.

[0225] **C. IGF2 표적화 펩타이드 서열**

[0226] 하나의 구현예에서, rAAV 게놈은 GAA 폴리펩타이드에 융합된 표적화 펩타이드(TP)를 암호화하는 이중 핵산을 포함한다. 일부 구현예에서, 표적화 펩타이드는 세포외 수용체에 대한 리간드이며, 여기서, 표적화 펩타이드는 표적 세포의 표면에 있는 수용체의 세포외 도메인에 결합하며, 수용체의 내재화 시 인간 리소솜에서 폴리펩타이드의 국소화를 허용한다. 하나의 구현예에서, 표적화 펩타이드는 양이온-비의존적 만노오스-6-포스페이트 수용체에 결합할 수 있는 유로키나제-유형 플라스미노겐 수용체 모이어티를 포함한다. 일부 구현예에서, 표적화 펩타이드는 IGF2 표적화 펩타이드의 하나 이상의 아미노산 서열을 포함한다.

[0227] 일부 구현예에서, 본원에 개시된 바와 같은 IGF2 표적화 펩타이드는 세포외 수용체에 대한 리간드의 적어도 일부를 포함하며, 예를 들어, IGF2 표적화 펩타이드는 인간 양이온-비의존적 만노오스-6-포스페이트 수용체(CI-MPR) 또는 IGF2 수용체에 결합한다.

[0228] IGF2는 별칭으로도 알려져 있다: 염색체 11 오픈 리딩 프레임 43, 인슐린-유사 성장 인자 2, IGF-II, FLJ44734; IGF2, 소마토메딘 A 및 프렙틴. 야생형 인간 IGF2 서열의 mRNA는 다음에 해당한다:

[0229] GCTTACGCCCCAGTGAGACCCTGTGCGCGGGAGCTGGTGGACACCTCCAGTTCGTCTGTGGGACCGCGGCTTCTACTTCAGCAGGCCCGCAAGCCGTGTGAGCCGTCGACGCCGTGGCATCGTTGAGGAGTGTGTTCCGCGAGCTGTGACCTGGCCCTCTGGAGACGTACTGTGCTACCCCGCCAAGTCCGAG (서열 번호 1). 전장 IGF2 단백질(IGF2 표적화 서열 포함)은 NM_000612.6의 핵산 서열에 의해 암호화되며 전장 IGF2 단백질 NP_000603.1을 암호화한다.

[0230] 성숙한 인간 IGF2 표적화 펩타이드는 아래에 나타낸다:

[0231] A Y R P S E T L C G G E L V D T L Q F V C G D R G F Y F S R P A S R V S R R S R G I V E E C C F R S C D L A L L E T Y C A T P A K S E (서열 번호 5)

[0232] 인간 IGF2의 암호화 서열은 또한 그 전체가 본원에 참조로 포함된 미국 특허 8,492,388(예를 들어, 도 2 참조)에 개시되어 있다. IGF2 단백질은 아미노 말단에 24 아미노산 신호 펩타이드와 89 아미노산 카복시 말단 영역이 둘 모두 번역 후 제거되는 프리프로단백질로서 합성되며 문헌(O'Dell et al. (1998) Int. J. Biochem Cell Biol. 30(7):767-71)에서 검토되었다. 성숙한 단백질은 67개 아미노산이다. 성숙한 IGF2의 리슈마니아(*Leishmania*) 코돈 최적화된 버전은 미국 특허 8,492,388(예를 들어, 8,492,388의 도 3 참조)에 개시되어 있다 (Langford et al. (1992) Exp. Parasitol. 74(3):360-1). 성숙한 폴리펩타이드의 아미노산 1-7 또는 2-7의 결

실($\Delta 1-7$), 티로신에서 류신으로의 잔기 27의 변경(Y27L) 또는 두 돌연변이($\Delta 1-7$, Y27L 또는 $\Delta 2-7$, Y27L)를 포함하는 추가 카세트는 하기에 기술된 바와 같이 원하는 수용체에만 특이성을 갖는 IGF-2 카세트를 생성하기 위해 제조하였다. 따라서, 일부 구현예에서, IGF2 표적화 서열은 하기 중 어느 하나로부터 선택될 수 있으며: 야생형, Y27L, $\Delta 1-7$, $\Delta 2-7$ 및 Y27L- $\Delta 1-7$, Y27L- $\Delta 2-7$, V43M, Y27L-V43M, Y27L- $\Delta 1-7$ -V43M, Y27L- $\Delta 2-7$ -V43M IGF2 변이체는 본원에 사용하기 위해 포함된다.

[0233] 본원의 방법 및 조성물에 사용하기 위한 예시적인 IGF2 표적화 펩타이드는 2019년 11월 19일에 출원된 미국 가출원 62,937,556 및 2019년 11월 15일에 출원된 국제 출원 PCT/US19/61653 및 2019년 11월 15일에 출원된 국제 출원 PCT/US19/61701에 개시되어 있으며, 이들 각각은 그 전체가 본원에 참조로 포함된다.

[0234] 일부 구현예에서, 본원의 방법 및 조성물에 사용하기 위한 IGF2 표적화 펩타이드는 그 전체가 본원에 포함된 US 출원 2019/0343968에 개시된 바와 같은 하기 중 어느 하나의 변형을 가질 수 있다: E6R, F26S, Y27L, V43L, F48T, R49S, S50I, A54R, L55R, K65R. 일부 구현예에서, IGF2 표적화 펩타이드는 E6R, F26S, Y27L, V43L, F48T, R49S, S50I, A54R, L55R 및 K65R로부터 선택된 하나 이상의 변형에 추가로 V43M의 변형을 갖는다. 일부 구현예에서, IGF2 표적화 펩타이드는 하기로부터 선택된 하나 이상의 변형에 추가로 $\Delta 1-7$ 또는 $\Delta 2-7$ 변형을 갖는다: E6R, F26S, Y27L, V43L, F48T, R49S, S50I, A54R, L55R 및 K65R. 일부 구현예에서, IGF2 표적화 펩타이드는 $\Delta 1-7$ 또는 $\Delta 2-7$ 변형, V43M 변형, 및 하기로부터 선택된 하나 이상의 변형을 갖는다: E6R, F26S, Y27L, V43L, F48T, R49S, S50I, A54R, L55R 및 K65R.

[0235] 특정 구현예에서, IGF2 표적화 펩타이드는 발린 43에서 변형을 포함하고, 여기서, 발린은 아미노산 43에서 변역 개시가 시작되도록 met(V43M)로 변형된다. 표적화 펩타이드 또는 IGF2 표적화 펩타이드로서 본원에서 사용을 위해 포함되는 V43M의 변형을 갖는 IGF2 표적화 펩타이드는 양이온-비의존적 만노오스-6-포스페이트 수용체에 결합한다. 대안적인 구현예에서, IGF2 표적화 펩타이드는 V43이 Met(즉, IGF2- $\Delta 1-42$ (서열 번호 8) 또는 IGF2-V43M(서열 번호 9))로 변경된 IGF2의 델타 1-42이다.

[0236] 일부 구현예에서, rAAV 게놈은 IGF2-GAA 융합 단백질을 암호화하는 핵산을 포함하며, 여기서, 성숙한 IGF2 표적화 펩타이드(서열 번호 5) 또는 IGF2 표적화 펩타이드 변이체(예를 들어, 서열 번호 6(IGF2- $\Delta 2-7$); 서열 번호 7(IGF2- $\Delta 1-7$); 서열 번호 8(IGF2- $\Delta 1-42$), 서열 번호 9(IGF2-V43M)) 또는 서열 번호 5-9와 적어도 85%, 또는 90% 또는 95% 서열 동일성을 갖는 서열을 암호화하는 핵산은 GAA 단백질을 암호화하는 핵산의 5' 말단에 융합되며, 다양한 세포 유형에 의해 흡수되고 리소솜으로 수송될 수 있는 융합 단백질(예를 들어, IGF2-GAA 융합 폴리펩타이드)이 생성된다. 대안적으로, 전구체 IGF2 폴리펩타이드를 암호화하는 핵산은 GAA 유전자의 3' 말단에 융합될 수 있으며; 전구체는 성숙한 IGF2 폴리펩타이드를 생성하기 위해 포유동물 세포에서 절단되는 카복시-말단 부분을 포함하지만, IGF2 표적화 펩타이드는 바람직하게는 생략 (또는 GAA 유전자의 5' 말단으로 이동)된다. 이 방법은 단백질이 단리되면 추가 변형이 필요하지 않기 때문에 단순성 및 비용 효율성을 포함하여 글리코실화를 포함하는 방법에 비해 많은 이점이 있다.

[0237] 일부 구현예에서, 본원에서 사용하기 위해 포함되는 IGF2 표적화 펩타이드는, 각각 그 전체가 본원에 참조로 포함된 미국 특허 7,785,856 및 9,873,868에 기재되어 있다.

[0238] (i) *GF2의 결실 돌연변이체*:

[0239] 일부 구현예에서, IGF2 표적화 펩타이드는 그 전체가 본원에 참조로 포함된, 2019년 11월 15일에 출원된 국제 출원 PCT/US19/61701에 개시된 바와 같은 변형된 또는 절단된 IGF2 표적화 펩타이드(IGF2의 결실 돌연변이체로도 지칭됨)이다. 예를 들어, 일부 구현예에서, IGF2 표적화 펩타이드는 V43M 변형 및 또한 아미노산 1-42로부터 하나 이상의 아미노산의 임의의 결실을 포함한다. 예를 들어, 본원에 개시된 방법 및 조성물의 일부 구현예에서, IGF2 표적화 펩타이드는 V43M을 포함하며, 서열 번호 5의 $\Delta 1-3$, $\Delta 1-4$, $\Delta 1-5$, $\Delta 1-6$, $\Delta 1-8$, $\Delta 1-9$, $\Delta 1-10$, $\Delta 1-11$, $\Delta 1-12$, $\Delta 1-13$, $\Delta 1-14$, $\Delta 1-15$, $\Delta 1-16$, $\Delta 1-17$, $\Delta 1-18$, $\Delta 1-19$, $\Delta 1-20$, $\Delta 1-21$, $\Delta 1-22$, $\Delta 1-23$, $\Delta 1-24$, $\Delta 1-25$, $\Delta 1-26$, $\Delta 1-27$, $\Delta 1-28$, $\Delta 1-29$, $\Delta 1-30$, $\Delta 1-31$, $\Delta 1-32$, $\Delta 1-33$, $\Delta 1-34$, $\Delta 1-35$, $\Delta 1-36$, $\Delta 1-37$, $\Delta 1-38$, $\Delta 1-39$, $\Delta 1-40$, $\Delta 1-41$ 또는 $\Delta 1-42$ 중 어느 하나로부터 선택된 하나 이상의 결실을 추가로 포함하며, 여기서, 서열 번호 5의 잔기 43은 메티오닌(V43M)이다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, IGF2 표적화 펩타이드는 V43M을 포함하며, $\Delta 1-7$ 결실(IGF2- $\Delta 1-7$, V43M)을 추가로 포함한다.

[0240] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 리소솜 IGF2 표적화 펩타이드는 서열 번호 5의 $\Delta 2-3$, $\Delta 2-4$, $\Delta 2-5$, $\Delta 2-6$, $\Delta 2-8$, $\Delta 2-9$, $\Delta 2-10$, $\Delta 2-11$, $\Delta 2-12$, $\Delta 2-13$, $\Delta 2-14$, $\Delta 2-15$, $\Delta 2-16$, $\Delta 2-18$, $\Delta 2-19$, $\Delta 2-20$, $\Delta 2-21$, $\Delta 2-22$, $\Delta 2-23$, $\Delta 2-24$, $\Delta 2-25$, $\Delta 2-26$, $\Delta 2-27$, $\Delta 2-28$, $\Delta 2-29$, $\Delta 2-30$, $\Delta 2-31$, $\Delta 2-32$,

△2-33, △2-34, △2-35, △2-36, △2-37, △2-38, △2-39, △2-40, △2-41 또는 △2-42 중 어느 하나로부터 선택된 하나 이상의 변형을 추가로 포함하며, 여기서, 서열 번호 5의 잔기 43은 메티오닌(V43M)이다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, IGF2 표적화 펩타이드는 V43M을 포함하며, △2-7 결실(IGF2-△2-7,V43M)을 추가로 포함한다.

[0241] 일부 구현예에서, GAA-폴리펩타이드에 대한 융합을 위한 IGF2 표적화 펩타이드는 IGF2의 아미노산 8-28 및 41-61을 포함할 수 있다. 일부 구현예에서, 아미노산의 이러한 스트레치는 직접 연결되거나 링커에 의해 분리될 수 있다. 대안적으로, 아미노산 8-28 및 41-61은 별도의 폴리펩타이드 사슬에 제공될 수 있다. 일부 구현예에서, IGF2의 아미노산 8-28 또는 이의 보존적 치환 변이체는 GAA 폴리펩타이드에 융합되어 rAVV 벡터로부터의 IGF2-GAA 융합 단백질을 발현할 수 있으며, 별도의 rAAV 벡터는 IGF2 아미노산 41-61 또는 이의 보존적 치환 변이체를 발현할 수 있다.

[0242] IGF2 표적화 펩타이드의 적절한 제시 및 폴딩을 용이하게 하기 위해, IGF2 단백질의 더 긴 부분이 사용될 수 있다. 예를 들어, 아미노산 잔기 1-67, 1-87, 또는 전체 전구체 형태를 포함하는 IGF2 표적화 펩타이드가 사용될 수 있다.

[0243] 일부 구현예에서, IGF2 표적화 펩타이드는 다음 중 어느 하나의 IGF2 표적화 펩타이드를 암호화하는 핵산 서열이다: 서열 번호 5의 야생형 성숙한 인간 인슐린-유사 성장 인자 II(IGF2)의 잔기 1에 이어 잔기 8-67(즉, 서열 번호 6; 즉, IGF2-델타 2-7); 서열 번호 5의 야생형 성숙한 인간 인슐린-유사 성장 인자 II(IGF2)의 잔기 8-67(즉, 서열 번호 7; IGF2-델타 1-7) 또는 서열 번호 5의 야생형 성숙한 인간 인슐린-유사 성장 인자 II(IGF2)의 잔기 43-67(즉, IGF2-V43M(서열 번호 9) 또는 IGF-델타 1-42(서열 번호 8)).

[0244] 본원에 개시된 방법 및 조성물의 일부 구현예에서, IGF2 표적화 펩타이드는 서열 번호 2(즉, IGF2-델타 2-7); 서열 번호 3(즉, IGF2-델타 1-7) 또는 서열 번호 4(즉, IGF2-V43M) 중 어느 하나 또는 이와 적어도 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 서열을 포함하는 임의의 핵산 서열로부터 선택된 핵산 서열이다.

[0245] 본원에 개시된 방법 및 조성물의 일부 구현예에서, IGF2(V43M) 서열은 서열 번호 65 (IGF2D2-7V43M) 또는 서열 번호 65와 적어도 85%, 또는 90%, 또는 95% 또는 96%, 또는 97%, 또는 98% 또는 99% 또는 100% 동일성을 갖는 아미노산 서열, 또는 서열 번호 66(IGF△1-7V43M) 또는 서열 번호 66과 적어도 85%, 또는 90%, 또는 95% 또는 96%, 또는 97%, 또는 98% 또는 99% 또는 100% 동일성을 갖는 아미노산 서열 중 어느 하나의 IGF2(V43M) 서열을 암호화하는 핵산 서열이다.

[0246] [표 3] IGF2 표적화 펩타이드를 암호화하는 예시적인 핵산 서열

IGF2 표적화 펩타이드	서열
IGF2-델타 2-7 (IGFA2-7)	GCTCTGTGCGGGCGGGAGCTGGTGGACACCCCTCCAGTTCGTCTGTGGGGA CCGCGGCTTCTACTTCAGCAGGCCCGCAAGCCGTGTGAGCCGTCGCAGCC GTGGCATCGTTGAGGAGTGCTGTTTCCGCAGCTGTGACCTGGCCCTCCTGG AGACGTACTGTGCTACCCCGCCAAGTCCGAG) (서열 번호 2)
IGF2-델타 1-7 (IGFA1-7)	CTGTGCGGGCGGGAGCTGGTGGACACCCCTCCAGTTCGTCTGTGGGGACCG CGGCTTCTACTTCAGCAGGCCCGCAAGCCGTGTGAGCCGTCGCAGCCGTG GCATCGTTGAGGAGTGCTGTTTCCGCAGCTGTGACCTGGCCCTCCTGGAG ACGTACTGTGCTACCCCGCCAAGTCCGAG (서열 번호 3)
IGF2-V43M	GCTTACCGCCCCAGTGAGACCCTGTGCGGGCGGGAGCTGGTGGACACCCCT CCAGTTCGTCTGTGGGACCGCGGCTTCTACTTCAGCAGGCCCGCAAGCC GTGTGAGCCGTCGCAGCCGTGGCATCATGGAGGAGTGCTGTTTCCGCAGC TGTGACCTGGCCCTCCTGGAGACGTACTGTGCTACCCCGCCAAGTCCGA G (서열 번호 4)

[0247]

[0248] 일부 구현예에서, IGF2 표적화 펩타이드의 적절한 제시 및 폴딩을 용이하게 하기 위해, IGF2 단백질의 더 긴 부분이 사용될 수 있다. 예를 들어, 아미노산 잔기 1-67, 1-87, 또는 전체 전구체 형태를 포함하는 IGF2 표적화

펩타이드가 사용될 수 있다.

- [0249] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 제조할 AAV는 신호 펩타이드-GAA(SP-GAA) 융합 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 분비 신호 펩타이드(SP)와 알파-글루코시다아제(GAA) 폴리펩타이드 사이에 위치한 IGF2 표적화 펩타이드를 추가로 포함한다.
- [0250] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 제조할 AAV 벡터는 인간 양이온-비의존적 만노오스-6-포스페이트 수용체(CI-MPR) 또는 IGF2 수용체에 결합하는 IGF2 표적화 펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 예를 들어, 이중 핵산 서열은 서열 번호 5의 아미노산 서열을 갖는 IGF2 표적화 펩타이드를 암호화하거나, IGF2 수용체에 결합하는 서열 번호 5에서 적어도 하나의 아미노 변형을 포함한다. 일부 구현예에서, 제조할 AAV 벡터는 V43M 아미노산 변형(서열 번호 8 또는 서열 번호 9) 또는 Δ 2-7(서열 번호 6) 또는 Δ 1-7(서열 번호 7), 또는 서열 번호 5-9와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 IGF2 펩타이드인 서열 번호 5에서 적어도 하나의 아미노 변형을 갖는 IGF2 표적화 펩타이드를 암호화하는 이중 핵산 서열을 포함한다.
- [0251] 본원에 개시된 방법 및 조성물의 일부 구현예에서, IGF2 표적화 펩타이드를 암호화하는 핵산은 서열 번호 2(IGF2- Δ 2-7), 서열 번호 3(IGF2- Δ 1-7), 또는 서열 번호 4(IGF2 V43M) 중 어느 하나, 또는 서열 번호 2, 3 또는 4 중 어느 하나와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열로부터 선택된다.
- [0252] .본원에 기재된 조성물 및 방법의 일부 구현예에서, IGF2 표적화 펩타이드는 서열 번호 5의 야생형 성숙한 인간 인슐린-유사 성장 인자 II(IGF2)의 잔기 1에 이어 잔기 8-67(즉, IGF2-텔타 2-7 또는 IGF2D2-7로, 서열 번호 6에 상응함); 서열 번호 5의 야생형 성숙한 인간 인슐린-유사 성장 인자 II(IGF2)의 잔기 8-67(즉, IGF2-텔타 1-7 또는 IGF2 Δ 1-7로, 서열 번호 7에 상응함) 또는 서열 번호 5의 야생형 성숙한 인간 인슐린-유사 성장 인자 II(IGF2)의 잔기 43-67(즉, IGF2 텔타 1-42 또는 IGF2 Δ 1-42로, 서열 번호 8에 상응함) 중 어느 하나를 암호화하는 핵산 서열이다. 본원에 기재된 조성물 및 방법의 일부 구현예에서, IGF2 표적화 펩타이드는 아미노산 잔기 43의 변형을 갖는 핵산 서열이고, 예를 들어, 잔기 43은 개시 코돈, 예를 들어 IGF2-V43M(서열 번호 9에 상응함)으로 변형된다.
- [0253] 본원에 기재된 조성물 및 방법의 일부 구현예에서, IGF2 표적화 펩타이드는 서열 번호 2(즉, IGF2-텔타 2-7); 서열 번호 3(즉, IGF2-텔타 1-7) 또는 서열 번호 4(즉, IGF2-V43M) 중 어느 하나를 포함하는 핵산 서열이다.
- [0254] 본원에 기재된 조성물 및 방법의 일부 구현예에서, GAA 폴리펩타이드 및 IGF2 표적화 펩타이드를 포함하는 융합 단백질은 야생형 성숙한 인간 인슐린-유사 성장 인자 II(IGF2) (서열 번호 5)의 잔기 1에 이어 잔기 8-67을 포함하는 IGF2 표적화 펩타이드에 부착된 인간 산 알파-글루코시다아제(GAA) 폴리펩타이드(서열 번호 10)의 아미노산 잔기 40-952 또는 잔기 70-952를 포함하고(즉 - 성숙한 인간 IGF2(서열 번호 5)의 잔기 2-7은 존재하지 않음), 여기서, IGF2 표적화 펩타이드는 인간 GAA의 아미노산 잔기 70(서열 번호 10)에 연결된다.
- [0255] 본원에 기재된 조성물 및 방법의 일부 구현예에서, GAA 폴리펩타이드 및 IGF2 표적화 펩타이드를 포함하는 융합 단백질은 야생형 성숙한 인간 인슐린-유사 성장 인자 II(IGF2) (서열 번호 5)의 잔기 8-67을 포함하는 IGF2 표적화 펩타이드에 부착된 인간 산 알파-글루코시다아제(GAA) 폴리펩타이드(서열 번호 10)의 아미노산 잔기 40-952 또는 잔기 70-952를 포함하고(여기서, 성숙한 인간 IGF2(즉, Y R P S E T; 서열 번호 63)의 잔기 17은 존재하지 않음), 여기서, IGF2 표적화 펩타이드는 인간 GAA의 아미노산 잔기 70(서열 번호 10)에 연결된다.
- [0256] 본원에 기재된 조성물 및 방법의 일부 구현예에서, GAA 폴리펩타이드 및 IGF2 표적화 펩타이드를 포함하는 융합 단백질은 야생형 성숙한 인간 인슐린-유사 성장 인자 II(IGF2) (서열 번호 5)의 잔기 43-67을 포함하는 변형된 IGF2 표적화 펩타이드에 부착된 인간 산 알파-글루코시다아제(GAA) 폴리펩타이드(서열 번호 10)의 아미노산 잔기 40-952 또는 잔기 70-952를 포함하고(즉 - 성숙한 인간 IGF2(즉, 서열 번호 5)의 잔기 1-42는 존재하지 않음), 여기서, IGF2 표적화 펩타이드는 인간 GAA의 아미노산 잔기 70(서열 번호 10)에 연결된다.
- [0257] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 제조할 AAV 벡터는 IGF2 펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서, IGF2 펩타이드 서열은 서열 번호 8 또는 서열 번호 9, 또는 서열 번호 8 또는 9와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 IGF2 펩타이드이다.
- [0258] (ii) 변형된 IGF2 표적화 펩타이드 및 IGF2 상동체

- [0259] 일부 구현예에서, IGF2를 암호화하는 핵산은 IGFBP에 대한 친화도를 감소시키고/시키거나, IGF-1 수용체에 대한 결합 친화도를 감소시켜 리소좀에 대한 표적화를 증가시키고 융합된 GAA-폴리펩타이드의 생체이용률을 증가시키도록 변형될 수 있다.
- [0260] IGF2 표적화 펩타이드는 바람직하게는 M6P 수용체에 특이적으로 표적화하고 결합한다. 특히 유용한 것은 IGF2 폴리펩타이드에 돌연변이가 있어서 CI-MPR/M6P 수용체에 높은 친화도로 결합하는 반면 다른 두 수용체에는 더 이상 상당한 친화도로 결합하지 않는 단백질을 생성하는 IGF2 표적화 펩타이드이다.
- [0261] IGF2(V43M) 표적화 펩타이드는 바람직하게는 M6P 수용체에 특이적으로 표적화된다. 특히 유용한 것은 IGF2 폴리펩타이드에 돌연변이가 있어서 CI-MPR/M6P 수용체에 높은 친화도로 결합하는 반면 다른 두 수용체에는 더 이상 상당한 친화도로 결합하지 않는 단백질을 생성하는 IGF2(V43M) 표적화 펩타이드이다.
- [0262] IGF2(V43M) 표적화 펩타이드는 또한 IGF2 작제물의 격리를 피하기 위해 혈청 IGF-결합 단백질(IGFBP) (Baxter (2000) *Am. J. Physiol Endocrinol Metab.* 278(6):967-76) 및 IGF-I 수용체에 대한 결합을 최소화하도록 변형될 수 있다. 많은 연구에서 IGF-결합 단백질에 결합하는 데 필요한 IGF-1 및 IGF2에 국소화된 잔기가 있다. 이들 잔기에 돌연변이가 있는 작제물은 M6P/IGF2 수용체에 대한 높은 친화도 결합의 보유 및 IGF-결합 단백질에 대한 감소된 친화도에 대해 스크리닝될 수 있다. 예를 들어, IGF2의 Phe 26을 Ser로 대체하는 것은 M6P/IGF2 수용체에 대한 결합에 영향을 미치지 않으면서 IGFBP-1 및 -6에 대한 IGF2의 친화도를 감소시키는 것으로 보고되었다(Bach et al. (1993) *J. Biol. Chem.* 268(13):9246-54). Phe 19에 대한 Ser 및 Glu 9에 대한 Lys와 같은 다른 치환도 유리할 수 있다. IGF2와 고도로 보존된 IGF-1 영역에서 개별적으로 또는 조합된 유사 돌연변이는 IGF-BP 결합을 크게 감소시킨다(Magee et al. (1999) *Biochemistry* 38(48): 15863-70).
- [0263] IGF2 표적화 펩타이드는 또한 IGF2 작제물의 격리를 피하기 위해 혈청 IGF-결합 단백질(IGFBP) 및 IGF-I 수용체에 대한 결합을 최소화하도록 변형될 수 있다.
- [0264] 일부 구현예에서, IGF2 표적화 펩타이드는 퓨린 내성, 즉 Arg-X-X-Arg 절단 부위를 인식하는 퓨린 프로테아제에 의한 분해에 내성이 되도록 변형된다. 이러한 IGF2 표적화 펩타이드는 그 전체가 본원에 참조로 포함된 미국 출원 22012/0213762에 개시되어 있다. 일부 구현예에서, 본원에 기재된 바와 같은 rAAV 게놈에서 사용하기 위한 퓨린 내성 IGF2 표적화 펩타이드는 서열 번호 5(wt IGF2 표적화 펩타이드)의 아미노산 30-40(예를 들어, 31-40, 32-40, 33-40, 34-40, 30-39, 31-39, 32-39, 34-37, 32-39, 33-39, 34-39, 35-39, 36-39, 37-40, 34-40)에 상응하는 영역 내에 돌연변이를 함유하며 임의의 다른 아미노산으로 치환되거나 결실될 수 있다. 예를 들어, 위치 34에서의 치환은 제1 절단 부위의 퓨린 인식에 영향을 미칠 수 있다. 각 인식 부위 내에 하나 이상의 추가 아미노산을 삽입하면 퓨린 절단 부위 중 하나 또는 둘 모두가 제거될 수 있다. 축퇴 위치에 하나 이상의 잔기를 결실하면 두 퓨린 절단 부위가 모두 제거될 수도 있다.
- [0265] 일부 구현예에서, 퓨린-내성 IGF2 표적화 펩타이드는 서열 번호 5의 Arg37(R37) 또는 Arg40(R40)에 상응하는 위치에서 아미노산 치환을 포함한다. 일부 구현예에서, 퓨린-내성 IGF2 표적화 펩타이드는 서열 번호 5의 Arg37 또는 Arg40 위치에 Lys(K) 또는 Ala(A) 치환을 포함한다. 위치 37 및 40 둘 모두에서 Lys 및/또는 Ala 돌연변이의 조합, 또는 Lys(K) 또는 Ala(A) 이외의 아미노산 치환을 포함한 다른 치환이 가능하다. 일부 구현예에서, 본원에 개시된 바와 같은 rAVV 게놈에 사용하기 위해 포함되는 IGF2 표적화 펩타이드는 IGF Δ 2-7-K37, 또는 IgF Δ 2-7-K40 또는 IGF Δ 1-7-K37 또는 IGF Δ 1-7-K40이며, 이는 IGF2 표적화 펩타이드가 각각 2-7 또는 1-7의 결실 및 위치 37에서 라이신(즉, R37K 변형) 또는 R40K 각각으로의 Arg (R) 잔기의 변형을 갖는다는 것을 나타낸다. 일부 구현예에서, 본원에 개시된 바와 같은 rAVV 게놈에 사용하기 위해 포함되는 IGF2 표적화 펩타이드는 IGF Δ 2-7-K37-K40 또는 IgF Δ 1-7-R37K-R40K이며, 이는 IGF2 표적화 펩타이드가 잔기 2-7 또는 잔기 1-7의 결실 및 위치 37 및 위치 40에서 리시닌으로의 R 잔기의 변형(R37K 및 R40K)을 갖는다는 것을 나타낸다. 일부 구현예에서, 본원에 개시된 바와 같은 rAVV 게놈에 사용하기 위해 포함되는 IGF2 표적화 펩타이드는 IGF Δ 2-7-R37A, 또는 IgF Δ 2-7-R40A 또는 IGF Δ 1-7-R37A 또는 IGF Δ 1-7-R40A, IGF Δ 2-7-R37A-R40A, 또는 IgF Δ 1-7-R37A-R40A 중 어느 하나로부터 선택된다. 본원에 개시된 바와 같은 rAVV 게놈에 사용하기 위해 포함되는 IGF2 표적화 펩타이드에 대한 예시적인 작제물은 그 전체가 본원에 참조로 포함된 미국 출원 2012/0213762에 개시되어 있다.
- [0266] 일부 구현예에서, 본 발명에 적합한 퓨린-내성 IGF2 표적화 펩타이드는 추가 돌연변이를 포함할 수 있다. 예를 들어, 서열 번호 5의 잔기의 최대 30% 이상이 변경될 수 있다(예를 들어, 최대 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30% 또는 그 이상의 잔기가 변경될 수 있음). 따라서, 본 발명에 적합한 퓨린-내성 IGF2 뮤테인은 서열 번호 5와 적어도 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90,

91, 92, 93, 94, 95, 96, 97, 98, 99와 동일함을 포함하여 적어도 70% 동일한 아미노산 서열을 가질 수 있다.

[0267] 더욱이, IGF2 표적화 펩타이드가 리소솜을 표적화하는 모이어티로서 M6P를 대체하기 때문에, 본원에 개시된 바와 같은 IGF2 표적화 펩타이드의 사용은 당업계에서 글리코실화 독립적 리소솜 표적화(GILT)로도 지칭된다. GILT 기술에 대한 세부사항은 그 개시내용 모두 본원에 참조로 포함된 미국 출원 공개 번호 2003/0082176, 2004/0006008, 2004/0005309, 2003/0072761, 2005/0281805, 2005/0244400, 및 국제 공개 WO 03/032913, WO 03/032727, WO 02/087510, WO 03/102583, WO 2005/078077에 기재되어 있다.

[0268] 본원에 개시된 바와 같은 방법 및 조성물에 사용하기 위한 IGF2 표적화 펩타이드의 아미노산 서열에 대한 다른 변형은 2019년 11월 19일에 출원된 미국 가출원 62,937,556 및 2019년 11월 15일에 출원된 PCT 출원 PCT/US19/61653에 개시되어 있으며, 상기 특허 둘 다 그 전체가 본원에 참조로 포함된다,

[0269] IGF2는 상대적으로 높은 친화도로 IGF2/M6P 및 IGF-1 수용체에 결합하고 인슐린 수용체에 대해서는 낮은 친화도로 결합한다. IGF2 잔기 48-50(Phe Arg Ser)을 인슐린(Thr Ser Ile)으로부터의 상응하는 잔기로 치환하거나, 잔기 54-55(Ala Leu)를 IGF-I(Arg Arg)로부터의 상응하는 잔기로 치환하면 IGF2/M6P 수용체에 대한 결합이 감소되지만 IGF-1 및 인슐린 수용체에 대한 결합은 유지된다(Sakano et al. (1991) J. Biol. Chem. 266(31):20626-35).

[0270] IGF2는 양이온-비의존적 M6P 수용체의 반복체 11에 결합한다. 실제로, 반복체 11만이 양이온-비의존적 M6P 수용체의 막관통 및 세포질 도메인에 융합된 미니수용체는 IGF2에 결합할 수 있고(전장 수용체의 친화도의 대략 10분의 1의 친화도로) IGF2의 내재화 및 리소솜으로의 전달을 매개할 수 있다(Grimme et al. (2000) J. Biol. Chem. 275(43):33697-33703). M6P 수용체의 도메인 11의 구조는 알려져 있다(Protein Data Base entries IGP0 and IGP3; Brown et al. (2002) EMBO J. 21(5):1054-1062). 추정되는 IGF2 결합 부위는 IGF2의 소수성 아미노산과 상호작용하는 것으로 여겨지는 소수성 포켓이다; IGF2의 후보 아미노산은 류신 8, 페닐알라닌 48, 알라닌 54 및 류신 55를 포함한다. 반복체 11이 IGF2 결합에 충분하지만 양이온-비의존적 M6P 수용체(예를 들어, 반복체 10-13 또는 1-15)의 더 큰 부분을 포함하는 작제물은 일반적으로 더 큰 친화도 및 증가된 pH 의존적으로 IGF2에 결합한다(예를 들어, 문헌(Linnell et al. (2001) J. Biol. Chem. 276(26):23986-23991) 참조).

[0271] IGF2 잔기 Tyr 27을 Leu로, 또는 Ser 26을 Phe로 치환하면 IGF-I 수용체에 대한 IGF2의 치환도가 각각 94배, 56배 및 4배 감소한다(Torres et al. (1995) J. Mol. Biol. 248(2):385-401). 인간 IGF2의 잔기 1-7을 결실시키면 인간 IGF-1 수용체에 대한 친화도가 30배 감소하고 래트 IGF2 수용체에 대한 친화도가 동시에 12배 증가한다(Hashimoto et al. (1995) J. Biol. Chem. 270(30):18013-8). IGF2(잔기 62-67)의 C-말단 절단은 또한 IGF-1 수용체에 대한 IGF2의 친화도를 5배까지 낮추는 것으로 보인다(Roth et al. (1991) Biochem. Biophys. Res. Commun. 181(2):907-14).

[0272] IGF2 잔기 페닐알라닌 26을 세린으로 치환하면 IGFBP 1-5에 대한 결합이 5-75배 감소한다(Bach et al. (1993) J. Biol. Chem. 268(13):9246-54). IGF2 잔기 48-50을 트레오닌-세린-이소류신으로 대체하면 대부분의 IGFBP에 대한 결합이 100배 초과 감소한다(Bach et al. (1993) J. Biol. Chem. 268(13):9246-54); 그러나, 이러한 잔기는 양이온-비의존적 만노오스-6-포스페이트 수용체에 결합하는 데에도 중요하다. IGF-I 수용체에 대한 결합을 분열하는 Y27L 치환은 IGFBP3 및 산 불완정성 소단위와의 삼원 복합체의 형성을 방해한다(Hashimoto et al. (1997) J. Biol. Chem. 272(44):27936-42); 이 삼원 복합체는 순환계에서 대부분의 IGF2를 차지한다(Yu et al. (1999) J. Clin. Lab Anal. 13(4):166-72). IGF2의 처음 6개 잔기의 결실도 IGFBP 결합을 방해한다(Luthi et al. (1992) Eur. J. Biochem. 205(2):483-90).

[0273] IGFBP와 IGF-I 상호작용에 대한 연구는 페닐알라닌 16에 대한 세린의 치환이 2차 구조에 영향을 미치지 않았지만 IGFBP 결합을 40배와 300배 사이로 감소시켰다는 것을 추가로 밝혀냈다(Magee et al. (1999) Biochemistry 38(48):15863-70). 글루타메이트 9를 라이신으로 변경하면 IGFBP 결합도 크게 감소했다. 또한, 이중 돌연변이 라이신 9/세린 16은 IGFBP에 대해 가장 낮은 친화도를 나타냈다. IGF-1과 IGF2의 이 영역 사이의 서열 보존은 유사한 돌연변이가 IGF2(글루타메이트 12 라이신/페닐알라닌 19 세린)에서 만들어질 때 유사한 효과가 관찰될 것임을 시사한다.

[0274] 일부 구현예에서, IGF2(V43M) 서열은 적어도 아미노산 48-55; 적어도 아미노산 8-28 및 41-61; 또는 적어도 아미노산 8-87, 또는 양이온-비의존적 만노오스-6-포스페이트 수용체에 결합하는 이의 서열 변이체(예를 들어, R68A) 또는 이의 절단된 형태(예를 들어, 위치 62로부터 C-말단이 절단됨)를 포함한다.

[0275] 본 발명의 또 다른 구현예에서, 표적화 펩타이드(예를 들어, IGF2 표적화 펩타이드)를 암호화하는 rAAV 게놈은

성숙한 70/76 kDaI 폴리펩타이드와 C-말단 도메인의 접합부, 예를 들어 위치 791에서 천연 GAA 암호화 서열 내로 삽입된다. 이것은 단일 키메라 폴리펩타이드를 생성한다. 일부 구현예에서, 프로테아제 절단 부위는 표적화 펩타이드(예를 들어, IGF2 표적화 펩타이드)의 다운스트림에 근접하여 삽입될 수 있다.

[0276] 하나의 구현예에서, 표적화 펩타이드, 예를 들어, 본원에 정의된 바와 같은 IGF2 표적화 펩타이드는 GAA 폴리펩타이드의 N- 또는 C-말단에 직접 융합된다. 또 다른 구현예에서, IGF2 표적화 펩타이드는 스페이서에 의해 GAA 폴리펩타이드의 N- 또는 C-말단에 융합된다. 하나의 특정 구현예에서, IGF2 표적화 펩타이드는 10-25개 아미노산의 스페이서에 의해 GAA 폴리펩타이드에 융합된다. 또 다른 구현예에서, IGF2 표적화 펩타이드는 글리신 잔기를 포함하는 스페이서에 의해 GAA 폴리펩타이드에 융합된다.

[0277] 일부 구현예에서, IGF2 표적화 펩타이드는 적어도 1, 2, 또는 3개 아미노산의 스페이서에 의해 GAA 폴리펩타이드에 융합된다. 일부 구현예에서, 스페이서는 아미노산 GAP 또는 Gly-Ala-Pro(서열 번호 31), 또는 이와 적어도 50% 동일한 아미노산 서열을 포함한다. 일부 구현예에서, 스페이서는 GGG 또는 GA 또는 AP, 또는 GP 또는 이의 변이체이다. 일부 구현예에서, 스페이서는 핵산 GGC GCG CCG(서열 번호 30)에 의해 암호화된다.

[0278] 일부 구현예에서, IGF2 표적화 펩타이드는 나선 구조를 포함하는 스페이서에 의해 GAA 폴리펩타이드에 융합된다. 또 다른 특정 구현예에서, IGF2 표적화 펩타이드는 서열 GGGTVGDDDDK(서열 번호 35)와 50% 이상 동일한 스페이서에 의해 GAA 폴리펩타이드에 융합된다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, 스페이서는 서열 번호 31(서열 번호 30의 핵산에 의해 암호화됨)이다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, 스페이서는 서열 번호 31, 서열 번호 32, 서열 번호 33, 서열 번호 34 또는 서열 번호 35 중 어느 하나, 또는 이와 적어도 적어도 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 서열로부터 선택된다.

[0279] **(iii) 양이온-비의존적 M6P 수용체(CI-MPR)에 결합하는 대안적인 표적화 펩타이드.**

[0280] 일부 구현예에서, 표적화 펩타이드는 리소솜 표적화 펩타이드 또는 단백질, 또는 만노오스-6-포스페이트-비의존적 방식으로 양이온 비의존적 M6P/IGF2 수용체(CI-MPR)에 결합하는 본원에 개시된 IGF2 표적화 펩타이드 이외의 다른 모이어티이다. CI-MPR은 또한 표적 펩타이드로서 사용될 수 있는 적어도 3개의 별개의 리간드에 대한 결합 부위를 포함한다. 본원에 개시된 바와 같이, IGF2 리간드는 주로 반복체 11과의 상호작용을 통해 약 pH 7.4에서 약 14 nM의 해리 상수로 CI-MPR에 결합한다. CI-MPR은 고분자량 O-글리코실화 IGF2 형태에 결합할 수 있다. 따라서, 일부 구현예에서, IGF2 표적화 펩타이드는 O-글리코실화를 포함하도록 전사후 변형될 수 있다.

[0281] 대안적인 구현예에서, CI-MPR에 결합하는 표적화 펩타이드 레티노산이다. 레티노산은 2.5 nM의 해리 상수로 수용체에 결합한다. 레티노산을 사용한 양이온-비의존적 M6P 수용체의 친화도 광표지는 수용체에 대한 IGF2 또는 M6P 결합을 방해하지 않으며, 이는 레티노산이 수용체의 별개의 부위에 결합함을 나타낸다. 수용체에 대한 레티노산의 결합은 세포질 소포에서 수용체의 더 많은 축적으로 수용체의 세포내 분포를 변경하고 또한 M6P 변형된 β -글루쿠로니다제의 흡수를 향상시킨다. 레티노산은 양이온-비의존적 M6P 수용체에 결합하는 능력을 방해하지 않으면서 이를 치료제에 연결하는 데 사용할 수 있는 광활성 모이어티를 갖고 있다.

[0282] 유로키나제형 플라스미노겐 수용체(uPAR)는 또한 9 μ M의 해리 상수로 CI-MPR에 결합한다. uPAR은 접착 분자로서 기능하고 플라스미노겐 및 TGF- β 의 단백질분해 활성화에서 기능하는 대부분의 세포 유형의 표면 상의 GPI-고정 수용체이다. CI-M6P 수용체에 대한 uPAR의 결합은 이를 리소솜으로 표적화하여 활성을 조절한다. 따라서, uPAR의 세포외 도메인, 또는 양이온-비의존적 M6P 수용체에 결합할 수 있는 이의 부분을 치료제에 융합시키는 것은 리소솜에 대한 제제의 표적화를 허용한다.

[0283] **D. GAA 폴리펩타이드의 스페이서 및 융합 접합**

[0284] GAA가 분비 신호 펩타이드(예를 들어, SS-GAA 융합 폴리펩타이드) 또는 표적 펩타이드(즉, SS-IGF2-GAA 폴리펩타이드 이중 융합 폴리펩타이드)와의 융합 단백질로서 발현되는 경우, 신호 펩타이드 또는 IGF2 표적화 펩타이드는 GAA 폴리펩타이드에 직접 융합될 수 있거나 링커에 의해 GAA 폴리펩타이드로부터 분리될 수 있다. 아미노산 링커(본원에서 "스페이서"로도 지칭됨)는 천연 단백질의 해당 위치에 나타나는 것 이외의 하나 이상의 아미노산을 포함한다. 스페이서는 일반적으로 2개의 단백질 모이어티 사이에 α -나선과 같은 구조를 삽입하거나 가요성을 갖도록 설계될 수 있다.

[0285] 따라서, 본원에 개시된 방법 및 조성물의 일부 구현예에서, 재조합 AAV 벡터는 IGF2-GAA 융합 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서 IGF2-GAA 융합 단백질은 적어도 1개의 아미노산 길이의 뉴클레오타이드 서열을 포함하는 스페이서를 추가로 포함하며, 이는 GAA 폴리펩타이드의 N-말단 및 IGF2 표적화 펩타

이드의 C-말단에 위치한다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, 재조합 AAV 벡터는 IGF2 표적화 펩타이드를 암호화하는 핵산과 GAA 폴리펩타이드를 암호화하는 핵산 사이에 위치한 적어도 1개 아미노산의 스페이서를 암호화하는 핵산을 포함하는 이중 핵산 서열을 포함한다.

[0286] 하나의 구현예에서, IGF2 표적화 펩타이드는 GAA 폴리펩타이드의 N- 또는 C-말단에 직접 융합된다. 또 다른 구현예에서, IGF2 표적화 펩타이드는 스페이서에 의해 GAA 폴리펩타이드의 N- 또는 C-말단에 융합된다. 하나의 특정 구현예에서, IGF2 표적화 펩타이드는 10-25개 아미노산의 스페이서에 의해 GAA 폴리펩타이드에 융합된다. 또 다른 특정 구현예에서, IGF2 표적화 펩타이드는 글리신 잔기를 포함하는 스페이서에 의해 GAA 폴리펩타이드에 융합된다. 또 다른 특정 구현예에서, IGF2 표적화 펩타이드는 나선 구조를 포함하는 스페이서에 의해 GAA 폴리펩타이드에 융합된다. 또 다른 특정 구현예에서, IGF2 표적화 펩타이드는 서열 GGGTVGDDDDK(서열 번호 35)와 적어도 50% 동일한 스페이서에 의해 GAA 폴리펩타이드에 융합된다.

[0287] 일부 구현예에서, 스페이서 또는 링커는 상대적으로 짧을 수 있으며, 예를 들어, 적어도 1, 2, 3, 4 또는 5개 아미노산, 또는 예컨대 서열 Gly-Ala-Pro(서열 번호 31) 또는 Gly-Gly-Gly-Gly-Gly-Pro(서열 번호 32)이거나, 예를 들어, 5-10개 아미노산 길이 또는 10-25개 아미노산 길이와 같이 더 길 수 있다. 예를 들어, 서열의 3-4개 카피의 가요성 반복 링커(GGGGS(서열 번호 33)) 및 서열의 2-5개 카피의 α -나선형 반복 링커(EAAAK(서열 번호 34))가 기술되었다(Arai et al. (2004) Proteins: Structure, Function and Bioinformatics 57:829-838).

[0288] IGF2 융합 단백질과 관련하여 또 다른 링커인 GGGTVGDDDDK(서열 번호 35)의 사용이 또한 보고되었으며(DiFalco et al.(1997) Biochem. J. 326:407-413) 사용을 위해 포함된다. 인간 혈청 단백질의 α -나선 부분을 포함하는 링커를 사용하여 링커 영역의 면역원성을 최소화할 수 있다.

[0289] 일부 구현예에서, 스페이서는 아미노산 GAP 또는 Gly-Ala-Pro(서열 번호 31)를 포함하는 아미노산 스페이서를 암호화하는 핵산 GGC GCG CCG(서열 번호 30)에 의해 암호화된다.

[0290] 신호 펩타이드(SS-GAA 융합 단백질을 생성하기 위해) 또는 표적화 펩타이드(예를 들어, SP-IGF2-GAA 이중 융합 폴리펩타이드를 생성하기 위해)와 융합하기 위한 GAA 폴리펩타이드에서 융합 접합 부위는 융합 단백질에서 각 폴리펩타이드의 적절한 폴딩 및 활성을 촉진하고 GAA 폴리펩타이드로부터 신호 펩타이드의 조기 분리를 방지하기 위해 주의하여 선택해야 한다.

[0291] 일부 구현예에서, IGF2 표적화 펩타이드는 나선 구조를 포함하는 스페이서에 의해 GAA 폴리펩타이드에 융합된다. 또 다른 특정 구현예에서, IGF2 표적화 펩타이드는 서열 GGGTVGDDDDK(서열 번호 35)와 적어도 50% 동일한 스페이서에 의해 GAA 폴리펩타이드에 융합된다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, 스페이서는 서열 번호 31(서열 번호 30의 핵산에 의해 암호화됨)이다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, 스페이서는 서열 번호 31, 서열 번호 32, 서열 번호 33, 서열 번호 34 또는 서열 번호 35 중 임의의 것으로부터 선택된다.

[0292] IGF2-GAA 융합 단백질을 생성하기 위한 4가지 예시적인 전략을 발생시킬 수 있으며, 둘 다 그 전체가 본원에 참조로 포함된 2019년 11월 19일에 출원된 가출원 62,937,556 및 2019년 11월 15일에 출원된 PCT/US19/61653에 개시되어 있다.

[0293] 일부 구현예에서, 표적화 펩타이드(예를 들어, IGF2 표적화 펩타이드)는 단백질의 발현, GAA 단백질의 촉매 활성, 및 실시예에서 본원에 기재된 바와 같은 IGF2 표적화 펩타이드에 의한 적절한 표적화를 허용하는 위치인 GAA의 아미노산 40 또는 아미노산 70에 직접적으로 또는 스페이서에 의해 융합될 수 있다. 대안적으로, 표적화 펩타이드(예를 들어, IGF2 표적화 펩타이드)는 성숙한 폴리펩타이드로부터 GAA의 C-말단 도메인을 분리하는 절단 부위에서 또는 그 근처에서 융합될 수 있다. 이것은 내부 표적화 펩타이드(예를 들어, IGF2 표적화 펩타이드)를 갖는 GAA 단백질의 합성을 허용하며, 이는 절단 부위의 위치에 따라 임의로 절단되어 표적화 도메인으로부터 성숙한 폴리펩타이드 또는 C-말단 도메인을 유리시킬 수 있다. 대안적으로, 성숙한 폴리펩타이드는 발현 자체물의 오픈 리딩 프레임에 C-말단 서열을 혼입하지 않고 약 위치 791에서 융합 단백질로서 합성될 수 있다.

[0294] IGF2 표적화 펩타이드의 폴딩을 용이하게 하기 위해, 융합 접합부에 인접한 GAA 아미노산 잔기가 변형될 수 있다. 예를 들어, GAA 시스테인 잔기가 표적화 펩타이드(예를 들어, IGF2 표적화 펩타이드)의 적절한 폴딩을 방해할 수 있기 때문에, 말단 GAA 시스테인 952가 결실되거나 C-말단 표적화 펩타이드(예를 들어, IGF2 표적화 펩타이드)를 수용하기 위해 세린으로 치환될 수 있다. 표적화 펩타이드(예를 들어, IGF2 표적화 펩타이드)는 또한 최종 Cys952 직전에 융합될 수 있다. 끝에서 두 번째(penultimate) cys938은 최종 Cys952의 세린으로의 돌

연변이와 함께 프롤린으로 변경될 수 있다.

[0295] E. CS 서열

[0296] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 재조합 AAV 벡터는 GAA 폴리펩타이드를 암호화하는 핵산의 3' 및 3' ITR 서열의 5'에 위치한 콜라겐 안정성(CS) 서열을 추가로 포함하는 이중 핵산 서열을 포함한다. 일부 구현예에서, 본원에 개시된 rAAV 게놈은 GAA 유전자의 3' 및 폴리A 신호의 5'에 위치하는 콜라겐 안정성 서열(CS 또는 CSS)을 임의로 포함할 수 있는 이중 핵산 서열을 포함한다. 일부 구현예에서, CS 서열은 본원에 개시된 바와 같은 3' UTR 서열로 대체될 수 있다.

[0297] 예시적인 콜라겐 안정성 서열은 CCCAGCCCACCTTTTCCCAA(서열 번호 65) 또는 이와 적어도 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 서열을 포함한다. 예시적인 콜라겐 안정성 서열은 P S P L F P의 아미노산 서열(서열 번호 66) 또는 이와 적어도 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 아미노산 서열을 가질 수 있다. CS 서열은 그 전체가 본원에 참조로 포함된 문헌(Holick and Liebhaver, Proc. Nat. Acad. Sci. 94: 2410-2414, 1997 (예를 들어, 도 3, p. 5205 참조))에 개시되어 있다.

[0298] F. 프로모터

[0299] 일부 구현예에서, 적절한 수준의 GAA 발현을 달성하기 위해, rAAV 유전자형은 간 특이적 프로모터(LSP)를 포함한다. LSP는 간에서 작동적으로 연결된 유전자의 발현을 가능하게 하고, 일부 구현예에서, LSP일 수 있고 유도할 수 있다. 일 구현예에서, LSP는 5' 업스트림에 위치하고 GAA 단백질을 암호화하는 이중 핵산 서열에 작동적으로 연결된다. 예시적인 간-특이적 프로모터는 본원에 개시되어 있으며, 예를 들어, 서열 번호 86, 91-96 또는 146-150을 포함하는 LSP, 또는 이의 기능적 변이체 또는 기능적 단편, 또는 본원의 표 4에 나열된 임의의 LSP, 또는 이의 기능적 단편 또는 기능적 변이체를 포함한다. 본원에 개시된 조성물 및 방법의 일부 구현예에서, 간-특이적 프로모터는 간-특이적 시스-조절 요소(CRE), 합성 간-특이적 시스-조절 모듈(CRM)을 포함하거나, 합성 간-특이적 프로모터는 서열 번호 270-341(CRM을 갖는 최소 LSP) 또는 서열 번호 342-430(본원의 표 4에 개시된 합성 간 특이적 근위 프로모터)으로부터 선택된다. 일부 구현예에서, rAAV 벡터 게놈은 전사 촉진에 일반적으로 활성인 포유동물 유전자로부터의 바이러스 프로모터 또는 프로모터와 같은 하나 이상의 작동적 프로모터를 포함할 수 있다.

[0300] (i) 합성 간-특이적 프로모터

[0301] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 프로모터는 간 특이적 프로모터이며, 본원에 개시된 표 4에 나열된 것들 또는 이의 기능적 변이체 및/또는 2019년 11월 19일에 출원된 미국 가출원 62,937,556의 표 4A 및 4B로부터 선택된 임의의 것 또는 이의 기능적 변이체를 포함하지만 이에 제한되지 않는 프로모터로부터 선택될 수 있다.

[0302] 트랜스티레틴 프로모터(TTR) (서열 번호 431) 및 SP0412(서열 번호 91) 및 SP0422(서열 번호 92)가 명세서 및 실시예에서 예시적인 간 특이적 프로모터(실시예 1, 12 및 13 참조)로서 사용되지만, 당업자는 TTR을 본원의 표 4에 개시된 바와 같은 임의의 간 특이적 프로모터 또는 이의 기능적 변이체 및/또는 2019년 11월 19일에 출원된 미국 가출원 62,937,556의 표 4A 및 4B로부터 선택된 임의의 것 또는 이의 기능적 변이체로 쉽게 대체할 수 있다. 간-특이적 프로모터는 간-특이적 시스-조절 요소(CRE), 합성 간-특이적 시스-조절 모듈(CRM) 또는 2019년 11월 19일에 출원된 미국 가출원 62,937,556의 표 4A 및 4B에 개시된 합성 간-특이적 프로모터 또는 이의 기능적 변이체를 포함한다.

[0303] 표 4는 예시적인 간-특이적 프로모터를 나타낸다. 본원에 개시된 비교적 작은 크기의 간-특이적 프로모터는 벡터의 페이로드의 최소량을 차지하기 때문에 유리하다. 이는 AAV 기반 벡터와 같이 용량이 제한된 벡터에서 LSP를 사용할 때 특히 중요하다.

[0304] [표 4] 본원에 개시된 바와 같은 방법 및 조성물에 사용하기 위한 서열 번호에 의해 확인된 예시적인 LSP

표 4: 예시적인 LSP					
서열 번호	LSP 명칭	서열 번호	LSP 명칭	서열 번호	LSP 명칭
270	CRM_SP0107	330	CRM_SP0396	384	SP0269 (LVR_132_V1)
271	CRM_SP0109	331	CRM_SP0397	385	SP0270 (LVR_132_V2)
273	CRM_SP0111	332	CRM_SP0398	386	SP0271 (LVR_133_A1)
274	CRM_SP0112	333	CRM_SP0399	387	SP0272 (LVR_133_V1)
275	CRM_SP0113	334	CRM_SP0403	388	SP0273 (LVR_133_V2)
276	CRM_SP0115	335	CRM_SP0404	389	SP0256
277	CRM_SP0116	336	CRM_SP0405	390	SP0257
278	CRM_SP0121	337	CRM_SP0406	391	SP0258
279	CRM_SP0124	338	CRM_SP0407	392	SP0259
280	CRM_SP0127 (CRM_LVR_127)	339	CRM_SP0409	393	SP0264
281	CRM_SP0127A1 (CRM_LVR_127_A1)	340	CRM_SP0411	394	SP0265(LVR_SP1 31_A1)
282	CRM_SP0127V1 (CRM_LVR_127_V1)	86	CRM_SP0412	395	SP0266(LVR_SP1 31_V1)
283	CRM_SP0127V2 (CRM_LVR_127_V2)	341	CRM_SP0413	396	SP0267(LVR_SP1 31_V2)
284	CRM_SP0128	342	SP0107	397	SP0268 (LVR_132_A1)
285	CRM_SP0131 (CRM_LVR_131)	343	SP0109	398	SP0269 (LVR_132_V1)
286	CRM_SP0132 (CRM_LVR_132)	344	SP0111	399	SP0270 (LVR_132_V2)
287	CRM_SP0133 (CRM_LVR_133)	345	SP0112	400	SP0271 (LVR_133_A1)
288	CRM_SP0155	346	SP0113	401	SP0272 (LVR_133_V1)
289	CRM_SP0158	347	SP0115	402	SP0273 (LVR_133_V2)
290	CRM_SP0163	348	SP0116	403	SP0368
291	CRM_SP0236	349	SP0121	404	SP0373
292	CRM_SP0239	350	SP0124	405	SP0378
293	CRM_SP0240	351	SP0127 (LVR_SP127)	406	SP0379

[0305]

294	CRM_SP0241	352	SP0127A1 (LVR_SP127_A1)	407	SP0380
295	CRM_SP0242	353	SP0127V1 (LVR_SP127_V1)	408	SP0381
296	CRM_SP0243	354	SP0127V2 (LVR_SP127_V2)	409	SP0272 (LVR_133_V1)
297	CRM_SP0244	355	SP0128	410	SP0273 (LVR_133_V2)
298	CRM_SP0246	356	SP0131 (LVR_SP131)	411	SP0368
299	CRM_SP0247	357	SP0132 (LVR_SP132)	412	SP0373
300	CRM_SP0248	358	SP0133 (LVR_SP133)	413	SP0378
301	CRM_SP0249	359	SP0155	414	SP0379
302	CRM_SP0250	360	SP0158	415	SP0380
303	CRM_SP0251	361	SP0163	416	SP0381
304	CRM_SP0252	362	SP0236	417	SP0384
305	CRM_SP0253	93	SP0239	418	SP0388
306	CRM_SP0254	95	SP0240	419	SP0396
307	CRM_SP0255	363	SP0241	420	SP0397
308	CRM_SP0256	364	SP0242	421	SP0398
309	CRM_SP0257	365	SP0243	422	SP0399
310	CRM_SP0258	366	SP0244	423	SP0403
311	CRM_SP0259	96	SP0246	424	SP0404
312	CRM_SP0264	367	SP0247	425	SP0405
313	CRM_SP0265 (CRM_LVR_131_A1)	368	SP0248	426	SP0406
314	CRM_SP0266 (CRM_LVR_131_V1)	369	SP0249	427	SP0407
315	CRM_SP0267 (CRM_LVR_131_V2)	370	SP0250	428	SP0409
316	CRM_SP0268 (CRM_LVR_132_A1)	371	SP0251	429	SP0411
317	CRM_SP0269 (CRM_LVR_132_V1)	372	SP0252	91	SP0412
318	CRM_SP0270 (CRM_LVR_132_V2)	373	SP0253	92	SP0422
319	CRM_SP0271 (CRM_LVR_133_A1)	374	SP0254	146	SP0265-UTR
320	CRM_SP0272 (CRM_LVR_133_V1)	375	SP0255	147	SP0239-UTR
321	CRM_SP0273 (CRM_LVR_133_V2)	376	SP0256	148	SP0240-UTR
322	CRM_SP0368	377	SP0257	149	SP0246-UTR
323	CRM_SP0373	378	SP0258	150	SP0131-A1-UTR
324	CRM_SP0378	379	SP0259	430	SP0413
325	CRM_SP0379	380	SP0264	431	TTR 프로모터

[0306]

326	CRM_SP0380	94	SP0265 (LVR_SP131_A1)	432	LP1
327	CRM_SP0381	381	SP0266(LVR_SP131_V1)	433	CMV-IE
328	CRM_SP0384	382	SP0267(LVR_SP131_V2)	434	CBA
329	CRM_SP0388	383	SP0268 (LVR_132_A1)	435	TBG 프로모터

[0307]

[0308]

(ii) 간-특이적 프로모터의 기능적 변이체

[0309]

일부 구현예에서, 본원에 개시된 바와 같은 방법 및 조성물에 유용한 합성 간-특이적 프로모터는 본원에 정의된 바와 같은 이중-특이적 또는 삼중-특이적 프로모터이다. 예시적인 예로서, 간 이중-특이적 프로모터는 간 및 하나의 다른 조직, 예를 들어, 근육에서 활성이다. 추가로, 간 이중-특이적 프로모터의 또 다른 예시적인 예는 간 및 하나의 다른 조직, 예를 들어, 뇌에서 활성이다. 간의 예시적인 예로서, 삼중-특이적 프로모터는 간 및 2개의 다른 조직, 예를 들어, 근육 및 뇌에서 활성이다. 추가로, 간 삼중-특이적 프로모터의 또 다른 예시적인 예는 간 및 예를 들어 콩팥 및 근육과 같은 2개의 다른 조직에서 활성이다.

- [0310] 일부 구현예에서, 서열 번호 86, 91-96, 146-150, 270-430 중 어느 하나와 적어도 50%, 60%, 70%, 80%, 90% 또는 95% 동일한 합성 간 특이적 프로모터는 간에서 우선적으로 활성이고 또한 제2 유형의 세포 또는 조직, 예를 들어, 근육 또는 CNS에서 더 적은 정도(예를 들어, 총 발현의 <50%, 또는 약 49-40%, 또는 약 39-30%, 또는 약 29-20% 또는 약 19-10% 또는 <10%)로 활성인 공급원 조절 핵산 서열을 포함한다.
- [0311] 일부 구현예에서, 프로모터는 시스-조절 요소(CRE) CRE0051(서열 번호 97) 및 CRE0042(서열 번호 104), 또는 이의 기능적 변이체의 조합을 포함하는 합성 간-특이적 프로모터이다. 이의 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다. 전형적으로, CRE는 프로모터 요소에 작동적으로 연결된다. 일부 바람직한 구현예에서, 간-특이적 프로모터는 CRE0051(서열 번호 97), CRE0042(서열 번호 104)에 이어서 프로모터 요소(순서는 당업계에서 통상적인 바와 같이 업스트림에서 다운스트림 방향으로 주어짐)의 순서로 상기 CRE 또는 이의 기능적 변이체를 포함한다.
- [0312] 프로모터 요소는 임의의 적합한 근위 프로모터 또는 최소 프로모터일 수 있다. 일부 구현예에서, 프로모터 요소는 최소 프로모터이다. 프로모터가 근위 프로모터인 경우, 근위 프로모터가 간-특이적인 것이 일반적으로 바람직하다.
- [0313] 일부 바람직한 구현예에서, 프로모터 요소는 CRE0059(서열 번호 110), 또는 이의 기능적 변이체이다. CRE0059는 아래에 추가로 논의된 바와 같은 근위 프로모터이다.
- [0314] 따라서, 하나의 구현예에서, 프로모터는 다음 조절 요소를 포함한다: CRE0051(서열 번호 97), CRE0042(서열 번호 104) 및 CRE0059(서열 번호 110), 또는 이의 기능적 변이체.
- [0315] CRE0051(서열 번호 97)의 기능적 변이체는 CRE0051과 상이한 서열을 갖는 조절 요소이지만, 간-특이적 CRE로서 활성을 실질적으로 유지한다. 당업자는 필수 전사 인자(TF)에 결합하고 발현을 향상시키는 능력을 유지하면서 CRE의 서열을 변화시키는 것이 가능하다는 것을 인식할 것이다. 기능적 변이체는 이들이 CRE를 실질적으로 비기능적으로 만들지 않는 경우 참조 CRE와 비교하여 치환, 삭제 및/또는 삽입을 포함할 수 있다.
- [0316] 일부 구현예에서, CRE0051의 기능적 변이체는 프로모터에서 CRE0051 대신에 치환될 때 이의 활성을 실질적으로 유지하는 CRE로 볼 수 있다. 예를 들어, CRE0051 대신에 치환된 CRE0051의 기능적 변이체를 포함하는 간-프로모터는 바람직하게는 이의 활성의 80%, 보다 바람직하게는 이의 활성의 90%, 보다 바람직하게는 이의 활성의 95%, 더욱 더 바람직하게는 이의 활성의 100%를 유지한다. 예를 들어, 프로모터 SP0412(서열 번호 91)를 예로 고려하면, SP0412에서 CRE0051은 CRE0051의 기능적 변이체로 대체될 수 있고, 프로모터는 실질적으로 이의 활성을 유지한다. 활성 유지는 참조 프로모터의 제어 하에 적합한 리포터의 발현을 동등한 조건 하에 치환된 CRE를 포함하는 다른 동일한 프로모터와 비교함으로써 평가될 수 있다.
- [0317] 일부 구현예에서, CRE0051의 기능적 변이체는 CRE0051과 동일한 간-특이적 TF에 대한 전사 인자 결합 부위(TFBS)를 포함한다. CRE0051에 존재하는 간-특이적 TFBS는 이들이 존재하는 순서대로 나열되어 있다: HNF1(서열 번호 98), HNF4(서열 번호 99), HNF3(서열 번호 100), HNF1'(서열 번호 101) 및 HNF3'(서열 번호 102), 표 5 참조. 따라서, CRE0051의 기능적 변이체는 바람직하게는 이러한 TFBS를 모두 포함한다. 바람직하게는, 이들은 CRE0051에 존재하는 것과 동일한 순서로, 즉 HNF1(서열 번호 98), HNF4(서열 번호 99), HNF3(서열 번호 100), HNF1'(서열 번호 101) 및 HNF3'(서열 번호 102)의 순서로 존재한다. 시스-조절 요소가 프로모터 및 유전자와 관련될 때, 이 순서는 바람직하게는 업스트림에서 다운스트림 방향으로 고려된다(즉, 전사 개시 부위(TSS)로부터 원위에서 TSS에 대한 근위 방향으로). 인접한 TFBS 사이에 스페이서 서열에 제공될 수 있다. 일부 구현예에서, TFBS는 이들이 기능적으로 남아 있는 경우 적합하게 중첩될 수 있는데, 즉 중첩 서열은 둘 다 발현을 조절하는 데 필요한 정도로 각각의 TF에 결합할 수 있다.
- [0318] 일부 구현예에서, CRE0051(서열 번호 97)의 기능적 변이체는 다음 TFBS 서열을 포함한다: GTTAATTTTAAA(HNF1)(서열 번호 98), GTGGCCCTTGG(HNF4)(서열 번호 99), TGTTCG(HNF3)(서열 번호 100), TGGTAATAATCTCA(HNF1')(서열 번호 101)에 이어서 ACAACA(HNF3)(서열 번호 102), 이에 상보적인 서열, 또는 이들 각각의 TF에 결합하는 능력을 유지하는 이들 TFBS 서열의 기능적 변이체. 이는 CRE0051과 동일한 순서, 즉 상기 제시된 순서로 존재할 수 있다. TFBS와 관련된 서열 가변성이 있고, 주어진 TFBS에 대해, 전형적으로 어느 정도의 편차가 전형적으로 존재하는 컨센서스 서열이 있다는 것은 당업계에 잘 알려져 있다. TFBS에서 발생하는 변이에 대한 추가 정보는, 주어진 뉴클레오타이드가 전형적으로 컨센서스 서열의 주어진 위치에서 발견되는 빈도를 나타내는 위치가중치 매트릭스(PWM: positional weight matrix)를 사용하여 설명할 수 있다. TF 컨센서스 서열 및 관련된 PWM에 대한 세부 정보는, 예를 들어, Jaspar 또는 Transfac 데이터베이스(<http://jaspar.genereg.net/> 및

<http://gene-regulation.com/pub/databases.html>)에서 찾을 수 있다. 이 정보를 통해 당업자는 CRE 기능을 유지하고 일부 경우에는 심지어 증가시키는 방식으로 CRE의 임의의 주어진 TFBS에서 서열을 변형할 수 있다.

- [0319] 일부 구현예에서, CRE0051의 기능적 변이체는 다음 서열:
- [0320] GTTAATTTTAAA-Na-GTGGCCCTTGG-Nb-TGTTGC-Nc-TGGTTAATAATCTCA-Nd-ACAAACA(서열 번호 103), 또는 이와 적어도 70%, 80%, 90%, 95% 또는 99% 동일한 서열을 포함하며, 여기서, Na, Nb, Nc, 및 Nd는 임의의 스페이스 서열을 나타낸다. 존재하는 경우, Na는 임의로 10 내지 26개의 뉴클레오타이드, 바람직하게는 14 내지 22개의 뉴클레오타이드, 보다 바람직하게는 18개의 뉴클레오타이드 길이를 갖는다. 존재하는 경우, Nb는 임의로 8 내지 22개의 뉴클레오타이드, 바람직하게는 12 내지 20개의 뉴클레오타이드, 보다 바람직하게는 16개의 뉴클레오타이드 길이를 갖는다. 존재하는 경우, Nc는 임의로 1 내지 10개의 뉴클레오타이드, 바람직하게는 1 내지 5개의 뉴클레오타이드, 보다 바람직하게는 2개의 뉴클레오타이드 길이를 갖는다. 존재하는 경우, Nd는 적합하게는 1 내지 13개의 뉴클레오타이드, 바람직하게는 2 내지 9개의 뉴클레오타이드 길이, 보다 바람직하게는 5개의 뉴클레오타이드 길이를 갖는다.
- [0321] 일부 구현예에서, CRE는 서열 번호 98-102 또는 이의 기능적 변이체로 이루어진다.
- [0322] CRE 또는 이의 기능적 변이체는 이중 가닥 폴리뉴클레오타이드의 어느 한 가닥에 제공될 수 있으며 어느 방향으로든 제공될 수 있음에 주목할 것이다. 이와 같이, 서열 번호 97-102의 상보적 및 역상보적 서열 또는 이의 기능적 변이체는 본 발명의 범위 내에 속한다. 서열 번호 97 또는 10에 따른 서열 또는 이의 기능적 변이체를 포함하는 단일 가닥 핵산도 본 발명의 범위 내에 속한다.
- [0323] 일부 구현예에서, CRE0051(서열 번호 97) 또는 이의 기능적 변이체를 포함하거나 이로 이루어진 CRE는 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 125개 이하의 뉴클레오타이드 또는 100개 이하의 뉴클레오타이드 길이를 갖는다.
- [0324] 일부 구현예에서, CRE0042(서열 번호 104) 또는 이의 기능적 변이체를 포함하거나 이로 이루어진 CRE는 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 125개 이하의 뉴클레오타이드, 또는 100개 이하의 뉴클레오타이드 길이를 갖는다. 이의 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0325] CRE0042(서열 번호 104)의 기능적 변이체는 CRE0042와는 다른 서열을 가진 조절 요소이지만 간-특이적 CRE로서 이들의 활성을 실질적으로 유지한다. 당업자는 필수 전사 인자(TF)에 결합하고 발현을 향상시키는 능력을 유지하면서 CRE의 서열을 변화시키는 것이 가능하다는 것을 인식할 것이다. 기능적 변이체는 이들이 CRE를 실질적으로 비기능적으로 만들지 않는 경우 참조 CRE와 비교하여 치환, 삭제 및/또는 삽입을 포함할 수 있다.
- [0326] 일부 구현예에서, CRE0042(서열 번호 104)의 기능적 변이체는 프로모터에서 CRE0042 대신에 치환될 때 이의 활성을 실질적으로 유지하는 CRE로 볼 수 있다. 예를 들어, CRE0042 대신에 치환된 CRE0042의 기능적 변이체를 포함하는 프로모터는 (CRE0042(서열 번호 104)를 포함하는 참조 프로모터와 비교하여) 바람직하게는 이의 활성의 80%, 보다 바람직하게는 이의 활성의 90%, 보다 바람직하게는 이의 활성의 95%, 더욱 더 바람직하게는 이의 활성의 100%를 유지한다. 예를 들어, 프로모터 SP0412를 예로서 고려하면, SP412(서열 번호 91)에서 CRE0042(서열 번호 104)는 CRE0042의 기능적 변이체로 대체될 수 있으며, 프로모터는 실질적으로 이의 활성을 유지한다. 활성 유지는 참조 프로모터의 제어 하에 적합한 리포터의 발현을 동등한 조건 하에 치환된 CRE를 포함하는 다른 동일한 프로모터와 비교함으로써 평가될 수 있다.
- [0327] 일부 구현예에서, CRE0042(서열 번호 104)의 기능적 변이체(서열 번호 104)가 CRE0042와 동일한 간-특이적 TF에 대한 TFBS를 포함하는 것이 바람직하다. CRE0042에 존재하는 간-특이적 TFBS는 이들이 존재하는 순서대로 나열되어 있다: HNF-3(서열 번호 106), C/EBP(서열 번호 107), HNF-4(서열 번호 108) 및 C/EBP'(서열 번호 109). 따라서, CRE0042의 기능적 변이체는 바람직하게는 이러한 TFBS 모두를 포함한다. 바람직하게는, 이들은 CRE0042에 존재하는 것과 동일한 순서, 즉 HNF-3, C/EBP, HNF-4 및 이어서 C/EBP'의 순서로 존재한다. 시스-조절 요소가 프로모터 및 유전자와 관련될 때, 이 순서는 바람직하게는 업스트림에서 다운스트림 방향으로 고려된다(즉, 전사 개시 부위(TSS)로부터 원위에서 TSS에 대한 근위 방향으로). 인접한 TFBS 사이에 스페이스 서열에 제공될 수 있다. 일부 구현예에서, TFBS는 이들이 기능적으로 남아 있는 경우 적합하게 중첩될 수 있는데, 즉 중첩 서열은 둘 다 각각의 TF에 결합할 수 있다.
- [0328] 일부 구현예에서, CRE0042(서열 번호 104)의 기능적 변이체는 다음 TFBS 서열을 포함한다: GTTCAACATG(HNF-3)(서열 번호 106), CTAATACTCTG(C/EBP)(서열 번호 107), TGCAAGGTCAT(HNF-4)(서열 번호 108), 및

TTACTCAACA(C/EBP) (서열 번호 109) 및 이에 상보적인 서열, 또는 이들 각각의 TF에 결합하는 능력을 유지하는 이들 TFBS 서열의 기능적 변이체. 이는 CRE0042와 동일한 순서, 즉 상기 제시된 순서로 존재할 수 있다. 상기 논의된 바와 같이, TFBS와 관련된 서열 가변성이 있고, 주어진 TFBS에 대해, 전형적으로 어느 정도의 편차가 전형적으로 존재하는 컨센서스 서열이 있다는 것은 당업계에 잘 알려져 있다.

[0329] 본 발명의 일부 구현예에서, CRE0042의 기능적 변이체는 다음 서열:

[0330] GTTCAACATG-Na-CTAATACTCTG-Nb-TGCAAGGGTCAT-Nc-TTACTCAACA(서열 번호 105) 또는 이와 적어도 70%, 80%, 90%, 95% 또는 99% 동일한 서열을 포함하며, 여기서, Na, Nb 및 Nc는 임의의 스페이스 서열을 나타낸다. 존재하는 경우, Na는 임의로 1 내지 10개의 뉴클레오타이드, 바람직하게는 1 내지 5개의 뉴클레오타이드, 보다 바람직하게는 2개의 뉴클레오타이드 길이를 갖는다. 존재하는 경우, Nb는 임의로 1 내지 10개의 뉴클레오타이드, 바람직하게는 2 내지 6개의 뉴클레오타이드, 보다 바람직하게는 4개의 뉴클레오타이드 길이를 갖는다. 존재하는 경우, Nc는 임의로 8 내지 23개의 뉴클레오타이드, 바람직하게는 10 내지 20개의 뉴클레오타이드, 보다 바람직하게는 15개의 뉴클레오타이드 길이를 갖는다.

[0331] 본 발명의 일부 구현예에서, 시스-조절 인핸서 요소는 CRE0042(서열 번호 104) 또는 이의 기능적 변이체로 이루어진다.

[0332] CRE 또는 이의 기능적 변이체는 이중 가닥 폴리뉴클레오타이드의 어느 한 가닥에 제공될 수 있으며 어느 방향으로든 제공될 수 있음에 주목할 것이다. 이와 같이, 서열 번호 104 또는 105의 상보적 및 역상보적 서열 또는 이의 기능적 변이체는 본 발명의 범위 내에 속한다. 서열 번호 104 또는 105에 따른 서열 또는 이의 기능적 변이체를 포함하는 단일 가닥 핵산도 본 발명의 범위 내에 속한다.

[0333] 일부 구현예에서, CRE0042(서열 번호 104) 또는 이의 기능적 변이체를 포함하거나 이로 이루어진 CRE는 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 125개 이하의 뉴클레오타이드, 100개 이하의 뉴클레오타이드, 또는 80개 이하의 뉴클레오타이드 길이를 갖는다.

[0334] 일부 구현예에서, CRE0059(서열 번호 110) 또는 이의 기능적 변이체를 포함하거나 이로 이루어진 CRE는 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 125개 이하의 뉴클레오타이드, 또는 100개 이하의 뉴클레오타이드 길이를 갖는다. 이의 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.

[0335] 상기 논의된 바와 같이, CRE0059(서열 번호 110)의 기능적 변이체는 간-특이적 프로모터 요소로 작용하는 CRE0059의 능력을 실질적으로 유지한다. 예를 들어, CRE0059의 기능적 변이체가 간-특이적 프로모터 SP0412로 치환되는 경우, 변형된 프로모터는 SP0412(서열 번호 91)의 활성의 적어도 80%, 보다 바람직하게는 이의 활성의 적어도 90%, 보다 바람직하게는 이의 활성의 적어도 95%, 더욱 더 바람직하게는 이의 활성의 100%를 유지한다. 적합하게는, CRE0059의 기능적 변이체는 서열 번호 110과 적어도 70%, 80%, 90%, 95% 또는 99% 동일성을 갖는 서열을 포함한다.

[0336] CRE0059는 근위 프로모터이고 TSS의 업스트림에 있는 간-특이적 TF, 즉 HNF1에 대한 TFBS를 포함한다. 따라서, CRE0059의 기능적 변이체는 바람직하게는 TSS의 업스트림에 HNF1에 대한 TFBS를 포함한다.

[0337] 일부 구현예에서, CRE0059의 기능적 변이체는 서열 번호 110과 적어도 70% 동일 (바람직하게는 서열 번호 110과 적어도 80%, 90%, 95% 또는 99% 동일)하고 HNF1에 대한 TFBS(서열 번호 111)를 포함하고 HNF1에 대한 상기 TFBS의 다운스트림에 서열 번호 112와 적어도 80%, 90%, 95% 또는 완전히 동일한 TSS 서열(p1@SERPINA1 또는 p1@AFP로 치칭됨)을 포함하는 서열을 포함한다.

[0338] 일부 구현예에서, CRE0059의 기능적 변이체는 서열 번호 110과 적어도 70%, 80%, 90%, 95% 또는 99% 동일성을 갖고 위치 24-36에서 또는 그 근처에서 HNF1에 대해 서열 번호 111을 포함하는 TFBS를 추가로 포함하고; 위치 73-93에서 또는 그 근처에서 서열 번호 112와 적어도 80%, 90%, 95% 또는 완전히 동일한 TSS 서열을 포함하는 서열을 포함하며, 상기 위치는 서열 번호 110을 참조하여 넘버링된다. 본 문맥에서 ~에서 또는 그 근처에서란, 적합하게는 서열 번호 110과 관련하여 인용된 위치의 10, 5, 4, 3, 2, 또는 1 뉴클레오타이드 이내를 의미한다. 적합한 TFBS 서열은 서열 번호 111 및 서열 번호 112이지만, 대안적인 TFBS 서열이 사용될 수 있다.

[0339] 일부 구현예에서, CRE0059(서열 번호 110) 또는 이의 기능적 변이체를 포함하거나 이로 이루어진 프로모터 요소는 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 125개 이하의 뉴클레오타이드, 110개 이하의 뉴클레오타이드, 또는 95개 이하의 뉴클레오타이드 길이를 갖는다.

- [0340] 일부 구현예에서, 본원에 개시된 바와 같은 방법 및 조성물에 유용한 간-특이적 프로모터는 서열 번호 91 또는 이의 기능적 변이체를 포함하거나 이로 이루어진다. 일부 구현예에서, 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다. 서열 번호 91에 따른 서열을 갖는 프로모터는 SP0412로 지칭된다. 일부 구현예에서는 SP0412 프로모터가 특히 바람직하다. 이 프로모터는 강력하고 또한 매우 짧기 때문에 일부 상황에서는 유리하다.
- [0341] **a. SP0265(SP131A1로도 알려져 있음) 및 이의 변이체**
- [0342] 일부 구현예에서, 프로모터는 CRE CRE0051(서열 번호 97), CRE0058(서열 번호 113), CRE0065(서열 번호 117), 및 CRE0066(서열 번호 122), 또는 이의 기능적 변이체의 조합을 포함하는 합성 간-특이적 프로모터이다. 전형적으로, CRE는 프로모터 요소에 작동적으로 연결된다. 일부 바람직한 구현예에서, 간-특이적 프로모터는 CRE0051, CRE0058, CRE0065, CRE0066 및 이어서 프로모터 요소(업스트림에서 다운스트림 방향으로)의 순서로 상기 CRE 또는 이의 기능적 변이체를 포함한다.
- [0343] 프로모터 요소는 임의의 적합한 근위 또는 최소 프로모터일 수 있다. 일부 바람직한 구현예에서, 프로모터 요소는 최소 프로모터이다. 프로모터가 근위 프로모터인 경우, 근위 프로모터가 간-특이적인 것이 일반적으로 바람직하다.
- [0344] 일부 바람직한 구현예에서, 프로모터 요소는 CRE0052(G6PC로도 지칭됨) (서열 번호 126)이다. CRE0052는 최소 프로모터(코어 프로모터로도 지칭함)이다.
- [0345] 일부 구현예에서, 간-특이적 프로모터는 다음 조절 요소 (또는 이의 기능적 변이체)를 포함한다: CRE0051, CRE0058, CRE0065, CRE0066에 이어서 CRE0052(서열 번호 126).
- [0346] CRE0051의 서열(서열 번호 97) 및 이의 변이체는 상기 제시되어 있다.
- [0347] 일부 구현예에서, CRE0058(서열 번호 113) 또는 이의 기능적 변이체를 포함하거나 이로 이루어진 CRE는 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 125개 이하의 뉴클레오타이드, 100개 이하의 뉴클레오타이드, 또는 80개 이하의 뉴클레오타이드 길이를 갖는다. 이의 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0348] CRE0058(서열 번호 113)의 기능적 변이체는 CRE0058과는 다른 서열을 가진 조절 요소이지만 간-특이적 CRE로서 이들의 활성을 실질적으로 유지한다. 당업자는 필요한 TF에 결합하고 발현을 향상시키는 능력을 유지하면서 CRE의 서열을 변화시키는 것이 가능하다는 것을 인식할 것이다. 기능적 변이체는 이들이 CRE를 비기능적으로 만들지 않는 경우 참조 CRE와 비교하여 치환, 결실 및/또는 삽입을 포함할 수 있다.
- [0349] 일부 구현예에서, CRE0058(서열 번호 113)의 기능적 변이체는 프로모터에서 CRE0058 대신에 치환될 때 이의 활성을 실질적으로 유지하는 CRE로 볼 수 있다. 예를 들어, CRE0058 대신에 치환된 CRE0058의 기능적 변이체를 포함하는 프로모터는 (CRE0058(서열 번호 113)을 포함하는 참조 프로모터와 비교하여) 바람직하게는 이의 활성의 80%, 보다 바람직하게는 이의 활성의 90%, 보다 바람직하게는 이의 활성의 95%, 더욱 더 바람직하게는 이의 활성의 100%를 유지한다. 예를 들어, 프로모터 SP0265(서열 번호 94)를 예로서 고려하면, SP0265에서 CRE0058은 CRE0058의 기능적 변이체로 대체될 수 있으며, 프로모터는 실질적으로 이의 활성을 유지한다. 활성 유지하는 참조 프로모터의 제어 하에 적합한 리포터의 발현을 동등한 조건 하에 치환된 CRE를 포함하는 다른 동일한 프로모터와 비교함으로써 평가될 수 있다.
- [0350] 일부 구현예에서, CRE0058(서열 번호 113)의 기능적 변이체가 CRE0058과 동일한 간-특이적 전사 인자(TF)에 대한 전사 인자 결합 부위(TFBS)를 포함하는 것이 바람직하다. CRE0058에 존재하는 간-특이적 TFBS는 이들이 존재하는 순서대로 나열되어 있다: HNF4(서열 번호 115) 및 c/EBP(서열 번호 116). 따라서, CRE0058의 기능적 변이체는 바람직하게는 이러한 TFBS 모두를 포함한다. 바람직하게는, 이들은 CRE0058에 존재하는 것과 동일한 순서, 즉 HNF4에 이어서 c/EBP의 순서로 존재한다. CRE가 프로모터 및 유전자와 관련될 때, 이 순서는 바람직하게는 업스트림에서 다운스트림 방향으로 고려된다(즉, 전사 개시 부위(TSS)로부터 원위에서 TSS에 대한 근위 방향으로). 인접한 TFBS 사이에 스페이서 서열에 제공될 수 있다. 일부 구현예에서, TFBS는 이들이 기능적으로 남아 있는 경우 적합하게 중첩될 수 있는데, 즉 중첩 서열은 둘 다 각각의 TF에 결합할 수 있다.
- [0351] 일부 구현예에서, CRE0058(서열 번호 113)의 기능적 변이체는 다음 TFBS 서열을 포함한다: CGCCCTTTGACCC(HNF4) (서열 번호 115) 및 GACCTTTTGCAATCCTGG(c/EBP) (서열 번호 116), 이에 상보적인 서열, 또는 이들 각각의 TF에 결합하는 능력을 유지하는 이들 TFBS 서열의 기능적 변이체. 이는 CRE0058과 동일한 순

서, 즉 상기 제시된 순서로 존재할 수 있다. 상기 논의된 바와 같이, TFBS와 관련된 서열 가변성이 있고, 주어진 TFBS에 대해, 전형적으로 어느 정도의 편차가 전형적으로 존재하는 컨센서스 서열이 있다는 것은 당업계에 잘 알려져 있다.

- [0352] 일부 구현예에서, CRE0058의 기능적 변이체는 서열: GCGCCCTTTGGACCTTTTGCAATCCTGG(서열 번호 114), 또는 이와 적어도 70%, 80%, 90%, 95% 또는 99% 동일한 서열을 포함한다.
- [0353] 일부 구현예에서, CRE는 서열 번호 113 또는 114 또는 이의 기능적 변이체로 이루어진다.
- [0354] CRE 또는 이의 기능적 변이체는 이중 가닥 폴리뉴클레오타이드의 어느 한 가닥에 제공될 수 있으며 어느 방향으로든 제공될 수 있음에 주목할 것이다. 이와 같이, 서열 번호 113 또는 114의 상보적 및 역상보적 서열 또는 이의 기능적 변이체는 본 발명의 범위 내에 속한다. 서열 번호 113 또는 114에 따른 서열 또는 이의 기능적 변이체를 포함하는 단일 가닥 핵산도 본 발명의 범위 내에 속한다.
- [0355] 일부 구현예에서, CRE0058 또는 이의 기능적 변이체를 포함하거나 이로 이루어진 CRE는 120개 이하의 뉴클레오타이드, 80개 이하의 뉴클레오타이드, 60개 이하의 뉴클레오타이드, 또는 40개 이하의 뉴클레오타이드 길이를 갖는다.
- [0356] 일부 구현예에서, CRE0065(서열 번호 117) 또는 이의 기능적 변이체를 포함하거나 이로 이루어진 CRE는 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 125개 이하의 뉴클레오타이드, 100개 이하의 뉴클레오타이드, 또는 80개 이하의 뉴클레오타이드 길이를 갖는다. 이의 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0357] CRE0065(서열 번호 117)의 기능적 변이체는 CRE0065와는 다른 서열을 가진 조절 요소이지만, 간-특이적 CRE로서 이들의 활성을 실질적으로 유지한다. 당업자는 필요한 TF에 결합하고 발현을 향상시키는 능력을 유지하면서 CRE의 서열을 변화시키는 것이 가능하다는 것을 인식할 것이다. 기능적 변이체는 이들이 CRE를 비기능적으로 만들지 않는 경우 참조 CRE와 비교하여 치환, 결실 및/또는 삽입을 포함할 수 있다.
- [0358] 일부 구현예에서, CRE0065의 기능적 변이체는 프로모터에서 CRE0065 대신에 치환될 때 이의 활성을 실질적으로 유지하는 CRE로 볼 수 있다. 예를 들어, CRE0065 대신에 치환된 CRE0065의 기능적 변이체를 포함하는 프로모터는 (CRE0065를 포함하는 참조 프로모터와 비교하여) 바람직하게는 이의 활성의 80%, 보다 바람직하게는 이의 활성의 90%, 보다 바람직하게는 이의 활성의 95%, 더욱 더 바람직하게는 이의 활성의 100%를 유지한다. 예를 들어, 프로모터 SP0265(서열 번호 94)를 예로서 고려하면, SP0265에서 CRE0065는 CRE0065의 기능적 변이체로 대체될 수 있으며, 프로모터는 실질적으로 이의 활성을 유지한다. 활성 유지는 참조 프로모터의 제어 하에 적합한 리포터의 발현을 동등한 조건 하에 치환된 CRE를 포함하는 다른 동일한 프로모터와 비교함으로써 평가될 수 있다.
- [0359] 일부 구현예에서, CRE0065의 기능적 변이체가 CRE0065와 동일한 간-특이적 TF에 대한 TFBS를 포함하는 것이 바람직하다. CRE0065에 존재하는 간-특이적 TFBS는 이들이 존재하는 순서대로 나열되어 있다: RXR 알파 (서열 번호 119), HNF3 (서열 번호 120) 및 HNF3 (서열 번호 121). 따라서, CRE0065의 기능적 변이체는 바람직하게는 이러한 TFBS 모두를 포함한다. 바람직하게는, 이들은 CRE0065에 존재하는 것과 동일한 순서, 즉 RXR 알파, HNF3에 이어서 HNF3의 순서로 존재한다. 시스-조절 요소가 프로모터 및 유전자와 관련될 때, 이 순서는 바람직하게는 업스트림에서 다운스트림 방향으로 고려된다(즉, 전사 개시 부위(TSS)로부터 원위에서 TSS에 대한 근위 방향으로). 인접한 TFBS 사이에 스페이서 서열에 제공될 수 있다. 일부 구현예에서, TFBS는 이들이 기능적으로 남아 있는 경우 적합하게 중첩될 수 있는데, 즉 중첩 서열은 둘 다 각각의 TF에 결합할 수 있다.
- [0360] 일부 구현예에서, CRE0065의 기능적 변이체는 다음 TFBS 서열을 포함한다: ACTGAACCCTTGACCCCTGCCCT(RXR 알파)(서열 번호 119), CTGTTTGCCC(HNF3) (서열 번호 120), 및 CTATTTGCCC(HNF3) (서열 번호 121), 이에 상보적인 서열, 또는 이들 각각의 TF에 결합하는 능력을 유지하는 이들 TFBS 서열의 기능적 변이체. 이는 CRE0065와 동일한 순서, 즉 상기 제시된 순서로 존재할 수 있다. 상기 논의된 바와 같이, TFBS와 관련된 서열 가변성이 있고, 주어진 TFBS에 대해, 전형적으로 어느 정도의 편차가 전형적으로 존재하는 컨센서스 서열이 있다는 것은 당업계에 잘 알려져 있다.
- [0361] 일부 구현예에서, CRE0065의 기능적 변이체는 다음 서열:
- [0362] ACTGAACCCTTGACCCCT-Na-CTGTTTGCCC-Nb-TATTGCCC(서열 번호 118), 또는 이와 적어도 70%, 80%, 90%, 95% 또는 99% 동일한 서열을 포함하며, 여기서, Na 및 Nb는 임의의 스페이서 서열을 나타낸다. 존재하는 경우, Na는 임

의로 14 내지 30개의 뉴클레오타이드, 바람직하게는 18 내지 26개의 뉴클레오타이드, 보다 바람직하게는 22개의 뉴클레오타이드 길이를 갖는다. 존재하는 경우, Nb는 임의로 1 내지 10개의 뉴클레오타이드, 바람직하게는 2 내지 6개의 뉴클레오타이드, 보다 바람직하게는 4개의 뉴클레오타이드 길이를 갖는다. 일부 구현예에서, CRE는 서열 번호 117 또는 118, 또는 이의 기능적 변이체로 이루어진다.

[0363] CRE 또는 이의 기능적 변이체는 이중 가닥 폴리뉴클레오타이드의 어느 한 가닥에 제공될 수 있으며 어느 방향으로든 제공될 수 있음에 주목할 것이다. 이와 같이, 서열 번호 117 또는 118의 상보적 및 역상보적 서열 또는 이의 기능적 변이체는 본 발명의 범위 내에 속한다. 서열 번호 117 또는 118에 따른 서열 또는 이의 기능적 변이체를 포함하는 단일 가닥 핵산도 본 발명의 범위 내에 속한다.

[0364] 일부 바람직한 구현예에서, CRE0065 또는 이의 기능적 변이체를 포함하거나 이로 이루어진 CRE는 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 125개 이하의 뉴클레오타이드, 90개 이하의 뉴클레오타이드, 또는 72개 이하의 뉴클레오타이드 길이를 갖는다.

[0365] 일부 구현예에서, CRE0066(서열 번호 122) 또는 이의 기능적 변이체를 포함하거나 이로 이루어진 CRE는 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 125개 이하의 뉴클레오타이드, 100개 이하의 뉴클레오타이드, 또는 80개 이하의 뉴클레오타이드 길이를 갖는다. 이의 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.

[0366] CRE0066(서열 번호 122)의 기능적 변이체는 CRE0066과는 다른 서열을 가진 조절 요소이지만, 간-특이적 CRE로서 이들의 활성을 실질적으로 유지한다. 당업자는 필요한 TF에 결합하고 발현을 향상시키는 능력을 유지하면서 CRE의 서열을 변화시키는 것이 가능하다는 것을 인식할 것이다. 기능적 변이체는 이들이 CRE를 비기능적으로 만들지 않는 경우 참조 CRE와 비교하여 치환, 결실 및/또는 삽입을 포함할 수 있다.

[0367] 일부 구현예에서, CRE0066의 기능적 변이체는 프로모터에서 CRE0066 대신에 치환될 때 이의 활성을 실질적으로 유지하는 CRE로 볼 수 있다. 예를 들어, CRE0066 대신에 치환된 CRE0066의 기능적 변이체를 포함하는 프로모터는 (CRE0066(서열 번호 122)을 포함하는 참조 프로모터와 비교하여) 바람직하게는 이의 활성의 80%, 보다 바람직하게는 이의 활성의 90%, 보다 바람직하게는 이의 활성의 95%, 더욱 더 바람직하게는 이의 활성의 100%를 유지한다. 예를 들어, 프로모터 SP0265(서열 번호 94)를 예로서 고려하면, SP0265에서 CRE0066은 CRE0066의 기능적 변이체로 대체될 수 있으며, 프로모터는 실질적으로 이의 활성을 유지한다. 활성 유지는 참조 프로모터의 제어 하에 적합한 리포트의 발현을 동등한 조건 하에 치환된 CRE를 포함하는 다른 동일한 프로모터와 비교함으로써 평가될 수 있다.

[0368] 일부 구현예에서, CRE0066의 기능적 변이체가 CRE0066과 동일한 간-특이적 전사 인자(TF)에 대한 전사 인자 결합 부위(TFBS)를 포함하는 것이 바람직하다. CRE0066에 존재하는 간-특이적 TFBS는 이들이 존재하는 순서대로 나열되어 있다: HNF4G(서열 번호 124) 및 FOS::JUN(서열 번호 125). 따라서, CRE0066의 기능적 변이체는 바람직하게는 이러한 TFBS 모두를 포함한다. 바람직하게는, 이들은 CRE0066에 존재하는 것과 동일한 순서, 즉 HNF4G에 이어서 FOS::JUN의 순서로 존재한다. 시스-조절 요소가 프로모터 및 유전자와 관련된 때, 이 순서는 바람직하게는 업스트림에서 다운스트림 방향으로 고려된다(즉, 전사 개시 부위(TSS)로부터 원위에서 TSS에 대한 근위 방향으로). 스페이서 서열은 인접한 TFBS 사이에 제공될 수 있다. 일부 구현예에서, TFBS는 이들이 기능적으로 남아 있는 경우 적합하게 중첩될 수 있는데, 즉 중첩 서열은 둘 다 각각의 TF에 결합할 수 있다.

[0369] 일부 구현예에서, CRE0066(서열 번호 122)의 기능적 변이체는 다음 TFBS 서열을 포함한다: GCAGGGCAAAGTGCA(HNF4G)(서열 번호 124) 및 GATGACTCAG(FOS::JUN)(서열 번호 125), 이에 상보적인 서열, 또는 이들 각각의 TF에 결합하는 능력을 유지하는 이들 TFBS 서열의 기능적 변이체. 이는 CRE0066과 동일한 순서, 즉 상기 제시된 순서로 존재할 수 있다. 상기 논의된 바와 같이, TFBS와 관련된 서열 가변성이 있고, 주어진 TFBS에 대해, 전형적으로 어느 정도의 편차가 전형적으로 존재하는 컨센서스 서열이 있다는 것은 당업계에 잘 알려져 있다.

[0370] 일부 구현예에서, CRE0066(서열 번호 122)의 기능적 변이체는 서열: GCAGGGCAAAGTGCA-Na-GATGACTCAG(서열 번호 123) 또는 이와 적어도 70%, 80%, 90%, 95% 또는 99% 동일한 서열을 포함하며, 여기서, Na는 임의의 스페이서 서열을 나타낸다. 존재하는 경우, Na는 임의로 10 내지 28개의 뉴클레오타이드, 바람직하게는 14 내지 24개의 뉴클레오타이드, 보다 바람직하게는 19개의 뉴클레오타이드 길이를 갖는다.

[0371] 일부 구현예에서, CRE는 CRE0066 또는 이의 기능적 변이체로 이루어진다.

[0372] CRE 또는 이의 기능적 변이체는 이중 가닥 폴리뉴클레오타이드의 어느 한 가닥에 제공될 수 있으며 어느 방향으로

로든 제공될 수 있음에 주목할 것이다. 이와 같이, 서열 번호 122 또는 123의 상보적 및 역상보적 서열 또는 이의 기능적 변이체는 본 발명의 범위 내에 속한다. 서열 번호 122 또는 123에 따른 서열 또는 이의 기능적 변이체를 포함하는 단일 가닥 핵산도 본 발명의 범위 내에 속한다.

- [0373] 일부 바람직한 구현예에서, CRE0066 또는 이의 기능적 변이체를 포함하거나 이로 이루어진 CRE는 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 125개 이하의 뉴클레오타이드, 100개 이하의 뉴클레오타이드, 또는 87개 이하의 뉴클레오타이드 길이를 갖는다.
- [0374] 일부 구현예에서, 프로모터는 프로모터 요소 CRE0052(G6PC로도 지칭됨) (서열 번호 126) 또는 이의 기능적 변이체 또는 기능적 단편을 포함한다. 이의 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0375] CRE0052(서열 번호 126)의 기능적 변이체는 간-특이적 프로모터 요소로서 작용하는 CRE0052의 능력을 실질적으로 유지한다. 예를 들어, CRE0052의 기능적 변이체가 간-특이적 프로모터 SP0265로 치환되는 경우, 변형된 프로모터는 SP0265의 활성의 적어도 80%, 보다 바람직하게는 이의 활성의 적어도 90%, 보다 바람직하게는 이의 활성의 적어도 95%, 더욱 더 바람직하게는 이의 활성의 100%를 유지한다.
- [0376] 하나의 구현예에서, 간-특이적 프로모터는 서열 번호 94 또는 이의 기능적 변이체를 포함한다. 일부 구현예에서, 기능적 변이체는 서열 번호 94와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다. 서열 번호 94에 따른 서열을 갖는 프로모터는 SP0265(SP131A1 또는 LVR131_A1로도 알려져 있음)로 지칭된다. 일부 구현예에서는 서열 번호 94를 포함하거나 이로 이루어진 프로모터가 특히 바람직하다.
- [0377] 일부 구현예에서, 간-특이적 프로모터는 서열 번호 94이고 다음 구성요소를 포함한다: CRE0051(서열 번호 97); CRE0058(서열 번호 113); CRE0065(서열 번호 117), CRE0066(서열 번호 122), CRE0052(서열 번호 126); 또는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있는 서열 번호 97, 113, 117, 122 또는 126의 기능적 변이체.
- [0378] **b. SP0239 및 이의 변이체**
- [0379] 일부 구현예에서, 프로모터는 다음 CRE: CRE0018(서열 번호 151), CRE0051(서열 번호 97), CRE0058(서열 번호 113), CRE0065(서열 번호 117) 및 CRE0066(서열 번호 122), 또는 이의 기능적 변이체를 포함하는 합성 간-특이적 프로모터이다. 전형적으로, CRE는 프로모터 요소에 작동적으로 연결된다. 일부 바람직한 구현예에서, 간-특이적 프로모터는 CRE0018, CRE0051, CRE0058, CRE0065, CRE0066 및 이어서 프로모터 요소(업스트림에서 다운스트림 방향으로)의 순서로 상기 CRE 또는 이의 기능적 변이체를 포함한다.
- [0380] 프로모터 요소는 임의의 적합한 근위 또는 최소 프로모터일 수 있다. 일부 바람직한 구현예에서, 프로모터 요소는 CRE0052(G6PC로도 지칭됨)이다. CRE0052는 최소 프로모터(코어 프로모터로도 지칭됨)이다.
- [0381] 일부 구현예에서, 간-특이적 프로모터는 다음 요소 (또는 이의 기능적 변이체)를 포함한다: CRE0018, CRE0051, CRE0058, CRE0065, CRE0066 및 이어서 CRE0052.
- [0382] CRE0051, CRE0058, CRE0065, 및 CRE0066의 서열 및 프로모터 요소 CRE0052, 및 이의 기능적 변이체가 상기 제시되어 있다.
- [0383] CRE0018은 서열 번호 151의 서열 또는 이의 기능적 변이체 또는 기능적 단편을 갖는다. 이의 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0384] CRE0018(서열 번호 151)의 기능적 변이체는 CRE0018과는 다른 서열을 가진 조절 요소이지만 간-특이적 CRE로서 이들의 활성을 실질적으로 유지한다. 당업자는 필요한 TF에 결합하고 발현을 향상시키는 능력을 유지하면서 CRE의 서열을 변화시키는 것이 가능하다는 것을 인식할 것이다. 기능적 변이체는 이들이 CRE를 실질적으로 비기능적으로 만들지 않는 경우 참조 CRE와 비교하여 치환, 결실 및/또는 삽입을 포함할 수 있다.
- [0385] 일부 구현예에서, CRE0018의 기능적 변이체는 프로모터에서 CRE0018 대신에 치환될 때 이의 활성을 실질적으로 유지하는 CRE로 볼 수 있다. 예를 들어, CRE0018 대신에 치환된 CRE0018의 기능적 변이체를 포함하는 프로모터는 (CRE0018을 포함하는 참조 프로모터와 비교하여) 바람직하게는 이의 활성의 80%, 보다 바람직하게는 이의 활성의 90%, 보다 바람직하게는 이의 활성의 95%, 더욱 더 바람직하게는 이의 활성의 100%를 유지한다. 예를 들어, 프로모터 SP0239를 예로서 고려하면, SP0239에서 CRE0018은 CRE0018의 기능적 변이체로 대체될 수 있으며,

프로모터는 실질적으로 이의 활성을 유지한다. 활성 유지는 참조 프로모터의 제어 하에 적합한 리포터의 발현을 동등한 조건 하에 치환된 CRE를 포함하는 다른 동일한 프로모터와 비교함으로써 평가될 수 있다.

[0386] 일부 구현예에서, CRE0018의 기능적 변이체가 CRE0018과 동일한 간-특이적 TF에 대한 TFBS를 포함한다. CRE0018에 존재하는 간-특이적 TFBS는 이들이 존재하는 순서대로 나열되어 있다: IRF(서열 번호 129), NF1(서열 번호 130), HNF3(서열 번호 131), HBLF(서열 번호 132), RXRa(서열 번호 133), EF-C(서열 번호 134), NF1(서열 번호 135), 및 c/EBP(서열 번호 136). 따라서, CRE0018의 기능적 변이체는 바람직하게는 이러한 TFBS 모두를 포함한다. 바람직하게는, 이들은 CRE0018에 존재하는 것과 동일한 순서, 즉 IRF, NF1, HNF3, HBLF, RXRa, EF-C, NF1, 및 이어서 c/EBP의 순서로 존재한다. CRE가 프로모터 및 유전자와 관련될 때, 이 순서는 바람직하게는 업스트림에서 다운스트림 방향으로 고려된다(즉, 전사 개시 부위(TSS)로부터 원위에서 TSS에 대한 근위 방향으로). 인접한 TFBS 사이에 스페이서 서열에 제공될 수 있다. 일부 구현예에서, TFBS는 이들이 기능적으로 남아 있는 경우 적합하게 중첩될 수 있는데, 즉 중첩 서열은 둘 다 각각의 TF에 결합할 수 있다.

[0387] 일부 구현예에서, CRE0018의 기능적 변이체는 다음 TFBS 서열을 포함한다: CTTTCACTTTC(IRF) (서열 번호 129), TCGCCAA(NF1) (서열 번호 130), TGTGTAAACA(HNF3) (서열 번호 131), TGTAACAATA(HBLF) (서열 번호 132), CTGAACCTTTACCC(RXRa) (서열 번호 133), GTTGCCCGCAAC(EF-C) (서열 번호 134), CAGGTCTGTGCCAAG(NF1) (서열 번호 135), TGCCAAGTGTTC(c/EBP) (서열 번호 136), 이에 상보적인 서열, 또는 서열 번호 129-136의 이들 각각의 TF에 결합하는 능력을 유지하는 이들 TFBS 서열의 기능적 변이체. 이는 CRE0018과 동일한 순서, 즉 상기 제시된 순서로 존재할 수 있다. 상기 논의된 바와 같이, TFBS와 관련된 서열 가변성이 있고, 주어진 TFBS에 대해, 전형적으로 어느 정도의 편차가 전형적으로 존재하는 컨센서스 서열이 있다는 것은 당업계에 잘 알려져 있다.

[0388] 본 발명의 일부 구현예에서, CRE0018의 기능적 변이체는 서열: CTTTCACTTTCGCCAA-Na-TGTGTAAACAATA-Nb-CTGAACCTTTACCC-Nc-GTTGCCCGCAAC-Nd-CAGGTCTGTGCCAAGTGTTC(서열 번호 128), 또는 이와 적어도 70%, 80%, 90%, 95% 또는 99% 동일한 서열을 포함하며, 여기서, Na, Nb, Nc, 및 Nd는 임의의 스페이서 서열을 나타낸다. 존재하는 경우, Na는 임의로 10 내지 20개의 뉴클레오타이드, 바람직하게는 13 내지 17개의 뉴클레오타이드, 보다 바람직하게는 15개의 뉴클레오타이드 길이를 갖는다. 존재하는 경우, Nb는 임의로 1 내지 10개의 뉴클레오타이드, 바람직하게는 1 내지 5개의 뉴클레오타이드, 보다 바람직하게는 1개의 뉴클레오타이드 길이를 갖는다. 존재하는 경우, Nc는 임의로 1 내지 10개의 뉴클레오타이드, 바람직하게는 1 내지 5개의 뉴클레오타이드, 보다 바람직하게는 1개의 뉴클레오타이드 길이를 갖는다. 존재하는 경우, Nd는 적합하게는 1 내지 10개의 뉴클레오타이드 길이, 바람직하게는 2 내지 8개의 뉴클레오타이드 길이, 보다 바람직하게는 3개의 뉴클레오타이드 길이를 갖는다.

[0389] 본 발명의 일부 구현예에서, CRE는 서열 번호 127 또는 128 또는 이의 기능적 변이체로 이루어진다.

[0390] CRE 또는 이의 기능적 변이체는 이중 가닥 폴리뉴클레오타이드의 어느 한 가닥에 제공될 수 있으며 어느 방향으로든 제공될 수 있음에 주목할 것이다. 이와 같이, 서열 번호 128 또는 129의 상보적 및 역상보적 서열 또는 이의 기능적 변이체는 본 발명의 범위 내에 속한다. 서열 번호 128 또는 129에 따른 서열 또는 이의 기능적 변이체를 포함하는 단일 가닥 핵산도 본 발명의 범위 내에 속한다.

[0391] 일부 구현예에서, CRE0018(서열 번호 151) 또는 이의 기능적 변이체를 포함하거나 이로 이루어진 CRE는 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 125개 이하의 뉴클레오타이드, 또는 103개 이하의 뉴클레오타이드 길이를 갖는다.

[0392] 하나의 구현예에서, 간-특이적 프로모터는 서열 번호 93 또는 이의 기능적 변이체를 포함하거나 이로 이루어진다. 서열 번호 93에 따른 서열을 갖는 프로모터는 SP0239로 지칭된다. SP0239의 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.

[0393] 따라서, 일부 구현예에서, 간-특이적 프로모터는 SP0239(서열 번호 93)이고 다음 구성요소를 포함하거나: CRE0018(서열 번호 151), CRE0051(서열 번호 97), CRE0058(서열 번호 113), CRE0065(서열 번호 117) 및 CRE0066(서열 번호 122), 및 CRE0052(서열 번호 126); 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.

[0394] c. SP0240 및 이의 변이체

[0395] 일부 구현예에서, 프로모터는 프로모터 요소에 작동적으로 연결된 CRE0018을 포함하는 합성 간-특이적 프로모터이다. 일부 바람직한 구현예에서, 간-특이적 프로모터는 CRE0018을 포함하거나 프로모터 요소의 바로 업스트림

에 있다.

- [0396] 프로모터 요소는 임의의 적합한 근위 또는 최소 프로모터일 수 있다. 일부 바람직한 구현예에서, 프로모터 요소는 CRE0006(서열 번호 137)이다. CRE0006은 간-특이적 근위 프로모터이다.
- [0397] 일부 구현예에서, 간-특이적 프로모터는 다음 요소 (또는 이의 기능적 변이체)를 포함한다: CRE0018 및 이어서 CRE0006.
- [0398] CRE0018의 서열 및 이의 변이체는 상기 제시되어 있다.
- [0399] CRE0006은 근위 프로모터이고 TSS의 업스트림의 간-특이적 TF에 대한 TFBS를 포함한다. 나열된 순서로, CRE0006에 존재하는 간-특이적 TFBS는 HNF4(서열 번호 138), RXRa(서열 번호 139), HNF4(서열 번호 140), c/EBP(서열 번호 141), 및 HNF3(서열 번호 142), 및 임의로 p1@VTN(서열 번호 143)이다. 따라서, CRE0006의 기능적 변이체는 바람직하게는 이러한 TFBS를 포함한다. 바람직하게는, 이들은 CRE0006에 존재하는 것과 동일한 순서, 즉 HNF4, c/EBP, HNF3, 및 HNF3의 순서로 존재한다. 일부 구현예에서, TFBS는 이들이 기능적으로 남아 있는 경우 적합하게 중첩되는데, 즉 중첩 서열은 둘 다 각각의 TF에 결합할 수 있다.
- [0400] p1@VTN(서열 번호 143)은 유전자 발현의 캡 분석(CAGE)에 의해 결정된 바와 같이 CRE0006의 전사 개시 부위(TSS)를 나타낸다.
- [0401] 일부 구현예에서, CRE0006의 기능적 변이체는 서열 번호 137과 적어도 70% 동일 (바람직하게는 서열 번호 25와 적어도 80%, 90%, 95% 또는 99% 동일)하고 HNF4, RXRa, HNF4, c/EBP, 및 HNF3에 대한 TFBS를 포함하고, 바람직하게는 상기 TFBS의 다운스트림에 HNF4, RXRa, HNF4, c/EBP, 및 HNF3에 대한 TFBS와 적어도 80%, 90%, 95% 또는 완전히 동일한 TSS 서열을 포함하는 서열을 포함한다.
- [0402] 일부 구현예에서, CRE0006의 기능적 변이체는 서열 번호 137과 적어도 70%, 80%, 90%, 95% 또는 99% 동일성을 갖고 다음 TFBS: 위치 25-37에서 또는 그 근처에서 HNF4 (서열 번호 138); 위치 73-83에서 또는 그 근처에서 RXRa (서열 번호 139); 위치 74-86에서 또는 그 근처에서 HNF4(서열 번호 140); 위치 123-136에서 또는 그 근처에서 c/EBP (서열 번호 141); 및 위치 129-137에서 또는 그 근처에서 HNF3(서열 번호 142)을 추가로 포함하고; 위치 166-196에서 또는 그 근처에서 서열 번호 143과 적어도 80%, 90%, 95% 또는 완전히 동일한 TSS 서열을 포함하는 서열을 포함하며, 상기 위치는 서열 번호 137을 참조하여 넘버링된다. 본 문맥에서 ~에서 또는 그 근처에서란, 적합하게는 서열 번호 137과 관련하여 인용된 위치의 10, 5, 4, 3, 2, 또는 1 뉴클레오타이드 이내를 의미한다. 적합한 TFBS 서열은 서열 번호 138-142이지만, 대안적인 TFBS 서열이 사용될 수 있다.
- [0403] 하나의 구현예에서, 간-특이적 프로모터는 서열 번호 95 또는 이의 기능적 변이체를 포함하거나 이로 이루어진다. 서열 번호 95에 따른 서열을 갖는 프로모터는 SP0240으로 지칭된다. SP0240의 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0404] **d. SP0246 및 이의 변이체**
- [0405] 일부 구현예에서, 프로모터는 다음 CRE: CRE0051, CRE0058, 및 CRE0065, 또는 이의 기능적 변이체를 포함하는 합성 간-특이적 프로모터이다. 전형적으로, CRE는 프로모터 요소에 작동적으로 연결된다. 일부 바람직한 구현예에서, 간-특이적 프로모터는 CRE0051, CRE0058, 및 CRE0065 및 이어서 프로모터 요소(업스트림에서 다운스트림 방향으로)의 순서로 상기 CRE 또는 이의 기능적 변이체를 포함한다.
- [0406] 프로모터 요소는 임의의 적합한 근위 또는 최소 프로모터일 수 있다. 일부 바람직한 구현예에서, 프로모터 요소는 CRE0052(G6PC로도 지칭됨)이다. CRE0052는 최소 프로모터(코어 프로모터로도 지칭됨)이다.
- [0407] 일부 구현예에서, 간-특이적 프로모터는 다음 요소 (또는 이의 기능적 변이체)를 포함한다: CRE0051, CRE0058, CRE0065, 및 이어서 CRE0052. CRE0051, CRE0058, CRE0065 및 프로모터 요소 CRE0052의 서열, 및 이의 기능적 변이체가 상기 제시되어 있다.
- [0408] 하나의 구현예에서, 간-특이적 프로모터는 서열 번호 96 또는 이의 기능적 변이체를 포함하거나 이로 이루어진다. 서열 번호 96에 따른 서열을 갖는 프로모터는 SP0246으로 지칭된다. SP0246의 기능적 변이체는 서열 번호 96과 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0409] **e. SP0131 및 이의 변이체**
- [0410] 일부 구현예에서, 프로모터는 다음 CRE: CRE0058, CRE0065 및 CRE0066, 또는 이의 기능적 변이체를 포함하는 합

성 간-특이적 프로모터이다. 전형적으로, CRE는 프로모터 요소에 작동적으로 연결된다. 일부 바람직한 구현예에서, 간-특이적 프로모터는 CRE0058, CRE0065, CRE0066 및 이어서 프로모터 요소(업스트림에서 다운스트림 방향으로)의 순서로 상기 CRE 또는 이의 기능적 변이체를 포함한다.

[0411] 프로모터 요소는 임의의 적합한 근위 또는 최소 프로모터일 수 있다. 일부 바람직한 구현예에서, 프로모터 요소는 CRE0052(G6PC로도 지칭됨)이다. CRE0052는 최소 프로모터(코어 프로모터로도 지칭됨)이다.

[0412] CRE0058, CRE0065, 및 CRE0066 및 프로모터 요소 CRE0052의 서열, 및 이의 기능적 변이체가 상기 제시되어 있다.

[0413] 서열 번호 141 또는 이의 기능적 변이체. 서열 번호 141에 따른 서열을 갖는 프로모터가 SP0131로 지칭된다. SP0131의 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.

[0414] **f. 복합 프로모터:**

[0415] 일부 구현예에서, 상기 제시된 바와 같은 간-특이적 프로모터는 하나 이상의 추가 조절 서열에 작동적으로 연결된다. 추가 조절 서열은, 예를 들어, 추가 조절 서열에 작동적으로 연결되지 않은 간-특이적 프로모터와 비교하여 발현을 향상시킬 수 있다. 일반적으로, 추가 조절 서열은 간-특이적 프로모터의 특이성을 실질적으로 감소시키지 않는 것이 바람직하다.

[0416] 예를 들어, 간-특이적 프로모터는 UTR(예를 들어, 5' 및/또는 3' UTR), 인트론 등을 암호화하는 서열에 작동적으로 연결될 수 있다.

[0417] 일부 구현예에서, 간-특이적 프로모터는 UTR, 예를 들어, 5' UTR을 암호화하는 서열에 작동적으로 연결된다. 5' UTR은 유전자 발현을 조절할 수 있는 다양한 요소를 포함할 수 있다. 천연 유전자에서 5' UTR은 전사 개시 부위에서 시작하여 암호화 영역의 개시 코돈 이전 뉴클레오타이드 하나에서 끝난다. 본원에서 언급된 5' UTR은 전체 자연적으로 발생하는 5' UTR일 수 있거나 자연적으로 발생하는 5' UTR의 일부일 수 있다는 점에 주목해야 한다. 5'UTR은 부분적으로 또는 전체적으로 합성될 수도 있다. 진핵생물에서, 5' UTR의 중간 길이는 대략 150 nt이지만 일부 경우에 상당히 더길 수 있다. 5' UTR에서 찾을 수 있는 조절 서열은 다음을 포함되지만 이에 제한되지 않는다:

[0418] mRNA의 안정성 또는 번역에 영향을 미칠 수 있는 단백질에 대한 결합 부위;

[0419] 리보스위치;

[0420] 번역 개시를 촉진하거나 억제하는 서열; 및

[0421] 5' UTR 내의 인트론은 유전자 발현 및 mRNA 익스포트의 조절과 연결되어 있다.

[0422] 일부 구현예에서, 상기 제시된 바와 같은 간-특이적 프로모터는 CMV 주요 즉시 유전자(CMV-IE 유전자)로부터 유래된 5' UTR을 암호화하는 서열에 작동적으로 연결된다. 예를 들어, CMV-IE 유전자로부터의 5' UTR은 적합하게는 CMV-IE 유전자 엑손 1 및 CMV-IE 유전자 엑손 1, 또는 이의 부분을 포함한다. 일부 경우에, 프로모터 요소는 5' UTR에 대한 연결의 관점에서 변형될 수 있으며, 예를 들어, 프로모터 요소에서 전사 개시 부위(TSS)의 다운스트림에 있는 서열은 제거될 수 있다(예를 들어, 5' UTR로 대체됨).

[0423] CMV-IE 5'UTR은 본원에 참조로 포함된 문헌[Simari, *et al.*, Molecular Medicine 4: 700-706, 1998 "Requirements for Enhanced Transgene Expression by Untranslated Sequences from the Human Cytomegalovirus Immediate-Early Gene"]에 기재되어 있다. Simari 등에 논의된 CMV-IE 5' UTR 서열의 변형은 또한 참조로 포함된 WO2002/031137에 제시되어 있으며, 이에 개시된 조절 서열도 사용될 수 있다. 프로모터와 조합하여 사용될 수 있는 다른 UTR은 당업계, 예를 들어, 참조로 포함된 문헌(Leppek, K., Das, R. & Barna, M. "Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them". Nat Rev Mol Cell Biol 19, 158-174 (2018))에 알려져 있다.

[0424] 일부 구현예에서, 5' UTR을 암호화하는 서열은 서열 번호 145 또는 이의 기능적 변이체를 포함한다. 일부 구현예에서, 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다. 서열 번호 145는 CMV-IE 5' UTR을 암호화한다.

[0425] 일부 구현예에서, 5' UTR은 단백질 번역 개시 부위로서 기능하는 핵산 모티프, 예를 들어, 생성된 mRNA에서 코

작 서열을 정의하는 서열을 포함한다. 예를 들어, 일부 구현예에서, 5' UTR을 암호화하는 서열은 이의 3' 말단에서 또는 그 근처에서 서열 모티프 GCCACC(서열 번호 153)를 포함한다. 당업계에 알려진 바와 같이 다른 코작 서열 또는 다른 단백질 번역 개시 부위가 사용될 수 있다(예를 들어, 모두 본원에 참조로 포함된 문헌(Marilyn Kozak, "Point Mutations Define a Sequence Flanking the AUG Initiator Codon That Modulates Translation by Eukaryotic Ribosomes" Cell, Vol. 44, 283-292, January 31, 1986; Marilyn Kozak "At Least Six Nucleotides Preceding the AUG Initiator Codon Enhance Translation in Mammalian Cells" J. Mol. Biol. (1987) 196, 947-950; Marilyn Kozak "An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs" Nucleic Acids Research. Vol. 15 (20) 1987)). 단백질 번역 개시 부위(예를 들어, 코작 서열)는 바람직하게는 개시 코돈에 바로 인접하게 위치한다.

- [0426] 일부 구현예에서, 5' UTR을 암호화하는 서열은 서열 번호 438 또는 이의 기능적 변이체를 포함한다. 일부 구현예에서, 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다. 이러한 5' UTR은 CMV-IE 5' UTR의 3' 말단에서 코작 서열을 정의하는 서열 번호 153의 6개의 뉴클레오타이드를 포함한다.
- [0427] 일부 구현예에서, 상기 논의된 바와 같은 SP0412 프로모터 또는 이의 변이체는 5' UTR을 암호화하는 서열에 연결되어 복합 프로모터/5' UTR 조절 작제물을 제공한다. 본원에서, 이러한 복합 프로모터/5' UTR 작제물은 단순히 "복합 프로모터"로 지칭되며, 일부 경우에 간결성을 위해 단순히 "프로모터"로 지칭될 수 있다.
- [0428] 일부 구현예에서, 복합 프로모터는 서열 번호 92 또는 이의 기능적 변이체를 포함하거나 이로 이루어진다. 일부 구현예에서, 기능적 변이체는 서열 번호 92와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0429] 이 복합 프로모터는 상기 논의된 CMV-IE 유전자(서열 번호 145) 및 GCCACC(서열 번호 153) 코작 서열로부터의 5' UTR을 암호화하는 서열에 작동으로 연결된 SP0412를 포함한다. 이 (복합) 프로모터는 SP0422(서열 번호 92)로 지칭된다. 일부 구현예에서 SP0422는 바람직한 간 특이적 프로모터이다. 상기 논의된 바와 같이, 5' UTR은 단백질 번역 개시 부위로서 기능하는 핵산 모티프, 예를 들어, 코작 서열을 정의하는 서열을 적합하게 포함한다. 상기 서열에서, 5' UTR은 이의 3' 말단에서 서열 모티프 GCCACC(서열 번호 153)를 포함하지만, 이 서열 모티프를 생략하거나 대안적인 서열을 사용할 수 있다.
- [0430] 일부 구현예에서, 상기 논의된 바와 같은 SP0265 프로모터 또는 이의 변이체는 5' UTR을 암호화하는 서열에 연결되어 복합 프로모터(SP0236-5'UTR)를 제공한다.
- [0431] 일부 구현예에서, 복합 프로모터는 서열 번호 146 또는 이의 기능적 변이체를 포함하거나 이로 이루어진다. 일부 구현예에서, 기능적 변이체는 서열 번호 146과 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0432] 이 복합 프로모터는 CMV-IE 유전자(서열 번호 145) 및 GCCACC(서열 번호 153) 코작 서열로부터 5' UTR에 작동적으로 연결된 SP0265(서열 번호 94)를 포함한다. 이 (복합) 프로모터는 SP0420으로 지칭된다. 이 프로모터에서, CRE0052 프로모터 요소에서 TSS의 다운스트림에 있는 짧은 서열은 CMV-IE로부터 5' UTR로부터의 서열로 대체되었다. 따라서, 이 프로모터는 실제로 일부 서열이 제거된 CRE0052에 대한 변형을 갖는 SP0265의 마이너(minor) 변이체를 포함한다. 일부 구현예에서는 SP0420이 바람직하다. 상기 논의된 바와 같이, 5' UTR은 단백질 번역 개시 부위로서 기능하는 핵산 모티프, 예를 들어, 코작 서열을 정의하는 서열을 적합하게 포함한다. 상기 서열에서, 5' UTR은 이의 3' 말단에서 서열 모티프 GCCACC(서열 번호 153)를 포함하지만, 이 서열 모티프를 생략하거나 대안적인 서열을 사용할 수 있다.
- [0433] 일부 구현예에서, 상기 논의된 바와 같은 SP0239 프로모터 또는 이의 변이체는 5' UTR을 암호화하는 서열에 연결되어 복합 프로모터(SP0239-UTR)를 제공한다.
- [0434] 일부 구현예에서, 복합 프로모터는 서열 번호 147 또는 이의 기능적 변이체를 포함하거나 이로 이루어진다. 일부 구현예에서, 기능적 변이체는 서열 번호 147과 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0435] 이 복합 프로모터/5' UTR 작제물은 CMV-IE 유전자 및 GCCACC(서열 번호 153) 코작 서열로부터 5' UTR에 작동적으로 연결된 SP0239를 포함한다. 이 (복합) 프로모터는 SP0421로 지칭된다. 다시, 이 프로모터에서, CRE0052 프로모터 요소에서 TSS의 다운스트림에 있는 짧은 서열은 CMV-IE로부터 5' UTR로부터의 서열로 대체되었다. 따라서, 이 프로모터는 실제로 일부 서열이 제거된 CRE0052에 대한 변형을 갖는 SP0239의 마이너 변이체를 포함한다.

다. 일부 구현예에서는 SP0421이 바람직하다. 상기 논의된 바와 같이, 5' UTR은 단백질 번역 개시 부위로서 기능하는 핵산 모티프, 예를 들어, 코작 서열을 정의하는 서열을 적합하게 포함한다. 상기 서열에서, 5' UTR은 이의 3' 말단에서 서열 모티프 GCCACC(서열 번호 153)를 포함하지만, 이 서열 모티프를 생략하거나 대안적인 서열을 사용할 수 있다.

- [0436] 일부 구현예에서, 상기 논의된 바와 같은 SP0240 프로모터 또는 이의 변이체는 5' UTR을 암호화하는 서열에 연결되어 복합 프로모터를 제공한다.
- [0437] 일부 구현예에서, 복합 프로모터는 서열 번호 148 또는 이의 기능적 변이체를 포함하거나 이로 이루어진다. 일부 구현예에서, 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0438] 이 복합 프로모터/5' UTR 작제물은 CMV-IE 유전자 및 GCCACC(서열 번호 153) 코작 서열로부터 5' UTR에 작동적으로 연결된 SP0240을 포함한다. 이 (복합) 프로모터는 SP0240-UTR로 지칭된다. 다시, 이 프로모터에서, CRE0006 프로모터 요소에서 TSS의 다운스트림에 있는 짧은 서열은 CMV-IE로부터 5' UTR로부터의 서열로 대체되었다. 따라서, 이 프로모터는 실제로 일부 서열이 제거된 CRE0006에 대한 변형을 갖는 SP0240의 마이너 변이체를 포함한다. 일부 구현예에서는 SP0240-UTR이 바람직하다. 상기 논의된 바와 같이, 5' UTR은 단백질 번역 개시 부위로서 기능하는 핵산 모티프, 예를 들어, 코작 서열을 정의하는 서열을 적합하게 포함한다. 상기 서열에서, 5' UTR은 이의 3' 말단에서 서열 모티프 GCCACC(서열 번호 153)를 포함하지만, 이 서열 모티프를 생략하거나 대안적인 서열을 사용할 수 있다.
- [0439] 일부 구현예에서, 상기 논의된 바와 같은 SP0246 프로모터 또는 이의 변이체는 5' UTR을 암호화하는 서열에 연결되어 복합 프로모터를 제공한다.
- [0440] 일부 구현예에서, 복합 프로모터는 서열 번호 149(SP0246-UTR) 또는 이의 기능적 변이체를 포함하거나 이로 이루어진다. 일부 구현예에서, 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0441] 이 복합 프로모터/5' UTR 작제물은 CMV-IE 유전자 및 GCCACC(서열 번호 153) 코작 서열로부터 5' UTR에 작동적으로 연결된 SP0246을 포함한다. 이 (복합) 프로모터는 SP0246-UTR로 지칭된다. 다시, 이 프로모터에서, CRE0052 프로모터 요소에서 TSS의 다운스트림에 있는 짧은 서열은 CMV-IE로부터 5' UTR로부터의 서열로 대체되었다. 따라서, 이 프로모터는 실제로 일부 서열이 제거된 CRE0052에 대한 변형을 갖는 SP0246의 마이너 변이체를 포함한다. 일부 구현예에서는 SP0246-UTR이 바람직하다. 상기 논의된 바와 같이, 5' UTR은 단백질 번역 개시 부위로서 기능하는 핵산 모티프, 예를 들어, 코작 서열을 정의하는 서열을 적합하게 포함한다. 상기 서열에서, 5' UTR은 이의 3' 말단에서 서열 모티프 GCCACC(서열 번호 153)를 포함하지만, 이 서열 모티프를 생략하거나 대안적인 서열을 사용할 수 있다.
- [0442] 일부 구현예에서, 상기 논의된 바와 같은 SP0131_A1 프로모터 또는 이의 변이체는 5' UTR을 암호화하는 서열에 연결되어 복합 프로모터를 제공한다.
- [0443] 일부 구현예에서, 복합 프로모터는 서열 번호 150(SP0131 A1-UTR) 또는 이의 기능적 변이체를 포함하거나 이로 이루어진다. 일부 구현예에서, 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0444] 이 복합 프로모터/5' UTR 작제물은 CMV-IE 유전자 및 GCCACC(서열 번호 153) 코작 서열로부터 5' UTR에 작동적으로 연결된 SP0131을 포함한다. 이 (복합) 프로모터는 SP0131-UTR로 지칭된다. 다시, 이 프로모터에서, CRE0052 프로모터 요소에서 TSS의 다운스트림에 있는 짧은 서열은 CMV-IE로부터 5' UTR로부터의 서열로 대체되었다. 따라서, 이 프로모터는 실제로 일부 서열이 제거된 CRE0052에 대한 변형을 갖는 SP0131의 마이너 변이체를 포함한다. 일부 구현예에서는 SP0131-UTR이 바람직하다. 상기 논의된 바와 같이, 5' UTR은 단백질 번역 개시 부위로서 기능하는 핵산 모티프, 예를 들어, 코작 서열을 정의하는 서열을 적합하게 포함한다. 상기 서열에서, 5' UTR은 이의 3' 말단에서 서열 모티프 GCCACC(서열 번호 153)를 포함하지만, 이 서열 모티프를 생략하거나 대안적인 서열을 사용할 수 있다.
- [0445] 일부 구현예에서, 간-특이적 프로모터는 SP0412(서열 번호 91)이고 다음 구성요소를 포함하거나: CRE0051(서열 번호 97), CRE0067(서열 번호 152), CRE0059(서열 번호 110) 및 코작 서열(서열 번호 153); 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.

- [0446] 일부 구현예에서, 간-특이적 프로모터는 SP0422(서열 번호 9)이고 다음 구성요소를 포함하거나: CRE0051(서열 번호 97), CRE0067(서열 번호 152), CRE0059(서열 번호 110), CMV-IE 5'UTR(서열 번호 153) 및 코작 서열(서열 번호 153, 기능적 변이체는 이와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 서열을 가질 수 있다.
- [0447] (iii) 합성 간-특이적 프로모터의 기능적 변이체
- [0448] 일부 구현예에서, 간-특이적 프로모터의 기능적 변이체는 프로모터에서 참조 프로모터 요소 대신에 치환될 때 이의 활성을 실질적으로 유지하는 프로모터 요소로 볼 수 있다. 예를 들어, 본원의 표 4의 주어진 프로모터의 기능적 변이체, 또는 서열 번호 86(CRM 0412), 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422), 서열 번호 93(SP0239), 서열 번호 94(SP0265, SP131_A1로도 지칭됨), 서열 번호 95(SP0240) 또는 서열 번호 96(SP0246), 또는 서열 번호 146(SP0265-UTR), 서열 번호 147(SP0239-UTR), 서열 번호 148(SP0240-UTR), 서열 번호 149(SP0246-UTR) 또는 서열 번호 150(SP0131-A1-UTR)으로부터 나열된 임의의 프로모터, 또는 이의 기능적 단편 또는 변이체, 또는 서열 번호 270-341 또는 342-430, 또는 이의 기능적 단편 또는 이의 변이체로부터 선택된 임의의 LSP를 포함하는, 간-특이적 프로모터의 기능적 변이체로서, 여기서, 기능적 변이체 또는 기능적 단편은 바람직하게는 (비변형된 프로모터 요소를 포함하는 변경되지 않은 프로모터 서열과 비교하여) 변경되지 않은 프로모터의 활성의 적어도 35%, 또는 적어도 40%, 또는 적어도 45%, 또는 적어도 50%, 또는 적어도 55%, 또는 적어도 60%, 또는 적어도 70% 또는 적어도 80%, 보다 바람직하게는 이의 활성의 적어도 90%, 보다 바람직하게는 이의 활성의 적어도 95%, 보다 더 바람직하게는 이의 활성의 100%를 유지한다. 간-특이적 프로모터 활성을 평가하기 위한 적합한 검정은, 예를 들어, 실시예 12 및 13에서와 같이 본원에 개시되어 있다.
- [0449] 일부 구현예에서, 본원의 표 4에 개시된 간-특이적 프로모터, 또는 서열 번호 86(CRM 0412), 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422), 서열 번호 93(SP0239), 서열 번호 94(SP0265, SP131_A1로도 지칭됨), 서열 번호 95(SP0240) 또는 서열 번호 96(SP0246), 또는 서열 번호 146(SP0265-UTR), 서열 번호 147(SP0239-UTR), 서열 번호 148(SP0240-UTR), 서열 번호 149(SP0246-UTR) 또는 서열 번호 150(SP0131-A1-UTR)으로부터 선택된 임의의 프로모터, 또는 서열 번호 270-341 또는 342-430으로부터 선택된 임의의 LSP의 기능적 변이체 또는 기능적 단편은 원래의 비변형된 서열과 적어도 약 75% 서열 동일성, 또는 적어도 약 80% 서열 동일성, 적어도 약 90% 서열 동일성, 적어도 약 95% 서열 동일성, 적어도 약 98% 서열 동일성 및 또한 상응하는 비변형된 프로모터 서열의 프로모터 활성의 적어도 35%, 또는 이의 프로모터 활성의 적어도 약 45%, 또는 이의 프로모터 활성의 적어도 약 50%, 또는 이의 프로모터 활성의 적어도 약 60%, 또는 프로모터 활성의 적어도 약 75%, 또는 프로모터 활성의 적어도 약 80%, 또는 프로모터 활성의 적어도 약 85%, 또는 프로모터 활성의 적어도 약 90%, 또는 프로모터 활성의 적어도 약 95%를 갖는다.
- [0450] 예를 들어, 서열 번호 258(SP0412) 또는 서열 번호 92(SP0422)의 기능적 변이체 또는 기능적 단편은 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422)와 적어도 약 75% 서열 동일성, 또는 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422)와 적어도 약 80% 서열 동일성, 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422)와 적어도 약 95% 서열 동일성, 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422)와 적어도 약 98% 서열 동일성, 및 또한 각각 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422)의 상응하는 비변형된 프로모터 서열의 프로모터 활성의 적어도 35%, 또는 프로모터 활성의 적어도 약 45%, 또는 프로모터 활성의 적어도 약 50%, 또는 프로모터 활성의 적어도 약 60%, 또는 프로모터 활성의 적어도 약 75%, 또는 프로모터 활성의 적어도 약 80%, 또는 프로모터 활성의 적어도 약 85%, 또는 프로모터 활성의 적어도 약 90%, 또는 프로모터 활성의 적어도 약 95%를 갖는다.
- [0451] 일부 구현예에서, 본원에 개시된 간-특이적 프로모터의 기능적 변이체는 비변형된 프로모터 서열과 유의한 수준의 서열 동일성을 유지한다. 적합한 기능적 변이체는 비변형된 프로모터 서열과 적어도 60% 동일한 서열, 보다 바람직하게는 비변형된 간-특이적 프로모터 서열과 적어도 70%, 80%, 90%, 95% 또는 99% 동일한 서열을 포함한다.
- [0452] 일부 구현예에서, 본원에 개시된 간-특이적 프로모터의 기능적 단편은 비변형된 프로모터 서열과 유의한 수준의 서열 동일성을 유지한다. 적합한 기능적 단편은 비변형된 프로모터 서열과 적어도 60% 동일한 서열, 보다 바람직하게는 비변형된 간-특이적 프로모터 서열과 적어도 70%, 80%, 90%, 95% 또는 99% 동일한 서열을 포함한다.
- [0453] 일부 구현예에서, 프로모터 요소의 기능적 변이체는 프로모터에서 참조 프로모터 요소 대신에 치환될 때 이의 활성을 실질적으로 유지하는 프로모터 요소로 볼 수 있다. 예를 들어, 주어진 프로모터 요소의 기능적 변이체

를 포함하는 간-특이적 프로모터는 (비변형된 프로모터 요소를 포함하는 참조 프로모터와 비교하여) 바람직하게는 이의 활성의 적어도 80%, 보다 바람직하게는 이의 활성의 적어도 90%, 보다 바람직하게는 이의 활성의 적어도 95%, 보다 더 바람직하게는 이의 활성의 100%를 유지한다. 간-특이적 프로모터 활성을 평가하기 위한 적합한 검정은 본원에, 예를 들어, 실시예 12 및 13에 개시되어 있다.

[0454] 본원의 표 4에 개시된 바와 같은 간-특이적 프로모터, 또는 서열 번호 270-341 또는 342-430으로부터 선택된 임의의 LSP의 서열은 실질적인 활성 손실 없이 변경될 수 있다는 점에 주목해야 한다. 따라서, 아래에 논의된 간-특이적 프로모터의 기능적 변이체는 간-특이적 프로모터의 활성에 상당히 해로운 변형을 피하는 경우 본원의 표 4에 개시된 간-특이적 프로모터, 또는 서열 번호 270-341 또는 342-430으로부터 선택된 임의의 LSP의 서열을 변형함으로써 제조될 수 있다. 본 개시내용에 제공된 정보의 관점에서, 기능적 변이체를 제공하기 위한 본원의 표 4에 개시된 간-특이적 프로모터, 또는 서열 번호 270-341 또는 342-430으로부터 선택된 임의의 LSP의 변형은 간단하다. 더욱이, 본 개시내용은 임의의 주어진 간-특이적 프로모터 변이체의 기능성을 단순히 평가하기 위한 방법론을 제공한다. 각 간-특이적 프로모터에 대한 기능적 변이체는 아래에 논의된다.

[0455] 본 발명의 일부 구현예에서, 합성 간-특이적 프로모터는 서열 번호 86(CRM 0412), 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422), 서열 번호 93(SP0239), 서열 번호 94(SP0265, SP131_A1로도 지칭됨), 서열 번호 95(SP0240) 또는 서열 번호 96(SP0246), 또는 서열 번호 146(SP0265-UTR), 서열 번호 147(SP0239-UTR), 서열 번호 148(SP0240-UTR), 서열 번호 149(SP0246-UTR) 또는 서열 번호 150(SP0131-A1-UTR)으로부터 나열된 임의의 프로모터, 또는 서열 번호 270-341 또는 342-430으로부터 선택된 임의의 LSP, 또는 이들 중 임의의 것의 기능적 변이체로 이루어진 군으로부터 서열을 포함한다. 적합하게는, 임의의 상기 간-특이적 프로모터의 기능적 변이체는 참조 합성 간-특이적 프로모터와 70% 이상 동일한 서열, 보다 바람직하게는 참조 합성 간-특이적 프로모터와 적어도 80%, 90%, 95% 또는 99% 동일한 서열을 포함한다.

[0456] 일부 구현예에서, 이의 기능적 변이체는 적합하게는 표 4에 나열된 서열 중 어느 하나와 적어도 60%, 70%, 80%, 90%, 95% 또는 99% 동일한 서열을 포함할 수 있다. 추가로 또는 대안적으로, 표 4에 나열된 서열 중 어느 하나의 기능적 변이체는 적합하게는 엄격한 조건 하에 참조 서열에 하이브리드화하는 서열을 포함한다. 표 4에 나열된 서열 중 어느 하나의 기능적 변이체는 본원에 제공된 하나 이상의 서열이 상기 정의된 바와 같은 기능적 변이체로 대체되고/되거나 이에 제공된 서열의 순서가 변경된 변이체를 포함한다.

[0457] 일부 구현예에서, 표 4에 나열된 간-특이적 프로모터 서열 중 어느 하나의 기능적 변이체는 적어도 하나 이상의 뉴클레오타이드가 치환되고 이의 활성을 실질적으로 유지하는 경우 간-특이적 프로모터로 볼 수 있다. 예를 들어, 표 4에 나열된 간-특이적 프로모터 서열 중 어느 하나의 기능적 변이체를 포함하는 간-특이적 프로모터는 (참조 프로모터 서열과 비교하여) 바람직하게는 이의 활성의 80%, 보다 바람직하게는 이의 활성의 90%, 보다 바람직하게는 이의 활성의 95%, 더욱 더 바람직하게는 이의 활성의 100%를 유지한다. 예를 들어, LSP가 예로서 SP0412(서열 번호 91)를 포함하는 핵산 서열을 포함하는 경우, SP0412(예를 들어, 서열 번호 91)에 있는 뉴클레오타이드의 일부는 이의 기능적 변이체로 대체될 수 있으며, 간 특이적 SP0412 프로모터는 이의 활성을 실질적으로 유지한다. 활성 유지는 참조 프로모터의 제어 하에 적합한 리포터의 발현을 동등한 조건 하에 치환된 핵산을 포함하는 달리 동일한 프로모터와 비교함으로써 평가될 수 있다. 간-특이적 프로모터 활성을 평가하기 위한 적합한 검정은 본원에, 예를 들어, 실시예 12 및 13에 개시되어 있다.

[0458] 본원에 개시된 조성물 및 방법의 일부 구현예에서, 본원의 표 4에 개시된 합성 간-특이적 프로모터, 또는 서열 번호 86(CRM 0412), 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422), 서열 번호 93(SP0239), 서열 번호 94(SP0265, SP131_A1로도 지칭됨), 서열 번호 95(SP0240) 또는 서열 번호 96(SP0246), 또는 서열 번호 146(SP0265-UTR), 서열 번호 147(SP0239-UTR), 서열 번호 148(SP0240-UTR), 서열 번호 149(SP0246-UTR) 또는 서열 번호 150(SP0131-A1-UTR)으로부터 나열된 임의의 프로모터, 또는 서열 번호 270-341 또는 342-430으로부터 선택된 또는 임의의 LSP는 700, 600, 500, 450, 400, 350, 300, 250 또는 200 또는 150개 이상의 뉴클레오타이드 길이를 갖는 이의 기능적 변이체이다.

[0459] 본원에 개시된 조성물 및 방법의 일부 구현예에서, 본원의 표 4에 개시된 합성 간-특이적 프로모터, 서열 번호 86(CRM 0412), 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422), 서열 번호 93(SP0239), 서열 번호 94(SP0265, SP131_A1로도 지칭됨), 서열 번호 95(SP0240) 또는 서열 번호 96(SP0246), 또는 서열 번호 146(SP0265-UTR), 서열 번호 147(SP0239-UTR), 서열 번호 148(SP0240-UTR), 서열 번호 149(SP0246-UTR) 또는 서열 번호 150(SP0131-A1-UTR)으로부터 나열된 임의의 프로모터, 또는 서열 번호 270-341 또는 342-430으로부터 선택된 임의의 LSP는 최소 프로모터에 작동적으로 연결된 합성 간-특이적 시스-조절 요소(CRE) 또는 시스-조절

모듈(CRM)을 포함한다. 본 발명에 사용하기에 적합한 최소 프로모터의 예는 CMV-최소 프로모터, MinTk 최소 프로모터, 및 LVR_CRE0052_G6PC 최소 프로모터(서열 번호 126)를 포함하지만 이에 제한되지 않는다. 특정 구현예에서, 최소 프로모터는 서열 번호 145의 서열, 서열 번호 145와 적어도 60%, 70%, 80%, 90%, 95% 또는 99% 동일한 서열을 포함하는 CMV-IE 프로모터이다. 본원에 개시된 바와 같은 방법 및 조성물에 사용하기 위한 CMV-IE를 포함하는 예시적인 프로모터는 서열 번호 92(SP0422), 서열 번호 146(SP0265-UTR), 서열 번호 147(SP0239-UTR), 서열 번호 148(SP0240-UTR), 서열 번호 149(SP0246-UTR) 또는 서열 번호 150(SP0131-A1-UTR)으로부터 선택될 수 있지만 이에 제한되지 않는다.

[0460] 본원에 개시된 조성물 및 방법의 일부 구현예에서, 본원의 표 4에 개시된 합성 간-특이적 프로모터, 예를 들어, 서열 번호 86(CRM 0412), 서열 번호 91(SP0412) 또는 서열 번호 92(SP0422), 서열 번호 93(SP0239), 서열 번호 94(SP0265, SP131_A1로도 지칭됨), 서열 번호 95(SP0240) 또는 서열 번호 96(SP0246), 또는 서열 번호 146(SP0265-UTR), 서열 번호 147(SP0239-UTR), 서열 번호 148(SP0240-UTR), 서열 번호 149(SP0246-UTR) 또는 서열 번호 150(SP0131-A1-UTR)으로부터 나열된 임의의 프로모터, 또는 서열 번호 270-341 또는 342-430으로부터 선택된 임의의 LSP는 대상체의 간 또는 간 세포에서 유전자의 발현을 LP-1 프로모터(서열 번호 432)에 비해 적어도 20%, 적어도 40%, 적어도 60%, 적어도 80%, 적어도 100%, 적어도 200%, 적어도 300%, 적어도 500%, 적어도 1000% 이상 증가시킬 수 있다.

[0461] 본원에 개시된 조성물 및 방법의 일부 구현예에서, 본원의 표 4에 개시된 합성 간-특이적 프로모터, 또는 서열 번호 86, 91-96, 146-150으로부터 선택된 임의의 LSP 프로모터, 또는 서열 번호 270-341 또는 342-430으로부터 선택된 임의의 LSP, 합성 간-특이적 프로모터는 간-특이적 전이유전자 발현을 촉진할 수 있으며, TTR 프로모터(서열 번호 431)의 활성의 적어도 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 125%, 150%, 175%, 200%, 250%, 300%, 350% 또는 400%인 간 세포에서의 활성을 갖는다.

[0462] 본원에 개시된 조성물 및 방법의 일부 구현예에서, 본원의 표 4에 개시된 합성 간-특이적 프로모터, 또는 서열 번호 86, 91-96, 146-150으로부터 선택된 임의의 LSP 프로모터, 또는 서열 번호 270-341 또는 342-430으로부터 선택된 임의의 LSP, 합성 간-특이적 프로모터는 간-특이적 전이유전자 발현을 촉진할 수 있으며, TBG 프로모터(서열 번호 435)의 활성의 적어도 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 125%, 150%, 175%, 200%, 250%, 300%, 350% 또는 400%인 간 세포에서의 활성을 갖는다.

[0463] G. 인트론 서열

[0464] 일부 구현예에서, rAAV 유전자형은 프로모터 서열의 3' 및 분비 신호 펩타이드의 5'에 위치한 인트론 서열을 포함한다. 인트론 서열은 mRNA 안정성, 핵 밖으로 mRNA 수송 및/또는 발현된 GAA 융합 폴리펩타이드(예를 들어, SS-GAA 융합 폴리펩타이드 또는 SS-IGF2-GAA 폴리펩타이드)의 발현 및/또는 조절 중 하나 이상을 증가시키는 역할을 한다. 대안적인 구현예에서, rAAV 유전자형은 인트론 서열을 포함하지 않는다.

[0465] 일부 구현예에서, 인트론 서열은 MVM 인트론 서열, 예를 들어, 서열 번호 13의 인트론 서열 또는 이와 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 뉴클레오타이드 서열 동일성을 갖는 핵산 서열이지만 이에 제한되지 않는다.

[0466] 일부 구현예에서, 인트론 서열은 HBB2 인트론 서열, 예를 들어, 서열 번호 14의 인트론 서열 또는 이와 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 뉴클레오타이드 서열 동일성을 갖는 핵산 서열이지만 이에 제한되지 않는다.

[0467] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 재조합 AAV 벡터는 분비 신호 펩타이드를 암호화하는 서열의 5' 및 프로모터의 3'에 위치한 인트론 서열을 추가로 포함하는 이중 핵산 서열을 포함한다. 일부 구현예에서, 인트론 서열은 MVM 서열 또는 HBB2 서열을 포함하고, 여기서, MVM 서열은 서열 번호 13의 핵산 서열, 또는 서열 번호 13과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하고, HBB2 서열은 서열 번호 14의 핵산 서열, 또는 서열 번호 14와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열을 포함한다.

[0468] 일부 구현예에서, rAAV 유전자형은 인간 베타 글로빈 b2 (또는 HBB2) 인트론, FIX 인트론, 닭 베타-글로빈 인트론, 및 SV40 인트론으로 이루어진 군으로부터 선택된 인트론 서열을 포함한다. 일부 구현예에서, 인트론은 임의로 그 전체가 본원에 참조로 포함된 W02015/162302에 개시된 변형된 HBB2 인트론(예를 들어, W02018046774A1에서 서열 번호 17 참조): 변형된 FIX 인트론(예를 들어, W02018046774A1에서 서열 번호 19 참조), 또는 변형된 닭 베타-글로빈 인트론(예를 들어, W02018046774A1에서 서열 번호 21 참조), 또는 변형된 HBB2 또는 FIX 인트론

과 같은 변형된 인트론이다.

[0469] H. 폴리-A

[0470] 일부 구현예에서, rAAV 벡터 계놈은 하나의 구현예에서 GAA 융합 폴리펩타이드(예를 들어, SS-GAA 융합 폴리펩타이드 또는 SS-IGF2-GAA 폴리펩타이드)를 암호화하는 이중 핵산 유전자로부터 3' 및 다운스트림에 위치한 적어도 하나의 폴리-A 테일을 포함한다. 일부 구현예에서, 폴리A 신호는 본원에 정의된 바와 같은 안정성 서열 또는 CS 서열의 3'이다. hGH 폴리 A, synpA 폴리A 등을 포함하지만 이에 제한되지 않는 임의의 폴리A 서열이 사용될 수 있다. 일부 구현예에서, 폴리A는 합성 폴리A 서열이다. 일부 구현예에서, rAAV 벡터 계놈은 2개의 폴리-A 테일, 예를 들어, hGH 폴리A 서열 및 또 다른 폴리A 서열을 포함하며, 여기서, 스페이서 핵산 서열은 2개의 폴리A 서열 사이에 위치한다. 일부 구현예에서, rAAV 계놈은 GAA 융합 폴리펩타이드(예를 들어, SS-GAA 융합 폴리펩타이드 또는 SS-IGF2-GAA 폴리펩타이드)를 암호화하는 핵산의 3', 또는 대안적으로 CS 서열의 3'에 다음 요소를 포함한다: 제1 폴리A 서열, 스페이서 핵산 서열(100 bp와 400 bp 사이의, 또는 약 250 bp), 제2 폴리A 서열, 스페이서 핵산 서열, 및 3' ITR. 일부 구현예에서, 제1 및 제2 폴리A 서열은 hGH 폴리A 서열이고, 일부 구현예에서, 제1 및 제2 폴리A 서열은 합성 폴리A 서열이다. 일부 구현예에서, 제1 폴리A 서열은 hGH 폴리A 서열이고, 제2 폴리A 서열은 합성 서열, 또는 그 반대도 마찬가지며 - 즉, 대안적인 구현예에서, 제1 폴리A 서열은 합성 폴리A 서열이고, 제2 폴리A 서열은 hGH 폴리A 서열이다. 예시적인 폴리A 서열은, 예를 들어, 서열 번호 15 (hGH 폴리A 서열), 또는 서열 번호 15와 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 뉴클레오타이드 서열 동일성을 갖는 폴리 A 핵산 서열이다. 일부 구현예에서, 사용을 위해 포함된 hGHpoly 서열은 그 전체가 본원에 참조로 포함된 문헌(Anderson et al. J. Biol. Chem 264(14); 8222-8229, 1989(예를 들어, p. 8223, 제2 컬럼, 제1 단락 참조))에 기재되어 있다.

[0471] 일부 구현예에서, 폴리-A 테일은 이중 유전자에 대한 전사체를 포함하는 rAAV 벡터 계놈으로부터 전사된 RNA 전사체를 안정화하도록 조작될 수 있으며, 이는 하나의 구현예에서 GAA이며, 대안적인 구현예에서, 폴리-A 테일은 불안정하게 만드는 요소를 포함하도록 조작될 수 있다.

[0472] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 재조합 AAV 벡터는 GAA 유전자를 암호화하는 핵산의 3' 및 3' ITR 서열의 5'에 위치한 적어도 하나의 폴리A 서열을 포함한다. 일부 구현예에서, 폴리 A는 전장 폴리A(f1-폴리A) 서열이다. 일부 구현예에서, 폴리A는 절단된 폴리A 서열로, 예를 들어, 도. 5g를 참조한다.

[0473] 일 구현예에서, 폴리-A 테일은 폴리-A 테일의 길이를 변경하여 불안정 요소가 되도록 조작될 수 있다. 일 구현예에서, 폴리-A 테일은 연장되거나 단축될 수 있다. 일부 구현예에서, 3' 비번역된 영역은 GAA 3' UTR(서열 번호 85) 또는 3' UTR(서열 번호 77)을 포함한다.

[0474] 또 다른 구현예에서, 불안정화 요소는 이중 유전자를 암호화하는 miRNA가 결합하는 RNA 전사체를 사일링시키는 (번역을 억제하고 분해를 촉진하는) 마이크로RNA(miRNA)이다. 일 구현예에서, 폴리-A 테일 내의 시드 영역의 추가 또는 결실은 단백질, 예컨대 GAA 단백질 또는 변형된 GAA 폴리펩타이드의 발현을 증가 또는 감소시킬 수 있다.

[0475] 또 다른 구현예에서, 시드 영역은 또한 이중 유전자와 폴리-A 테일 사이에 위치한 3' 비번역된 영역으로 조작될 수 있다. 추가의 구현예에서, 불안정화제는 siRNA일 수 있다. siRNA의 암호화 영역은 rAAV 벡터 계놈에 포함될 수 있으며, 일반적으로 폴리-A 테일의 3'의 다운스트림에 위치한다.

[0476] 본원에 개시된 방법 및 방법의 모든 양태에서, rAAV 계놈은 또한 스테퍼 DNA 핵산 서열을 포함할 수 있다. 예시적인 스테퍼 DNA 서열은 서열 번호 71, 또는 이와 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 뉴클레오타이드 서열 동일성을 갖는 핵산 서열이다. 도 7-8 및 도 9a-9e에 나타난 바와 같이, 스테퍼 서열은, 예를 들어, 폴리 A 테일의 3에 위치하며, '3 ITR 서열의 5'에 위치한다. 일부 구현예에서, 스테퍼 DNA 서열은 역방향으로 합성 폴리아데닐화 신호를 포함한다.

[0477] 일부 구현예에서, 스테퍼 핵산 서열("스페이서" 핵산 단편으로도 지칭됨, 도 7-8 참조)은 폴리A 서열과 3' ITR 사이에 위치할 수 있다(즉, 스테퍼 핵산 서열은 폴리A 서열의 3' 및 3' ITR의 5'에 위치함) (예를 들어, 도 7-8 참조). 이러한 스테퍼 핵산 서열은 약 30 bp, 50 pb, 75 bp, 100 bp, 150 bp, 200 bp, 250 bp, 300 bp 또는 300 bp보다 길 수 있다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, 스테퍼 핵산 단편은 20 bp와 50 bp 사이, 50 bp와 100 bp 사이, 100 bp와 200 bp 사이, 200 bp와 300 bp 사이, 300 bp와 500 bp 사이, 또는 20 bp와 500 bp 사이의 임의의 정수이다. 예시적인 스테퍼 (또는 스페이서) 핵산 서열은 서열 번호 16, 서열 번호 71 또는 서열 번호 78, 또는 서열 번호 16 또는 서열 번호 71 또는 서열 번호 78과 서열 적어도 약 70, 71,

72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% 동일한 핵산을 포함한다.

[0478] I. AAV ITR

[0479] 본원에 개시된 바와 같은 rAAV 게놈은 바람직한 특성을 갖고 ITR을 포함하는 벡터의 활성 및 이에 대한 세포 반응을 조절하도록 설계될 수 있는 AAV ITR을 포함한다. 또 다른 구현예에서, AAV ITR은 바람직한 특성을 갖고 본원에 참조로 포함된 미국 특허 번호 9,447,433에 기재된 것을 포함하여 1개 또는 2개의 합성 ITR을 포함하는 벡터의 활성 및 이에 대한 세포 반응을 조작하도록 설계될 수 있는 합성 AAV ITR이다.

[0480] 또 다른 구현예에서, ITR은 자연적으로 발생하는 ITR, 예를 들어, AAV2로부터의 ITR2에 비해 변형된 전사 활성을 나타낸다. ITR2 서열은 본질적으로 프로모터 활성을 갖는 것으로 알려져 있다. 또한 본질적으로 폴리(A) 서열과 유사한 종결 활성을 갖는다. 본 발명의 최소 기능적 ITR은 비록 ITR2에 비해 감소된 수준이지만, 실시예에 나타난 바와 같이 전사 활성을 나타낸다. 따라서, 일부 구현예에서, ITR은 전사에 대해 기능적이다. 다른 구현예에서, ITR은 전사에 결합이 있다. 특정 구현예에서, ITR은 전사 절연체로서 작용할 수 있으며, 예를 들어, 벡터가 숙주 염색체 내로 통합될 때 벡터에 존재하는 트랜스제닉 카세트의 전사를 방지할 수 있다.

[0481] 본 발명의 하나의 양태는 적어도 하나의 합성 AAV ITR을 포함하는 rAAV 벡터 게놈에 관한 것이며, 여기서, ITR에서 하나 이상의 전사 인자 결합 부위의 뉴클레오타이드 서열은 ITR2와 같은 자연적으로 발생하는 AAV ITR의 서열에 비해 결실 및/또는 치환된다. 일부 구현예에서, 하나 이상의 전사 인자 결합 부위가 결실 및/또는 치환되는 최소 기능적 ITR이다. 일부 구현예에서, 적어도 1개의 전사 인자 결합 부위, 예를 들어, 적어도 5개 이상 또는 10개 이상의 전사 인자 결합 부위, 예를 들어, 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 또는 21개의 전사 인자 결합 부위가 결실 및/또는 치환된다.

[0482] 또 다른 구현예에서, 본원에 기재된 바와 같은 rAAV 벡터 게놈을 포함하는 rAAV 벡터는 적어도 하나의 합성 AAV ITR을 포함하는 폴리뉴클레오타이드를 포함하며, 여기서, ITR에서 전사 개시 부위에서 또는 그 근처에서 전형적으로 발생하는 하나 이상의 CpG 섬(시토신 염기 다음에 구아닌 염기(CpG)가 바로 이어지며, 이러한 배열의 시토신은 메틸화되는 경향이 있음)은 결실 및/또는 치환된다. 일 구현예에서, CpG 섬 수의 결실 또는 감소는 rAAV 벡터의 면역원성을 감소시킬 수 있다. 이는 CpG 섬에서 발생하는 rAAV 벡터 DNA 서열에 대한 TLR-9 결합의 감소 또는 완전한 억제의 결과이다. CpG 모티프의 메틸화가 전사 사일런싱을 초래한다는 것도 잘 알려져 있다. ITR에서 CpG 모티프의 제거는 TLR-9 인식 감소 및/또는 메틸화 감소를 초래하여 전이유전자 사일런싱을 감소시킬 것으로 예상된다. 일부 구현예에서, 하나 이상의 CpG 섬이 삭제 및/또는 치환되는 최소 기능적 ITR이다. 일 구현예에서, AAV ITR2는 16개 CpG 섬을 포함하는 것으로 알려져 있으며, 그 중 1개 이상 또는 16개 모두가 결실될 수 있다.

[0483] 일부 구현예에서, 적어도 1개의 CpG 모티프, 예를 들어, 적어도 4개 이상 또는 8개 이상의 CpG 모티프, 예를 들어, 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 또는 16개의 CpG 모티프가 결실 및/또는 치환된다.

[0484] 또 다른 구현예에서, 합성 ITR은 표 7에 나열된 뉴클레오타이드 서열 중 하나를 포함하거나 이로 본질적으로 이루어지거나 이로 이루어진다. 다른 구현예에서, 합성 ITR은 표 7에 나열된 뉴클레오타이드 서열 중 임의의 하나와 적어도 80% 동일한, 예를 들어, 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 뉴클레오타이드 서열을 포함하거나 이로 본질적으로 이루어지거나 이로 이루어진다. 일부 구현예에서, ITR은 도 1에서 변형된 ITR 서열을 개시하는 문헌[Samulski *et al.*, 1983, Cell, 33; 135-143 (그 전체가 본원에 참조로 포함된 "Samulski *et al.*, 1983"으로 지칭됨)]의 도 1에 개시된 서열이다. 일부 구현예에서, ITR 서열은 문헌[Samulski *et al.*, 1993]에 개시된 도 1의 ITR 서열 중 어느 하나와 적어도 80% 동일한, 예를 들어, 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 뉴클레오타이드 서열을 포함하거나 이로 이루어진다. 일부 구현예에서, ITR은 문헌[Samulski *et al.*, 1993]에 개시된 도 1(9bp가 결여됨)의 중간 패널에 개시된 우측 pSM 609의 ITR 서열과 적어도 80% 동일한, 예를 들어, 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 또는 99.5% 동일한 뉴클레오타이드 서열을 포함하거나 이로 이루어진다. 일부 구현예에서, ITR은 서열 번호 441-444 중 어느 하나의 ITR 서열과 적어도 80% 동일한, 예를 들어, 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 또는 99.5% 동일한 뉴클레오타이드 서열을 포함한다. 일부 구현예에서, ITR 서열, 예를 들어, 우측 ITR (또는 3' ITR)은 서열 번호 442, 또는 서열 번호 442와 적어도 80% 동일한, 예를 들어, 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 또는 99.5% 동일한 뉴클레오타이드 서열이다. 일부 구현예에서, ITR 서열, 예를 들어, 좌측 ITR (또는 5' ITR)은 서열 번호 441, 또는 서열 번호 441과 적어도 80% 동일한, 예를 들어, 적어도 85%,

90%, 95%, 96%, 97%, 98%, 또는 99% 또는 99.5% 동일한 뉴클레오타이드 서열이다.

[0485]

[표 7] 예시적인 합성 ITR 서열

MH-257	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGCTGAGGCAATTTGATAAAAAATCGTCAAATTATAAACAGGCTTTGCGCTGTTTAGCCTCAGTGAGCGAGCGAGCGCGAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTTCCT (서열 번호 36)
MH-258	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGCTGAGGGATAAAAAATCCAGGCTTTGCGCTGCCTCAGTGAGCGAGCGAGCGCGCGAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTTCCT (서열 번호 37)
MH 델타 258	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGCTGAGGGATAAAAAATCCAGGCTTTGCGCTGCCTCAGTGAGCGAGCGAGCGCGCGAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTTCCT (서열 번호 38)
MH 텔로미어-1-ITR	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGGGATTGGGATTGCGCGCTCGCTCGCGGGATTGGGATTGGGATTGGGATTGGGATTGGGATTGATAAAAAATCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCGCGAGCGAGCGCGCAATCCCAATCCCAAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTTCCT (서열 번호 39)
MH 텔로미어-2 ITR	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGGATTGGGATTGGGATTGGGATTGGGATTGGGATTGATAAAAAATCAATCCCAATCCCAATCCCAATCCCAATCCCGCGAGCGAGCGCGAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTTCCTAAGCTTATTAT A (서열 번호 40)
MH PolIII 258 ITR	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGCTGAGGGCGCCTATAAAGATAAAAAATCCAGGCTTTGCGCTGCCTCAGTTAGCGAGCGAGCGCGCGAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTTCCT (서열 번호 41)
MH 258 델타 D 보존적	CTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGGATAAAAAATCCAGGCTTTGCGCTGCCTCAGTGAGCGAGCGAGCGCGAGAGAGGGAGTGGCCAACTCCATCACTAG (서열 번호 42)

[0486]

5' ITR	TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACC AAAGGTGCCCCGACGCCCGGGCTTTGCCCCGGGCGGCTCAGTGAGCGAG CGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCT (서열 번호 161)
3' ITR	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCG CTCACTGAGGCCGGGCGACCAAAGGTGCCCCGACGCCCGGGCTTTGCCC GGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA (서열 번호 165)
L-ITR	CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCCGGGCGTC GGGCGACCTTTGGTTCGCCCGGCTCAGTGAGCGAGCGAGCGCGCAGAGA GGGAGTGGCCAACTCCATCACTAGGGGTTCT (서열 번호 441)
R-ITR	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCG CTCACTGAGGCCGGGCGACCAAAGGTGCCCCGACGCCCGGGCTTTGCCC GGGCGGCCTCAGTGAGCGAGCGAGCGCGCAG (서열 번호 442)
ITR (145bp)	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCG CTCACTGAGGCCGGGCGACCAAAGGTGCCCCGACGCCCGGGCTTTGCCC GGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA (서열 번호 443)
ITR (145bp – 1983, 9bp 결여)	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCG CTCACTGAGGCCGGGCGACCAAAGGTGCCCCGACGCCCGGGCTTTGCCC GGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGG (서열 번호 444)

[0487]

[0488]

J. 예시적인 rAAV 게놈 요소:

[0489]

(i) AAV-LVR412

[0490]

일부 구현예에서, rAAV는 그 게놈에 5'에서 3' 방향으로, 5' AAV2-ITR(서열 번호 161), 스테퍼 DNA(서열 번호 162), SP0412 LSP(서열 번호 91), Gluc에 대한 코작 서열(서열 번호 153), GAA 폴리펩타이드를 암호화하는 핵산 서열(서열 번호 182), 3' UTR(서열 번호 77), 폴리A 서열(서열 번호 164), 3' AAV-ITR 서열(서열 번호 165)을 포함하는 작제물을 포함한다. 예시적인 작제물은 서열 번호 154(LVR412_AskBioEU)로서 나타낸다. 상이한 AAV 혈청형에 대한 임의의 ITR 서열 또는 서열 번호 441-444 중 어느 하나로부터 선택된 ITR에 대해 서열 번호 161 및 165의 ITR 서열을 쉽게 변경할 수 있을 뿐만 아니라, 각각 서열 77 및 서열 번호 164 대신 상이한 UTR 서열 또는 폴리A 서열을 사용할 수 있다.

[0491]

일부 구현예에서, rAAV는 그 게놈에 5'에서 3' 방향으로, 5' AAV2-ITR(서열 번호 161), 스테퍼 DNA(서열 번호 162), SP0412 LSP(서열 번호 91), GAA 폴리펩타이드를 암호화하는 핵산 서열(서열 번호 182), 3' UTR(서열 번호 77), 폴리A 서열(서열 번호 166), 3' AAV-ITR 서열(서열 번호 165)을 포함하는 작제물을 포함한다. 예시적인 작제물은 서열 번호 155(SsAAV_LVR412WT-hGAA_AskBio_CHATHAM_백본_Ask)로 나타낸다. 상이한 AAV 혈청형에 대한 임의의 ITR 서열 또는 서열 번호 441-444 중 어느 하나로부터 선택된 ITR에 대해 서열 번호 161 및 165의 ITR 서열을 쉽게 변경할 수 있을 뿐만 아니라, 각각 서열 번호 77 및 서열 번호 164 대신 상이한 UTR 서열 또는 폴리A 서열을 사용할 수 있다.

[0492]

AAV2-LVP422:

[0493]

일부 구현예에서, rAAV는 그 게놈에 5'에서 3' 방향으로, 5' AAV2-ITR(서열 번호 161), 스테퍼 DNA, SP0422 LSP(서열 번호 92), Gluc에 대한 코작 서열(서열 번호 153), GAA 폴리펩타이드를 암호화하는 핵산 서열(서열 번호 55), 콜라겐 안정성 서열(서열 번호 65), 폴리A 서열(서열 번호 164), 3' AAV-ITR 서열(서열 번호 165)을 포함하는 작제물을 포함한다. 예시적인 작제물은 서열 번호 156(LVR412스테퍼)으로 나타낸다. 상이한 AAV 혈청형에 대한 임의의 ITR 서열 또는 서열 번호 441-444 중 어느 하나로부터 선택된 ITR에 대해 서열 번호 161 및 165의 ITR 서열을 쉽게 변경할 수 있을 뿐만 아니라, 각각 서열 번호 65 및 서열 번호 164 대신 상이한 콜라겐 안정성 서열 또는 폴리A 서열을 사용할 수 있다.

[0494]

일부 구현예에서, rAAV는 그 게놈에 5'에서 3' 방향으로, 5' AAV2-ITR(서열 번호 161), 스테퍼 DNA, SP0422 LSP(서열 번호 92), Gluc에 대한 코작 서열(서열 번호 153), GAA 폴리펩타이드를 암호화하는 핵산 서열(서열 번호 182), 3' UTR 서열(서열 번호 77), AATAA 정지 신호를 임의로 포함하는 폴리A 서열(서열 번호 164), 3'

AAV-ITR 서열(서열 번호 165)을 포함하는 작제물을 포함한다. 예시적인 작제물은 서열 번호 157(LVR422AskBio EU 작제물)로서 나타난다. 상이한 AAV 혈청형에 대한 임의의 ITR 서열 또는 서열 번호 441-444 중 어느 하나로 부터 선택된 ITR에 대해 서열 번호 161 및 165의 ITR 서열을 쉽게 변경할 수 있을 뿐만 아니라, 각각 서열 번호 77 및 서열 번호 164 대신 상이한 UTR 서열 또는 폴리A 서열을 사용할 수 있다.

[0495] 상이한 AAV 혈청형에 대한 임의의 ITR 서열 또는 서열 번호 441-444의 ITR에 대해 서열 번호 160, 159, 155, 158, 156의 rAAV 게놈에서 ITR 서열을 쉽게 변경할 수 있을 뿐만 아니라, 본원에 개시된 바와 같은 상이한 UTR 서열 또는 폴리A 서열을 사용할 수 있다.

[0496] 일부 구현예에서, rAAV 벡터 또는 rAAV 게놈은 서열 번호 170 (동족 GAA신호 서열 및 H199R, R223H 변형을 갖는 GAA 폴리펩타이드), 또는 서열 번호 171(동족 GAA신호 서열 및 H199R, H201L 및 R223H 변형을 갖는 GAA 폴리펩타이드)을 포함하는 GAA 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함한다. 서열 번호 170의 GAA 폴리펩타이드는 서열 번호 182의 핵산 서열에 의해 암호화된다. 따라서, 일부 구현예에서, rAAV 벡터는 H199R, R223H 변형을 포함하는 변형된 GAA 폴리펩타이드를 암호화하는 서열 번호 182의 핵산을 포함한다. 서열 번호 171의 GAA 폴리펩타이드는 서열 번호 182의 핵산 서열에 의해 암호화되며, 여기서, 서열 번호 182의 염기쌍(bp) 667-669는 CAC로부터: UUA, UUG, CUU, CUC CUA, CUG 중 어느 하나로 변경되거나(히스타딘(H)에서 류신(L) 아미노산 변경을 초래함); 서열 번호 182의 bp 668은 A에서 U로 변경된다(히스타딘(H)에서 류신(L) 아미노산 변경을 초래함). 따라서, 일부 구현예에서, rAAV 벡터는 서열 번호 182의 핵산을 포함하며, 여기서, 서열 번호 182의 bp 667-669는 CAC로부터: UUA, UUG, CUU, CUC CUA, CUG 중 어느 하나로 변경되거나; 서열 번호 182의 bp 668은 H199R, H201L 및 R223H 변형을 포함하는 변형된 GAA 폴리펩타이드를 암호화하는 A에서 U로 변경된다.

[0497] 일부 구현예에서, rAAV 벡터 또는 rAAV 게놈은 서열 번호 172(동족 신호 펩타이드가 IgG신호 서열 및 H199R, R223H 변형으로 대체된 GAA 폴리펩타이드), 또는 서열 번호 172, 또는 서열 번호 173과 적어도 85% 서열 동일성을 갖는 서열(동족 신호 펩타이드가 wt IL2신호 서열 및 H199R 및 R223H 변형으로 대체된 GAA 폴리펩타이드), 또는 서열 번호 173 또는 서열 번호 174에 대해 적어도 85% 서열 동일성을 갖는 서열(동족 신호 펩타이드가 mut IL3 신호 서열 및 H199R 및 R223H 변형으로 대체된 GAA 폴리펩타이드) 또는 서열 번호 174와 적어도 85% 서열 동일성을 갖는 서열 중 어느 하나로부터 선택된 GAA 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함한다.

[0498] 일부 구현예에서, rAAV 벡터 또는 rAAV 게놈은 서열 번호 182를 포함하는 이중 핵산 서열을 포함하며, 여기서, 서열 번호 182의 bp 1-81은 서열 번호 172(H199R 및 R223H 변형을 갖는 IgG 리더-GAA)의 GAA 폴리펩타이드를 암호화하는 서열 번호 177의 핵산(IgG신호 서열)으로 대체된다. 일부 구현예에서, rAAV 벡터는 서열 번호 182를 포함하는 이중 핵산 서열을 포함하며, 여기서, 서열 번호 182의 bp 668은 A에서 U로 변경되고, 서열 번호 182의 bp 1-81은 서열 번호 177의 핵산(IgG 신호 펩타이드), 또는 서열 번호 172의 GAA 폴리펩타이드(H199R, H201L 및 R223H 변형을 갖는 IgG 리더-GAA)를 암호화하는, 이와 적어도 85% 서열 동일성을 갖는 서열로 대체된다.

[0499] 일부 구현예에서, rAAV 벡터 또는 rAAV 게놈은 서열 번호 182를 포함하는 이중 핵산 서열을 포함하며, 여기서, 서열 번호 182의 bp 1-81은 서열 번호 179의 핵산(wt IL2 신호 펩타이드), 또는 서열 번호 173의 GAA 폴리펩타이드(H199R, R223H 변형을 갖는 wt IL2 신호 펩타이드-GAA)를 암호화하는 이와 적어도 85% 서열 동일성을 갖는 서열로 대체된다. 일부 구현예에서, rAAV 벡터는 서열 번호 182를 포함하는 이중 핵산 서열을 포함하며, 여기서, 서열 번호 182의 bp 668은 A에서 U로 변경되고, 서열 번호 182의 bp 1-81은 서열 번호 179의 핵산(wt IL2 신호 펩타이드), 또는 서열 번호 173의 GAA 폴리펩타이드(H199R, H201L R223H 변형을 갖는 wt IL2 신호 펩타이드-GAA)를 암호화하는, 이와 적어도 85% 서열 동일성을 갖는 서열로 대체된다.

[0500] 일부 구현예에서, rAAV 벡터는 서열 번호 182를 포함하는 이중 핵산 서열을 포함하며, 여기서, 서열 번호 182의 bp 1-81은 서열 번호 181의 핵산(mut IL2 신호 펩타이드), 또는 서열 번호 174의 GAA 폴리펩타이드(H199R, R223H 변형을 갖는 mut IL2 신호 펩타이드-GAA)를 암호화하는, 이와 적어도 85% 서열 동일성을 갖는 서열로 대체된다. 일부 구현예에서, rAAV 벡터는 서열 번호 182를 포함하는 이중 핵산 서열을 포함하며, 여기서, 서열 번호 182의 bp 668은 A에서 U로 변경되고, 서열 번호 182의 bp 1-81은 서열 번호 181의 핵산(mut IL2 신호 펩타이드), 또는 서열 번호 174의 GAA 폴리펩타이드(H199R, H201L 및 R223H 변형을 갖는 mut IL2 신호 펩타이드-GAA)를 암호화하는, 이와 적어도 85% 서열 동일성을 갖는 서열로 대체된다.

[0501] III. 벡터 및 비리온

[0502] 하나의 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터(rAAV 비리온으로도 지칭됨)는 캡시드 단백질, 및 캡시

드 단백질에서 rAAV 게놈을 포함한다. 폼페병을 치료하는 데 사용되는 rAAV 비리온의 rAAV 캡시드는 그 전체가 본원에 참조로 포함된 2019년 11월 19일에 출원된 62,937,556에 개시된 표 1에 나열된 것들 중 어느 하나 또는 이의 임의의 조합이다.

- [0503] 하나의 구현예에서, 폼페병을 치료하는 데 사용되는 rAAV 비리온의 rAAV 캡시드는 국제 출원 WO2020/102645 및 WO2020/102667에 개시된 표 1에 나열된 것들 중 하나이며, 이들 각각은 그 전체가 본원에 참조로 포함된다. 하나의 구현예에서, 폼페병을 치료하는 데 사용되는 rAAV 비리온의 rAAV 캡시드는 AAV8 캡시드이다. 하나의 구현예에서, rAAV 벡터는 rAAV8 벡터이다.
- [0504] 하나의 구현예에서, 본원에 개시된 바와 같은 AAV 벡터(rAAV 비리온으로도 지칭됨)는 그 전체가 본원에 참조로 구체적으로 포함된 WO2019/241324에 개시된 것들 중 임의의 것으로부터의 캡시드 단백질을 포함한다. 일부 구현예에서, rAAV 벡터는 간 특이적 캡시드, 예를 들어, 그 전체가 본원에 참조로 포함된 WO2019/241324에 개시된 바와 같은 XL32 및 XL32.1로부터 선택된 간 특이적 캡시드를 포함한다. 일부 구현예에서, rAAV 벡터는 그 전체가 본원에 참조로 포함된 WO2019/241324에 개시된 바와 같은 AAVXL32 또는 AAVXL32.1이다.
- [0505] 본원에 기재된 rAAV 벡터에서 AAV 캡시드로서 사용될 수 있는 예시적인 키메라 또는 변이체 캡시드 단백질은 그 전체가 본원에 참조로 구체적으로 포함된, 2019년 11월 19일에 출원된 미국 가출원 62,937,556의 표 2로부터 선택될 수 있거나, 야생형 캡시드 단백질 및/또는 현재 알려져 있거나 나중에 확인되는 다른 키메라 또는 변이체 캡시드 단백질과의 임의의 조합과 함께 사용될 수 있으며 각각은 본원에 포함된다. 일부 구현예에서, 사용을 위해 포함된 rAAV 벡터는, 예를 들어, 그 전체가 본원에 참조로 포함된 9,012,224 및 US 7,892,809에 개시된 바와 같은 키메라 벡터이다.
- [0506] 일부 구현예에서, rAAV 벡터는 미국 출원 US2018/0371496 및 PCT/US18/22725에 개시된 바와 같은 반수체 rAAV 벡터이거나, 예를 들어, 7/31/2018에 출원된 PCT/US2018/044632 및 미국 출원 16/151,110에 개시된 바와 같은 배수체 rAAV 벡터이며, 이들 각각은 그 전체가 본원에 참조로 포함된다. 일부 구현예에서, rAAV 벡터는 그 전체가 본원에 참조로 포함된 9,012,224 및 WO 2017/106236에 개시된 바와 같은 rAAV3 벡터이다.
- [0507] 특정 구현예에서, rAAV는 그 전체가 본원에 참조로 포함된 WO2019/241324에 개시된 바와 같은 AAVXL32 또는 AAVXL32.1 AAV 벡터이다. 일부 구현예에서, rAAV 벡터는 그 전체가 본원에 참조로 포함된 WO2019241324A1 또는 국제 특허 출원 PCT/US2019/036676에 개시된 캡시드를 포함한다. 일부 구현예에서, AAV 벡터는 AAV8 벡터 또는 AAV8 캡시드 단백질을 포함하는 합리적인 반수체이다. 일부 구현예에서, 재조합 AAV 벡터는 키메라 AAV 벡터, 반수체 AAV 벡터, 하이브리드 AAV 벡터 또는 배수체 AAV 벡터이다. 일부 구현예에서, 재조합 AAV 벡터는 합리적인 반수체 벡터, 모자이크 AAV 벡터, 화학적으로 변형된 AAV 벡터, 또는 임의의 AAV 혈청형, 예를 들어, 이들 각각은 그 전체가 본원에 참조로 포함된 국제 출원 WO2020/102645 및 WO2020/102667에 개시된 표 1에 개시된 임의의 AAV 혈청형으로부터의 AAV 벡터이다.
- [0508] 일부 구현예에서, AAV 벡터는 WO2019241324A1에 개시된 바와 같은 서열 번호 1-3 중 어느 하나의 뉴클레오타이드 서열과 적어도 90% 동일한 핵산 AAV 캡시드 암호화 서열; 또는 (b) WO2019241324A1에 개시된 바와 같은 서열 번호 4-6 중 어느 하나를 암호화하는 뉴클레오타이드 서열에 의해 암호화된 캡시드를 포함한다. 일부 구현예에서, AAV 캡시드는 AAV 벡터 게놈 및 본 발명의 AAV 캡시드를 포함하는 AAV 입자와 함께, WO2019241324A1에 개시된 바와 같은 서열 번호 4-6 중 어느 하나와 적어도 90% 동일한 아미노산 서열을 포함한다.
- [0509] 하나의 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터는 야생형 캡시드 단백질 및/또는 현재 알려져 있거나 나중에 확인되는 다른 키메라 또는 변이체 캡시드 단백질과 임의의 조합으로 본 발명의 AAV 캡시드에 포함될 수 있는 키메라 또는 변이체 캡시드 단백질을 설명하는(설명 목적을 위해, 11486254는 미국 특허 출원 번호 11/486,254에 상응하며 다른 생물학적 서열 파일은 유사한 방식으로 관독되어야 함) USPTO 허여된 특허 및 공개 출원의 파일 래퍼에 나열된 하기 생물학적 서열 파일 중 임의의 것과 관련된 캡시드 단백질을 포함한다: 11486254, 11932017, 12172121, 12302206, 12308959, 12679144, 13036343, 13121532, 13172915, 13583920, 13668120, 13673351, 13679684, 14006954, 14149953, 14192101, 14194538, 14225821, 14468108, 14516544, 14603469, 14680836, 14695644, 14878703, 14956934, 15191357, 15284164, 15368570, 15371188, 15493744, 15503120, 15660906, 및 15675677.
- [0510] 일 구현예에서, 본 발명의 AAV 캡시드 단백질 및 바이러스 캡시드는, 예를 들어, 참조로 포함되는 국제 특허 공개 WO 00/28004에 기재된 바와 같은 또 다른 바이러스, 임의로 또 다른 파르보바이러스 또는 AAV로부터의 캡시드 서브유닛의 전부 또는 일부를 포함할 수 있다는 점에서 키메라일 수 있다. 일부 구현예에서, rAAV 벡터 게

높은 본원에 포함된 미국 특허 번호 8,784,799에 기재된 바와 같은 단일 가닥 또는 단량체 듀플렉스이다.

- [0511] 추가 구현예로서, 본 발명의 AAV 캡시드 단백질 및 바이러스 캡시드는 참조로 포함된 미국 출원 US2018/0371496에 기재된 바와 같은 단일 AAV 캡시드에서 VP1, VP2 및 VP3 AAV 혈청형의 상이한 조합을 포함할 수 있다는 점에서 배수체(반수체로도 지칭됨)일 수 있다.
- [0512] 일 구현예에서, 본원에 개시된 바와 같은 폼페병의 치료에 유용한 rAAV 벡터는 AAV3b 캡시드이다. 사용을 위해 포함된 AAV3b 캡시드는 2017/106236, 및 9,012,224 및 7,892,809, 및 2019년 11월 15일에 출원된 국제 출원 PCT/US19/61653, 및 국제 출원 WO2020/102645 및 WO2020/102667에 기재되어 있으며, 이들 각각은 그 전체가 본원에 참조로 포함된다.
- [0513] 일부 구현예에서, AAV3b 캡시드는 서열 번호 44를 포함한다. 일 구현예에서, 폼페병의 치료에 사용된 AAV 캡시드는 서열 번호 44에 기재된 AAV 캡시드로부터 전체 또는 부분적으로 유래된 변형된 AAV 캡시드일 수 있다. 일부 구현예에서, 서열 번호 44에 기재된 AAV3b 캡시드로부터의 아미노산은 상이한 AAV 혈청형의 또 다른 캡시드로부터의 아미노산일 수 있거나 이로 치환될 수 있으며, 여기서, 치환 및/또는 삽입된 아미노산은 임의의 AAV 혈청형으로부터 유래할 수 있고, 자연적으로 발생하는 또는 부분적으로 또는 완전히 합성된 아미노산을 포함할 수 있다.
- [0514] 또 다른 구현예에서, 폼페병의 치료에 사용되는 AAV 캡시드는 AAV3b265D 캡시드이다. 이러한 특정 구현예에서, AAV3b265D 캡시드는 AAV3b 캡시드의 아미노산 G265의 D265로의 대체를 통한 AAV3b 캡시드의 2중 측 루프의 아미노산 서열의 변형을 포함한다. 일부 구현예에서, AAV3b265D 캡시드는 서열 번호 46을 포함한다. 그러나, 본 발명의 변형된 바이러스 캡시드는 서열 번호 46에 기재된 AAV 캡시드로 제한되지 않는다. 일부 구현예에서, 서열 번호 46에 기재된 AAV3b265D로부터의 아미노산은 상이한 혈청형의 AAV로부터 캡시드로부터의 아미노산일 수 있거나 이로 치환될 수 있으며, 여기서, 치환 및/또는 삽입된 아미노산은 임의의 AAV 혈청형으로부터 유래할 수 있고, 자연적으로 발생하는 또는 부분적으로 또는 완전히 합성된 아미노산을 포함할 수 있다.
- [0515] 또 다른 구현예에서, 본원에 개시된 바와 같은 폼페병의 치료에 유용한 rAAV 벡터는 AAV3b265D549A 캡시드이다. 이러한 특정 구현예에서, AAV3b265D549A 캡시드는 AAV3b 캡시드의 아미노산 G265의 D265로의 대체 및 AAV3b 캡시드의 아미노산 T549의 A549로의 대체를 통한 AAV3b 캡시드의 2중 측 루프의 아미노산 서열의 변형을 포함한다. 일부 구현예에서, AAV3b265D549A 캡시드는 서열 번호 50을 포함한다. 그러나, 본 발명의 변형된 바이러스 캡시드는 서열 번호 50에 기재된 AAV 캡시드로 제한되지 않는다. 일부 구현예에서, 서열 번호 50에 기재된 AAV3b265D549A로부터의 아미노산은 상이한 혈청형의 AAV로부터 캡시드로부터의 아미노산일 수 있거나 이로 치환될 수 있으며, 여기서, 치환 및/또는 삽입된 아미노산은 임의의 AAV 혈청형으로부터 유래할 수 있고, 자연적으로 발생하는 또는 부분적으로 또는 완전히 합성된 아미노산을 포함할 수 있다.
- [0516] 또 다른 구현예에서, 본원에 개시된 바와 같은 폼페병의 치료에 유용한 rAAV 벡터는 AAV3b549A 캡시드이다. 이러한 특정 구현예에서, AAV3b549A 캡시드는 AAV3b 캡시드의 아미노산 T549의 A549로의 대체를 통한 AAV3b 캡시드의 2중 측 루프의 아미노산 서열의 변형을 포함한다. 일부 구현예에서, AAV3b549A 캡시드는 서열 번호 52를 포함한다. 그러나, 본 발명의 변형된 바이러스 캡시드는 서열 번호 52에 제시된 AAV 캡시드로 제한되지 않는다. 일부 구현예에서, 서열 번호 52에 기재된 AAV3b549A로부터의 아미노산은 상이한 혈청형의 AAV로부터 캡시드로부터의 아미노산일 수 있거나 이로 치환될 수 있으며, 여기서, 치환 및/또는 삽입된 아미노산은 임의의 AAV 혈청형으로부터 유래할 수 있고, 자연적으로 발생하는 또는 부분적으로 또는 완전히 합성된 아미노산을 포함할 수 있다.
- [0517] 또 다른 구현예에서, 본원에 개시된 바와 같은 폼페병의 치료에 유용한 rAAV 벡터는 AAV3bQ263Y 캡시드이다. 이러한 특정 구현예에서, AAV3bQ263Y 캡시드는 AAV3b 캡시드의 아미노산 Q263의 Y263으로의 대체를 통한 AAV3b 캡시드의 2중 측 루프의 아미노산 서열의 변형을 포함한다. 일부 구현예에서, AAV3b549A 캡시드는 서열 번호 54를 포함한다. 그러나, 본 발명의 변형된 바이러스 캡시드는 서열 번호 54에 기재된 AAV 캡시드로 제한되지 않는다. 일부 구현예에서, 서열 번호 54에 기재된 AAV3bQ263Y로부터의 아미노산은 상이한 혈청형의 AAV로부터 캡시드로부터의 아미노산일 수 있거나 이로 치환될 수 있으며, 여기서, 치환 및/또는 삽입된 아미노산은 임의의 AAV 혈청형으로부터 유래할 수 있고, 자연적으로 발생하는 또는 부분적으로 또는 완전히 합성된 아미노산을 포

함할 수 있다.

[0518] 또 다른 구현예에서, 본원에 개시된 바와 같은 폼페병의 치료에 유용한 rAAV 벡터는 AAV3bSASTG 혈청형이거나, AAV3bSASTG 캡시드를 포함한다. 이러한 특정 구현예에서, AAV3bSASTG 캡시드는 SASTG 돌연변이를 포함하는 아미노산 서열의 변형을 포함하고, 특히 AAV3b 캡시드는 AAV3b 캡시드의 유사한 위치에 이러한 변형을 도입함으로써 AAV2 Q263A/T265 하위변이체와 유사하도록 변형되었다(문헌(Messina EL, et al., Adeno-associated viral vectors based on serotype 3b use components of the fibroblast growth factor receptor signaling complex for efficient transduction. Hum. Gene Ther. 2012 Oct; 23(10):1031-4, Piacentino III, Valentino, et al. "X-linked inhibitor of apoptosis protein-mediated attenuation of apoptosis, using a novel cardiac-enhanced adeno-associated viral vector." Human gene therapy 23.6 (2012): 635-646)에 개시된 바와 같으며, 이들 둘 다는 그 전체가 본원에 참조로 포함됨). 따라서, 일부 구현예에서, 본원에 개시된 바와 같은 폼페병의 치료에 유용한 rAAV 벡터는 AAV3bSASTG 혈청형이거나, AAV3b Q263A/T265 캡시드를 포함하는 AAV3bSASTG 캡시드를 포함한다. 일부 구현예에서, AAV3bSASTG로부터의 아미노산은 상이한 혈청형의 AAV로부터 캡시드로부터의 아미노산일 수 있거나 이로 치환될 수 있으며, 여기서, 치환 및/또는 삽입된 아미노산은 임의의 AAV 혈청형으로부터 유래할 수 있고, 자연적으로 발생하는 또는 부분적으로 또는 완전히 합성된 아미노산을 포함할 수 있다.

[0519] 세포 내로의 이들의 도입을 촉진하기 위해, 본 발명에 유용한 rAAV 벡터 계놈은 다음을 포함하는 제조함 핵산 작제물이다: (1) 발현될 이중 서열(하나의 구현예에서, GAA 폴리펩타이드를 암호화하는 폴리뉴클레오타이드) 및 (2) 이중 유전자의 통합 및 발현을 촉진하는 바이러스 서열 요소. 바이러스 서열 요소는 AAV 캡시드로의 DNA의 복제 및 패키징(예를 들어, 기능적 ITR)을 위해 시스템에서 요구되는 AAV 벡터 계놈의 서열을 포함할 수 있다. 일 구현예에서, 이중 유전자는 폼페병을 앓고 있는 환자에서 GAA-결핍을 교정하는데 유용한 GAA를 암호화한다. 일 구현예에서, 이러한 rAAV 벡터 계놈은 또한 마커 또는 리포터 유전자를 포함할 수 있다. 일 구현예에서, rAAV 벡터 계놈은 AAV3b 야생형(WT) 시스템 유전자 중 하나 이상이 전체 또는 부분적으로 대체되거나 결실될 수 있지만, 기능적 플랭킹 ITR 서열을 보유할 수 있다.

[0520] 하나의 구현예에서, 폼페병의 치료에 유용한 본원에 개시된 rAAV 벡터는 AAV3b 캡시드에 의해 캡슐화된 본원에 개시된 바와 같은 rAAV 계놈을 포함한다. 일부 구현예에서, 폼페병의 치료에 유용한 본원에 개시된 바와 같은 rAAV 벡터는 다음으로부터 선택된 임의의 AAV3b 캡시드에 의해 캡슐화된 본원에 개시된 바와 같은 rAAV 계놈을 포함한다: AAV3b 캡시드(서열 번호 44); AAV3b265D 캡시드(서열 번호 46), AAV3b ST(S663V+T492V) 캡시드(서열 번호 48), AAV3b265D549A 캡시드(서열 번호 50); AAV3b549A 캡시드(서열 번호 52); AAV3bQ263Y 캡시드(서열 번호 54), 또는 AAV3bSASTG(즉, Q263A/T265) 캡시드.

[0521] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터는 다음 중 어느 하나의 핵산 서열: AAV_LVR412_EU(서열 번호 154), ssAAV_LVR412WT-hGAA_AskBio_CHATHAM(서열 번호 155), AAV-LVR412스터퍼(서열 번호 156), AAV_LVR422_EU(서열 번호 157), AAV-LVR422스터퍼(서열 번호 158), ssAAV_LVR412_WT-hGAA CHATHAM(서열 번호 159), ssAAV_LSP_WT-hGAA-CHATHAM(서열 번호 160), 또는 이와 적어도 80%, 85%, 90%, 95% 또는 98% 동일성을 갖는 핵산 서열을 포함한다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, 서열 번호 154-160 중 어느 하나의 핵산 서열을 포함하는 rAAV 벡터는 본원에 개시된 바와 같은 변형된 GAA 핵산 서열에 의해 대체된 wtGAA 서열을 가질 수 있다.

[0522] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터는 다음 중 어느 하나의 핵산 서열: 서열 번호 57(AAT-V43M-wtGAA(델타1-69aa)); 서열 번호 58(ratFN1-IGF2V43M-wtGAA(델타1-69aa)); 서열 번호 59(hFN1-IGF2V43M-wtGAA(델타1-69aa)); 서열 번호 60(AAT-IGF2D2-7-wtGAA(델타 1-69)); 서열 번호 61(FN1rat-IGF Δ 2-7-wtGAA(델타 1-69)); 서열 번호 62(hFN1-IGF Δ 2-7-wtGAA(델타 1-69)), 또는 이와 적어도 80%, 85%, 90%, 95% 또는 98% 동일성을 갖는 핵산 서열을 포함한다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, rAAV 벡터는 다음 중 어느 하나의 핵산 서열: AAT_hIGF2-V43M_wtGAA_del1-69_스터퍼.V02(서열 번호 79); FIBrat_hIGF2-V43M_wtGAA_del1-69_스터퍼.V02(서열 번호 80); FIBhum_hIGF2-V43M_wtGAA_del1-69_스터퍼.V02(서열 번호 81); AAT_GILT_wtGAA_del1-69_스터퍼.V02(서열 번호 82); FIBrat_GILT_wtGAA_del1-69_스터퍼.V02(서열 번호 83); FIBhum_GILT_wtGAA_del1-69_스터퍼.V02(서열 번호 84) 또는 이와 적어도 80%, 85%, 90%, 95% 또는 98% 동일성을 갖는 핵산 서열을 포함한다.

[0523] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터는 다음 중 어느 하나의 핵산 서열: AAV_LVR412_EU(서열 번호 154), ssAAV_LVR412WT-hGAA_AskBio_CHATHAM(서열 번호 155), AAV-

LVR412스터퍼(서열 번호 156), AAV_LVR422_EU(서열 번호 157), AAV-LVR422_스터퍼(서열 번호 158), ssAAV_LVR412_WT-hGAA CHATHAM(서열 번호 159), ssAAV_LSP_WT-hGAA-CHATHAM(서열 번호 160), 또는 이와 적어도 80%, 85%, 90%, 95% 또는 98% 동일성을 갖는 핵산 서열을 포함한다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, 서열 번호 154-160 중 어느 하나의 핵산 서열을 포함하는 rAAV 벡터는 wtGAA 서열이 본원에 개시된 바와 같은 변형된 GAA 핵산 서열로 대체될 수 있다.

[0524] IV. 최적화된 rAAV 벡터 계능

[0525] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 최적화된 rAAV 벡터 계능은 프로모터, ITR, 폴리-A 테일, 이중 유전자의 발현을 증가 또는 감소시킬 수 있는 요소를 암호화하는 핵산 서열, 및 하나의 구현예에서 생체내 GAA 단백질의 발현을 위해 코돈 최적화된 핵산 서열(즉, coGAA 또는 코돈 최적화된 GAA) 및 임의로 번역원성을 감소시키는 하나 이상의 요소를 포함하여, 본원에 개시된 임의의 요소 및 임의의 조합으로부터 생성된다. 이러한 최적화된 rAAV 벡터 계능은 rAAV 벡터 계능이 형질도입되고 발현될 조직 및 세포에 대해 친화성을 갖는 임의의 AAV 캡시드와 함께 사용될 수 있다.

[0526] 일부 구현예에서, rAAV 계능에는 일반적으로 본원에 개시된 바와 같은 간-특이적 프로모터의 업스트림에 위치하는 AAV P5 프로모터가 결여되어 있다. 일반적으로, P5 프로모터는 AAV 복제 동안 AAV rep/cap 단백질의 발현을 제어한다. 일부 구현예에서, 이 P5 프로모터 단편은 스트레스 반응(Watson 2017)에도 관여하는 전사 인자 2(ATF2)를 활성화하는 소포체(ER)/골지 스트레스(Sampieri 2019)에 의해 활성화될 수 있는, 예측된 전사 인자 결합 부위, 예를 들어, 사이클릭 AMP-반응성 요소-결합 단백질 3(CREB3)을 함유하는 본원에 개시된 바와 같은 rAAV 벡터에 존재한다. 핵 수용체 서브패밀리 1 그룹 I 구성원 2(NR1I2) (Pregnane X 수용체 [PXR]로도 알려져 있음)는 간에 풍부한 것으로 알려져 있으며 프리그난 스테로이드, 리팜핀 및 텍사메타손(NR1I2_HGNC)을 포함한 기타 분자에 의해 활성화된다(NR1I2_HGNC) (Xing 2020). 따라서, 일부 구현예에서, rAAV 계능에서 AAV P5 프로모터의 단편은 GAA 카세트의 의도된 성능에 영향을 미치지 않으면서 제거된다. 일부 구현예에서, rAAV 벡터는 또한 폴리A 신호와 3' ITR 사이에 위치한 RNA 폴리머라제 II 종결 서열을 포함한다. 예시적인 말단 서열은 2개의 종결 코돈 및 1개의 제한 부위(예를 들어, XhoI)가 TAG를 대체하고 hGAA의 마지막 암호화 아미노산의 바로 다운스트림에 위치하고 3' UTR의 바로 업스트림에 위치하는 서열 번호 439이다.

[0527] AAV3b 캡시드 변형

[0528] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터에 사용하기 위한 AAV3b 캡시드는 둘 다 그 전체가 본원에 참조로 포함된 문헌(Nienaber et al., Hum. Gen Ther, 2012, 23(10); 1031-42 and Piacentino III, Valentino, et al. "X-linked inhibitor of apoptosis protein-mediated attenuation of apoptosis, using a novel cardiac-enhanced adeno-associated viral vector." Human gene therapy 23.6 (2012): 635-646]에 개시된 AAV3b 캡시드(서열 번호 44); AAV3b265D 캡시드(서열 번호 46), AAV3b ST (S663V+T492V) 캡시드(서열 번호 48), AAV3b265D549A 캡시드(서열 번호 50); AAV3b549A 캡시드(서열 번호 52); AAV3bQ263Y 캡시드(서열 번호 54) 또는 AAV3bSASTG 캡시드(즉, Q263A/T265 돌연변이를 포함하는 AAV3b 캡시드) 중 어느 하나와 예를 들어, 약 75% 내지 약 100%, 약 80% 내지 약 100%, 약 85% 내지 약 100%, 약 90% 내지 약 100%, 약 95% 내지 약 100%, 약 75% 내지 약 99%, 약 80% 내지 약 99%, 약 85% 내지 약 99%, 약 90% 내지 약 99%, 약 95% 내지 약 99%, 약 75% 내지 약 97%, 약 80% 내지 약 97%, 약 85% 내지 약 97%, 약 90% 내지 약 97%, 또는 약 95% 내지 약 97% 범위의 아미노산 동일성을 갖는다. 이 구현예의 또 다른 양태에서, AAV3b로부터 유래된 AAV는 문헌(Nienaber et al., Hum. Gen Ther, 2012, 23(10); 1031-42, 및 Piacentino III, Valentino, et al. "X-linked inhibitor of apoptosis protein-mediated attenuation of apoptosis, using a novel cardiac-enhanced adeno-associated viral vector." Human gene therapy 23.6 (2012): 635-646)에 개시된 바와 같은 AAV3b 캡시드(서열 번호 44); AAV3b265D 캡시드(서열 번호 46), AAV3b ST (S663V+T492V) 캡시드(서열 번호 48), AAV3b265D549A 캡시드(서열 번호 50); AAV3b549A 캡시드(서열 번호 52); AAV3bQ263Y 캡시드(서열 번호 54), AAV3bSASTG 캡시드(즉, Q263A/T265 돌연변이를 포함하는 AAV3b 캡시드)에 대한 임의의 아미노산 서열과 예를 들어, 약 75% 내지 약 100%, 약 80% 내지 약 100%, 약 85% 내지 약 100%, 약 90% 내지 약 100%, 약 95% 내지 약 100%, 약 75% 내지 약 99%, 약 80% 내지 약 99%, 약 85% 내지 약 99%, 약 90% 내지 약 99%, 약 95% 내지 약 99%, 약 75% 내지 약 97%, 약 80% 내지 약 97%, 약 85% 내지 약 97%, 약 90% 내지 약 97%, 또는 약 95% 내지 약 97% 범위의 아미노산 동일성을 갖지만, 캡시드는 여전히 기능적으로 활성인 AAV 단백질이다.

[0529] 본원에 개시된 방법 및 조성물의 일부 구현예에서, AAV 혈청형(예를 들어, AAV3b)은 그 전체가 본원에 참조로 포함된 문헌(Messina EL, et al., Adeno-associated viral vectors based on serotype 3b use components of

the fibroblast growth factor receptor signaling complex for efficient transduction. Hum. Gene Ther. 2012 Oct; 23(10):1031-42)에 기재된 바와 같은 SASTG 돌연변이를 포함한다.

[0530] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터에 사용하기 위한 AAV3b 캡시드는 AAV3b 캡시드(서열 번호 44); AAV3b265D 캡시드(서열 번호 46), AAV3b ST(S663V+T492V) 캡시드(서열 번호 48), AAV3b265D549A 캡시드(서열 번호 50); AAV3b549A 캡시드(서열 번호 52); AAV3bQ263Y 캡시드(서열 번호 54), 또는 AAV3bSASTG 캡시드(즉, Q263A/T265 돌연변이를 포함하는 AAV3b 캡시드) (문헌(Nienaber et al., Hum. Gen Ther, 2012, 23(10); 1031-42)에 개시된 바와 같음)에 대한 임의의 아미노산 서열에 대해 예를 들어, 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 또는 50개의 연속 아미노산 결실, 부가 및/또는 치환; 또는 AAV3b 캡시드(서열 번호 44); AAV3b265D 캡시드(서열 번호 46), AAV3b ST (S663V+T492V) 캡시드(서열 번호 48), AAV3b265D549A 캡시드(서열 번호 50); AAV3b549A 캡시드(서열 번호 52); AAV3bQ263Y 캡시드(서열 번호 54), AAV3bSASTG 캡시드(즉, Q263A/T265 돌연변이를 포함하는 AAV3b 캡시드) (문헌(Nienaber et al., Hum. Gen Ther, 2012, 23(10); 1031-42)에 개시된 바와 같음)에 대한 임의의 아미노산 서열에 대해 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 또는 50개의 연속 아미노산 결실, 부가 및/또는 치환을 갖는다. 또 다른 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터에서 사용하기 위한 AAV3b 캡시드는 AAV3b 캡시드(서열 번호 44); AAV3b265D 캡시드(서열 번호 46), AAV3b ST (S663V+T492V) 캡시드(서열 번호 48), AAV3b265D549A 캡시드(서열 번호 50); AAV3b549A 캡시드(서열 번호 52); AAV3bQ263Y 캡시드(서열 번호 54), 또는 AAV3bSASTG 캡시드(즉, Q263A/T265 돌연변이를 포함하는 AAV3b 캡시드) [문헌(Nienaber et al., Hum. Gen Ther, 2012, 23(10); 1031-42)에 개시된 바와 같음]에 대한 임의의 아미노산 서열에 대해 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 또는 50개의 연속 아미노산 결실, 부가 및/또는 치환을 갖지만, 여전히 기능적으로 활성화된 AAV이다.

[0531] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터에서 사용하기 위한 AAV3b 캡시드는 AAV3b 캡시드(서열 번호 44); AAV3b265D 캡시드(서열 번호 46), AAV3b ST (S663V+T492V) 캡시드(서열 번호 48), AAV3b265D549A 캡시드(서열 번호 50); AAV3b549A 캡시드(서열 번호 52); AAV3bQ263Y 캡시드(서열 번호 54), AAV3bSASTG 캡시드(즉, Q263A/T265 돌연변이를 포함하는 AAV3b 캡시드) (문헌(Nienaber et al., Hum. Gen Ther, 2012, 23(10); 1031-42)에 개시된 바와 같음)에 대한 임의의 아미노산 서열과 예를 들어, 약 75% 내지 약 100%, 약 80% 내지 약 100%, 약 85% 내지 약 100%, 약 90% 내지 약 100%, 약 95% 내지 약 100%, 약 75% 내지 약 99%, 약 80% 내지 약 99%, 약 85% 내지 약 99%, 약 90% 내지 약 99%, 약 95% 내지 약 99%, 약 75% 내지 약 97%, 약 80% 내지 약 97%, 약 85% 내지 약 97%, 약 90% 내지 약 97%, 또는 약 95% 내지 약 97% 범위의 아미노산 동일성을 갖는다. 더 추가의 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터에 사용하기 위한 AAV3b 캡시드는 AAV3b 캡시드(서열 번호 44); AAV3b265D 캡시드(서열 번호 46), AAV3b ST(S663V+T492V) 캡시드(서열 번호 48), AAV3b265D549A 캡시드(서열 번호 50); AAV3b549A 캡시드(서열 번호 52); AAV3bQ263Y 캡시드(서열 번호 54), AAV3bSASTG 캡시드(즉, Q263A/T265 돌연변이를 포함하는 AAV3b 캡시드) (문헌(Nienaber et al., Hum. Gen Ther, 2012, 23(10); 1031-42)에 개시된 바와 같음)에 대한 임의의 아미노산 서열과 예를 들어, 약 75% 내지 약 100%, 약 80% 내지 약 100%, 약 85% 내지 약 100%, 약 90% 내지 약 100%, 약 95% 내지 약 100%, 약 75% 내지 약 99%, 약 80% 내지 약 99%, 약 85% 내지 약 99%, 약 90% 내지 약 99%, 약 95% 내지 약 99%, 약 75% 내지 약 97%, 약 80% 내지 약 97%, 약 85% 내지 약 97%, 약 90% 내지 약 97%, 또는 약 95% 내지 약 97% 범위의 아미노산 동일성을 갖지만, 여전히 기능적으로 활성화된 AAV 단백질이다.

[0532] V. 치료 방법

[0533] A. 폼페병

[0534] 본원에 개시된 바와 같은 GAA 단백질을 발현하는 재조합 AAV는 폼페병 및 기타 글리코젠 저장 질환(GSD)을 치료하는 방법에 사용될 수 있다. 폼페병은 에너지로 사용되는 저장된 형태의 당인 글리코젠을 분해하는 데 필요한 효소 산 알파-글루코시다아제(GAA)의 결핍에 의해 야기된 희귀 유전 질환이다. 폼페병은 글리코젠 축적 질환 II형, GSD II, II형 글리코젠 축적 질환, 글리코젠증 II형, 산성 말타아제 결핍증, 알파-1,4-글루코시다아제 결핍증, 심근비대증 글리코젠 확산, 및 심장 형태의 전신 글리코젠증으로도 알려져 있다. 글리코젠의 축적은 전

신에 걸쳐 진행성 근육 약화(근병증)를 야기하고 다양한 신체 조직, 특히 심장, 골격근, 간, 호흡기 및 신경계에 영향을 미친다.

[0535] 폼페병의 임상 증상을 나타내는 것은 발병 연령과 잔류 GAA 활성에 따라 크게 달라질 수 있다. 잔류 GAA 활성은 글리코겐 축적의 양 및 조직 분포 둘 모두뿐만 아니라 질환의 중증도와 상관관계가 있다. 영아-발병형 폼페병(정상 GAA 활성의 1% 미만)은 가장 심각한 형태이며 근육긴장저하, 전신성 근력 저하, 및 비대성 심근병증, 및 심장 및 기타 근육 조직에서의 글리코겐 대량 축적을 특징으로 한다. 사망은 보통 심폐 기능 부전으로 인해 출생 후 1년 이내에 발생한다. 청소년-발병(정상 GAA 활성의 1-10%) 및 성인-발병(정상 GAA 활성의 10-40%) 폼페병은 발병 연령, 임상 증상 및 질환 진행의 더 큰 변화와 함께 임상적으로 더 이질적이다. 청소년- 및 성인-발병 폼페병은 일반적으로 심각한 심장 침범이 없고 발병 연령이 늦고 질환 진행이 느린 것이 특징이지만, 결국 호흡 또는 사지 근육 침범은 심각한 이환율과 사망률을 초래한다. 기대 수명은 다양할 수 있지만 일반적으로 호흡 부전으로 인해 사망이 발생한다.

[0536] 본원에 개시된 방법 및 조성물의 임의의 구현예에서, 폼페병 치료에 적합한 GAA 효소는 선형 올리고사카라이드에서 α1-4 연결을 절단하는 능력을 유지하는 야생형 인간 GAA, 또는 이의 단편 또는 서열 변이체를 포함한다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, GAA 단백질은 야생형 GAA 핵산 서열, 예를 들어, 서열 번호 11 또는 서열 번호 72에 의해 암호화된다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, GAA 단백질은 코돈 최적화된 GAA 핵산 서열에 의해, 예를 들어 (1) 생체내 향상된 발현, (2) CpG 섬 감소 또는 (3) 선천성 면역 반응 감소 중 어느 하나 이상에 대해 암호화된다. 본원에 개시된 방법 및 조성물의 일부 구현예에서, GAA 단백질은 코돈 최적화된 GAA 핵산 서열, 예를 들어, 서열 번호 73, 서열 번호 74, 서열 번호 75 및 서열 번호 76 또는 서열 번호 182 중 어느 하나로부터 선택된 임의의 핵산 서열, 또는 서열 번호 73, 서열 번호 74, 서열 번호 75, 서열 번호 76 또는 서열 번호 182와 적어도 60%, 또는 70%, 또는 80%, 85% 또는 90% 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열에 의해 암호화된다.

[0537] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 본원에 기재된 바와 같은 rAAV 벡터는 대상체의 간을 형질도입하고 hGAA 폴리펩타이드를 혈액으로 분비하며, 이는 융합된 IGF2-서열의 도움으로 hGAA 폴리펩타이드가 세포에 의해 흡수되어 리소솜으로 수송되는 환자 조직을 관류하며, 여기서, 효소는 효소 결핍으로 인해 리소솜에 축적된 물질을 제거하는 역할을 한다. 리소솜 효소 대체 요법이 효과적이기 위해서는 치료 효소가 저장 결함이 나타나는 조직의 적절한 세포에 있는 리소솜으로 전달되어야 한다.

[0538] B. 생체의 세포에서 GAA 수준 조절

[0539] 본원에 기재된 바와 같은 핵산, 벡터, 및 비리온은 세포에서 GAA의 수준을 조절하는 데 사용될 수 있다. 상기 방법은 2개의 AAV ITR 사이에 삽입된 GAA를 암호화하는 폴리뉴클레오타이드를 포함하는 핵산을 포함하는 조성물을 세포에 투여하는 단계를 포함한다. 세포는 본 발명의 핵산이 투여될 수 있는 임의의 동물로부터 유래할 수 있다. GAA 결핍이 있는 대상체로부터의 포유동물 세포(예를 들어, 인간, 개, 고양이, 돼지, 양, 마우스, 래트, 토끼, 소, 염소 등)는 본 발명에 사용하기 위한 전형적인 표적 세포이다. 일부 구현예에서, 세포는 간 세포 또는 심근 세포(myocardial cell), 예를 들어 심근세포(myocardiocyte)이다.

[0540] 일 구현예에서, rAAV 벡터로 형질도입된 세포의 생체의 전달이 본원에 개시되어 있다. 추가의 구현예에서, 생체의 유전자 전달을 사용하여 본원에 개시된 바와 같은 rAAV 벡터로 형질도입된 세포를 숙주 내로 다시 이식할 수 있다. 추가의 구현예에서, 생체의 줄기 세포(예를 들어, 중간엽 줄기 세포) 요법을 사용하여 본원에 개시된 바와 같은 rAAV 벡터로 형질도입된 세포를 숙주 내로 세포를 다시 이식할 수 있다. 또 다른 구현예에서, 적합한 생체의 프로토콜은 여러 단계를 포함할 수 있다.

[0541] 일부 구현예에서, 표적 조직(예를 들어, 근육, 간 조직)의 분절이 대상체로부터 수확될 수 있으며, 본원에 기재된 rAAV 벡터는 GAA-암호화 핵산을 숙주의 세포로 형질도입하는데 사용된다. 이러한 유전적으로 변형된 세포는 숙주에 다시 이식될 수 있다. 정맥 주사, 복강내 주사, 피하 주사 또는 표적 조직으로의 동일 반응계내 주사를 포함하여 세포를 숙주로 재도입하기 위해 여러 접근법이 사용될 수 있다. 본원에 기재된 바와 같은 rAAV 벡터로 형질도입되거나 감염된 변형된 생체의 세포의 미세캡슐화는 본 발명 내에서 사용될 수 있는 또 다른 기술이다. 자가 및 동종이체 세포 이식이 본 발명에 따라 사용될 수 있다.

[0542] 또 다른 추가 구현예에서, 약제학적으로 허용되는 담체에 본원에 개시된 바와 같은 GAA 발현 세포를 치료학적 유효량으로 대상체에게 투여하는 단계를 포함하여, 대상체에서 GAA 결핍을 치료하는 방법이 본원에 개시되어 있다. 일부 구현예에서, 대상체는 인간이다.

[0543] C. 대상체에서 GAA 활성 증가

[0544] 본원에 기재된 바와 같은 핵산, 벡터, 및 비리온은 대상체, 예를 들어, 인간 대상체, 또는 폼페병이 있거나 폼페병에 걸릴 위험이 있는 대상체에서 기능적 GAA 폴리펩타이드의 수준을 조절하는 데 사용될 수 있다. 상기 방법은 2개의 AAV ITR 사이에 삽입된 GAA를 암호화하는 이중 핵산을 포함하는, 본원에 기재된 바와 같은 rAAV 게놈을 포함하는 rAAV 벡터를 포함하는 조성물을 대상체에게 투여하는 단계를 포함하며, 여기서, hGAA는 본원에 기재된 바와 같은 신호 펩타이드, 및 임의로 본원에 개시된 바와 같은 IGF2 표적화 펩타이드에 연결된다. 대상체는 임의의 동물일 수 있으며, 예를 들어, 포유동물(예를 들어, 인간, 개, 고양이, 돼지, 양, 마우스, 래트, 토끼, 소, 염소 등)이 적합한 대상체이다. 본 발명의 방법 및 조성물은 GAA-결핍 인간 대상체에 특히 적용 가능하다.

[0545] 또한, 본원에 기재된 핵산, 벡터, 및 비리온은 임의의 적절한 방법에 의해 임의의 적절한 제형으로 인간을 포함한 동물에게 투여될 수 있다. 예를 들어, 본원에 개시된 방법 및 조성물의 임의의 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터, 또는 rAAV 게놈은 경구, 직장, 경점막, 비강내, 흡입(예를 들어, 에어로졸을 통해), 협측(예를 들어, 설하), 질, 척추강내, 안내, 경피, 자궁내(또는 난소 내), 비경구(예를 들어, 정맥내, 피하, 피내, 근육내[골격, 횡격막 및/또는 심장 근육에 대한 투여 포함], 피내, 흉막내, 뇌내 및 관절내), 국소(예를 들어, 기도 표면 및 경피 투여를 포함한 피부 및 점막 표면 둘 모두에), 림프관내 등 뿐만 아니라 직접적인 조직 또는 기관 주사(예를 들어, 간, 골격근, 심장 근육, 횡격막 근육 또는 뇌) 또는 원하는 투여 경로 및 표적이 되는 조직에 따라 다른 비경구 경로에 의한 투여를 포함하여 동물에 직접 도입될 수 있다.

[0546] 일부 구현예에서, rAAV 벡터 또는 rAAV 게놈이 간 특이적 프로모터의 제어 하에 있거나 이에 작동적으로 연결된 GAA를 암호화하는 핵산 서열을 포함하기 때문에, 본원에 개시된 바와 같은 방법 및 조성물은 정맥내 또는 근육내 주사를 통해 투여될 수 있으며, 여기서 rAAV 벡터 또는 rAAV 게놈은 간으로 이동하여 GAA 단백질을 발현할 것이다.

[0547] 본원에 개시된 방법 및 조성물의 임의의 구현예에서, 근육에 대한 투여는 정맥내 투여, 동맥내 투여 및/또는 복강내 투여를 포함하는 임의의 적합한 방법에 의해 이루어질 수 있다. 예시적인 투여 방식은 경구, 직장, 경점막, 비강내, 흡입(예를 들어, 에어로졸을 통해), 협측(예를 들어, 설하), 질, 척추강내, 안내, 경피, 자궁내(또는 난소 내), 비경구(예를 들어, 정맥내, 피하, 피내, 근육내[골격, 횡격막 및/또는 심장 근육에 대한 투여 포함], 피내, 흉막내, 뇌내 및 관절내), 국소(예를 들어, 기도 표면 및 경피 투여를 포함한 피부 및 점막 표면 둘 모두에), 림프관내 등 뿐만 아니라 직접적인 조직 또는 기관 주사(예를 들어, 간, 골격근, 심장 근육, 횡격막 근육 또는 뇌)를 포함한다. 임의의 주어진 경우에 가장 적합한 경로는 치료 및/또는 예방되는 병태의 특징 및 중증도 및 사용되는 특정 벡터의 특징에 따라 달라질 것이다.

[0548] 본원에 개시된 방법 및 조성물의 임의의 구현예에서, 본 발명에 따른 골격근에 대한 투여는 사지(예를 들어, 위팔, 아래팔, 윗다리 및/또는 아랫다리), 등, 목, 머리(예를 들어, 혀), 흉부, 복부, 골반/회음 및/또는 손가락의 골격근에 대한 투여를 포함하지만 이에 제한되지 않는다. 주사될 수 있는 적합한 골격근은 그 전체가 본원에 참조로 포함된 2019년 11월 19일에 출원된 미국 가출원 62,937,556에 개시되어 있다.

[0549] 본원에 개시된 방법 및 조성물의 임의의 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터 및/또는 rAAV 게놈은 대상체의 골격근, 간, 횡격막, 늑골 및/또는 심장 근육 세포에 투여된다. 예를 들어, 기존의 주사기와 바늘을 사용하여 rAAV 비리온 현탁액을 동물에 주사할 수 있다. 주사에 의한 rAAV 벡터 및/또는 rAAV 게놈의 비경구 투여는, 예를 들어, 볼루스 주사 또는 연속 주입에 의해 수행될 수 있다. 주사용 제형은 방부제가 첨가된 단위 투여 형태, 예를 들어, 앰플 또는 다중 투여 용기로 제공될 수 있다. 조성물은 유성 또는 수성 비히클 중의 현탁액, 용액 또는 에멀전과 같은 형태를 취할 수 있으며, 현탁제, 안정화제 및/또는 분산제와 같은 약제학적 제형용 제제를 함유할 수 있다. 대안적으로, 본원에 개시된 바와 같은 rAAV 벡터 및/또는 rAAV 게놈은 사용 전에 적합한 비히클, 예를 들어, 멸균 발열원 없는 물과의 구성을 위해 분말 형태(예를 들어, 동결건조됨)일 수 있다.

[0550] 특정 구현예에서, 다양한 간격, 예를 들어, 매시간, 매일, 매주, 매월, 매년 등의 기간에 걸쳐 원하는 수준의 유전자 발현을 달성하기 위해 1회 초과 투여(예를 들어, 2회, 3회, 4회, 5회, 6회, 7회, 8회, 9회, 10회 등, 또는 그 이상의 투여)가 사용될 수 있다. 투여량은 단일 투여량 또는 누적 투여량(연속 투여량)일 수 있으며, 당업자에 의해 쉽게 결정될 수 있다. 예를 들어, 질환 또는 장애의 치료는 유효 용량의 본원에 개시된 약제학적 조성물 바이러스 벡터의 1회 투여를 포함할 수 있다. 대안적으로, 질환 또는 장애의 치료는, 예를 들어, 1일 1회, 1일 2회, 1일 3회, 며칠마다 1회 또는 매주 1회와 같은 기간 범위에 걸쳐 수행되는 유효 용량의 바이러스

스 백터의 다중 투여를 포함할 수 있다.

- [0551] 투여 시기는 개인의 증상의 중증도와 같은 인자에 따라 개인마다 다를 수 있다. 예를 들어, 유효 용량의 본원에 개시된 바이러스 백터는 무기한 기간 동안 또는 개체가 더 이상 치료를 필요로 하지 않을 때까지 6개월마다 1회 개체에게 투여될 수 있다. 당업자는 개체의 병태가 치료 과정 전반에 걸쳐 모니터링될 수 있으며 유효량의 투여되는 본원에 개시된 바이러스 백터가 이에 따라 조정될 수 있음을 인식할 것이다.
- [0552] 주사제는 액체 용액 또는 현탁액, 주사 전 액체 중의 용액 또는 현탁액에 적합한 고체 형태로서, 또는 에멀전으로서 통상적인 형태로 제조될 수 있다. 대안적으로, 본 발명의 바이러스 백터 및/또는 바이러스 캡시드를 전신 방식보다는 국소 방식으로, 예를 들어, 데포(depot) 또는 서방성 제형으로 투여할 수 있다. 또한, 바이러스 백터 및/또는 바이러스 캡시드는 외과적으로 이식 가능한 기질에 부착되어 전달될 수 있다(예를 들어, 미국 특허 공개 번호 US-2004-0013645-A1에 기재된 바와 같음). 본원에 개시된 바이러스 백터 및/또는 바이러스 캡시드는 임의의 적합한 수단에 의해, 임의로 대상체가 흡입하는 바이러스 백터 및/또는 바이러스 캡시드로 구성된 호흡 가능한 입자의 에어로졸 현탁액을 투여함으로써 대상체의 폐에 투여될 수 있다. 호흡 가능한 입자는 액체 또는 고체일 수 있다. 바이러스 백터 및/또는 바이러스 캡시드를 포함하는 액체 입자의 에어로졸은 당업자에게 알려져 바와 같이, 압력-구동 에어로졸 분무기 또는 초음파 분무기와 같은 임의의 적합한 수단에 의해 생성될 수 있다. 예를 들어, 미국 특허 번호 4,501,729를 참조한다. 바이러스 백터 및/또는 캡시드를 포함하는 고체 입자의 에어로졸은 마찬가지로 제약 분야에 알려진 기술에 의해 임의의 고체 미립자 약제 에어로졸 발생기를 사용하여 생성될 수 있다.
- [0553] 일부 구현예에서, 본원에 개시된 바와 같은 rAAV 백터 및/또는 rAAV 게놈은 본원에 개시된 rAAV 백터를 용해시키기에 충분한 양으로 용매, 에멀전 또는 기타 희석제에서 제형화될 수 있다. 이 구현예의 다른 양태에서, 본원에 개시된 바와 같은 rAAV 백터 및/또는 rAAV 게놈은 용매, 에멀전 또는 희석제에서, 예를 들어, 약 90% (v/v) 미만, 약 80% (v/v) 미만, 약 70% (v/v) 미만, 약 65% (v/v) 미만, 약 60% (v/v) 미만, 약 55% (v/v) 미만, 약 50% (v/v) 미만, 약 45% (v/v) 미만, 약 40% (v/v) 미만, 약 35% (v/v) 미만, 약 30% (v/v) 미만, 약 25% (v/v) 미만, 약 20% (v/v) 미만, 약 15% (v/v) 미만, 약 10% (v/v) 미만, 약 5% (v/v) 미만, 또는 약 1% (v/v) 미만의 양으로 제형화될 수 있다. 다른 양태에서, 본원에 개시된 바와 같은 rAAV 백터 및/또는 rAAV 게놈은 용매, 에멀전 또는 기타 희석제를 예를 들어, 약 1% (v/v) 내지 90% (v/v), 약 1% (v/v) 내지 70% (v/v), 약 1% (v/v) 내지 60% (v/v), 약 1% (v/v) 내지 50% (v/v), 약 1% (v/v) 내지 40% (v/v), 약 1% (v/v) 내지 30% (v/v), 약 1% (v/v) 내지 20% (v/v), 약 1% (v/v) 내지 10% (v/v), 약 2% (v/v) 내지 50% (v/v), 약 2% (v/v) 내지 40% (v/v), 약 2% (v/v) 내지 30% (v/v), 약 2% (v/v) 내지 20% (v/v), 약 2% (v/v) 내지 10% (v/v), 약 4% (v/v) 내지 50% (v/v), 약 4% (v/v) 내지 40% (v/v), 약 4% (v/v) 내지 30% (v/v), 약 4% (v/v) 내지 20% (v/v), 약 4% (v/v) 내지 10% (v/v), 약 6% (v/v) 내지 50% (v/v), 약 6% (v/v) 내지 40% (v/v), 약 6% (v/v) 내지 30% (v/v), 약 6% (v/v) 내지 20% (v/v), 약 6% (v/v) 내지 10% (v/v), 약 8% (v/v) 내지 50% (v/v), 약 8% (v/v) 내지 40% (v/v), 약 8% (v/v) 내지 30% (v/v), 약 8% (v/v) 내지 20% (v/v), 약 8% (v/v) 내지 15% (v/v), 또는 약 8% (v/v) 내지 12% (v/v) 범위의 양으로 포함할 수 있다.
- [0554] 본원에 개시된 방법 및 조성물의 임의의 구현예에서, AAV3b 캡시드(서열 번호 44); AAV3b265D 캡시드(서열 번호 46), AAV3b ST (S663V+T492V) 캡시드(서열 번호 48), AAV3b265D549A 캡시드(서열 번호 50); AAV3b549A 캡시드(서열 번호 52); AAV3bQ263Y 캡시드(서열 번호 54) 또는 AAV3bSASTG 캡시드(즉, Q263A/T265 돌연변이를 포함하는 AAV3b 캡시드)로부터 선택된 임의의 AAV3b 캡시드에 의해 캡슐화된 것을 포함하지만 이에 제한되지 않는 임의의 혈청형의 본원에 개시된 바와 같은 rAAV 백터 및/또는 rAAV 게놈은 치료학적 유효량의 치료 화합물을 포함할 수 있다. 일 구현예에서, 본원에 사용된 바와 같이, 제한 없이, 용어 "유효량"은 "치료학적 유효량", "유효 용량" 또는 "치료학적 유효 용량"과 동의어이다. 일 구현예에서, 폐폐병을 치료하기 위한 본원에 개시된 치료 화합물의 유효성은, 제한 없이, 하나 이상의 임상 증상 및/또는 폐폐병과 관련된 생리학적 지표에 기반한 개체의 개선을 관찰함으로써 결정될 수 있다. 일 구현예에서, 폐폐병과 관련된 증상의 개선은 동시 요법에 대한 필요성 감소로 나타날 수 있다.
- [0555] 본원에 개시된 바와 같은 rAAV 백터 및/또는 rAAV 게놈의 전달을 쉽게 하기 위해, 이는 담체 또는 부형제와 혼합될 수 있다. 사용될 수 있는 담체 및 부형제는 식염수(특히 멸균, 발열원이 없는 식염수), 식염수 완충액(예를 들어, 구연산염 완충액, 인산염 완충액, 아세트산염 완충액 및 중탄산염 완충액), 아미노산, 요소, 알코올, 아스코르브산, 인지질, 단백질(예를 들어, 혈청 알부민), EDTA, 염화나트륨, 리포솜, 만니톨, 소르비톨, 및 글리세롤을 포함한다. USP 등급 담체 및 부형제는 인간 대상체에게 비리온을 전달하는 데 특히 유용하다.

- [0556] 이전에 설명된 제형에 추가하여, 본원에 개시된 바와 같은 rAAV 벡터 및/또는 rAAV 게놈도 테포 제제로서 제형화될 수 있다. 이러한 장기 작용 제형은 이식(예를 들어, 피하 또는 근육내) 또는 IM 주사에 의해 투여될 수 있다. 따라서, 예를 들어, 본원에 개시된 바와 같은 rAAV 벡터 및/또는 rAAV 게놈은 적합한 중합체 또는 소수성 물질(예를 들어, 허용되는 오일 중 에멀전으로서) 또는 이온 교환 수지와 함께 또는 난용성 유도체로 제형화될 수 있다.
- [0557] 본원에 개시된 방법 및 조성물의 임의의 구현예에서, 상기 방법은 대상체에서 GAA 결핍으로부터 초래된 폼페병의 치료에 관한 것이며, 여기서, 본원에 개시된 바와 같은 rAAV 벡터 및/또는 rAAV 게놈은 폼페병을 앓고 있는 환자에게 투여되고, 투여 후 GAA는 간 세포로부터 분비되며, 골격근 조직, 심장 근육 조직, 횡격막 근육 조직 또는 이들의 조합에서 세포에 의해 분비된 GAA의 흡수가 있고, 여기서, 분비된 GAA의 흡수는 조직(들)에서 리소좀 글리코겐 저장을 감소시킨다. 일부 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터 및/또는 rAAV 게놈은 캡시드에 캡슐화되며, 예를 들어, AAV3b 캡시드(서열 번호 44); AAV3b265D 캡시드(서열 번호 46), AAV3b ST (S663V+T492V) 캡시드(서열 번호 48), AAV3b265D549A 캡시드(서열 번호 50); AAV3b549A 캡시드(서열 번호 52); AAV3bQ263Y 캡시드(서열 번호 54)로부터 선택된 임의의 AAV3b 캡시드에 의해 캡슐화된다.
- [0558] 특정 구현예에서, 적어도 약 10^2 내지 약 10^8 개 세포 또는 적어도 약 10^3 내지 약 10^6 개 세포가 약제학적으로 허용되는 담체에서 용량당 투여될 것이다. 추가 구현예에서, 대상체에게 투여될 바이러스 벡터 및/또는 캡시드의 투여량은 투여 모드, 치료 및/또는 예방할 질환 또는 병태, 개별 대상체의 병태, 특정 바이러스 벡터 또는 캡시드, 전달될 핵산 등에 의존하며, 일상적인 방식으로 결정될 수 있다. 치료 효과를 달성하기 위한 예시적인 용량은 적어도 약 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} 개 형질도입 단위, 임의로 약 10^8 - 10^{13} 개 형질도입 단위의 역가이다.
- [0559] 또 다른 양태에서, 본원에는 핵산이 세포 내로 도입되고 발현되어 GAA를 생성하기 위한 조건 하에 세포를 본원에 개시된 바와 같은 rAAV 벡터 및/또는 rAAV 게놈과 접촉시키는 단계를 포함하여, GAA를 암호화하는 핵산을 세포에 투여하는 방법이 개시되어 있다. 일부 구현예에서, 세포는 배양된 세포이다. 일부 구현예에서, 세포는 생체내 세포이다. 일부 구현예에서, 세포는 포유동물 세포이다. 일부 구현예에서, GAA를 암호화하는 핵산을 세포에 투여하는 방법은 세포 배양 배지 내로 분비된 GAA를 수집하는 단계를 추가로 포함한다.
- [0560] **D. 포유동물의 운동뉴런 기능 증가**
- [0561] 본원에 개시된 방법 및 조성물의 임의의 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터 및/또는 rAAV 게놈은 폼페병 및/또는 불충분한 GAA 수준을 갖는 포유동물에서 횡격막 신경 활성을 증가시키기 위한 조성물 및 방법에 유용하다. 예를 들어, 본원에 개시된 바와 같은 rAAV 벡터 및/또는 rAAV 게놈, 예를 들어, 캡시드에 캡슐화된, 예를 들어, AAV3b 캡시드(서열 번호 44); AAV3b265D 캡시드(서열 번호 46), AAV3b ST (S663V+T492V) 캡시드(서열 번호 48), AAV3b265D549A 캡시드(서열 번호 50); AAV3b549A 캡시드(서열 번호 52); AAV3bQ263Y 캡시드(서열 번호 54)로부터 선택된 임의의 AAV3b 캡시드에 의해 캡슐화된 rAAV 벡터 및/또는 rAAV 게놈은 중추 신경계(예를 들어, 뉴런)에 투여될 수 있다. 또 다른 구현예에서, 횡격막 (또는 다른 근육)으로부터 횡격막 신경 또는 다른 운동 뉴런으로의 GAA를 암호화하는 본원에 개시된 바와 같은 rAAV 벡터 및/또는 rAAV 게놈의 역행 수송은 폼페병의 생화학적 및 생리학 적정치를 초래할 수 있다. 이러한 동일한 원리는 다른 신경퇴행성 질환에 적용될 수 있다.
- [0562] 일 구현예에서, AAV8 또는 AAV3, 또는 AAV3b (AAV3b 혈청형 AAV3b265D, AAV3b265D549A, AAV3b549A, AAV3bQ263Y, AAV3bSASTG(즉, Q263A/T265 돌연변이를 포함하는 AAV3b 캡시드) 혈청형을 포함하지만 이에 제한되지 않음)를 포함한 표 1에 기재된 바와 같은 임의의 혈청형의 rAAV GAA 작제물은 하기 증상 중 임의의 하나 이상을 감소시킬 수 있다: (i) 다리, 몸통 및/또는 팔을 포함한 환자의 하지에서의 쇠약감, (ii) 숨가쁨 (shortness of breath), 힘든 운동 시간, 폐 감염, 척추의 큰 만곡, 수면 동안 호흡 곤란, 간 비대, 허 비대 및/또는 관절 뻣뻣함, (iii) 폼페병을 앓고 있는 환자에서 동일한 치료를 받지 않은 환자와 비교하여 예를 들어, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90% 또는 적어도 95% 까지. 이 구현예의 다른 양태에서, 임의의 혈청형의 AAV GAA는 하기 증상 중 임의의 하나 이상을 감소시킬 수 있다: (i) 다리, 몸통 및/또는 팔을 포함한 환자의 하지에서의 쇠약감, ii) 숨가쁨, 힘든 운동 시간, 폐 감염, 척추의 큰 만곡, 수면 동안 호흡 곤란, 간 비대, 허 비대 및/또는 관절 뻣뻣함, (iii) 폼페병을 앓고 있는 환자에서 동일한 치료를 받지 않은 환자와 비교하여, 예를 들어, 약 10% 내지 약 100%, 약 20% 내지

약 100%, 약 30% 내지 약 100%, 약 40% 내지 약 100%, 약 50% 내지 약 100%, 약 60% 내지 약 100%, 약 70% 내지 약 100%, 약 80% 내지 약 100%, 약 10% 내지 약 90%, 약 20% 내지 약 90%, 약 30% 내지 약 90%, 약 40% 내지 약 90%, 약 50% 내지 약 90%, 약 60% 내지 약 90%, 약 70% 내지 약 90%, 약 10% 내지 약 80%, 약 20% 내지 약 80%, 약 30% 내지 약 80%, 약 40% 내지 약 80%, 약 50% 내지 약 80%, 또는 약 60% 내지 약 80%, 약 10% 내지 약 70%, 약 20% 내지 약 70%, 약 30% 내지 약 70%, 약 40% 내지 약 70%, 또는 약 50% 내지 약 70% 까지.

[0563] 본원에 개시된 방법 및 조성물의 임의의 구현예에서, 폼페병과 관련된 적어도 하나의 증상, 또는 폼페병과 관련된 적어도 하나의 부작용이 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 또는 적어도 95%까지 감소되고, 폼페병과 관련된 적어도 하나의 증상의 중증도, 또는 적어도 하나의 유해한 부작용이 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 또는 적어도 95%까지 감소된다. 또 다른 구현예에서, 폼페병과 관련된 적어도 하나의 증상, 또는 폼페병과 관련된 적어도 하나의 유해한 부작용이 약 10% 내지 약 100%, 약 20% 내지 약 100%, 약 30% 내지 약 100%, 약 40% 내지 약 100%, 약 50% 내지 약 100%, 약 60% 내지 약 100%, 약 70% 내지 약 100%, 약 80% 내지 약 100%, 약 10% 내지 약 90%, 약 20% 내지 약 90%, 약 30% 내지 약 90%, 약 40% 내지 약 90%, 약 50% 내지 약 90%, 약 60% 내지 약 90%, 약 70% 내지 약 90%, 약 10% 내지 약 80%, 약 20% 내지 약 80%, 약 30% 내지 약 80%, 약 40% 내지 약 80%, 약 50% 내지 약 80%, 또는 약 60% 내지 약 80%, 약 10% 내지 약 70%, 약 20% 내지 약 70%, 약 30% 내지 약 70%, 약 40% 내지 약 70%, 또는 약 50% 내지 약 70%까지 감소된다.

[0564] E. 면역억제

[0565] 본원에 개시된 방법 및 조성물의 임의의 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터 또는 rAAV 게놈이 투여되는 대상체는 또한 면역억제제를 투여받는다. 다양한 방법이 AAV가 투여되는 환자의 면역 반응의 면역억제를 초래하는 것으로 알려져 있다. 당업계에 알려진 방법은 프로테아좀 억제제와 같은 면역억제제를 환자에게 투여하는 단계를 포함한다. 둘 모두 본원에 참조로 포함된, 예를 들어, 미국 특허 번호 9,169,492 및 미국 특허 출원 번호 15/796,137에 개시된 바와 같은 당업계에 알려진 하나의 이러한 프로테아좀 억제제는 보르테조미이다. 또 다른 구현예에서, 면역억제제는, 예를 들어, 항체 생성 세포의 제거 또는 억제제를 통해 면역 반응을 억제할 수 있는 폴리클로날, 모노클로날, scfv 또는 기타 항체 유래 분자를 포함하는 항체일 수 있다. 추가 구현예에서, 면역억제 요소는 짧은 헤어핀 RNA(shRNA)일 수 있다. 이러한 구현예에서, shRNA의 암호화 영역은 rAAV 카세트에 포함되며 일반적으로 폴리-A 테일의 3' 다운스트림에 위치한다. shRNA는 사이토카인, 성장 인자 (전환 성장 인자 $\beta 1$ 및 $\beta 2$, TNF 및 기타 공개적으로 알려진 것 포함)와 같은 면역자극제의 발현을 감소 또는 제거하기 위해 표적화될 수 있다.

[0566] VI. GAA를 다른 리소좀 효소로 대체.

[0567] 일부 구현예에서, 본원에 개시된 임의의 방법 및 조성물에서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 리소좀 효소의 핵산 서열을 대체할 수 있다. 본원에 개시된 바와 같은 rAAV 벡터 또는 rAAV 게놈에 의해 발현되기에 적합한 리소좀 효소는 포유동물 리소좀에 축적된 물질을 감소시킬 수 있거나 하나 이상의 리소좀 축적 질환 증상을 구제 또는 개선할 수 있는 임의의 효소를 포함한다. 적합한 리소좀 효소는 야생형 또는 변형된 리소좀 효소를 둘 모두 포함하고 재조합 또는 합성 방법을 사용하여 생성하거나 천연 공급원으로부터 정제될 수 있다. 예시적인 리소좀 효소는 표 5A 또는 표 6A에 나열되어 있다.

[0568] [표 5A] 예시적인 리소좀 축적 질환(LSD) 및 관련 효소 결함

질환	효소 결함	물질 축적
A. 글리코겐증 장애		
폼페병	산- α 1,4-글리코시다아제	글리코겐 α 1-4 연결된 올리고사카라이드
B. 당지질증 장애		
GM1 강글리오시드증	β -갈락토시다제	GM1 강글리오사이드
테이-삭스병	β -헥소사미니다제 A	GM2 강글리오사이드
AB 변이체	단백질	
잔트호프병	β -헥소사미니다제 A & B	GM2 강글리오사이드
파브리병	α -갈락토시다제 A	글로보사이드
고셔병	글루코세레브로시다제	글루코실세라미
이색성 백질영양증	아릴설파타제 A	설페이트
크라베병	갈락토실세라미다제	갈락토세레브로사이드
니만-피크, A & B 형	산성 스팅고미엘리나제	스�팅고미엘린
니만-피크, D 형	알려져 있지 않음	스�팅고미엘린
파버병	산성 세라미다제	세라마이드
월만병	산성 리파제	클레스테릴 에스테르
C. 점액다당류 장애		
혈러 증후군 (MPS IH)	α -L-이두로니다제	헤파란 & 더마탄 설페이트
샤이 증후군 (MPS IS)	α -L-이두로니다제	헤파란 & 더마탄 설페이트
혈러 샤이에 (MPS IH/S)	α -L-이두로니다제	헤파란 & 더마탄 설페이트
헌터 증후군 (MPS II)	이두로네이트 설파타제	헤파란 & 더마탄 설페이트
산필리포 A (MPS IIIA)	헤파란 N-설파타제	헤파란 설페이트
산필리포 B (MPS IIIB)	α -N- 아세틸글루코사미니다제	헤파란 설페이트
산필리포 C (MPS IIIC)	아세틸-CoA- 글루코사미니드 아세틸트랜스퍼라제	헤파란 설페이트
산필리포 D (MPS IIID)	N-아세틸글루코사민-6- 설파타제	헤파란 설페이트
모르quio A (MPS IVA)	갈락토사민-6-설파타제	케라탄 설페이트
모르quio B (MPS IVB)	β -갈락토시다제	케라탄 설페이트
마르토-라미 (MPS VI)	아릴설파타제 B	더마탄 설페이트
슬라이 증후군 (MPS VII)	β -글루쿠로니다제	

[0569]

D. 올리고사카라이드/당단백질 장애		
α-만노시드증	α-만노시드증	만노오스/올리고사카라이드
β-만노시드증	β-만노시드증	만노오스/올리고사카라이드
푸코스축적증	α-L-푸코시다아제	푸코실 올리고사카라이드
아스파르틸글루코사민뇨증	N-아스파르틸-β-글루코사미니다제	아스파르틸글루코사민 아스파라긴
시알리도증 (점액지질증 I)	α-뉴라미니다제	시알릴올리고사카라이드
갈락토시알리도증 (골드버그 증후군)	리소좀 보호 단백질 결핍증	시알릴올리고사카라이드
선들러병	α-N-아세틸-갈락토사미니다제	
E. 리소좀 효소 수송 장애		
점액지질증 II (I-세포 질환)	N-아세틸글루코사민-1-포스포트랜스퍼라제	헤파란 설페이트
점액지질증 III (유사-홀러 다발이상증)	MLII와 동일	
F. 리소좀 막 수송 장애		
시스틴증	시스틴 수송 단백질	유리 시스틴
살라병	시알산 수송 단백질	유리 시알산 및 글루쿠론산
유아 시알산 축적 질환	시알산 수송 단백질	유리 시알산 및 글루쿠론산
G. 기타		
배튼병 (청소년 신경세포 세로이드 지방갈색소증)	알려져 있지 않음	리포푸신
유아 신경세포 세로이드 지방갈색소증	팔미토일-단백질 티오에스테라제	리포푸신
점액지질증 IV	알려져 있지 않음	강글리오사이드 & 히알루론산
프로사포신	사포신 A, B, C 또는 D	

[0570]

[0571]

본원에 개시된 조성물 및 방법의 일부 구현예에서, 하나의 특히 바람직한 리소좀 효소는 글루코세레브로시다제이며, 이는 현재 Genzyme에 의해 재조합적으로 생산 및 제조되고 고셔병(Gaucher's Disease)에 대한 효소 대체 요법에 사용된다. 현재, 재조합 효소는 대식세포 계통의 세포에 특이적으로 단백질을 표적으로 하는 노출된 만노오스 잔기로 제조된다. 제1형 고셔 환자에서 일차적 병리는 글루코세레브로시다제를 축적하는 대식세포로 인한 것이지만, 글루코세레브로시다제를 다른 세포 유형에 전달하는 것은 치료적 이점이 있을 수 있다. 본 발명을 사용하여 글루코세레브로시다제를 리소좀에 표적화하는 것은 체제를 여러 세포 유형에 표적화할 것이고 다른 체제에 비해 치료적 이점을 가질 수 있다. 본원에 개시된 조성물 및 방법의 일부 구현예에서, 본원에 개시된 방법에서 치료되는 리소좀 질환은 폼페가 아니다. 본원에 개시된 조성물 및 방법의 일부 구현예에서, 표적화 벡터 또는 rAAV 벡터 내의 핵산에 의해 암호화되는 리소좀 효소는 GAA가 아니다.

[0572]

본 발명의 방법 및 조성물은 임의의 치료제를 생산하고 아세포 구획으로 전달하는 데 유용하지만, 본 발명은 대사 질환을 치료하기 위한 유전자 생성물을 전달하는 데 특히 유용하다.

[0573]

일부 구현예에서, 리소좀 축적 질환(LSD)을 치료하기 위한 리소좀 효소는 **표 5A**에 나타내었다. 일부 구현예에서, 리소좀 효소는 **표 6A**에 나타낸 골지체 또는 ER 결함과 관련이 있다. 바람직한 구현예에서, 본원에 기재된 바와 같은 리소좀 효소를 암호화하는 바이러스 벡터는 동일한 리소좀 효소 유전자의 결함을 앓고 있는 환자에게 전달된다. 대안적인 구현예에서, 리소좀 효소 유전자의 기능적 서열 또는 중 변이체가 사용된다. 추가 구현예에서, 리소좀 효소 유전자 결함을 구제할 수 있는 상이한 효소를 암호화하는 유전자는 본 발명의 방법과 다르다.

[0574] [표 6A]

표 6A: 골지체 및 ER 질환		
질환명	유전자 및 효소 결함	특징
엘러스-단로스 증후군 VI형	PLOD1 리실 하이드록실라제	콜라겐의 리실 하이드록실화 결함; ER 루멘에 위치함
Ia형 글리코겐 축적 질환	글루코스 6 포스파타제	간, 콩팥, 및 장 점막에 글리코겐의 과도한 축적을 야기함; 효소는 막관통이지만 활성 부위는 ER 루멘임
선천성 글리코실화 장애		
CDG Ic	ALG6 α 1,3 글루코실트랜스퍼라제	N-글리코실화 ER 루멘의 결함
CDG Id	ALG3 α 1,3 만노실트랜스퍼라제	N-글리코실화 ER 막관통 단백질의 결함
CDG IIa	MGAT2 N-아세틸글루코사미닐-트랜스퍼라제 II	N-글리코실화 골지체 막관통 단백질의 결함
CDG IIb	GCS1 α 1,2-글리코시다아제 I	단백질분해에 의해 방출 가능한 루멘 측매 도메인과 결합된 N 글리코실화 ER 막의 결함.

[0575]

[0576] 본원에 개시된 방법 및 조성물의 일부 구현예에서, 표적화 벡터 또는 rAAV는 표 5B 또는 표 6B의 서열 중 어느 하나의 단백질을 발현한다.

[0577] [표 5B] 예시적인 리소좀 축적 질환(LSD) 및 표적화 벡터 또는 rAAV 벡터에 의해 발현될 단백질, 및 단백질을 암호화하는 핵산 서열

질환	효소 결함	표적화 벡터 또는 rAAV 에 의한 발현을 위한 핵산 서열	효소 결함을 암호화하는 단백질 서열
A. 글리코겐증 장애			
폼페병	산-a1,4-글리코시다아제 (GAA)(WT)	서열 번호 11>> wt 전장 hGAA (비코돈 최적화됨); NP_000143.2	서열 번호 10 (aa wt 전장 hGAA aa (비코돈 최적화됨); NM_000152.4), 또는 서열 번호 170, 171, 172, 173, 174
	산-a1,4-글리코시다아제 (hGAA)	서열 번호 72 또는 서열 번호 182>> >hGAA	
	산-a1,4-글리코시다아제 (hGAA 3X)	서열 번호 73 >hGAA_3X	
	hGAA hGAA_코돈_최적화된_No1	서열 번호 74 >hGAA_코돈_최적화된_No1	
	>hGAA_코돈_최적화된_No2	서열 번호 75 >hGAA_코돈_최적화된_No2	
	>hGAA_코돈_최적화된_No3	서열 번호 76 >hGAA_코돈_최적화된_No3	
B. 당지질증 장애			
GM1 강글리오시드증 (GLB1 결핍증)	β -갈락토시다아제 (GLB1)	서열 번호 227 GLB1 (NM_000404.4)	서열 번호 185 NP_000395.3
테이-삭스병	β -헥소사미니다제 A (HEXA)	서열 번호 228 NM_000520.6 (HEXA)	서열 번호 186 NP_000511.2 (HEXA)
GM2-강글리오시드증, AB 변이체	β -헥소사미니다제 A (HEXA) 및 HEXB	서열 번호 229 NM_000520.6 (HEXA)	서열 번호 186 NP_000511.2 (HEXA)

[0578]

잔트호프병	β -헥소사미니다제 A & β -헥소사미니다제 B (HEXA 및 HEXB)	서열 번호 229; NM_000520.6 (HEXA) 서열 번호 230; NM_000521.4 (HEXB)	서열 번호 186 NP_000511.2 (HEXA) 서열 번호 187; NP_000512.2(HEX B)
파브리병	α -갈락토시다제 A (GLA)	서열 번호 231 NM_000169.3 (GLA)	서열 번호 188; NP_000160.1 (GLA)
고셔병	글루코세레브로시다제 (GBA)	서열 번호 232 NM_000157.4 (GBA)	서열 번호 189; NP_000148.2 (GBA)
이색성 백질영양증	아릴설파타제 A (ARSA)	서열 번호 233 NM_000487.6(ARSA)	서열 번호 190
크라베병 (구상세포 백질영양증으로도 불림)	갈락토실세라미다제 (GALC)	서열 번호 234	서열 번호 191; NP_000144.2
니만-피크, A & B 형	산성 스팅고미엘리나제 (SMPD1)	서열 번호 235 NM_000543.5 (SMPD1)	서열 번호 192; NP_000534.3 (SMPD1)
니만-피크, C1 형	NPC 세포내 콜레스테롤 수송체 1 (NPC1)	서열 번호 236 NM_000271.5 (NPC1)	서열 번호 193; NP_000262.2 (NPC1)
니만-피크, C2 형	NPC 세포내 콜레스테롤 수송체 2 (NPC2)	서열 번호 237 NM_006432.5 (NPC2)	서열 번호 194; NP_006423.1 (NPC2)
파버병 (파버 지방육아증증)	산성 세라미다제 (ASAH1)(N-아실스핑고신 아미도하이드롤라제로도 알려져 있음)	서열 번호 238 NM_004315.6 (ASAH1)	서열 번호 195; NP_004306.3 (ASAH1)
월만병 (리소좀 산성 리파제 결핍증으로도 알려져 있음)	리소좀 산성 리파제 (LIPA) (리파제 A 로도 알려져 있음)	서열 번호 239 NM_000235.4 (LIPA)	서열 번호 196; NP_000226.2 (LIPA)
C. 점액다당류 장애			
점액다당증 I 형 (MPS I) (3 가지 MPS I 형 포함: 혈리 증후군(MPS IH);	α -L-이두로니다제 (IDUA)	서열 번호 240; NM_000203.5 (IDUA)	서열 번호 197; NP_000194.2 (IDUA)

[0579]

샤이 증후군 (MPS IS) 혈리 샤이에 (MPS IH/S)			
점액다당증 유형 II (MPS II), (헌터 증후군으로도 알려져 있음)	이두로네이트 설파타제 (IDS)	서열 번호 241; NM_000202.8 (IDS)	서열 번호 198; NP_000193.1 (IDS)
산필리포 A (MPS IIIA)	헤파란 N-설파타제 (N-설포글루코사민 설포하이드롤라제로도 지칭됨) (SGSH)	서열 번호 242; NM_000199.5 (SGSH)	서열 번호 199; NP_000190.1 (SGSH)
산필리포 B (MPS IIIB)	α -N-아세틸글루코사미니다제 (NAGLU)	서열 번호 243; NM_000263.4 (NAGLU)	서열 번호 200 NP_000254.2 (NAGLU)
산필리포 C (MPS IIIC)	아세틸-CoA-글루코사미니다아세틸트랜스퍼라제 (헤파란-알파-글루코사미니다 N-아세틸트랜스퍼라제로도 지칭됨) (중중 줄여서 N-아세틸트랜스퍼라제) HGSNAT)	서열 번호 244; NM_152419.3 (HGSNAT)	서열 번호 201; NP_689632.2 (HGSNAT)
산필리포 D (MPS IIID)	N-아세틸글루코사민-6-설파타제 (GNS)(글루코사민 (N-아세틸)-6-설파타제로도 지칭됨)	서열 번호 245; NM_002076.4 (GNS)	서열 번호 202; NP_002067.1 (GNS)
점액다당증 유형 IVA (MPS IVA) (모르퀴오 A 증후군으로도 알려져 있음)	갈락토사민-6-설파타제 (GALNS)	서열 번호 246; NM_000512.5 (GALNS)	서열 번호 203; NP_000503.1 (GALNS)
점액다당증 유형 IVB (MPS IVB), (모르퀴오 B 증후군으로도 알려져 있음)	β -갈락토시다제 (GLB1)	서열 번호 227; GLB1 (NM_000404.4)	서열 번호 185 NP_000395.3

[0580]

점액다당증 유형 VI (MPS VI), (마르토-라미 증후군으로도 알려져 있음)	아릴설파타제 B (ARSB)	서열 번호 247; NM_000046.5 (ARSB)	서열 번호 204 NP_000037.2 (ARSB)
점액다당증 유형 VII (MPS VII), (슬라이 증후군으로도 알려져 있음)	β -글루루로니다제 (GUSB)	서열 번호 248 NM_000181.4 (GUSB)	서열 번호 205 NP_000172.2 (GUSB)
<i>D. 올리고사카라이드/당단백질 장애</i>			
α -만노시드증	α -만노시드증 (MAN2B1)	서열 번호 249 NM_000528.4 (MAN2B1)	서열 번호 206 NP_000519.2 (MAN2B1)
β -만노시드증	β -만노시드증 (MANBA)	서열 번호 250 NM_005908.4 (MANBA)	서열 번호 207 NP_005899.3 (MANBA)
푸코스축적증	α -L-푸코시다아제 (FUCA1)	서열 번호 251 NM_000147.4 (FUCA1)	서열 번호 208 NP_000138.2 (FUCA1)
아스파르틸글루코사민 뇨증	N-아스파르틸- β -글루코사미니다제 (AGA)(아스파르틸글루코사미니다제 (ASRG)로도 알려져 있음 또는 N(4)-(베타-N-아세틸글루코사미닐)-L-아스파라기나제 또는 글리코실라스파라기나제 (AGU)	서열 번호 252 NM_000027.4 (AGA)	서열 번호 209 NP_000018.2 (AGA)
시알리도증 (점액지질증 I)	뉴라미니다제 1 (α -뉴라미니다제로도 지칭됨) (NEU1)	서열 번호 253 NM_000434.4 (NEU1)	서열 번호 210 NP_000425.1 (NEU1)
갈락토시알리도증 (베타-갈락토시다제 결핍증을 동반한 골드버그 증후군 또는 리소좀 보호 단백질 결핍증 또는 PPCA	카텅신 A (CTSA)(보호 단백질/카텅신 A 또는 PPCA 또는 PPGB, 또는 GSL 로도 알려져 있음)	서열 번호 254 NM_000308.4 (CTSA)	서열 번호 211 NP_000299.3 (CTSA)

[0581]

결핍증 또는 뉴라미니다제 결핍증으로도 알려져 있음)			
신들러병 (NAGA 결핍증 또는 알파-갈락토시다제 B 결핍증으로도 지칭됨)	α -N-아세틸-갈락토사미니다제 (NAGA)(알파-갈락토시다제 B 또는 GALB 로도 지칭됨)	서열 번호 255 NM_000262.3 (NAGA)	서열 번호 212 NP_000253.1 (NAGA)
E. 리소좀 효소 수송 장애			
점액지질증 II (I-세포 질환)	N-아세틸글루코사민-1-포스포트랜스퍼라제 (GNPTAB) (GlcNAc-1-포스포트랜스퍼라제로도 지칭됨)	서열 번호 256 NM_024312.5 (GNPTAB)	서열 번호 213 NP_077288.2 (GNPTAB)
점액지질증 III (유사 홀러 다발이상증)	MLII와 동일	서열 번호 256	서열 번호 213
F. 리소좀 막 수송 장애			
시스틴증	시스티노신 (리소좀 시스틴 수송체로도 지칭됨) (CTNS)	서열 번호 257 NM_004937.3 (CTNS)	서열 번호 214 NP_004928.2 (CTNS)
시알산 축적 질환 (살라병은 덜 심각한 형태)	용질 운반체 패밀리 17 구성원 5 (SLC17A5) (시알린 또는 시알산 수송 단백질 또는 SAISD, SD, ISSD, NSD 로도 알려져 있음)	서열 번호 258 NM_012434.5 (SLC17A5)	서열 번호 2015 NP_036566.1 (SLC17A5)
유아 시알산 축적 질환 (ISSD)	용질 운반체 패밀리 17 구성원 5 (SLC17A5) (시알린 또는 시알산 수송 단백질 또는 SAISD, SD, ISSD, NSD 로도 알려져 있음)	서열 번호 1258	서열 번호 215
G. 기타			
CLN3 질환	바테닌 (CLN3 리소좀/엔도좀 막관통 단백질 또는	서열 번호 259 NM_000086.2 (CLN3)	서열 번호 216 NP_000077.1 (CLN3)

(배튼병 (청소년 신경세포 세로이드 지방갈색소증) 포함)	바테닌으로도 지칭됨) (CLN3)		
유아 신경세포 세로이드 지방갈색소증 (또는 유아 배튼병) (CLN1 질환)	팔미토일-단백질 티오에스테라제 1(PPT1 유전자)	서열 번호 260 NM_000310.3 (PPT1)	서열 번호 217 NP_000301.1 (PPT1)
점액지질증 IV	뮤코리핀-1 (MCOLN1)	서열 번호 261 NM_020533.3 (MCOLN1)	서열 번호 218 NP_065394.1(MCOLN1)
프로사포신 결핍증 (이색성 백질영양증 또는 PSAP 돌연변이와 관련됨)	프로사포신 (PSAP). PSAP 단백질은 사포신 A, B, C, D 로도 불리는 4 개의 더 작은 단백질의 전구체임)	서열 번호 262 PSAP (NM_002778.4)	서열 번호 219 NP_002769.1 (PSAP)

[표 6B] 예시적인 리소좀 축적 질환(LSD) 및 표적화 벡터 또는 rAAV 벡터에 의해 발현될 단백질, 및 단백질을 암호화하는 핵산 서열.

질환	효소 결함	핵산 서열	단백질 서열
엘리스-단로스 증후군 유형 VI	PLOD1 리실 하이드록실라제 (PLOD1) (프로콜라겐-리신, 2-옥소글루타레이트 5-다옥시게나제 1)	서열 번호 263 NM_000302.4	서열 번호 220 NP_000293.2
유형 Ia 글리코겐 축적 질환 (GSDIa)	글루코스 6 포스파타제 촉매 서브유닛 (G6PC)	서열 번호 264 NM_000151.4 (G6PC)	서열 번호 221 NP_000142.2 (G6PC)
유형 Ib 글리코겐 축적 질환 (GSDIb)	용질 운반체 패밀리 37 구성원 4 (글루코스 6-포스페이트트랜스로카제 단백질 또는 G6PT1 또는 GSD1b 로도 알려져 있음) (SLC37A4)	서열 번호 265 NM_001467.6 (SLC37A4)	서열 번호 222 NP_001458.1 (SLC37A4)
글리코실화의 선천성 장애			
ALG6-선천성 글리코실화 장애 (ALG6-CDG) (글리코실화 Ic 형 또는 CDG Ic 의 선천성 장애로도 알려져 있음)	ALG6 α 1,3 글루코실트랜스퍼라제 (ALG6)	서열 번호 266 NM_013339.4 (ALG6)	서열 번호 223 NP_037471.2 (ALG6)
비정상적인 N-글리코실화를 특징으로 하는 글리코실화 Id 형 (CDG-Id)의 선천성 장애	ALG3 α 1,3 만노실트랜스퍼라제 (ALG3)	서열 번호 267 NM_005787.6 (ALG3)	서열 번호 224 NP_005778.1 (ALG3)
글리코실화 2A 의 선천성 장애 (CDG2A 또는 CDG IIa)	MGAT2 (N-아세틸글루코사미닐-트랜스퍼라제 II (알파-1,6-만노실-당단백질 2-베타-N-아세틸글루코사미닐트랜스퍼라제 또는 GNT-	서열 번호 268 NM_002408.4 (MGAT2)	서열 번호 225 NP_002399.1 (MGAT2)

	II 로도 지칭됨)을 암호화함)		
IIb 형 선천성 글리코실화 장애 (CDGIIb)	GCS1 α 1,2-글리코시다아제 I (GDG2B 또는 GCS1) (만노실-올리고사카라이드 글리코시다아제 또는 MOGS 로도 알려져 있음)	서열 번호 269 NM_006302.3 (MOGS)	서열 번호 226 NP_006293.2 (MOGS)

VII. 투여

일부 구현예에서, 본원에 기재된 기술은 본원에 개시된 바와 같은 rAAV 벡터 또는 rAAV 계통을 리소좀 질환 또는 장애가 있는 대상체, 예를 들어, 폼페병이 있는 대상체에게 투여하기 위한 방법 및 조성물에 관한 것이다. 일부 구현예에서, 상기 방법은 rAAV 벡터를 포함하는 rAAV 벡터의 균질한 집단을 포함하는 조성물을 투여하는 단계를 포함하고, 여기서, rAAV 벡터는 AAV3 또는 AAV8 벡터, 또는 본원에 개시된 바와 같은 AAV3 또는 AAV8로부터의 적어도 하나의 캡시드 단백질을 포함하는 반수체 rAAV 벡터이다. 일부 구현예에서, 대상체는 본원에 개시된 바와 같은 상이한 rAAV 벡터의 콕테일, 예를 들어, 간 세포를 표적화하거나 형질도입하는 rAAV 벡터 및 근육 세포를 표적화하거나 형질도입하는 rAAV 벡터를 포함하는 조성물을 투여받는다. 예시적인 구현예에서, 대상

체는 본원에 개시된 바와 같은 2개 이상의 상이한 rAAV 벡터의 콕테일을 포함하는 조성물, 예를 들어, 본원에 개시된 바와 같은 AAV3 벡터 및 AAV8 벡터를 포함하는 조성물을 투여받으며, 여기서, 각각의 AAV 벡터는 본원에 개시된 바와 같은 LSP에 작동적으로 연결된 GAA 폴리펩타이드를 암호화하는 핵산을 포함한다.

[0589] 일부 구현예에서, 대상체는, 예를 들어, 동시에 또는 후속적으로 본원에 개시된 바와 같은 2개 이상의 상이한 rAAV 벡터를 공동-투여받는다. 예시적인 목적으로만, 대상체는 간 세포를 표적으로 하거나 형질도입하는 rAAV 벡터를 포함하는 조성물을 투여받을 수 있으며, 여기서, 대상체는 또한 근육 세포를 표적으로 하거나 형질도입하는 rAAV 벡터를 포함하는 조성물을 공동-투여받는다. 예시적인 구현예에서, 대상체는 본원에 개시된 바와 같은 rAAV 벡터를 포함하는 조성물, 예를 들어, 본원에 개시된 바와 같은 LSP에 작동적으로 연결된 GAA 폴리펩타이드를 암호화하는 핵산을 포함하는, 본원에 개시된 바와 같은 AAV3 벡터 또는 이의 변이체를 포함하는 조성물(여기서, 대상체는 또한 본원에 개시된 바와 같은 상이한 rAAV 벡터를 포함하는 조성물을 공동-투여받는다), 예를 들어, 대안적으로, 예를 들어, 본원에 개시된 바와 같은 LSP에 작동적으로 연결된 GAA 폴리펩타이드를 암호화하는 핵산을 포함하는 AAV8 벡터 또는 반수체 AAV 벡터를 포함하는 조성물(여기서, LSP는 상이한 프로모터, 예를 들어, 근육 특이적 프로모터로 대체된다)을 공동-투여받는다.

[0590] 대상체에게 투여될 본원에 개시된 바와 같은 rAAV 벡터 또는 rAAV 게놈의 투여량은 투여 모드, 치료 및/또는 예방할 품폐병 또는 기타 병태의 중증도, 개별 대상체의 병태, 특정 바이러스 벡터 또는 캡시드, 및 전달될 핵산 등에 의존하며, 일상적인 방식으로 결정될 수 있다. 치료 효과를 달성하기 위한 예시적인 용량은 적어도 약 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} 형질도입 단위, 임의로 약 10^8 내지 약 10^{13} 형질도입 단위의 역가이다.

[0591] 추가 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터 또는 rAAV 게놈을 대상체에게 투여하면 2시간, 3시간, 4시간, 5시간, 6시간, 7시간, 8시간, 9시간, 10시간, 11시간, 12시간, 13시간, 14시간, 15시간, 16시간, 17시간, 18시간, 19시간, 20시간, 21시간, 22시간, 23시간, 1일, 2일, 3일, 4일, 5일, 6일, 7일, 1주, 2주, 3주, 4주, 1개월, 2개월, 3개월, 4개월 이상의 순환 반감기를 갖는 GAA 단백질이 생성된다.

[0592] 일 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터 또는 rAAV 게놈을 대상체에게 투여하는 기간은 1일, 2일, 3일, 4일, 5일, 6일, 7일, 8일, 9일, 10일, 11일, 12일, 13일, 14일, 3주, 4주, 5주, 6주, 7주, 8주, 9주, 10주, 11주, 12주, 4개월, 5개월, 6개월, 7개월, 8개월, 9개월, 10개월, 11개월, 12개월 이상이다. 추가 구현예에서, 투여가 중단되는 기간은 1일, 2일, 3일, 4일, 5일, 6일, 7일, 8일, 9일, 10일, 11일, 12일, 13일, 14일, 3주, 4주, 5주, 6주, 7주, 8주, 9주, 10주, 11주, 12주, 4개월, 5개월, 6개월, 7개월, 8개월, 9개월, 10개월, 11개월, 12개월 이상이다.

[0593] 또 다른 구현예에서, 품폐병의 치료를 위해 본원에 개시된 바와 같은 rAAV 벡터 또는 rAAV 게놈의 투여는, 예를 들어, 적어도 0.5 파운드, 적어도 1 파운드, 적어도 1.5 파운드, 적어도 2 파운드, 적어도 2.5 파운드, 적어도 3 파운드, 적어도 3.5 파운드, 적어도 4 파운드, 적어도 4.5 파운드, 적어도 5 파운드, 적어도 5.5 파운드, 적어도 6 파운드, 적어도 6.5 파운드, 적어도 7 파운드, 적어도 7.5 파운드, 적어도 8 파운드, 적어도 8.5 파운드, 적어도 9 파운드, 적어도 9.5 파운드, 적어도 10 파운드, 적어도 10.5 파운드, 적어도 11 파운드, 적어도 11.5 파운드, 적어도 12 파운드, 적어도 12.5 파운드, 적어도 13 파운드, 적어도 13.5 파운드, 적어도 14 파운드, 적어도 14.5 파운드, 적어도 15 파운드, 적어도 20 파운드, 적어도 25 파운드, 적어도 30 파운드, 적어도 50 파운드의 체중 증가를 초래한다.

[0594] 또 다른 구현예에서, 품폐병의 치료를 위해 본원에 개시된 바와 같은 임의의 혈청형의 AAV GAA는, 예를 들어, 0.5 파운드 내지 50 파운드, 0.5 파운드 내지 30 파운드, 0.5 파운드 내지 25 파운드, 0.5 파운드 내지 20 파운드, 0.5 파운드 내지 15 파운드, 0.5 파운드 내지 10 파운드, 0.5 파운드 내지 7.5 파운드, 0.5 파운드 내지 5 파운드, 1 파운드 내지 15 파운드, 1 파운드 내지 10 파운드, 1 파운드 내지 7.5 파운드, 1 파운드 내지 5 파운드, 2 파운드 내지 10 파운드, 2 파운드 내지 7.5 파운드의 체중 증가를 초래한다.

[0595] A. 약제학적 조성물

[0596] 본원에 개시된 바와 같은 투여 방법에 사용하기 위한 본원에 개시된 바와 같은 rAAV 벡터는 약제학적으로 허용되는 부형제, 즉 하나 이상의 약제학적으로 허용되는 담체 물질 및/또는 첨가제, 예를 들어 완충제, 담체, 부형제, 안정화제 등과 함께 약제학적 조성물로 제형화될 수 있다. 약제학적 조성물은 키트의 형태로 제공될 수 있다. 본원에 개시된 바와 같은 투여 방법 및 이의 용도에 사용하기 위한 본원에 개시된 바와 같은 rAAV 벡터를 포함하는 약제학적 조성물은 당업계에 알려져 있다.

- [0597] 따라서, 본 발명의 추가 양태는 본원에 개시된 바와 같은 투여 방법에 사용하기 위한 본원에 개시된 바와 같은 rAAV 벡터를 포함하는 약제학적 조성물을 제공한다. 활성 성분(예를 들어, 본원에 개시된 바와 같은 rAAV 벡터), 약제학적으로 허용되는 부형제, 및/또는 본 개시내용에 따른 약제학적 조성물 내의 임의의 추가 성분의 상대적인 양은 치료되는 대상체의 실체, 크기 및/또는 병태에 따라, 그리고 조성물이 투여되는 경로에 따라 추가로 달라질 수 있다. 예를 들어, 조성물은 0.1% (w/w)와 99% (w/w) 사이의 활성 성분을 포함할 수 있다. 예로서, 조성물은 0.1%와 100% 사이, 예를 들어, 5%와 50% 사이, 1%와 30% 사이, 5%와 80% 사이, 적어도 80% (w/w) 활성 성분을 포함할 수 있다.
- [0598] 약제학적 조성물은 (1) 안정성을 증가시키고; (2) 세포 형질감염 또는 형질도입을 증가시키기 위해; (3) 페이로드의 서방출 또는 지연 방출을 허용하기 위해; (4) 생체분포를 변경하기 위해(예를 들어, 바이러스 입자를 특정 조직 또는 세포 유형에 표적화하기 위해); (5) 암호화된 단백질의 번역을 증가시키고; (6) 암호화된 단백질의 방출 프로파일을 변경하고/하거나 (7) 본 발명의 페이로드의 조절 가능한 발현을 허용하기 위해 하나 이상의 부형제 또는 희석제를 사용하여 제형화될 수 있다. 일부 구현예에서, 약제학적으로 허용되는 부형제는 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 순수할 수 있다. 일부 구현예에서, 부형제는 인간 및 수의학적 사용을 위해 승인된다. 일부 구현예에서, 부형제는 미국 식품의약국의 승인을 받을 수 있다. 일부 구현예에서, 부형제는 제약 등급일 수 있다. 일부 구현예에서, 부형제는 미국 약전(USP), 유럽 약전(EP), 영국 약전 및/또는 국제 약전의 표준을 충족할 수 있다. 본원에 사용된 바와 같은 부형제는 원하는 특정 투여 형태에 적합한, 임의의 모든 용매, 분산 매질, 희석제, 또는 기타 액체 비히클, 분산 또는 현탁 보조제, 표면 활성제, 등장화제, 증점제 또는 유화제, 보존제 등을 포함하지만 이에 제한되지는 않는다. 약제학적 조성물을 제형화하기 위한 다양한 부형제 및 조성물을 제조하기 위한 기술이 당업계에 알려져 있다(그 전체가 본원에 참조로 포함된 문헌(Remington: The Science and Practice of Pharmacy, 21 st Edition, A. R. Gennaro, Lippincott, Williams and Wilkins, Baltimore, MD, 2006) 참조). 임의의 통상적인 부형제 매질이 예컨대 임의의 바람직하지 않은 생물학적 효과를 생성하거나 그렇지 않으면 약제학적 조성물의 임의의 다른 성분(들)과 유해한 방식으로 상호작용함으로써 물질 또는 이의 유도체와 상용성일 수 없는 한을 제외하고는, 통상적인 부형제 매질의 사용은 본 개시내용의 범위 내에서 고려될 수 있다.
- [0599] 본원에 개시된 바와 같은 투여 방법에 사용하기 위한 본원에 개시된 바와 같은 rAAV 벡터는 하나 이상의 다른 치료제, 예방제, 연구 또는 진단제와 조합하여 사용될 수 있다. 이러한 전달 방법이 본 발명의 범위 내에 있더라도, "와 조합하여"란 제제가 동시에 투여되고/되거나 함께 전달을 위해 제형화되어야 한다는 것을 암시하려는 의도는 아니다. 조성물은 하나 이상의 다른 원하는 치료제 또는 의료 절차와 동시에, 그 이전에, 또는 그 이후에 투여될 수 있다. 일부 구현예에서, 하나의 치료(예를 들어, 유전자 요법 벡터)의 전달은 두 번째(예를 들어, 하나 이상의 치료제)의 전달이 시작될 때 여전히 발생하므로 투여 측면에서 중복이 있다. 이것은 때때로 "동시" 또는 "동시 전달"로서 본원에서 언급된다. 다른 구현예에서, 한 치료의 전달은 다른 치료의 전달이 시작되기 전에 종료된다. 두 경우 중 일부 구현예에서, 치료는 조합 투여로 인해 더 효과적이다. 예를 들어, 2차 치료가 더 효과적이거나, 예를 들어, 2차 치료보다 적은 양으로 동등한 효과를 볼 수 있거나, 2차 치료가 1차 치료 없이 2차 치료가 투여된 경우에 나타나는 것보다 더 큰 정도로 증상을 감소시키거나, 1차 치료에서도 유사한 상황이 나타난다. 일부 구현예에서, 전달은 증상 또는 장애와 관련된 다른 파라미터의 감소가 다른 치료가 없을 때 전달된 한 치료로 관찰되는 것보다 더 큰 것이다. 두 치료 효과는 부분적으로 가산적이거나 완전히 가산적이거나 가산적인 것보다 클 수 있다. 전달은 전달된 1차 치료의 효과가 2차 치료가 전달될 때 여전히 감지 가능하도록 할 수 있다. 본원에 기재된 조성물 및 적어도 하나의 추가 치료제는 동일하거나 별개의 조성물로 동시에 또는 순차적으로 투여될 수 있다. 순차적 투여의 경우, 본원에 기재된 유전자 요법 벡터가 먼저 투여될 수 있고, 하나 초과의 치료제가 두 번째로 투여될 수 있거나, 투여 순서가 역전될 수 있다. 유전자 요법 벡터 및 하나 이상의 치료제는 활동성 장애 기간 동안, 또는 관해 또는 덜 활동적인 질환 기간 동안 투여될 수 있다. 유전자 요법 벡터는 또 다른 치료 전, 치료와 동시에, 치료 후에, 또는 장애의 관해 동안 투여될 수 있다.
- [0600] 조합하여 투여되는 경우, 본원에 개시된 바와 같은 투여 방법에 사용하기 위한 본원에 개시된 바와 같은 rAAV 벡터 및 하나 이상의 치료제(예를 들어, 제2 또는 제3 치료제), 또는 모두는 개별적으로, 예를 들어, 단일요법으로서 사용되는 각각의 양 또는 투여량보다 높거나, 더 낮거나, 동일한 양 또는 투여량으로 투여될 수 있다. 특정 구현예에서, 본원에 개시된 바와 같은 투여 방법에 사용하기 위한 본원에 개시된 바와 같은 rAAV 벡터의 투여된 양 또는 투여량 및 하나 이상의 치료제(예를 들어, 제2 또는 제3 제제) 또는 모두는 개별적으로 사용된 각각의 양 또는 투여량보다 더 낮다(예를 들어, 적어도 20%, 적어도 30%, 적어도 40%, 또는 적어도 50%). 다른 구현예에서, 원하는 효과(예를 들어, 심혈관 질환 또는 심장 질환의 치료)를 초래하는 본원에 개시된 바와 같은

투여 방법에 사용하기 위한 본원에 개시된 바와 같은 rAAV 벡터의 양 또는 투여량 및 하나 이상의 치료제(예를 들어, 제2 또는 제3 제제) 또는 모두는 동일한 치료 효과를 달성하기 위해 개별적으로 필요한 각각의 양 또는 투여량보다 더 낮다(예를 들어, 적어도 20%, 적어도 30%, 적어도 40%, 또는 적어도 50% 더 낮음).

[0601] 일부 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터의 투여 방법은 본원에 개시된 rAAV 벡터를 단독으로, 또는 추가 제제, 예를 들어, 본원에 개시된 바와 같은 면역 조절제와 조합하여 전달할 수 있다.

[0602] **B. 면역 조절제:**

[0603] 일부 구현예에서, 리소좀 질환, 예를 들어, 폼페를 치료하기 위해 본원에 기재된 바와 같은 AAV 벡터 및 AAV 케놈을 사용하는 방법 및 조성물은 면역 조절제를 투여하는 것을 추가로 포함한다. 일부 구현예에서, 면역 조절제는 rAAV 벡터 투여 시, rAAV 벡터 투여 전에, 또는 rAAV 벡터 투여 후에 투여될 수 있다.

[0604] 일부 구현예에서, 면역 조절제는 IdeS, IdeZ, IdeS/Z, Endo S, 또는 이들의 기능적 변이체와 같은 면역글로불린 분해 효소이다. 이러한 면역글로불린 분해 효소 및 이의 용도에 대한 참조의 비제한적인 예는, 이들 각각은 그 전체가 참조로 포함된 US 7,666,582, US 8,133,483, US 20180037962, US 20180023070, US 20170209550, US 8,889,128, W02010/057626, US 9,707,279, US 8,323,908, US 20190345533, US 20190262434, 및 W02020/016318 에 기재되어 있다.

[0605] 일부 구현예에서, 면역 조절제는 프로테아좀 억제제이다. 특정 양태에서, 프로테아좀 억제제는 보르테조미드이다. 구현예의 일부 양태에서, 면역 조절제는 보르테조미드 및 항 CD20 항체인 리툽시맙을 포함한다. 구현예의 다른 양태에서, 면역 조절제는 보르테조미드, 리툽시맙, 메토틱세이트, 및 정맥내 감마 글로불린을 포함한다. 이들 각각은 그 전체가 참조로 포함된 US 10,028,993, US 9,592,247, 및 US 8,809,282에 기재된 이러한 참고문헌의 비제한적인 예는 프로테아좀 억제제 및 리툽시맙, 메토틱세이트 및 정맥내 감마 글로불린과의 조합을 개시한다.

[0606] 대안적인 구현예에서, 면역 조절제는 NF-κB 경로의 억제제이다. 구현예의 특정 양태에서, 면역 조절제는 라파마이신 또는 기능적 변이체이다. US 10,071,114, US 20160067228, US 20160074531, US 20160074532, US 20190076458, US 10,046,064에 기재된 라파마이신 및 이의 용도를 개시하는 참조 문헌의 비제한적인 예는 그 전체가 포함된다. 구현예의 다른 양태에서, 면역 조절제는 면역억제제를 포함하는 합성 나노캐리어이다. 면역억제제, 합성 나노캐리어에 결합된 면역억제제, 라파마이신을 포함하는 합성 나노캐리어, 및/또는 톨로레제닉 합성 나노캐리어, 이들의 용량, 투여 및 용도의 참조에 대한 비제한적인 예는, 이들 각각은 그 전체가 참조로 포함된 US20150320728, US 20180193482, US 20190142974, US 20150328333, US20160243253, US 10,039,822, US 20190076522, US 20160022650, US 10,441,651, US 10,420,835, US 20150320870, US 2014035636, US 10,434,088, US 10,335,395, US 20200069659, US 10,357,483, US 20140335186, US 10,668,053, US 10,357,482, US 20160128986, US 20160128987, US 20200038462, US 20200038463에 기재되어 있다.

[0607] 일부 구현예에서, 면역 조절제는 이들 각각은 그 전체가 본원에 포함된 US20200038463, 미국 특허 9,006,254에 개시된 바와 같은 라파마이신(ImmTORTM 나노입자) (Kishimoto, et al., 2016, Nat Nanotechnol, 11(10): 890-899; Maldonado, et al., 2015, PNAS, 112(2): E156-165)을 포함하는 합성 나노캐리어이다. 일부 구현예에서, 면역 조절제는 조작된 세포, 예를 들어, 그 전체가 본원에 참조로 포함된 W02017192786에 개시된 바와 같은 SQZ 기술을 사용하여 변형된 면역 세포이다.

[0608] 일부 구현예에서, 면역 조절제는 폴리-ICLC, 1018 ISS, 알루미늄 염, Amplivax, AS15, BCG, CP-870,893, CpG7909, CyaA, dSLIM, GM-CSF, IC30, IC31, Imiquimod, ImuFact IMP321, IS Patch, ISS, ISCOMATRIX, JuvImmune, LipoVac, MF59, 모노포스포릴 지질 A, Montanide IMS 1312, Montanide ISA 206, Montanide ISA 50V, Montanide ISA-51, OK-432, OM-174, OM-197-MP-EC, ONTAK, PEPTEL, 벡터 시스템, PLGA 마이크로입자, 레시퀴모드, SRL172, Virosomes 및 기타 바이러스 유사 입자, YF-17D, VEGF 트랩, R848, 베타-글루칸, Pam3Cys, 및 Aquila의 QS21 자극제로 이루어진 군으로부터 선택된다. 또 다른 추가 구현예에서, 면역조절제 또는 애쥬번트는 폴리-ICLC이다

[0609] 일부 구현예에서, 면역 조절제는 클로로퀸(TLR 신호전달 억제제) 및 2-아미노퓨린(PKR 억제제)과 같은 세포에서 선천성 면역 반응을 억제하는 소분자이며, 또한 본원에 개시된 바와 같은 적어도 하나의 rAAV를 포함하는 조성물과 조합하여 투여될 수 있다. 상업적으로 입수 가능한 TLR-신호전달 억제제의 일부 비제한적인 예는 BX795, 클로로퀸, CLI-095, OxPAPC, 폴리믹신 B, 및 라파마이신(모두 INVIVOGENTM에서 입수 가능함)을 포함한다. 또한, 2-아미노퓨린, BX795, 클로로퀸 및 H-89와 같은 패턴 인식 수용체(PRR) (선천성 면역 신호전달에 관여)의

억제제는 또한 본원에 개시된 바와 같은 생체내 단백질 발현을 위해 본원에 개시된 바와 같은 적어도 하나의 rAAV 벡터를 포함하는 조성물 및 방법에서 사용될 수 있다.

- [0610] 일부 구현예에서, rAAV 벡터는 또한 NLRX1과 같은 선천성 면역의 음성 조절제를 암호화할 수 있다 따라서, 일부 구현예에서, rAAV 벡터는 또한 NLRX1, NS1, NS3/4A, 또는 A46R 중 하나 이상, 또는 임의의 조합을 임의로 암호화할 수 있다. 추가로, 일부 구현예에서, 본원에 개시된 바와 같은 적어도 하나의 rAAV 벡터를 포함하는 조성물은 또한 조직 또는 대상체에 의해 생성된 선천성 면역 반응을 피하기 위해 선천성 면역계의 억제제를 암호화하는 합성, 변형된-RNA를 포함할 수 있다.
- [0611] 일부 구현예에서, 본원에 개시된 바와 같은 투여 방법에 사용하기 위한 면역 조절제는 면역억제제이다. 본원에 사용된 바와 같이, "면역억제 약물 또는 제제"라는 용어는 정상적인 면역 기능을 억제하거나 방해하는 약제학적 제제를 포함하도록 의도된다. 본원에 개시된 방법에 적합한 면역억제제의 예는 미국 특허 공개 번호 2002/0182211에 개시된 바와 같은 CTLA4 및 B7 경로를 통한 T-세포와 B-세포의 커플링을 방해하는 제제와 같은 T-세포/B-세포 공동자극 경로를 억제하는 제제를 포함한다. 하나의 구현예에서, 면역억제제는 사이클로스포린 A이다. 다른 예는 미오페닐레이트 모페틸, 라파마이신, 및 항홍선 글로불린을 포함한다. 하나의 구현예에서, 면역억제 약물은 본원에 개시된 바와 같은 적어도 하나의 rAAV 벡터를 포함하는 조성물로 투여되거나, 또는 별도의 조성물로 투여될 수 있지만, 본원에 개시된 바와 같은 투여 방법에 따라 적어도 하나의 rAAV 벡터를 포함하는 조성물의 투여와 동시에, 그 전에 또는 그 후에 투여될 수 있다. 면역억제 약물은 투여 경로에 적합한 제형으로 투여되고 원하는 치료 효과를 달성하기에 충분한 용량으로 대상체에게 투여된다. 일부 구현예에서, 면역억제 약물은 본원에 개시된 바와 같은 rAAV 벡터에 대한 내성을 유도하기에 충분한 시간 동안 일시적으로 투여된다.
- [0612] 본원에 개시된 방법 및 조성물의 임의의 구현예에서, 본원에 개시된 바와 같은 rAAV 벡터 또는 rAAV 게놈이 투여되는 대상체는 또한 면역억제제를 투여받는다. 다양한 방법이 AAV가 투여되는 환자의 면역 반응의 면역억제를 초래하는 것으로 알려져 있다. 당업계에 알려진 방법은 프로테아좀 억제제와 같은 면역억제제를 환자에게 투여하는 단계를 포함한다. 둘 모두 본원에 참조로 포함된, 예를 들어, 미국 특허 번호 9,169,492 및 미국 특허 출원 번호 15/796,137에 개시된 바와 같은 당업계에 알려진 하나의 이러한 프로테아좀 억제제는 보르테조미이다. 일부 구현예에서, 면역억제제는, 예를 들어, 항체 생성 세포의 제거 또는 억제를 통해 면역 반응을 억제할 수 있는 폴리클로날, 모노클로날, scfv 또는 기타 항체 유래 분자를 포함하는 항체일 수 있다. 추가 구현예에서, 면역억제 요소는 짧은 헤어핀 RNA(shRNA)일 수 있다. 이러한 구현예에서, shRNA의 암호화 영역은 rAAV 카세트에 포함되며 일반적으로 폴리-A 테일의 3' 다운스트림에 위치한다. shRNA는 사이토카인, 성장 인자(전환 성장 인자 $\beta 1$ 및 $\beta 2$, TNF 및 기타 공개적으로 알려진 것 포함)와 같은 면역자극제의 발현을 감소 또는 제거하기 위해 표적화될 수 있다.
- [0613] 이러한 면역 조절제의 사용은 한 사람이 수개월 및/또는 수년에 걸쳐 다중 투약(예를 들어, 다중 투여)를 사용할 수 있는 능력을 촉진한다. 이는 하기 논의된 바와 같은 다중 제제, 예를 들어, 다중 유전자를 암호화하는 rAAV 벡터, 또는 대상체에 다중 투여를 사용하는 것을 허용한다.
- [0614] *본원에 개시된 기술의 조성물 및 방법의 모든 양태는 다음의 넘버링된 단락 중 임의의 하나 이상에서 정의될 수 있다:*
- [0615] 1. 재조합 아데노바이러스 관련(AAV) 벡터로서, 그 게놈에 (a) 5' 및 3' AAV 역 말단 반복체(ITR) 서열 및 (b) 5' ITR과 3' ITR 사이에 위치하는, 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서, 이중 핵산은 간-특이적 프로모터에 작동적으로 연결되는 재조합 아데노바이러스 관련(AAV) 벡터.
- [0616] 2. 단락 1에 있어서, GAA 폴리펩타이드를 암호화하는 이중 핵산 서열은 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산의 5'에 위치한 신호 펩타이드를 추가로 포함하거나, GAA 폴리펩타이드를 암호화하는 이중 핵산 서열은 GAA 폴리펩타이드의 N-말단에 IGF2 표적화 펩타이드를 추가로 포함하는, 재조합 AAV 벡터.
- [0617] 3. 단락 1에 있어서, GAA 폴리펩타이드를 암호화하는 이중 핵산 서열은 분비 신호 펩타이드와 알파-글루코시다아제(GAA) 폴리펩타이드 사이에 위치한 IGF2 표적화 펩타이드를 추가로 포함하거나, GAA 폴리펩타이드를 암호화하는 이중 핵산 서열은 GAA 폴리펩타이드의 N-말단에 IGF2 표적화 펩타이드를 추가로 포함하는, 재조합 AAV 벡터.
- [0618] 4. 단락 1 내지 3 중 어느 하나에 있어서, AAV 게놈은 5'에서 3' 방향으로, (a) 5' ITR, (b) 간-특이적 프로모

터 서열, (c) 인트론 서열, (d) 분비 신호 펩타이드를 암호화하는 핵산, (e) 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산, (f) 폴리A 서열, 및 (g) 3' ITR을 포함하는, 재조합 AAV 벡터. 일부 구현예에서, 인트론 서열(c)은 부재이다.

- [0619] 5. 단락 1 내지 3 중 어느 하나에 있어서, AAV 게놈은 5'에서 3' 방향으로, (a) 5' ITR, (b) 간-특이적 프로모터 서열, (c) 인트론 서열, (d) IGF2 표적화 펩타이드를 암호화하는 핵산, (e) 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산, (f) 폴리A 서열, 및 (g) 3' ITR을 포함하는, 재조합 AAV 벡터. 일부 구현예에서, 인트론 서열(c)은 부재이다.
- [0620] 6. 단락 1 내지 6 중 어느 하나에 있어서, AAV 게놈은 5'에서 3' 방향으로, (a) 5' ITR, (b) 간-특이적 프로모터 서열, (c) 인트론 서열, (d) 분비 신호 펩타이드를 암호화하는 핵산, (e) IGF2 표적화 펩타이드를 암호화하는 핵산, (f) 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산, (g) 폴리A 서열, 및 (h) 3' ITR을 포함하는, 재조합 AAV 벡터. 일부 구현예에서, 인트론 서열은 부재이다.
- [0621] 7. 단락 1 내지 6 중 어느 하나에 있어서, 분비 신호 펩타이드는 AAT 신호 펩타이드, 피브로넥틴 신호 펩타이드(FN1), GAA 신호 펩타이드, 또는 분비 신호 활성을 갖는 이의 활성 단편으로부터 선택되는, 재조합 AAV 벡터. 예를 들어, 분비 신호 펩타이드를 암호화하는 핵산은 AAT 신호 펩타이드(예를 들어, 서열 번호 17), 피브로넥틴 신호 펩타이드(FN1) (예를 들어, 서열 번호 18-21), 동족 GAA 신호 펩타이드(서열 번호 175), hIGF2 신호 펩타이드(예를 들어, 서열 번호 22), IgG1 리더 펩타이드(서열 번호 177), wtIL2 리더 펩타이드(서열 번호 179), 돌연변이 IL2 리더 펩타이드(서열 번호 181) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 17-22, 175, 177, 179 또는 181과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산으로부터 선택될 수 있다.
- [0622] 8. 단락 1 내지 7 중 어느 하나에 있어서, IGF2 표적화 펩타이드는 인간 양이온-비의존적 만노오스-6-포스페이트 수용체(CI-MPR) 또는 IGF2 수용체에 결합하는, 재조합 AAV 벡터.
- [0623] 9. 단락 1 내지 5 중 어느 하나에 있어서, IGF2 표적화 펩타이드는 서열 번호 5를 포함하거나, CI-MPR 수용체에 대한 결합에 영향을 미치지 않거나 하나 이상의 IGF-결합 단백질(IGFBP, 예컨대 IGFBP1-6)에 대한 결합을 감소시키는 서열 번호 5의 적어도 하나의 아미노 변형을 포함하는, 재조합 AAV 벡터.
- [0624] 10. 단락 1 내지 9 중 어느 하나에 있어서, 서열 번호 5에서 적어도 하나의 아미노 변형은 V43M 아미노산 변형(서열 번호 8 또는 서열 번호 9) 또는 $\Delta 2-7$ (서열 번호 6) 또는 $\Delta 1-7$ (서열 번호 7)인, 재조합 AAV 벡터.
- [0625] 11. 단락 1 내지 10 중 어느 하나에 있어서, 간-특이적 프로모터는 본원의 표 4, 또는 2019년 11월 19일에 출원된 가출원 62,937,556의 표 4A 또는 4B에 나열된 임의의 서열의 프로모터, 또는 이의 기능적 변이체 또는 이의 기능적 단편인, 재조합 AAV 벡터.
- [0626] 12. 단락 1 내지 10 중 어느 하나에 있어서, 간-특이적 프로모터는 CRM_SP0412(서열 번호 86) 또는 SP0412(서열 번호 91), 또는 서열 번호 86 또는 서열 번호 91에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편을 포함하는, 재조합 AAV 벡터.
- [0627] 13. 단락 1 내지 12 중 어느 하나에 있어서, 간 특이적 프로모터는 하기 중 어느 하나로부터 선택되는, 재조합 AAV 벡터: (i) SP0422(서열 번호 92) 또는 서열 번호 92에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편; (ii) CRM_SP0239(서열 번호 87) 또는 SP0239(서열 번호 93) 또는 SP0238-UTR(서열 번호 147) 또는 서열 번호 87, 서열 번호 93 또는 서열 번호 147에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편; (iii) CRM_SP0265(SP0131_A1) (서열 번호 88) 또는 SP0265(LVR_SP0131_A1) (서열 번호 94) 또는 SP0265-UTR(서열 번호 146), 또는 서열 번호 88, 서열 번호 94 또는 서열 번호 146에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편; (iv) CRM_SP0240(서열 번호 89) 또는 SP0240(서열 번호 95) 또는 SP0240-UTR(서열 번호 148), 또는 서열 번호 89, 서열 번호 95 또는 서열 번호 148에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편; (v) CRM_SP0246(서열 번호 90) 또는 SP0246(서열 번호 96) 또는 SP0246-UTR(서열 번호 149), 또는 서열 번호 90, 서열 번호 96 또는 서열 번호 149에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편 중 어느 하나로부터 선택되는, 재조합 AAV 벡터.
- [0628] 14. 단락 1 내지 13 중 어느 하나에 있어서, 핵산 서열은 야생형 GAA 폴리펩타이드 또는 변형된 GAA 폴리펩타이드를 암호화하는, 재조합 AAV 벡터.
- [0629] 15. 단락 1 내지 14 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 인간 GAA 유전자 또는

인간 코돈 최적화된 GAA 유전자(coGAA) 또는 변형된 GAA 핵산 서열인, 재조합 AAV 벡터.

- [0630] 16. 단락 1 내지 15 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 생체내 향상된 발현을 위해 코돈 최적화된, 재조합 AAV 벡터.
- [0631] 17. 단락 1 내지 16 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 CpG 섬을 감소시키도록 코돈 최적화된, 재조합 AAV 벡터.
- [0632] 18. 단락 1 내지 17 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산은 선천성 면역 반응을 감소시키거나 CpG 섬을 감소시키거나, 또는 선천성 면역 반응을 감소시키고 선천성 면역 반응을 감소시키도록 코돈 최적화된, 재조합 AAV 벡터.
- [0633] 19. 단락 1 내지 18 중 어느 하나에 있어서, 핵산 서열은 서열 번호 10의 H201L, H199R 또는 R233H로부터 선택된 적어도 1개, 적어도 2개 또는 적어도 3개 모두의 아미노산 변형을 포함하는 GAA 폴리펩타이드를 암호화하는, 재조합 AAV 벡터.
- [0634] 20. 단락 1 내지 19 중 어느 하나에 있어서, 암호화된 폴리펩타이드는 GAA 폴리펩타이드의 아미노-말단 및 IGF2 표적화 펩타이드의 C-말단에 위치한 적어도 1개의 아미노산에 대한 뉴클레오타이드 서열을 포함하는 스페이서를 추가로 포함하는, 재조합 AAV 벡터.
- [0635] 21. 단락 1 내지 20 중 어느 하나에 있어서, IGF2 표적화 펩타이드를 암호화하는 핵산과 GAA 폴리펩타이드를 암호화하는 핵산 사이에 위치하는 적어도 1개의 아미노산의 스페이서를 암호화하는 핵산을 추가로 포함하는, 재조합 AAV 벡터.
- [0636] 22. 단락 1 내지 21 중 어느 하나에 있어서, GAA 유전자를 암호화하는 핵산의 3' 및 3' ITR 서열의 5'에 위치한 적어도 하나의 폴리A 서열을 추가로 포함하는, 재조합 AAV 벡터.
- [0637] 23. 단락 1 내지 22 중 어느 하나에 있어서, 이중 핵산 서열은 GAA 폴리펩타이드를 암호화하는 핵산의 3' 및 3' ITR 서열의 5'에 위치한 콜라겐 안정성(CS) 서열을 추가로 포함하는, 재조합 AAV 벡터.
- [0638] 24. 단락 1 내지 23 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산과 폴리A 서열 사이에 위치한 콜라겐 안정성(CS) 서열을 암호화하는 핵산을 추가로 포함하는, 재조합 AAV 벡터.
- [0639] 25. 단락 1 내지 24 중 어느 하나에 있어서, 분비 신호 펩타이드를 암호화하는 서열의 5' 및 프로모터의 3'에 위치한 인트론 서열을 추가로 포함하는, 재조합 AAV 벡터.
- [0640] 26. 단락 1 내지 25 중 어느 하나에 있어서, 인트론 서열은 MVM 서열, SV40 서열 또는 HBB2 서열을 포함하고, 여기서, MVM 서열은 서열 번호 13의 핵산 서열, 또는 서열 번호 13과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하며, HBB2 서열은 서열 번호 14의 핵산 서열, 또는 서열 번호 14와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는, 재조합 AAV 벡터.
- [0641] 27. 단락 1 내지 26 중 어느 하나에 있어서, ITR은 삽입, 결실 또는 치환을 포함하는, 재조합 AAV 벡터.
- [0642] 28. 단락 1 내지 27 중 어느 하나에 있어서, ITR에서 하나 이상의 CpG 섬이 제거되는, 재조합 AAV 벡터.
- [0643] 29. 단락 1 내지 28 중 어느 하나에 있어서,
- [0644] a. 이중 핵산 서열은 다음으로부터 선택되는 분비 신호 펩타이드를 암호화하며:
- [0645] 1. 피브로넥틴 신호 펩타이드(FN1) 또는 분비 신호 활성을 갖는 이의 활성 단편(예를 들어, FN1 신호 펩타이드는 서열 번호 18-21 중 어느 하나의 서열, 또는 서열 번호 18-21 중 어느 하나와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 가짐),
- [0646] 2. AAT 신호 펩타이드(예를 들어, 서열 번호 17), 또는 서열 번호 17 중 어느 하나와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열;
- [0647] 3. hIGF2 신호 펩타이드(예를 들어, 서열 번호 22), 또는 서열 번호 22 중 어느 하나와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열;
- [0648] 4. IgG1 리더 펩타이드(서열 번호 177), 또는 서열 번호 177 중 어느 하나와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열;

- [0649] 5. wt IL2 리더 펩타이드(서열 번호 179), 또는 서열 번호 179 중 어느 하나와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열;
- [0650] 6. 돌연변이 IL2 리더 펩타이드(서열 번호 181), 또는 서열 번호 181 중 어느 하나와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열, 및
- [0651] b. 이중 핵산 서열은 서열 번호 11, 서열 번호 72 또는 서열 번호 182, 또는 서열 번호 11, 서열 번호 72, 또는 서열 번호 182와 적어도 60%, 또는 70%, 또는 80%, 85% 또는 90% 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열로 이루어진 임의의 군으로부터 선택된 GAA 폴리펩타이드를 암호화하고, 임의로, 핵산 서열은 서열 번호 10의 H201L, H199R 또는 R233H로부터 선택된 적어도 1개, 적어도 2개 또는 적어도 3개 모두의 아미노산 변형을 갖는 GAA 폴리펩타이드를 암호화한다.
- [0652] 30. 단락 29에 있어서, 이중 핵산은 또한 서열 번호 5, 서열 번호 6, 서열 번호 7, 서열 번호 8 또는 서열 번호 9 중 어느 하나로부터 선택된 IGF2 표적화 펩타이드, 또는 서열 번호 5-9와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 IGF2 펩타이드를 암호화하는, 재조합 AAV 벡터.
- [0653] 31. 단락 1 내지 30 중 어느 하나에 있어서, 암호화된 분비 신호 펩타이드는 AAT 신호 펩타이드 또는 분비 신호 활성을 갖는 임의의 활성 단편이고(예를 들어, AAT 신호 펩타이드는 서열 번호 17의 서열, 또는 서열 번호 17과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 가짐), 이중 핵산 서열은 서열 번호 5, 서열 번호 6, 서열 번호 7, 서열 번호 8 또는 서열 번호 9 중 어느 하나로부터 선택된 IGF2 표적화 펩타이드, 또는 서열 번호 5-9와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 IGF2 펩타이드를 암호화하는, 재조합 AAV 벡터.
- [0654] 32. 단락 1 내지 31 중 어느 하나에 있어서, IGF2 표적화 펩타이드는 서열 번호 8 또는 서열 번호 9, 또는 서열 번호 8 또는 9와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 IGF2 펩타이드인, 재조합 AAV 벡터.
- [0655] 33. 단락 1 내지 32 중 어느 하나에 있어서, 재조합 AAV 벡터는 키메라 AAV 벡터, 반수체 AAV 벡터, 하이브리드 AAV 벡터, 배수체 AAV 벡터, 합리적인 반수체 벡터, 모자이크 AAV 벡터, 화학적으로 변형된 AAV 벡터, 또는 임의의 AVV 혈청형, 예를 들어, 국제 출원 W02020/102645 및 W02020/102667에 개시된 표 1에 나열된 임의의 혈청형으로부터의 AAV 벡터 중 어느 하나로부터 선택되는, 재조합 AAV 벡터.
- [0656] 34. 단락 1 내지 33 중 어느 하나에 있어서, 재조합 AAV 벡터는 2019년 11월 19일에 출원된 미국 가출원 62,937,556의 표 1, 또는 국제 출원 W02020/102645 및 W02020/102667에 나열된 것들 또는 이들의 임의의 조합에서 임의의 AAV 혈청형으로부터 선택된 캡시드 단백질을 포함하는, 재조합 AAV 벡터.
- [0657] 35. 단락 1 내지 34 중 어느 하나에 있어서, AAV 벡터는 AAV3 벡터, AAVXL32 벡터, AAVXL32.1 벡터, AAV8 벡터, 또는 적어도 하나의 AAV8 캡시드 단백질을 포함하는 반수체 AAV8 벡터로 이루어진 군으로부터의 혈청형으로부터 선택되는, 재조합 AAV 벡터.
- [0658] 36. 단락 1 내지 35 중 어느 하나에 있어서, AAV3b 혈청형은 265D, 549A, Q263Y 중 어느 하나로부터 선택된 캡시드 단백질에서 하나 이상의 돌연변이를 포함하는, 재조합 AAV 벡터.
- [0659] 37. 단락 1 내지 36 중 어느 하나에 있어서, AAV3b 혈청형은 AAV3b265D, AAV3b265D549A, AAV3b549A 또는 AAV3bQ263Y, 또는 AAV3bSASTG 중 어느 하나로부터 선택되는, 재조합 AAV 벡터.
- [0660] 38. 재조합 아테노바이러스 관련(AAV) 벡터로서, 그 게놈에
- [0661] a. 5' 및 3' AAV 역 말단 반복체(ITR) 서열, 및
- [0662] b. 5' ITR과 3' ITR 사이에 위치하는, 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서, 이중 핵산은 간 특이적 프로모터에 작동적으로 연결되며,
- [0663] 여기서, 재조합 AAV 벡터는 AAV8 혈청형 또는 AAV3b 혈청형의 캡시드 단백질을 포함하는, 재조합 아테노바이러스 관련(AAV) 벡터.
- [0664] 39. 단락 38에 있어서, GAA 폴리펩타이드는 GAA 폴리펩타이드의 N-말단에 위치한 분비 신호 펩타이드를 추가로 포함하는, 재조합 AAV 벡터.

- [0665] 40. 단락 38 또는 39에 있어서, GAA 폴리펩타이드를 암호화하는 이중 핵산 서열은 GAA 폴리펩타이드의 N-말단에 위치하거나 분비 신호 펩타이드와 알파-글루코시다아제(GAA) 폴리펩타이드 사이에 위치하는 IGF2 표적화 펩타이드를 추가로 포함하는, 재조합 AAV 벡터.
- [0666] 41. 단락 38에 있어서, AAV 계놈은 5'에서 3' 방향으로, 5' ITR, 간 특이적 프로모터 서열, 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산, 폴리A 서열, 및 3' ITR을 포함하는, 재조합 AAV 벡터.
- [0667] 42. 단락 38에 있어서, AAV 계놈은 5'에서 3' 방향으로, 5' ITR, 간 특이적 프로모터 서열, 분비 신호 펩타이드를 암호화하는 핵산, 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산, 폴리A 서열, 및 3' ITR을 포함하는, 재조합 AAV 벡터.
- [0668] 43. 단락 38에 있어서, AAV 계놈은 5'에서 3' 방향으로, 5' ITR, 간 특이적 프로모터 서열, 인트론 서열, 분비 신호 펩타이드를 암호화하는 핵산, 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산, 폴리A 서열, 및 3' ITR을 포함하는, 재조합 AAV 벡터.
- [0669] 44. 단락 38에 있어서, AAV 계놈은 5'에서 3' 방향으로, 5' ITR, 간 특이적 프로모터 서열, 인트론 서열, 분비 신호 펩타이드를 암호화하는 핵산, IGF2 표적화 펩타이드를 암호화하는 핵산, 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산, 폴리A 서열, 및 3' ITR을 포함하는, 재조합 AAV 벡터.
- [0670] 45. 단락 38에 있어서, AAV 계놈은 5'에서 3' 방향으로, 5' ITR, 간 특이적 프로모터 서열, 인트론 서열, 분비 신호 펩타이드를 암호화하는 핵산, IGF2 표적화 펩타이드를 암호화하는 핵산, 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산, 3' ITR 서열, 폴리A 서열, 및 3' ITR을 포함하는, 재조합 AAV 벡터.
- [0671] 46. 단락 34 내지 37 중 어느 하나에 있어서, 분비 신호 펩타이드는 분비 신호 펩타이드를 암호화하는 핵산으로부터 선택되며, AAT 신호 펩타이드(예를 들어, 서열 번호 17), 피브로넥틴 신호 펩타이드(FN1) (예를 들어, 서열 번호 18-21), 동족 GAA 신호 펩타이드(서열 번호 175), hIGF2 신호 펩타이드(예를 들어, 서열 번호 22), IgG1 리더 펩타이드(서열 번호 177), wt IL2 리더 펩타이드(서열 번호 179), 돌연변이 IL2 리더 펩타이드(서열 번호 181) 또는 분비 신호 활성을 갖는 이의 활성 단편, 예를 들어, 서열 번호 17-22, 175, 177, 179 또는 181 과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 핵산으로부터 선택될 수 있는, 재조합 AAV 벡터.
- [0672] 47. 단락 38 내지 46 중 어느 하나에 있어서, IGF2 표적화 펩타이드는 인간 양이온-비의존적 만노오스-6-포스페이트 수용체(CI-MPR) 또는 IGF2 수용체에 결합하는, 재조합 AAV 벡터.
- [0673] 48. 단락 38 내지 47 중 어느 하나에 있어서, IGF2 표적화 펩타이드는 서열 번호 5를 포함하거나, GF2 수용체 결합에 영향을 미치지 않거나 IGF 결합 단백질 IGFBP1-6 중 하나 이상에 대한 이의 결합을 감소시키는 서열 번호 5에서 적어도 하나의 아미노 변형을 포함하는, 재조합 AAV 벡터.
- [0674] 49. 단락 48에 있어서, 서열 번호 5에서 적어도 하나의 아미노 변형은 V43M 아미노산 변형(서열 번호 8 또는 서열 번호 9) 또는 Δ 2-7(서열 번호 6) 또는 Δ 1-7(서열 번호 7)인, 재조합 AAV 벡터.
- [0675] 50. 단락 38 내지 49 중 어느 하나에 있어서, 간 특이적 프로모터는 표 4, 또는 2019년 11월 19일에 출원된 가출원 62,937,556의 표 4A 또는 4B에 나열된 임의의 서열, 또는 이의 기능적 변이체 또는 기능적 단편으로부터 선택되는, 재조합 AAV 벡터.
- [0676] 51. 단락 38 내지 50 중 어느 하나에 있어서, 간 특이적 프로모터는 CRM_SP0412(서열 번호 86) 또는 SP0412(서열 번호 91) 중 어느 하나, 또는 서열 번호 86 또는 서열 번호 91에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편으로부터 선택되는, 재조합 AAV 벡터.
- [0677] 52. 단락 38 내지 50 중 어느 하나에 있어서, 간 특이적 프로모터는 다음 중 어느 하나로부터 선택되는, 재조합 AAV 벡터: (i) SP0422(서열 번호 92) 또는 서열 번호 92에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편; (ii) CRM_SP0239(서열 번호 87) 또는 SP0239(서열 번호 93) 또는 SP0238-UTR(서열 번호 147) 또는 서열 번호 87, 서열 번호 93 또는 서열 번호 147에 대해 적어도 60% 활성을 갖는 또는 이의 기능적 변이체 또는 기능적 단편; (iii) CRM_SP0265(SP0131_A1) (서열 번호 88) 또는 SP0265(LVR_SP0131_A1) (서열 번호 94) 또는 SP0265-UTR(서열 번호 146), 또는 서열 번호 88, 서열 번호 94 또는 서열 번호 146에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편; (iv) CRM_SP0240(서열 번호 89) 또는 SP0240(서열 번호 95) 또는 SP0240-UTR(서열 번호 148), 또는 서열 번호 89, 서열 번호 95 또는 서열 번호 148에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편; (v) CRM_SP0246(서열 번호 90) 또는 SP0246(서열 번호 96) 또

는 SP0246-UTR(서열 번호 149), 또는 서열 번호 90, 서열 번호 96 또는 서열 번호 149에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편.

- [0678] 53. 단락 38 내지 52 중 어느 하나에 있어서, 핵산 서열은 야생형 GAA 폴리펩타이드 또는 변형된 GAA 폴리펩타이드를 암호화하는, 재조합 AAV 벡터.
- [0679] 54. 단락 38 내지 53 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 인간 GAA 유전자 또는 인간 코돈 최적화된 GAA 유전자(coGAA) 또는 변형된 GAA 핵산 서열인, 재조합 AAV 벡터.
- [0680] 55. 단락 38 내지 54 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 생체내에서 향상된 발현을 위해 코돈 최적화된, 재조합 AAV 벡터.
- [0681] 56. 단락 38 내지 55 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 CpG 섬을 감소시키기 위해 코돈 최적화된, 재조합 AAV 벡터.
- [0682] 57. 단락 38 내지 56 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 선천성 면역 반응을 감소시키거나 CpG 섬을 감소시키거나, 또는 선천성 면역 반응을 감소시키고 CpG 섬을 감소시키도록 코돈 최적화된, 재조합 AAV 벡터.
- [0683] 58. 단락 38 내지 57 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 서열 번호 10의 H201L, H199R 또는 R233H로부터 선택된 적어도 1개, 적어도 2개 또는 적어도 3개 모두의 아미노산 변형을 포함하는 GAA 폴리펩타이드를 암호화하는, 재조합 AAV 벡터.
- [0684] 59. 단락 38 내지 58 중 어느 하나에 있어서, 인트론 서열은 MVM, HBB2 또는 SV40 인트론 서열 중 어느 하나로 부터 선택되며, 여기서, MVM 서열은 서열 번호 13의 핵산 서열, 또는 서열 번호 13과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하며, HBB2 서열은 서열 번호 14의 핵산 서열, 또는 서열 번호 14와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는, 재조합 AAV 벡터.
- [0685] 60. 단락 38 내지 59 중 어느 하나에 있어서, ITR은 삽입, 결실 또는 치환을 포함하는, 재조합 AAV 벡터.
- [0686] 61. 단락 60에 있어서, ITR에서 하나 이상의 CpG 섬이 제거된, 재조합 AAV 벡터.
- [0687] 62. 단락 38 내지 61 중 어느 하나에 있어서, 3' ITR은 서열 번호 442, 또는 서열 번호 442와 적어도 80% 동일한, 예를 들어, 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 또는 99.5% 동일한 뉴클레오타이드 서열을 포함하거나 이로 이루어지며, 5' ITR은 서열 번호 441 또는 서열 번호 441과 적어도 80% 동일한, 예를 들어, 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 또는 99.5% 동일한 뉴클레오타이드 서열을 포함하거나 이로 이루어진, 재조합 AAV 벡터.
- [0688] 63. 단락 38 내지 61 중 어느 하나에 있어서, 3' ITR은 서열 번호 165, 또는 서열 번호 165와 적어도 80% 동일한, 예를 들어, 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 또는 99.5% 동일한 뉴클레오타이드 서열을 포함하거나 이로 이루어지며, 5' ITR은 서열 번호 161, 또는 서열 번호 161과 적어도 80% 동일한, 예를 들어, 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 또는 99.5% 동일한 뉴클레오타이드 서열을 포함하거나 이로 이루어진, 재조합 AAV 벡터.
- [0689] 64. 단락 38 내지 63 중 어느 하나에 있어서, 분비 신호 펩타이드는 피브로넥틴 신호 펩타이드(FN1) 또는 분비 신호 활성을 갖는 이의 활성 단편이고(예를 들어, FN1 신호 펩타이드는 서열 번호 18-21 중 어느 하나의 서열, 또는 서열 번호 18-21 중 어느 하나와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 가짐), 이중 핵산 서열은 서열 번호 5, 서열 번호 6, 서열 번호 7, 서열 번호 8 또는 서열 번호 9 중 어느 하나로부터 선택된 IGF2 표적화 펩타이드, 또는 서열 번호 5-9와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 IGF2 펩타이드를 암호화하는, 재조합 AAV 벡터.
- [0690] 65. 단락 38 내지 64 중 어느 하나에 있어서, 암호화된 분비 신호 펩타이드는 AAT 신호 펩타이드 또는 분비 신호 활성을 갖는 이의 활성 단편이고(예를 들어, AAT 신호 펩타이드는 서열 번호 17의 서열, 또는 서열 번호 17과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 가짐), 이중 핵산 서열은 서열 번호 5, 서열 번호 6, 서열 번호 7, 서열 번호 8 또는 서열 번호 9 중 어느 하나로부터 선택된 IGF2 표적화 펩타이드, 또는 서열 번호 5-9와 적어도 약 75%, 또는 80%, 또는 85%, 또

는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 IGF2 펩타이드를 암호화하는, 재조합 AAV 벡터.

- [0691] 66. 단락 38 내지 65 중 어느 하나에 있어서, IGF2 표적화 펩타이드는 서열 번호 8 또는 서열 번호 9, 또는 서열 번호 8 또는 9와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 IGF2 펩타이드인, 재조합 AAV 벡터.
- [0692] 67. 단락 38 내지 66 중 어느 하나에 있어서, 간 특이적 프로모터는 CRM_SP0412(서열 번호 86) 또는 SP0412(서열 번호 91), 또는 서열 번호 86 또는 서열 번호 91에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편인, 재조합 AAV 벡터.
- [0693] 68. 단락 38 내지 66 중 어느 하나에 있어서, 간 특이적 프로모터는 SP0422(서열 번호 92), 또는 서열 번호 92에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편인, 재조합 AAV 벡터.
- [0694] 69. 단락 38 내지 66 중 어느 하나에 있어서, 간 특이적 프로모터는 CRM_SP0239(서열 번호 87) 또는 SP0239(서열 번호 93) 또는 SP0238-UTR(서열 번호 147), 또는 서열 번호 87, 서열 번호 93 또는 서열 번호 147에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편인, 재조합 AAV 벡터.
- [0695] 70. 단락 38 내지 66 중 어느 하나에 있어서, 간 특이적 프로모터는 CRM_SP0265(SP0131_A1) (서열 번호 88) 또는 SP0265(LVR_SP0131_A1) (서열 번호 94) 또는 SP0265-UTR(서열 번호 146), 또는 서열 번호 88, 서열 번호 94 또는 서열 번호 146에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편인, 재조합 AAV 벡터.
- [0696] 71. 단락 38 내지 66 중 어느 하나에 있어서, 간 특이적 프로모터는 CRM_SP0240(서열 번호 89) 또는 SP0240(서열 번호 95) 또는 SP0240-UTR(서열 번호 148), 또는 서열 번호 89, 서열 번호 95 또는 서열 번호 148에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편인, 재조합 AAV 벡터.
- [0697] 72. 단락 38 내지 66 중 어느 하나에 있어서, 간 특이적 프로모터는 CRM_SP0246(서열 번호 90) 또는 SP0246(서열 번호 96) 또는 SP0246-UTR(서열 번호 149), 또는 서열 번호 90, 서열 번호 96 또는 서열 번호 149에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편인, 재조합 AAV 벡터.
- [0698] 73. 약제학적으로 허용되는 담체에 이전 단락 중 어느 하나의 재조합 AAV 벡터를 포함하는 약제학적 조성물.
- [0699] 74. 핵산 서열로서,
- [0700] GAA 폴리펩타이드를 암호화하는 이중 핵산 서열에 작동적으로 연결된 간 특이적 프로모터를 포함하며, 여기서, 간 특이적 프로모터는 표 4, 또는 2019년 11월 19일에 출원된 미국 가출원 62,937,556의 표 4A 또는 4B에 개시된 서열 중 어느 하나, 또는 이의 기능적 변이체 또는 기능적 단편으로부터 선택되는, 프로모터.
- [0701] 75. 단락 74에 있어서, 이중 핵산 서열은 분비 신호 펩타이드를 암호화하는 핵산 및 GAA 폴리펩타이드를 암호화하는 핵산의 순서로 포함하는, 핵산 서열.
- [0702] 76. 단락 74에 있어서, 이중 핵산 서열은 분비 신호 펩타이드를 암호화하는 핵산, IGF2 표적화 펩타이드를 암호화하는 핵산 및 GAA 폴리펩타이드를 암호화하는 핵산의 순서로 포함하는, 핵산 서열.
- [0703] 77. 재조합 아데노바이러스 관련(rAAV) 벡터 계통에 대한 핵산 서열로서,
- [0704] a. 5' 및 3' AAV 역 말단 반복체(ITR) 핵산 서열 및
- [0705] b. 5' ITR 서열과 3' ITR 서열 사이에 위치하는, 알파-글루코시다아제(GAA) 폴리펩타이드를 포함하는 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서, 이중 핵산은 간-특이적 프로모터에 작동적으로 연결되고, 여기서, 간 특이적 프로모터는 표 4, 또는 2019년 11월 19일에 출원된 미국 가출원 62,937,556의 표 4A 또는 4B에 개시된 서열 중 어느 하나, 또는 이의 기능적 변이체 또는 기능적 단편으로 선택되는 핵산 서열,
- [0706] 78. 단락 75에 있어서, 핵산 서열은 분비 신호 및 알파-글루코시다아제(GAA) 폴리펩타이드를 포함하는 융합 폴리펩타이드를 암호화하는, 핵산 서열.
- [0707] 79. 단락 75에 있어서, 핵산 서열은 분비 신호, IGF2 표적화 펩타이드 및 알파-글루코시다아제(GAA) 폴리펩타이드를 포함하는 융합 폴리펩타이드를 암호화하는, 핵산 서열.
- [0708] 80. 핵산 서열로서,
- [0709] GAA 폴리펩타이드를 암호화하는 핵산을 포함하는 이중 핵산 서열에 작동적으로 연결된 간 특이적 프로모터를 포함하며, 여기서, 간 특이적 프로모터는 다음 중 어느 하나로부터 선택되는 핵산 서열:

- [0710] i. CRM_SP0412(서열 번호 86) 또는 SP0412(서열 번호 91), 또는 서열 번호 86 또는 서열 번호 91에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편,
- [0711] ii. SP0422(서열 번호 92), 또는 서열 번호 92에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편,
- [0712] iii. CRM_SP0239(서열 번호 87) 또는 SP0239(서열 번호 93) 또는 SP0238-UTR(서열 번호 147), 또는 서열 번호 87, 서열 번호 93 또는 서열 번호 147에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;
- [0713] iv. CRM_SP0265(SP0131_A1) (서열 번호 88) 또는 SP0265(LVR_SP0131_A1) (서열 번호 94) 또는 SP0265-UTR(서열 번호 146), 또는 서열 번호 88, 서열 번호 94 또는 서열 번호 146에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;
- [0714] v. CRM_SP0240(서열 번호 89) 또는 SP0240(서열 번호 95) 또는 SP0240-UTR(서열 번호 148), 또는 서열 번호 89, 서열 번호 95 또는 서열 번호 148에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편; 또는
- [0715] vi. CRM_SP0246(서열 번호 90) 또는 SP0246(서열 번호 96) 또는 SP0246-UTR(서열 번호 149), 또는 서열 번호 90, 서열 번호 96 또는 서열 번호 149에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편.
- [0716] 81. 단락 80에 있어서, 이중 핵산 서열은 분비 신호를 암호화하는 핵산 및 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산의 순서로 포함하는, 핵산 서열.
- [0717] 82. 단락 75에 있어서, 이중 핵산 서열은 분비 신호를 암호화하는 핵산, IGF2 표적화 펩타이드를 암호화하는 핵산 서열 및 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하는 핵산의 순서로 포함하는, 핵산 서열.
- [0718] 83. 제조합 아데노바이러스 관련(rAAV) 벡터 게놈에 대한 핵산 서열로서,
- [0719] (a) 5' 및 3' AAV 역 말단 반복체(ITR) 핵산 서열 및 (b) 5' ITR 서열과 3' ITR 서열 사이에 위치하는, 알파-글루코시다아제(GAA) 폴리펩타이드를 암호화하거나 분비 신호 펩타이드 및 알파-글루코시다아제(GAA) 폴리펩타이드를 포함하는 융합 폴리펩타이드를 암호화하는 이중 핵산 서열을 포함하며, 여기서, 이중 핵산은 간-특이적 프로모터에 작동적으로 연결되고, 간 특이적 프로모터는 다음 중 어느 하나로부터 선택되는 핵산 서열:
- [0720] i. CRM_SP0412(서열 번호 86) 또는 SP0412(서열 번호 91), 또는 서열 번호 86 또는 서열 번호 91에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편,
- [0721] ii. SP0422(서열 번호 92), 또는 서열 번호 92에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편,
- [0722] iii. CRM_SP0239(서열 번호 87) 또는 SP0239(서열 번호 93) 또는 SP0238-UTR(서열 번호 147), 또는 서열 번호 87, 서열 번호 93 또는 서열 번호 147에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;
- [0723] iv. CRM_SP0265(SP0131_A1) (서열 번호 88) 또는 SP0265(LVR_SP0131_A1) (서열 번호 94) 또는 SP0265-UTR(서열 번호 146), 또는 서열 번호 88, 서열 번호 94 또는 서열 번호 146에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편;
- [0724] v. CRM_SP0240(서열 번호 89) 또는 SP0240(서열 번호 95) 또는 SP0240-UTR(서열 번호 148), 또는 서열 번호 89, 서열 번호 95 또는 서열 번호 148에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편; 또는
- [0725] vi. CRM_SP0246(서열 번호 90) 또는 SP0246(서열 번호 96) 또는 SP0246-UTR(서열 번호 149), 또는 서열 번호 90, 서열 번호 96 또는 서열 번호 149에 대해 적어도 60% 활성을 갖는 이의 기능적 변이체 또는 기능적 단편.
- [0726] 84. 단락 74 내지 83 중 어느 하나에 있어서, GAA 폴리펩타이드 또는 융합 폴리펩타이드를 암호화하는 이중 핵산 서열은 분비 신호 펩타이드와 알파-글루코시다아제(GAA) 폴리펩타이드 사이에 위치한 IGF2 표적화 펩타이드를 추가로 포함하는, 핵산 서열.
- [0727] 85. 단락 74 내지 84 중 어느 하나에 있어서, 분비 신호를 암호화하는 핵산은 서열 번호 17, 22-26, 175, 177, 179 또는 181 중 어느 하나, 또는 서열 번호 17 또는 22-26, 또는 175, 177, 179 또는 181과 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열로부터 선택되는,

핵산 서열.

- [0728] 86. 단락 74 내지 85 중 어느 하나에 있어서, IGF2 표적화 펩타이드를 암호화하는 핵산은 서열 번호 2(IGF2-△2-7), 서열 번호 3(IGF2-△1-7), 또는 서열 번호 4(IGF2 V43M) 중 어느 하나, 또는 서열 번호 2, 3 또는 4 중 어느 하나와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열로부터 선택되는, 핵산 서열.
- [0729] 87. 단락 74 내지 86 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 인간 GAA 유전자 또는 인간 코돈 최적화된 GAA 유전자(coGAA) 또는 변형된 GAA 핵산 서열인, 핵산 서열.
- [0730] 88. 단락 74 내지 87 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 생체내 향상된 발현을 위해 코돈 최적화된, 핵산 서열.
- [0731] 89. 단락 74 내지 88 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 CpG 섬을 감소시키도록 코돈 최적화된, 핵산 서열.
- [0732] 90. 단락 74 내지 89 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 선천성 면역 반응을 감소시키거나 CpG 섬을 감소시키거나 선천성 면역 반응을 감소시키고 CpG 섬을 감소시키도록 코돈 최적화된, 핵산 서열.
- [0733] 91. 단락 74 내지 90 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산 서열은 서열 번호 10의 H201L 변형을 포함하는 GAA 폴리펩타이드를 암호화하는, 핵산 서열.
- [0734] 92. 단락 74 내지 91 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산은 서열 번호 11(전장 hGAA), 서열 번호 55(Dwight cDNA), 서열 번호 56(hGAA △1-66) 중 어느 하나, 또는 서열 번호 11, 55 또는 56 중 어느 하나와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열로부터 선택되는, 핵산 서열.
- [0735] 93. 단락 74 내지 92 중 어느 하나에 있어서, GAA 폴리펩타이드를 암호화하는 핵산은 서열 번호 74(코돈 최적화된 1), 서열 번호 75(코돈 최적화된 2), 및 서열 번호 76(코돈 최적화된 3) 중 어느 하나, 또는 서열 번호 74, 75 또는 76 중 어느 하나와 적어도 약 75%, 또는 80%, 또는 85%, 또는 90%, 또는 95%, 또는 98%, 또는 99% 서열 동일성을 갖는 핵산 서열로부터 선택되는, 핵산 서열.
- [0736] 94. 단락 74 내지 93 중 어느 하나에 있어서, 핵산은 서열 번호 57(AAT-V43M-wtGAA(델타1-69aa)); 서열 번호 58(ratFN1-IGF2V43M-wtGAA(델타1-69aa)); 서열 번호 59(hFN1-IGF2V43M-wtGAA(델타1-69aa)); 서열 번호 60(AAT-IGF2D2-7-wtGAA(델타 1-69)); 서열 번호 61(FN1rat-IGF△2-7-wtGAA(델타 1-69)); 서열 번호 62(hFN1-IGF△2-7-wtGAA(델타 1-69)), 서열 번호 79(AAT_hIGF2-V43M_wtGAA_del1-69_스터퍼.V02); 서열 번호 80(FIBrat_hIGF2-V43M_wtGAA_del1-69_스터퍼.V02); 서열 번호 81(FIBhum_hIGF2-V43M_wtGAA_del1-69_스터퍼.V02); 서열 번호 82(AAT_GILT_wtGAA_del1-69_스터퍼.V02); 서열 번호 83(FIBrat_GILT_wtGAA_del1-69_스터퍼.V02); 서열 번호 84(FIBhum_GILT_wtGAA_del1-69_스터퍼.V02) 중 어느 하나, 또는 서열 번호 57, 58, 59, 60, 61, 62, 79, 80, 81, 82, 83 또는 84와 적어도 80%, 85%, 90%, 95% 또는 98% 동일성을 갖는 핵산 서열로부터 선택되는, 핵산 서열.
- [0737] 95. 글리코겐 축적 질환 II형(GSD II, 폼페병, 산성 말타아제 결핍증) 또는 알파-글루코시다아제(GAA) 폴리펩타이드 결핍이 있는 대상체를 치료하는 방법으로서, 재조합 AAV 벡터, 또는 rAAV 게놈 또는 이전 단락 1 내지 95 중 어느 하나의 핵산 서열을 대상체에게 투여하는 단계를 포함하는 방법.
- [0738] 96. 단락 95에 있어서, GAA 폴리펩타이드는 대상체의 간으로부터 분리되고, 골격근 조직, 심장 근육 조직, 횡격막 근육 조직 또는 이들의 조합에 의해 분리된 GAA의 흡수가 있고, 여기서, 분리된 GAA의 흡수는 조직(들)에서 리소좀 글리코겐 저장의 감소를 초래하는, 방법.
- [0739] 97. 단락 95 또는 96에 있어서, 대상체에 투여하는 단계가 근육내, 피하, 척수내, 수조내, 척수강내, 정맥내 투여 중 어느 하나로부터 선택되는, 방법.
- [0740] 98. 단락 95 내지 97 중 어느 하나에 있어서, 재조합 AAV 벡터는 키메라 AAV 벡터, 반수체 AAV 벡터, 하이브리드 AAV 벡터 또는 배수체 AAV 벡터인, 방법.
- [0741] 99. 단락 95 내지 97 중 어느 하나에 있어서, 재조합 AAV 벡터는 합리적인 반수체 벡터, 모자이크 AAV 벡터, 화학적으로 변형된 AAV 벡터, 또는 임의의 AAV 혈청형으로부터의 AAV 벡터인, 방법.

- [0742] 100. 단락 95 내지 97 중 어느 하나에 있어서, 재조합 AAV 벡터는 AAVXL32 벡터 또는 AAVXL32.1 벡터 또는 AAV8 벡터, 또는 적어도 하나의 AAV8 캡시드 단백질을 포함하는 반수체 AAV8 벡터인, 방법.
- [0743] 101. 단락 100에 있어서, 재조합 AAV 벡터는 AAV8 벡터인, 방법.
- [0744] 102. 리소좀 축적 질환(LSD)이 있는 대상체를 치료하는 방법으로서, 상기 방법은 재조합 AAV 벡터, 또는 rAAV 게놈 또는 이전 단락 중 어느 한 하나의 핵산 서열 중 어느 하나를 대상체에게 투여하는 단계를 포함하고, 여기서, AAV 벡터는 표 5B 또는 표 6B에서 임의의 폴리펩타이드로부터 선택된 폴리펩타이드를 발현하는, 방법.
- [0745] 103. 단락 102에 있어서, 리소좀 축적 질환(LSD)은 표 5A 또는 표 6A에 나열된 것들 중 어느 하나로부터 선택되는, 방법.
- [0746] 104. 단락 102에 있어서, 재조합 AAV 벡터는 키메라 AAV 벡터, 반수체 AAV 벡터, 하이브리드 AAV 벡터 또는 배수체 AAV 벡터인, 방법.
- [0747] 105. 단락 102에 있어서, 재조합 AAV 벡터는 합리적인 반수체 벡터, 모자이크 AAV 벡터, 화학적으로 변형된 AAV 벡터, 또는 임의의 AAV 혈청형으로부터의 AAV 벡터인, 방법.
- [0748] 106. 단락 102에 있어서, 재조합 AAV 벡터는 AAVXL32 벡터 또는 AAVXL32.1 벡터 또는 AAV8 벡터, 또는 적어도 하나의 AAV8 캡시드 단백질을 포함하는 반수체 AAV8 벡터인, 방법.
- [0749] 107. 단락 74 내지 94 중 어느 하나의 핵산 서열을 포함하는 세포.
- [0750] 108. 단락 107에 있어서, 세포는 인간 세포인, 세포,
- [0751] 109. 단락 107 또는 108에 있어서, 세포는 비인간 세포 포유동물 세포인, 세포
- [0752] 110. 단락 107 내지 109 중 어느 하나에 있어서, 세포는 곤충 세포인, 세포.
- [0753] 111. 단락 1 내지 72 중 어느 하나의 재조합 AAV 벡터를 포함하는 세포.
- [0754] 112. 단락 1 내지 72 중 어느 하나의 재조합 AAV 벡터를 포함하는 숙주 동물,
- [0755] 113. 단락 112에 있어서, 숙주 동물은 포유동물인, 숙주 동물.
- [0756] 114. 단락 112 또는 113, 숙주 동물은 비인간 포유동물인, 숙주 동물.
- [0757] 115. 단락 113에 있어서, 숙주 동물은 인간인, 숙주 동물.
- [0758] 116. 단락 95 내지 단락 101 중 어느 하나의 방법에 사용하기 위한 단락 73의 약제학적 조성물.
- [0759] 117. 단락 107 내지 111 중 어느 하나의 세포를 포함하는 숙주 동물.
- [0760] 118. 단락 1 내지 72 중 어느 하나의 재조합 AAV 벡터를 포함하는 숙주 동물.
- [0761] 119. 단락 118에 있어서, 숙주 동물은 포유동물인, 숙주 동물.
- [0762] 120. 단락 118 또는 119에 있어서, 숙주 동물은 비인간 포유동물인, 숙주 동물.
- [0763] 121. 단락 119에 있어서, 숙주 동물은 인간인, 숙주 동물.
- [0764] **실시예**
- [0765] 다음의 비제한적인 실시예는 현재 고려되는 대표적인 구현예의 보다 완전한 이해를 쉽게 하기 위해 예시 목적으로만 제공된다. 이들 실시예는 AAV 비리온 및 rAAV 벡터가 사용될 수 있는 모든 가능한 상황의 단순한 서브세트인 것으로 의도된다. 따라서, 이들 실시예는 AAV 비리온 및 rAAV 벡터 및/또는 이의 방법 및 용도에 관한 것들을 포함하여 본 명세서에 기재된 임의의 구현예를 제한하는 것으로 해석되어서는 안 된다. 궁극적으로, AAV 비리온 및 벡터는 유전자 전달이 필요한 거의 모든 상황에서 사용될 수 있다.
- [0766] **실시예 1: rAAV 게놈의 작제**
- [0767] 수많은 rAAV 게놈은 깊은 클로닝 방법론을 사용하여 작제되었다. 하기 rAAV 게놈이 생성되었다: AAV_LVR412_EU(서열 번호 154), ssAAV_LVR412WT-hGAA_AskBio_CHATHAM(서열 번호 155), AAV-LVR412스터퍼(서열 번호 156), AAV_LVR422_EU(서열 번호 157), AAV-LVR422스터퍼(서열 번호 158), ssAAV_LVR412-WT-hGAA CHATHAM(서열 번호 159), ssAAV_LSP-WT-hGAA-CHATHAM(서열 번호 160), 서열 번호 57(AAT-V43M-wtGAA(델타1-

69aa)); 서열 번호 58(ratFN1-IGF2V43M-wtGAA(델타1-69aa)); 서열 번호 59(hFN1-IGF2V43M-wtGAA(델타1-69aa)); 서열 번호 60(AAT-IGF2D2-7-wtGAA(델타 1-69)); 서열 번호 61(FN1rat-IGF Δ 2-7-wtGAA(델타 1-69)); 서열 번호 62(hFN1-IGF Δ 2-7-wtGAA(델타 1-69)). 일부 rAAV 벡터는 wtGAA 폴리펩타이드를 암호화하는 핵산 서열을 포함하지만, wtGAA를 본원에 개시된 바와 같은 변형된 GAA 핵산 서열을 암호화하는 핵산 서열로 용이하게 대체할 수 있다.

[0768] 김슨 클로닝은 핵산 서열의 블록(예를 들어, 3개 블록)을 함께 클로닝하는 것을 포함한다. 일반적인 프로토콜은 다음과 같다: 다음 시약은 단일-튜브 반응으로 결합된다 (i) 김슨 조립 마스터 믹스(Gibson Assembly Master Mix) (엑소뉴클레아제, DNA 폴리머라제, DNA 리가제, 완충액) (ii) 15-25 bp의 상동 말단을 갖는 DNA 삽입물(블록 1-3) (도 6 참조) (iii) 가장 바깥쪽 DNA 삽입물에 대한 15-25 bp의 상동 말단을 갖는 선형화된 DNA 백본(도 6 참조). 반응은 50°C에서 15-60분 동안 배양된다. 반응 혼합물을 적격 세포로 형질전환하고 카나마이신 한천 플레이트에 플레이팅한다. 완전히 조립된 플라스미드 DNA의 미니프렙은 제한 분해 및/또는 콜로니 PCR 분석을 통해 스크리닝되고 DNA 시퀀싱 분석으로 확인된다. 확인된 클론은 맥시프렙 생산을 위해 확장되고 아테노바이러스 헬퍼, XX680 Kan 및 rAAV를 생성하기 위한 적절한 Rep/Cap 헬퍼와 함께 rAAV 생산자 세포주에서 일시적으로 형질감염된다. 도 6은 예시적인 rAAV 계놈을 생성하기 위한 클로닝 핵산 블록을 나타낸다.

[0769] 도 7-9는 wtGAA(Δ 1-69)가 예시적인 GAA 효소를 나타내지만, 이 핵산 서열은 생체내 향상된 발현을 위해 및/또는 면역 반응을 감소시키기 위해 및/또는 CpG 섬을 감소시키기 위해 및/또는 H201L 변형된 코돈 최적화된 GAA를 암호화하는 핵산 서열로 당업자에 의해 쉽게 대체될 수 있다. 따라서, 도 9a-9e는 야생형 GAA(wtGAA)를 갖는 예시적인 작제물을 나타내지만, wtGAA를 암호화하는 핵산을 본원에 개시된 바와 같은 변형된 GAA 핵산 서열, 예를 들어, 서열 번호 182를 암호화하는 핵산 서열로 쉽게 대체할 수 있다. 또한, 도 7-8에 예시된 클로닝 블록에는, IGF2(V43M) 또는 IGF2D2-7 표적화 펩타이드를 암호화하는 핵산 서열의 3' 및 GAA 효소를 암호화하는 핵산의 5'에 위치한 3 아미노산(3aa) 스페이스 핵산 서열, 및 스테퍼 핵산 서열, 폴리A 서열의 3' 및 3'ITR 서열의 5'에 위치한 스테퍼 서열(도 7-8에서 "스페이스" 서열로 지칭됨)인 rAAV 계놈의 생성을 나타낸다.

[0770] 실시예 2: rAAV 벡터 생성

[0771] rAAV 생성 세포주를 사용하여 rAAV 벡터를 생성하기 위해 rAAV 계놈을 캡시드에 패키징하였다. rAAV 벡터 작제의 원리를 증명하기 위해서만 사용된 캡시드는 AAV3b 캡시드였다.

[0772] rAAV 생성 세포주에서 rAAV 제조: 삼중 형질감염 기술을 사용하여 현탁 rAAV 생산자 세포주에서 rAAV를 제조하였으며, 이는 임상 등급 벡터를 제조하기 위해 확장될 수 있다. 대안적으로, 상이한 플라스미드를 사용될 수 있으며, 예를 들어, 1) pXX680 - ad 헬퍼 및 2) pXR3, Rep 및 Cap 및 3) 전이유전자 플라스미드 (ITR-전이유전자-ITR).

[0773] 실시예 1에서 생성된 rAAV 계놈은 그 전체가 본원에 참조로 포함된 미국 특허 9,441,206에 기술된 방법에 따라 rAAV 생성 세포주를 사용하여 rAAV 벡터를 생성하는 데 사용된다. 특히, rAAV 벡터 또는 rAAV 비리온은 (a) rAAV 생성 세포주에 AAV 발현 시스템을 제공하는 단계; (b) AAV 입자가 생성되는 조건 하에 세포를 배양하는 단계; 및 (c) 임의로 AAV 입자를 단리하는 단계를 포함하는 방법을 사용하여 생성된다. rAAV 벡터 생성을 위한 최적의 플라스미드 비율을 결정하기 위해 XX680, AAV rep/cap 헬퍼 및 TR 플라스미드의 다양한 플라스미드 비율을 사용하여 플라스미드의 삼중 형질감염 및 형질감염 각테일 부피의 비율이 최적화될 수 있다.

[0774] 일부 경우에, 세포는 AAV 입자가 생성되는 조건 하에 현탁액에서 배양된다. 또 다른 구현예에서, 세포는 동물 성분이 없는 조건에서 배양된다. 동물 성분이 없는 배지는 rAAV 생산자 세포주와 상용성인 임의의 동물 성분이 없는 배지(예를 들어, 무혈청 배지)일 수 있다. 예에는, 제한 없이, SFM4Transfx-293(Hyclone), Ex-세포 293(JRH Biosciences), LC-SFM(Invitrogen), 및 Pro293-S (Lonza)를 포함한다. AAV 입자의 복제 및 패키징에 충분한 조건은, 예를 들어, 본원에 기재된 rAAV 계놈의 복제에 충분한 AAV 서열의 존재 및 AAV 캡시드(예를 들어, AAV rep 서열 및 AAV 캡 서열) 및 아테노바이러스 및/또는 헤르페스바이러스로부터의 헬퍼 서열로의 캡시드 화일 수 있다.

[0775] 플라스미드 백본으로부터의 박테리아 DNA 서열은 재조합 AAV 벡터의 제조 동안 AAV 캡시드로 패키징되어 TLR9와의 상호작용을 통해 선천성 면역계의 활성화를 유도할 수 있다(Akira, 2006; Chadeuf, 2005; Wright, 2014). 다양한 기술을 사용하여 재조합 AAV 제조에서 플라스미드 백본 서열, 예를 들어, 확장성이 제한된 미니서클을 제거할 수 있다(Schnodt, 2016). 플라스미드 백본에서 박테리아 DNA 서열을 피하는 또 다른 방법은 재조합 AAV 벡터의 구체적 제조를 위한 도기본(doggy bone) DNA(dbDNA™)를 포함하지만 이에 제한되지 않는 다양한 DNA 복

제 기술을 포함하는 폐쇄 말단 선형 듀플렉스 DNA를 사용하는 것이다. dbDNA™와 같은 폐쇄 말단 선형 듀플렉스 DNA를 사용하면 박테리아 백본이 제거되며 백신 및 렌티바이러스를 생성하는 데 사용되었으며(Walters et al, 2014; Scott et al, 2015; Karda et al, 2019) DNA 백신 개발자에 의해 TLR9 반응을 촉발할 수 없는 것으로 나타났다.

[0776] 따라서, 대안적인 구현예에서, 본원에 개시된 바와 같은 방법 및 조성물에 사용하기 위한 rAAV 벡터의 생성은 그 전체가 본원에 참조로 포함된 미국 출원 2018/0037943 및 문헌(Karbowniczek et al., Bioinsights, 2017)에 개시된 바와 같은 도기본 기술(dbDNA™)을 포함하지만 이에 제한되지 않는 폐쇄 말단 선형 듀플렉스 DNA를 사용하여 수행될 수 있다. 간단히 말해서, 폐쇄 말단 선형 듀플렉스 DNA 기술을 사용하는 AAV 생성을 위한 플라스미드는 ITR, 프로모터 및 관심 있는 유전자를 포함할 수 있으며, 예를 들어, 본원 개시된 바와 같은 GAA는 56 bp 팔린드로믹 프로텔로머라제 인식 서열에 의해 플랭킹된다. 플라스미드가 변성되고 Phi29 DNA 폴리머라제 및 적절한 프라이머가 존재하는 경우 Phi29는 롤링 서클 증폭(RCA)을 시작하여 원래 작제물의 이중 가닥 칸카타머 반복체를 생성한다. 프로텔로머라제가 추가되면, 팔린드로믹 프로텔로머라제 인식 서열의 결합이 발생하고 절단-결합 반응이 발생하여 단량체 이중 가닥(ds) 선형 공유 폐쇄 DNA 작제물이 생성된다. 일반적인 제한 효소의 추가는 원하지 않는 DNA 플라스미드 백본 서열을 제거하고 엑소뉴클레아제 활성으로 분해하여 ITR, 프로모터 및 관심 있는 유전자를 암호화하는 dbDNA 서열을 단리하기 위해 크기를 분획화할 수 있는 dbDNA를 생성한다. dbDNA™ 기술과 같은 폐쇄 말단 선형 듀플렉스 DNA를 사용하는 rAAV 벡터를 생성하기 위한 예시적인 플라스미드는 하기 5'에서 3' 방향으로, 5'-프로텔로머라제 RS, 5'ITR, LSP 프로모터, hGAA, 3'UTR, hGH 폴리(A), 3'ITR, 3'-프로텔로머라제 RS(센스 가닥)를 포함하며, 여기서, 센스 가닥은 가닥(ds) 선형 공유 폐쇄 DNA 작제물에 대한 상보적 안티센스 가닥에 연결된다. 본원에 개시된 바와 같은 방법 및 조성물에 사용하기 위한 AAV 벡터의 제조를 위한 출발 물질로서의 폐쇄 말단 선형 듀플렉스 DNA, 예를 들어, 도기본 DNA(dbDNA™)의 사용은 AAV 벡터를 포함하는 플라스미드를 전파하는 데 사용되는 박테리아 백본을 제거하고 생성물이 톨형 수용체 9(TLR9) 반응을 촉발하지 못한다. 제조에 폐쇄 말단 선형 듀플렉스 DNA 기술을 사용하면 품폐병을 가진 환자에서 1.6E13 vg/kg의 용량에서 관찰되는 간 효소 상승 위험이 더욱 감소한다.

[0777] **실시예 3: rAAV 벡터 평가**

[0778] **전혈 제거.** 도 1은 상이한 AAV 혈청형(AAV3b, AAV3ST, AAV8, AAV9)의 3×10^{12} vg/kg을 3 kg의 혈청음성 수컷 마카크(macaque)에 정맥내 주사한 실험으로부터 유도된 결과를 나타낸다. 상이한 AAV 혈청형을 투여한 지 60일 후에 마카크를 안락사시켰다. 벡터 게놈은 전혈에서 검색되었으며 결과는 AAV3b가 일주일 이내에 제거되었고 희생 시 검출할 수 없었던 반면, AAV8 및 AAV9는 마카크가 희생되었을 때 전혈에서 여전히 검출 가능하다는 것을 나타내었다.

[0779] **간 특이적 벡터 효능:** 도 2는 3×10^{12} vg/kg의 상이한 AAV 혈청형(AAV3b, AAV3ST, AAV8, AAV9)을 3 kg의 혈청음성 수컷 마카크에 정맥내 주사한 실험으로부터 유도된 결과를 나타낸다. 상이한 AAV 혈청형을 투여한 지 60일 후에 마카크를 안락사시켰다. 벡터 게놈은 각각의 마카크의 간의 3개 엽 각각에서 정량화되었다. 정량 한계는 0.002 vg/dg이었다. 도 2에 제시된 결과에 기반하여, AAV3b는 강력한 간 벡터인 것으로 밝혀졌다. AAV3b는 AAV8보다 간 특이적이며 AAV9보다 더 신속하게 혈액으로부터 제거된다. AAV3ST 돌연변이는 어떠한 유의한 유익한 영향도 제공하지 않았다.

[0780] **실시예 4: 상청액으로의 GAA 분비 측정 및 GAA 흡수 검정**

[0781] **상청액에서 GAA 측정**

[0782] 따라서, 실시예 1에서 생성된 rAAV 게놈은 상청액으로의 GAA 폴리펩타이드의 분비에 대해 시험된다. 상청액에서 GAA 측정은 문헌[Kikuchi et al. (Kikuchi, Tateki, et al. "Clinical and metabolic correction of Pompe disease by enzyme therapy in acid maltase-deficient quail." The Journal of clinical investigation 101.4 (1998): 827-833.)]에 기재된 바와 같이 4-메틸-옴벨리페틸-알파-D-글루코시드(4-MU) 기질(4-MU 검정)을 사용하여 평가할 수 있다.

[0783] 간단히 말해서, rAAV 생산자 세포주는 rAAV 게놈 AAV_LVR412_EU(서열 번호 154), ssAAV_LVR412WT-hGAA_AskBio_CHATHAM(서열 번호 155), AAV-LVR412스터퍼(서열 번호 156), AAV_LVR422_EU(서열 번호 157), AAV-LVR422_스터퍼(서열 번호 158), ssAAV_LVR412_WT-hGAA CHATHAM(서열 번호 159), ssAAV_LSP_WT-hGAA-CHATHAM(서열 번호 160), 서열 번호 57(AAT-V43M-wtGAA(델타1-69aa)); 서열 번호 58(ratFN1-IGF2V43M-wtGAA(델타1-

69aa)); 서열 번호 59(hFN1-IGF2V43M-wtGAA(델타1-69aa)); 서열 번호 60(AAT-IGF2D2-7-wtGAA(델타 1-69)); 서열 번호 61(FN1rat-IGF Δ 2-7-wtGAA(델타 1-69)); 서열 번호 62(hFN1-IGF Δ 2-7-wtGAA(델타 1-69))로 형질감염될 수 있다. GAA 활성은 24시간에 걸쳐 초기 활성(t=0)의 %를 기반으로 측정된다. 샘플은 0, 3, 6 및 24시간에 형광 기질 4-MU- α -글루코스의 가수분해를 기반으로 하여 GAA 효소 활성에 대해 검정되었다. GAA 활성은 초기 활성의 %, 즉 잔류 활성으로 표시되었다.

[0784] 대안적으로, 수확 후, 배양 상청액은 HIC 크로마토그래피에 의해 부분적으로 정제되었다. 모든 샘플은 전기영동 전에 PNGase로 처리되었다. 세포에 의한 GAA 폴리펩타이드의 발현은 SDS-PAGE 및 면역블롯팅을 사용하여 평가될 수 있다.

[0785] 조직에서의 GAA 흡수 검정 및 GAA의 흡수 측정

[0786] 다음으로, 실시예 1 및 2에서 생성된 rAAV 게놈은 세포로의 흡수 활성의 보유에 대해 시험된다. 예를 들어, rAAV 생산자 세포주는 rAAV 게놈 AAV_LVR412_EU(서열 번호 154), ssAAV_LVR412WT-hGAA_AskBio_CHATHAM(서열 번호 155), AAV-LVR412스터퍼(서열 번호 156), AAV_LVR422_EU(서열 번호 157), AAV-LVR422_스터퍼(서열 번호 158), ssAAV_LVR412_WT-hGAA CHATHAM(서열 번호 159), ssAAV_LSP_WT-hGAA-CHATHAM(서열 번호 160), 서열 번호 57(AAT-V43M-wtGAA(델타1-69aa)); 서열 번호 58(ratFN1-IGF2V43M-wtGAA(델타1-69aa)); 서열 번호 59(hFN1-IGF2V43M-wtGAA(델타1-69aa)); 서열 번호 60(AAT-IGF2D2-7-wtGAA(델타 1-69)); 서열 번호 61(FN1rat-IGF Δ 2-7-wtGAA(델타 1-69)); 서열 번호 62(hFN1-IGF Δ 2-7-wtGAA(델타 1-69))로 형질감염될 수 있다.

[0787] 4-MU 검정(상기 기재된 바와 같음)은 포유동물 세포로의 rhGAA 흡수를 평가하기 위한 것일 수 있으며, 그 전체가 본원에 참고로 포함된 미국 특허 출원 US2009/0117091A1에 기재되어 있다. 실시예 1 및 2에서 생성된 rAAV 벡터 또는 rAAV 게놈은 10 mM 4-메틸움벨리페릴 α -D-글루코시다아제 기질(Sigma, 카탈로그 #M-9766)과 함께 123 mM 아세트산나트륨 pH 4.0을 함유하는 20 μ l 반응 혼합물에서 인큐베이션된다. 반응을 37°C에서 1시간 동안 인큐베이션하고 267 mM 탄산나트륨, 427 mM 글리신, pH 10.7을 함유하는 200 μ l 완충액으로 중단시켰다. 형광은 96웰 미세액가 플레이트에서 355 nm 여기 및 460 nm 필터로 측정하고 4-메틸움벨리페론(Sigma, 카탈로그 #M1381)으로부터 유래된 표준 곡선과 비교하였다. 1 GAA 4 MU 단위는 시간당 가수분해된 1 nmole 4-메틸움벨리페론으로 정의된다. 섬유아세포에서 예시적인 rAAV 게놈, 예를 들어, AAV_LVR412_EU(서열 번호 154), ssAAV_LVR412WT-hGAA_AskBio_CHATHAM(서열 번호 155), AAV-LVR412스터퍼(서열 번호 156), AAV_LVR422_EU(서열 번호 157), AAV-LVR422_스터퍼(서열 번호 158), ssAAV_LVR412_WT-hGAA CHATHAM(서열 번호 159), ssAAV_LSP_WT-hGAA-CHATHAM(서열 번호 160), 서열 번호 57(AAT-V43M-wtGAA(델타1-69aa)); 서열 번호 58(ratFN1-IGF2V43M-wtGAA(델타1-69aa)); 서열 번호 59(hFN1-IGF2V43M-wtGAA(델타1-69aa)); 서열 번호 60(AAT-IGF2D2-7-wtGAA(델타 1-69)); 서열 번호 61(FN1rat-IGF Δ 2-7-wtGAA(델타 1-69)); 서열 번호 62(hFN1-IGF Δ 2-7-wtGAA(델타 1-69))가 평가된다. IGF2-GAA 융합 폴리펩타이드 및/또는 SS-IGF2-GAA 이중 융합 폴리펩타이드의 효소 활성을 평가하고 그 자체로(즉, 이중 신호 펩타이드 또는 IGF2 표적화 펩타이드 없이)GAA(wtGAA) 폴리펩타이드와 비교한다.

[0788] 표적 세포에 들어가는 IGF2 태그된 또는 태그되지 않은 GAA의 능력을 입증하기 위해 세포 기반 흡수 검정을 수행할 수도 있다. 래트 L6 근모세포는 흡수 24시간 전에 24웰 플레이트에서 웰당 1×10^5 개 세포의 밀도로 플레이트된다. 실험 시작 시, 배지를 세포로부터 제거하고 실시예 1 및 2에서 생성된 rAAV 벡터를 함유하는 0.5 ml의 흡수 배지로 교체하였다. 흡수의 특이성을 입증하기 위해, 일부 웰에는 경쟁자 M6P(5 mM 최종 농도) 및/또는 IGF2(18 μ g/ml 최종 농도)가 추가로 포함되었다. 18시간 후, 배지를 세포에서 흡인하고, 세포를 PBS로 4회 세척하였다. 이어서, 세포를 200 μ l CellLytic MTM 용해 완충액으로 용해한다. 용해물은 4 MU 기질을 사용하여 상기 설명한 대로 GAA 활성에 대해 검정한다. Pierce BCATM 단백질 분석 키트를 사용하여 단백질을 측정한다.

[0789] 다른 세포주 및 근모세포 세포주가 사용될 수 있지만, 전형적인 흡수 실험은 CHO 세포에서 수행된다. 래트 L6 근모세포로의 GAA 폴리펩타이드의 흡수는 큰 물 과량의 M6P의 첨가에 의해 사실상 영향을 받지 않을 것으로 예상되는 반면, 흡수는 과량의 IGF2에 의해 상당히 제거될 것으로 예상된다. 대조적으로, wtGAA의 흡수는 과량의 M6P를 추가함으로써 상당히 없애지만 IGF2와의 경쟁에 의해 실질적으로 영향을 받지 않을 것으로 예상된다. 또한, IGF2V43M-wtGAA 및 IGF델타2-7wtGAA의 흡수는 과량의 IGF2에 의해 상당히 영향을 받지 않을 것으로 예상된다.

[0790] **실시예 5** 래트 L6 근모세포에서 GAA의 반감기

[0791] 래트 L6 근모세포에서 실시예 1 및 2에서 생성된 rAAV 벡터를 사용하여 상기 기재된 바와 같이 흡수 실험을 수행하였다(실시예 3 및 4 참조). 18시간 후, rAAV 벡터로 형질감염된 세포로부터의 배지를 흡인 제거하고 세포

를 PBS로 4회 세척하였다. 이때, 중복 웰을 용해시키고(시간 0) 용해물을 -80에서 동결시켰다. 그 후, 매일 중복 웰을 용해하고 분석을 위해 저장하였다. 14일 후, 모든 용해물을 GAA 활성에 대해 검정하여 반감기를 평가하고 일단 세포 내부에서 IGF2 태그된 GAA 효소가 태그되지 않은 GAA와 유사한 역학으로 지속되는지 평가하였다.

[0792] **실시예 6:** 흡수 후 GAA의 처리

[0793] 포유류 GAA는 전형적으로 문헌(Moreland et al. (2005) J. Biol. Chem., 280:6780-6791) 및 이에 포함된 참조 문헌에 의해 기재된 바와 같은 리소솜에서 순차적인 단백질분해 처리를 거친다. 처리된 단백질은 70 kDa, 20 kDa, 10 kDa의 펩타이드 및 일부 더 작은 펩타이드의 패턴을 생성한다. IGF2-GAA 융합 폴리펩타이드 및/또는 SS-IGF2-GAA 이중 융합 폴리펩타이드가 태그되지 않은 GAA와 유사하게 처리되는지 여부를 결정하기 위해, 상기 흡수 실험으로부터 용해물의 분취물을 70 kDa IGF2 펩타이드 및 IGF2 태그가 있는 더 큰 중간체를 인식하는 모노클로날 항체를 사용하여 웨스턴 블롯으로 분석하였다. 이 실험에서 확인된 폴리펩타이드의 유사한 프로파일은 일단 세포에 들어가면, IGF2 표적화 펩타이드가 손실되고 IGF2-GAA 폴리펩타이드는 태그되지 않은 GAA와 유사하게 처리되며, 이는 IGF2 표적화 펩타이드가 일단 세포 내부에 있으면 GAA의 거동에 거의 또는 전혀 영향을 미치지 않는다는 것을 입증한다.

[0794] **실시예 7:** 약동학

[0795] rAAV 벡터에 의해 생성된 IGF2-GAA 융합 폴리펩타이드 및/또는 SS-IGF2-GAA 이중 융합 폴리펩타이드의 약동학은 야생형 129 마우스에서 측정될 수 있다. 129마리 마우스에 실시예 1 및 2에서 생성된 rAAV 벡터를 주사하였다. 혈청 샘플을 주사 전 및 주사 후 15분, 30분, 45분, 60분, 90분, 120분, 4시간 및 8시간에 채취한다. 이어서, 동물을 희생시킨다. 혈청 샘플을 정량적 웨스턴 블롯으로 검정한다. IGF2-GAA 융합 폴리펩타이드 또는 SS-IGF2-GAA 이중 융합 폴리펩타이드를 발현하는 rAAV 벡터로부터의 GAA에 대한 반감기는 IGF2 융합된 GAA 폴리펩타이드가 순환으로부터 과도하게 빠르게 제거되는지를 결정하기 위해 평가된다.

[0796] **실시예 8:** GAA의 조직 반감기

[0797] 이 실험의 목적은 rAAV 벡터로부터 발현된 IGF2-GAA 융합 폴리펩타이드 또는 SS-IGF2-GAA 이중 융합 폴리펩타이드의 표적 조직에 도달하면 GAA 활성이 손실되는 속도를 결정하는 것이었다. 폼페 마우스 모델에서, MYOZYME®은 다양한 근육 조직에서 약 6-7일의 조직 반감기를 갖는 것으로 보인다(약물 평가 및 연구 센터 및 생물학적 제제 평가 및 연구 센터, 약리학 리뷰에 대한 출원 번호 125141/0).

[0798] 폼페 마우스[그 개시내용이 본원에 참조로 포함된 문헌(Raben (1998) JBC, 273:19086-19092)에 기재된 바와 같은 폼페 마우스 모델 6neo/6neo]의 경정맥에 실시예 1 및 2에서 생성된 rAAV 벡터를 주사한다. 이어서, 주사 1, 5, 10 및 15일 후에 마우스를 희생시킨다. 조직 샘플을 균질화하고 표준 절차에 따라 GAA 활성을 측정하였다. IGF2-GAA 융합 폴리펩타이드 및/또는 SS-IGF2-GAA 이중 융합 폴리펩타이드 및 태그되지 않은 GAA로부터의 GAA 활성의 조직 반감기는 상이한 조직(예를 들어, 대퇴사두근 조직, 심장 조직, 횡격막 조직; 및 간 조직)의 붕괴 곡선으로부터 계산되고, 각 조직의 반감기가 계산된다. 이것은 폼페 마우스의 세포 내부에 일단 들어가면 본원에 기재된 rAAV 벡터로부터 발현된 IGF2-GAA 융합 폴리펩타이드 및/또는 SS-IGF2-GAA 이중 융합 폴리펩타이드가 태그되지 않은 GAA와 유사한 역학으로 지속되는 것으로 나타나는지를 결정하기 위해 래트 L6 근모세포의 반감기와 비교할 수 있다. 또한, IGF2-GAA 융합 폴리펩타이드 및/또는 SS-IGF2-GAA 이중 융합 폴리펩타이드의 붕괴 역학에 대한 지식은 적절한 투여 간격을 설계하는 데 도움이 될 수 있다.

[0799] **실시예 9:** C2C12 마우스 근모세포의 리소솜으로의 IGF2-GAA 융합 폴리펩타이드

[0800] 및/또는 SS-IGF2-GAA 이중 융합 폴리펩타이드의 흡수

[0801] 폴리-라이신 코팅된 슬라이드(BD Biosciences)에서 성장한 C2C12 마우스 근모세포는 실시예 1 및 2에서 생성된 rAAV 벡터로 형질도입된다. 세포를 세척한 후, 세포를 이어서 성장 배지에서 1시간 동안 인큐베이션한 다음, 실온에서 15분 동안 메탄올로 고정하기 전에 D-PBS로 4회 세척하였다. 다음 인큐베이션은 모두 실온에서 이루어졌으며, 각각은 D-PBS에서 3회 세척으로 분리되었다. 슬라이드를 0.1% 트리톤 X-100으로 15분 동안 투과시킨 다음, 차단 완충액(D-PBS 중 10% 열-비활성화 말 혈청(Invitrogen)으로 차단된다. 슬라이드를 1차 마우스 모노클로날 항-GAA 항체 3A6-1F2(차단 완충액에서 1:5,000)와 함께 인큐베이션한 다음, 2차 토끼 항-마우스 IgG AF594 접합된 항체(Invitrogen A11032, 차단 완충액에서 1:200)와 함께 인큐베이션한다. FITC-접합된 래트 항-마우스 LAMP-1(BD Pharmingen 553793, 차단 완충액에서 1:50)을 인큐베이션한다. 슬라이드는 DAPI-함유 장착 용액(Invitrogen)으로 장착되고 플루오레세인 이소티오시아네이트, 텍사스 레드 및 DAPI 필터(Chroma

Technology)가 장착된 Nikon Eclipse 80i 현미경으로 관찰된다. MetaMorph 소프트웨어(Universal Imaging)로 제어되는 측광 캐스캐이드 카메라로 이미지를 캡처하고 Photoshop 소프트웨어(Adobe)를 사용하여 합할 수 있다. 항-GAA 항체에 의해 검출된 신호와 리소좀 마커, LAMP1에 대해 지시된 항체에 의해 검출된 신호의 공동-국소화는 IGF2-태그된 GAA가 리소좀으로 전달된다는 것을 입증하기 위해 평가될 수 있다.

[0802] 실시예 10 : 폼페 마우스 모델에서 rAAV 벡터의 치료 평가 및

[0803] 폼페 병리 역전

[0804] 실시예 1에서 생성된 rAAV 벡터는, 예를 들어, 그 전체가 본원에 참조로 포함된 문헌(Peng et al., "Reveglucosidase alfa (BMN 701), an IGF2-Tagged rhAcid α -Glucosidase, Improves Respiratory Functional Parameters in a Murine Model of Pompe Disease." Journal of Pharmacology and Experimental Therapeutics 360.2 (2017): 313-323))에 기재된 방법에 따라 폼페 마우스 모델에서 평가될 수 있다.

[0805] 임의의 폼페 마우스 모델을 사용하여 폼페 질환을 치료할 때 rAAV 벡터의 효과를 평가할 수 있다. 폼페의 한 마우스 모델은 분열된 GAA 마우스 모델을 설명하고 유아 및 성인 형태 둘 모두의 질환의 중요한 특징을 요약한 문헌(Raben et al., JBC, 1998; 273(30); 19086-19092)에 기재되어 있다. 다른 경우에, 폼페 마우스 모델(Sidman et al., 2008)과 산 α -글루코시다아제 유전자가 분열된 마우스 균주(B6;129-GAAtm1Rabn/J; Pompe)를 사용할 수 있다(Jackson Laboratory, Bar Harbor, ME). 폼페 마우스는 인간 성인 폼페병에서와 동일한 세포 및 임상 특성을 나타낸다(Raben et al., 1998). 동물은 12시간의 명암 주기로 유지되며, 신선한 물과 표준 실험용 음식이 임의로 제공된다.

[0806] 4.5-5개월령의 폼페 마우스는 본원에 기재된 rAAV 벡터를 투여받을 수 있고, 4주 이상 동안 투여 후 글리코겐 제거율에 대해 평가할 수 있다. 거시적 평가 후, 심장(좌심실), 대퇴사두근, 횡격막, 요근 및 가자미근(soleus muscle)을 수집하고, 칭량하고, 액체 질소에서 급속 냉동시키고, 글리코겐 유래 글루코스의 정량적 분석에 전에 -60 내지 90°C에 보관하였다. 근육은 세라믹 구체를 사용하여 얼음 위에서 완충액(0.2M NaOAc/0.5% NP40)에서 균질화되었다. 퍼옥시다제-글루코스 옥시다제 효소 반응 시스템(Sigma-Aldrich, St. Louis, MO)을 사용한 후속 비색 검출(430 nm, SpectraMax; Molecular Devices, Sunnyvale, CA)을 위해 글리코겐을 글루코스로 분해하기 위해 37°C에서 정화된 용해물에 아밀로글루코시다아제를 첨가하였다. 쌍을 이룬 샘플은 수확 시 글리코겐 형태가 아닌 내인성 조직 글루코스를 보정하기 위해 아밀로글루코시다아제 없이도 측정된다. 글루코스 값은 6점 표준 곡선으로부터 외삽되었다. 측정된 글루코스 농도(mg/ml)는 샘플의 글리코겐 농도에 비례하고 균질화 단계(조직 1 g당 5 μ l의 완충액 첨가)를 조정하여 mg 글리코겐/g 조직으로 변환된다.

[0807] 개별 마우스 근육 글리코겐 수준에 대한 본원에 기재된 rAAV 벡터의 효과는 Phoenix-WinNonlin 클래식 PD 모델링(Phoenix 빌드 버전 6.4; Certara, L.P., Princeton, NJ)을 사용하여 평가할 수 있다. 심장, 횡격막, 대퇴사두근, 요근, 및 가자미근에서 hGAA에 대한 결과를 얻을 수 있다. 약동학 분석을 위해, WT 마우스에 실시예 1에서 생성된 rAAV 벡터 및 투여 전, 투여 후 0.083, 0.5, 1, 2 및 4시간에 말단 심장 천자로서 수집된 혈액 샘플을 투여할 수 있다. 혈장 hGAA 농도는 LOQ가 100 ng/ml인 브리징 전기화학발광법을 사용하여 정량화할 수 있다. 간단히 말해서, 0.5 μ g/ml 루테튬-표지된 항-rhGAA(친화성 정제 염소 폴리클로날) 및 0.5 μ g/ml 비오틴-표지된 항-IGF2(MAB792; R&D Systems, Minneapolis, MN)는 완충액(Starting Block T20 (PBS); ThermoFisher Scientific, Sunnyvale, CA)에 1:10으로 희석된 K2EDTA 혈장 샘플과 합할 수 있으며 차단된 스트렙타비딘 검정 플레이트(Meso Scale Diagnostics, Rockville, MD)로 옮기기 전에 1시간 동안 인큐베이션한다. 30분 인큐베이션 후, 플레이트를 세척하고, 1 $\times$ 판독 완충액 T(Meso Scale Diagnostics)를 첨가하고, 전기화학발광 신호는 SECTOR Imager 2400(Meso Scale Diagnostics)에서 판독한다. hGAA 농도는 표준 곡선으로부터 외삽할 수 있다.

[0808] 대안적으로, 심장 및 횡격막 조직 균질물을 수확하고 형광 기질(4-MUG)을 사용하여 rhGAA 활성을 측정할 수 있다.

[0809] 본원의 실시예 1 및 2에서 생성된 rAAV 벡터를 사용하여 생성된 GAA 폴리펩티드의 치료 효과는 생체내에서 wt GAA와 비교될 수 있다. 실시예 1에 개시된 rAAV 벡터의 능력을 폼페 마우스의 골격근 조직으로부터 글리코겐을 제거하기 위해 태그되지 않은 wt GAA를 발현하는 능력과 비교하기 위한 연구가 수행될 수 있다(예를 들어, 폼페 마우스 모델 6neo/6neo 동물이 사용되었다(Raben (1998) JBC 273:19086-19092)). 폼페 마우스 그룹(5마리/그룹)은 실시예 1에서 생성된 wt GAA 또는 rAAV 벡터 또는 비히클의 2가지 용량 중 하나의 IV 주사를 투여받았다. 5마리의 미처리된 동물을 대조군으로 사용할 수 있으며, 식염수 용액을 매주 4회 투여받았다. 동물은 주사 2, 3 및 4시간 전에 1시간에 5 mg/kg의 경구 디펜히드로민을 투여받았다. 주사 후 1주일 후에 마우스를 희생시키

고 조직학적 및 생화학적 분석을 위해 조직(횡격막, 심장, 폐, 간, 가자미근, 대퇴사두근, 비복근, TA, EDL, 허)을 수확하였다. 조직 균질물에서 글리코젠 함량은 에이. 니거(A. niger) 아밀로글루코시다아제 및 Amplex 레드 글루코스 검정 키트를 사용하여 측정할 수 있으며, 표준 절차를 사용하여 상이한 조직 균질물에서 GAA 효소 수준을 평가하였다.

[0810] 조직 균질물에서 글리코젠 함량은 에이. 니거 아밀로글루코시다아제 및 Amplex® 레드 글루코스 검정 키트(Invitrogen)를 사용하여 본질적으로 문헌[Zhu et al. (2005) Biochem J., 389:619-628]에 기재된 바와 같이 측정할 수 있다.

[0811] 본원에 기재되고 실시예 1 및 2의 방법에 의해 생성된 rAAV 벡터 ss-IGF2-GAA rAAV는 wtGAA rAAV 벡터(즉, 이중 분비 신호 및 IGF2 표적화 펩타이드 중 하나가 없음) 및/또는 MYOZYME®보다 클 것으로 예상되는 IGF2-GAA rAAV(즉, 분비 신호 서열 없음)와 비교하여 폼페 마우스 모델에서 더 많은 분비에 이어 근육으로 흡수되고 더 큰 치료 효과를 가질 것으로 예상된다. 예를 들어, H201L 돌연변이가 있는 GAA를 포함하는, 본원에 개시된 바와 같은 변형된 GAA를 포함하는 rAAV 벡터가 비변형된 GAA, 예를 들어, wtGAA를 포함하는 rAAV보다 폼페 마우스 모델에서 치료학적으로 더 효과적인 것으로 추가로 예상된다. 확립된 폼페 모델을 감안할 때, 이러한 결과는 임상 적용되고 폼페병 치료에 대한 치료 효과와 상관관계가 있을 것으로 예상된다.

[0812] **실시예 11:** 생체내 글리코젠의 제거

[0813] 이 실험의 목적은 실시예 1 및 2에서 생성된 IGF2-GAA 융합 폴리펩타이드 및/또는 SS-IGF2-GAA 이중 융합 폴리펩타이드를 발현하는 rAAV 벡터의 단일 주사 후 폼페 마우스의 심장 조직에서 글리코젠이 제거되는 속도를 결정하는 것이다.

[0814] 폼페 마우스[그 개시내용이 본원에 참조로 포함된 문헌(Raben (1998) JBC, 273:19086-19092)에 기재된 바와 같은 폼페 마우스 모델 6neo/6neo]의 경정맥에 실시예 1 및 2에서 생성된 발현 rAAV 벡터를 주사한다. 주사 1, 5, 10 및 15일 후에 마우스를 희생시켰다. 심장 조직 샘플은 표준 절차에 따라 균질화하고 글리코젠 함량에 대해 분석한다. 이러한 조직 균질물에서 글리코젠 함량은 에이. 니거 아밀로글루코시다아제 및 Amplex® 레드 글루코스 검정 키트(Invitrogen)를 사용하여 본질적으로 문헌[Zhu et al. (2005) Biochem J., 389:619-628]에 기재된 바와 같이 측정한다. 마우스로부터의 심장 조직의 평가는 rAAV가 투여된 마우스와 비교하여 실시예 1 및 2에서 생성된 IGF2-GAA 융합 폴리펩타이드 및/또는 SS-IGF2-GAA 이중 융합 폴리펩타이드를 발현하는 rAAV 벡터가 투여된 마우스에서 글리코젠이 거의 완전한 제거가 있는지 결정할 수 있으며, 여기서, GAA는 본원에 기재된 바와 같이 IGF2 표적화 펩타이드 및/또는 SS에 융합되지 않았고, 글리코젠 함량의 작은 변화만이 최소 제거를 나타낼 것이다.

[0815] **실시예 12:** 예시적인 간 특이적 프로모터

[0816] 본원의 표 4로부터 선택되거나 본원의 표 4에 개시된 것들로부터 선택된 예시적인 간-특이적 프로모터는 pUC19와 본질적으로 동일한 성질을 갖는 백본을 갖는 벡터 내로 루시퍼라제 리porter 유전자에 이어 SV40 후기 폴리A 신호의 업스트림에서 클로닝되었다. 특히, 프로모터 SP0412 및 SP0422를 평가하는 실험을 rAAV 생성을 위한 작제물에 클로닝하였다. 플라스미드의 DNA 제제는 전사 활성을 평가하기 위해 Huh7(간세포 암종 세포주), HeLa(자궁경부암으로부터 유래한 불멸 세포주) 또는 HEK293(인간 배아 신장 세포)에 형질감염되었다. Huh-7 세포는 JCRB 세포 은행(JCRB0403)에서 공급했고, HeLa 및 HEK293은 ECACC 세포 은행에서 공급했다. 모든 세포주는 세포 은행의 권장 사항에 따라 성장 및 유지되었다.

[0817] 1:1.1의 DNA:FuGene HD 비율에서 FuGene HD 형질감염 시약(Promega #E2311)을 사용하여 48웰 플레이트에서 3중으로 형질감염을 수행하였다. 루시퍼라제 활성은 형질감염 24시간 후에 측정하였다. 세포를 인산염 완충 식염수(PBS)로 세척하고 100 μ l 수동 용해 완충액(Promega #E194A)에 용해시키고 -80°C에서 밤새 보관하였다. 루시퍼라제 활성은 96웰 평평한 바닥 고체 흰색 Microplate FluoroNunc 플레이트(ThermoFisher #236105)를 사용하여 10 μ l의 용해물에서 제조업체의 지침에 따라 루시퍼라제 리porter 1000 검정 시스템을 사용하여 정량화하고 발광은 FLUOstar Omega 플레이트 판독기(BMG Labtech) 기계에서 정량화하였다.

[0818] 상기 루시퍼라제 방법은 당업계에서 통상적이고 유사한 기술이, 예를 들어, 문헌[Alam and Cook, "Reporter Genes: Application to the Study of Mammalian Gene Transcription", ANALYTICAL BIOCHEMISTRY 188,245-254 (1990)]에 광범위하게 기재되어 있다.

[0819] 데이터

- [0820] 사용된 예시적인 간-특이적 프로모터의 서열은 본원의 표 4에 나타내었다. 간 세포에서 발현을 유도하는 이러한 합성 간-특이적 프로모터의 능력은 편재하는 CMV_{IE} 및 CBA 프로모터에 대해 그리고 알려진 간 특이적 프로모터 LP1에 대해서도 벤치마킹되었다. 본 발명에 따른 모든 합성 프로모터는 Huh7 세포에서 LP1 프로모터보다 더 높은 활성을 나타내었다(데이터는 나타내지 않음). 이들 프로모터가 비-간 유래 HEK293 및 HeLa 세포에서 역 스크리닝되었을 때, 유비쿼터스 활성 프로모터 CMV_{IE} 및 CBA와 비교하여 무시할만한 활성을 나타냈다(데이터는 나타내지 않음).
- [0821] 간-특이적 프로모터(즉, 표 4에 제시된 모든 프로모터)의 활성을 또한 본질적으로 상기 기재된 바와 같은 재료 및 방법을 사용하여 Huh7 세포에서 시험하였다. 그러나, 이 경우, TBG가 LP1보다 더 높고 더 일관된 시험관내 발현을 갖는 것으로 밝혀졌기 때문에, 간-특이적 프로모터의 활성을 프로모터 TBG(서열 번호 435)의 활성과 비교하였다. TBG는 매우 강력한 간-특이적 프로모터이므로 TBG보다 적은 발현을 나타내는 프로모터는 여전히 매우 유용할 수 있다는 점에 주목해야 한다. 특히, 표 4에 개시된 간-특이적 프로모터 또는 TBG보다 짧지만 여전히 높은 수준의 활성(예를 들어, 서열 번호 435의 TBG 활성의 15%, 보다 바람직하게는 25%, 50%, 또는 75% 이상)을 나타내는 이의 기능적 변이체가 특히 중요하다.
- [0822] 간 세포에 대한 간-특이적 프로모터의 특이성은 또한 실시예 2에 기재된 물질 및 방법을 사용하여 비-간 HEK293 세포를 사용하여 시험하였다. 간-특이적 프로모터의 활성은 CMV-IE(서열 번호 433)의 활성과 비교하여 표시된다(TBG 및 LP1은 간-특이적이므로 HEK293 세포에서 특히 활성이 아니다). HEK293 세포에서 시험된 간-특이적 프로모터의 특이성을 나타내는 그래프에서 '상대 활성'은 명명된 프로모터의 활성을 CMV-IE의 활성에 대한 비율로 표현하고, 여기서, 1은 CMV-IE 프로모터(서열번호 433)와 동일한 활성이고, 1 초과는 CMV-IE와 비교하여 더 높은 활성이고 1 미만은 서열 번호 433의 CMV-IE 프로모터와 비교하여 더 낮은 활성이다.
- [0823] **실시예 13:** 생체내 예시적인 간 특이적 프로모터로부터의 발현
- [0824] 특정 간-특이적 프로모터는 생체내 평가를 위한 예시적인 간-특이적 프로모터로서 선택되었다. 특히, 최소 프로모터 CRM_SP0239(서열 번호 87) 및 CRM_SP0412(서열 번호 86) 및 합성 간-특이적 프로모터 SP0239(서열 번호 93), SP0412(서열 번호 91) 및 SP0422(서열 번호 92)는 생체내에서 평가되었다(도 10 및 도 13a-13b 참조).
- [0825] **AAV 생성:**
- [0826] 본 발명에 따른 프로모터 서브세트의 활성을 생체내에서 시험하였다. 이 연구에 포함된 합성 프로모터는 SP0239, SP0244, SP0412 및 양성 대조군 LP1이었다. 사용된 리포터 유전자는 fLUC-T2A-EGFP, 즉 T2A 신호(양방향 자가-절단 펩타이드)를 통해 mEGFP(돌연변이 녹색 형광 단백질)에 융합된 fLUC(반딧불이 루시페라제)였다. pAAV_SYNP_Luc-T2A-GFP 지정 벡터는 ZsGreen1 리포터가 Luc-T2A-GFP 이중 리포터로 대체된 pAAV ZsGreen1(Clontech로부터 구매)로부터 유래된다. 모든 DNA 플라스미드는 제조업체 지침에 따라 QIAGEN Plasmid Mega 키트(Qiagen#12181, Germany)를 사용하여 제조되었다.
- [0827] rAAV 생산자 세포주는 둘베코의 변형된 이글 배지, 고 글루코스, 10% (v/v) 소 태아 혈청(Sigma# F7524,UK)이 보충된 GlutaMAX 보충제(Gibco)에서 145 mm의 조직 배양 처리된 배양 접시(Greiner Bio-One Ltd # 639160, UK)에서 배양하고, 37°C 및 5% CO₂에서 인큐베이션하였다. 세포 배양을 위한 다른 시약은 Invitrogen-UK로부터 구입하고 플라스틱 제품은 Life Technologies로부터 구입하였다.
- [0828] 이 연구에 사용된 모든 AAV 벡터는 AAV9 캡시드에서 위형이 지정되었다(pseudotyped). rAAV 생산자 세포주는 바이러스 증식(pDG9)을 가능하게 하는 헬퍼 기능을 암호화하는 플라스미드와 함께 리포터 유전자가 상이한 프로모터에 의해 제어되는 플라스미드로 공동-형질감염되었다. rAAV 생산자 세포주는 1:3(DNA:PEI)의 물비를 사용하여 스톡 농도(1 ug/ul)에서 폴리에틸렌이민(PEI) (Sigma-Aldrich# 764604, UK)을 사용하여 형질감염되었다.
- [0829] **AAV 정제 및 적정:**
- [0830] 형질감염 72시간 후, 세포를 용해시키고, 조 용해물을 여과한 다음, POROSTM CaptureSelectTM AAV 수지(Thermo Scientific, #A36739)를 함유하고 ÄKTAprime plus(고성능 액체 크로마토그래피-HPLC 시스템(GE Healthcare, #11001313))를 사용하여 HPLC 컬럼으로 정제하였다.
- [0831] 벡터 계층의 수는 QuantStudioTM 3 시스템 실시간 PCR(Thermo Fisher Scientific, UK)에서 Luna® Universal qPCR Master Mix(NEB# M3003, UK)의 제조업체 지침에 따라 정방향 프라이머(ACGCTGGGCTACTTGATC - 서열 번호 445), 역방향 프라이머(CGAGGAGGAGCTATTCTTG - 서열 번호 446) 및 프로브(TTTCGGGTCGTGCTCATG - 서열 번호

447)를 사용하여 표적 LUC 카세트에 대한 qPCR 적정에 의해 결정되었으며 데이터는 QuantStudio 설계 및 분석 소프트웨어 V1.4.1로 분석되었다.

[0832] 동물 절차:

[0833] 교배된 6주령 CD1 수컷 마우스를 Charles River-UK로부터 구입하였다. 이들은 일주일 동안 격리된 다음, 폐쇄된 환기 케이지로 옮겨졌고 최소 질환 시설에서 유지되었다. 이들은 케이지당 5마리의 마우스로 케이지에 가두었고 음식과 물을 자유롭게 섭취하여 체중으로 정규화되었다. 새로 수용된 마우스는 임의의 실험을 수행하기 전에 적응을 위해 일주일 더 주어졌다. 이 연구는 법적 내부부 권고; 규제, 윤리 및 라이선스 절차 하에; 그리고 동물(과학 절차)법 1986 및 University College London의 기관 지침에 따라 수행되었다.

[0834] 동물 주사:

[0835] AAV는 따뜻한 챔버(Thermo Fisher Scientific, UK)에서 의료용 공기(21% 산소) (Abbotts Laboratories, UK)에 공급된 2%-4% 이소플루란으로 마취된 8주령의 젊은 성체 수컷 CD1 마우스에 투여되었다. 인슐린 주사기: 27 G 1/2인치를 사용하여 마우스의 외측 꼬리 정맥에 1.0 ml(Fisher Scientific, UK)를 정맥내 주사하였다. 각 마우스에 200 μ l의 생리 식염수 용액의 최종 부피로 마우스당 8×10^4 AAV 바이러스 게놈의 AAV 벡터 용량을 주사한다. 이어서, 마우스를 케이지에 다시 넣기 전에 정상 온도로 되돌렸다.

[0836] 생물발광 이미징

[0837] 마우스를 매주 전신 생물발광 이미징에 적용했다. 적절한 경우, 마우스를 의료용 공기(21% 산소)로 공급된 2%-4% 이소플루란으로 마취하고 인슐린 주사기(Fisher Scientific, UK)를 사용하여 15 mg/mL의 D-루시페린 칼륨염 (Syd Labs # MB000102, US) 300 μ l를 복강내 주사하였다. D-루시페린 스톡은 생리 식염수(Gibco #14190-094, UK)에서 제조하였다. 냉각된 전하-결합 장치 카메라(IVIS Lumina II machine, Perkin Elmer, UK)를 사용하여 5분 후 마우스를 1초와 10초 사이에 이미지화하였다. 관심 있는 영역(ROI)은 IVIS Lumina Living image 4.5.5(Perkin Elmer)를 사용하여 측정되었으며 스테라디안당 제곱센티미터당 초당 광자(광자/초/cm²/sr)로 표현되었다.

[0838] 데이터

[0839] 이 연구의 결과는 도 10에 나타내었다. 결과는 각 그룹의 모든 시험된 동물에 대한 루시페라제 생물발광 강도, 총 플럭스(초당 광자)의 평균으로서 표현된다. 오차 막대는 평균의 표준 오차이다. '식염수' 그룹(n=10)에서, 동물에게 식염수만 주사했고 루시페라제 생물발광은 검출되지 않았다. 이 그룹은 음성 대조군이며 프로모터에 작동적으로 연결된 루시페라제가 주사되지 않은 경우 루시페라제 생물발광이 검출되지 않음을 나타낸다. 그룹 'LP1'(n=9)에서, 동물에 LP1 프로모터(서열 번호 432)에 작동적으로 연결된 루시페라제가 주사되었고 루시페라제 생물발광이 검출된다. 이 그룹은 양성 대조군이며 루시페라제가 LP1 프로모터의 제어 하에 발현되고 검출될 수 있음을 나타낸다.

[0840] 본 발명에 따른 간-특이적 프로모터의 활성을 시험하기 위해, 동물에 2개의 프로모터에 작동적으로 연결된 루시페라제를 포함하는 작제물을 주사하였다. 그룹 'SP0244'(n=8)에서, 루시페라아제는 SP0244 프로모터에 작동적으로 연결되어 있다. 그룹 'SP0239'(n=10)에서, 루시페라아제는 SP0239 프로모터(서열번호 93)에 작동적으로 연결된다.

[0841] 도 10으로부터 알 수 있는 바와 같이, 발현은 일반 간 특이적 프로모터 'LP1'(서열 번호 432)보다 더 높은 생물발광 강도로 나타낸 바와 같이 프로모터 'SP0244'(서열 번호: 366) 및 'SP0239'(서열 번호: 93)에서 가장 높았다. 따라서, 프로모터 SP0244 및 SP0239는 생체내에서 높은 활성을 나타내며 이들의 활성은 LP1의 활성보다 높다.

[0842] 도 13a 및 13b에서, GAA의 발현은 Huh7 세포 및 HEP2 세포에서 세포내 상이한 LSP; LSP NEW(서열 번호 160); 412 NEW(서열 번호 159); TTR NEW(서열 번호 155); 412-TTR, 422 스테퍼(서열 번호 158); 422 TTR; 412 스테퍼(서열 번호 156)를 포함하는 상이한 rAAV 벡터로부터 검출되었고, 이는 hGAA가 SP0412 및 SP0422 프로모터를 사용하여 유의한 수준으로 rAAV로부터 발현될 수 있음을 입증한다. 또한, 도 13a 및 13b에 나타낸 바와 같이, 임의의 상이한 AAV 벡터에 의한 GAA 발현은 일반적인 간 특이적 프로모터를 포함하는 AAV 벡터로부터의 GAA 발현보다 유의하게 더 높다. 중요하게는, 도 13a 및 13b는 도 13a 및 13b에서 "LSP SS"로 지칭되는 LP1 프로모터(서열 번호 432)를 포함하는 rAAV로부터의 hGAA 발현 수준과 비교하여 프로모터 412(서열 번호 91) 및 422(서열 번호 92)를 포함하는 rAAV 벡터로부터의 hGAA의 발현이 Huh7 및 HepG2 세포에서 hGAA의 높은 발현을 유도한다는

것을 나타낸다.

[0843] 이 실험은 실시예 12에서 얻은 시험관내 결과가 생체내에서 유의한 GAA 단백질 발현으로 달성될 수 있음을 입증한다. 다양한 강도를 갖는 다양한 간-특이적 프로모터가 본원에 개시된 바와 같은 방법 및 조성물에 사용될 수 있으며, 이는 질환, 예를 들어, 폼페병의 치료를 위한 치료 환경에서 원하는 수준의 간-특이적 발현을 제공하는 데 매우 유용할 수 있다.

[0844] **실시예 14: 폼페 환자의 임상 연구**

[0845] 3명의 대상체로 이루어진 2개의 코호트 각각은 LOPD가 있는 성인 대상체를 모집한 임상 연구 ACT-CS101에서 rAAV-LSP-hGAA를 투여받았다. 코호트 1 대상체는 1.6E12 vg/kg을 투여받았고 코호트 2 대상체는 1.6E13 vg/kg을 투여받았다. 모든 대상체는 효소 대체 요법(ERT) (알글루코시다아제 알파 [LUMIZYME®])으로부터 철회되었고; 코호트 1 대상체는 rAAV-LSP-hGAA 투여 후 6개월 동안 ERT에 남아 있는 반면, 코호트 2 대상체는 rAAV-LSP-hGAA 투여 시점에 ERT에서 철회되었다. AAV8-LSP-hGAA에 대한 작제물은 도 3b에 나타나며, AAV2 ITR을 포함하는 감염성, 비-복제 재조합 아데노-연관 바이러스(AAV) 혈청형 8 벡터, 또는 인간 산 알파-글루코시다아제(hGAA)를 발현하는 캡시드 8 및 예를 들어, AAV2로부터의 또 다른 캡시드 단백질을 포함하는 반수체 AAV 벡터(예를 들어, AAV2/8-LSP-hGAA)이다.

[0846] 예비 데이터는 2020년 6월 1일 현재 신속 보고 규칙을 충족하는 심각한 부작용이 없음을 나타낸다. 보고된 이상 반응은 대부분 예상/예견되었으며 중등도가 경증에서 중증도, 일시적이고 가역적이었다.

[0847] 코호트 1은 1.6E12 vg/kg을 투여받았는데 이는 3명의 대상체 모두가 내약성을 보였고, 이제 이들 모두는 투여 후 12개월이 지났고 생검 데이터를 곧 이용할 수 있다. 코호트 1(1.6E12 vg/kg)에 등록된 임의의 대상체에서 아스파르테이트 트랜스아미나제(AST) 또는 알라닌 트랜스아미나제(ALT)의 상승이 관찰되지 않았다. 모든 대상체는 4주 동안 60 mg/일로 시작하는 프레드니손으로 면역 예방을 받은 후 테이퍼링 용법(11주 동안 5 mg/주)을 받았다. 코호트 1에서 대상체는 캡시드 또는 산 알파-글루코시다아제(GAA)에 대해 양성 효소-결합 면역 흡수 반점(ELISPOT) 반응성을 나타내지 않았다. 모든 대상체는 각각의 기준 값보다 높은 혈청 GAA 농도를 나타냈다; 그러나 혈청 GAA 농도는 시간 경과에 따라 감소하고 있다. 아무도 ERT로 구조를 받을 필요가 없었다.

[0848] 코호트 2는 1.6E13 vg/kg을 투여받았다(2019년 8월 중순, 9월 중순 및 10월 초에 투여). AST 및 ALT의 무증상 상승은 46일(004), 41일(005) 및 55일(006)차에 시작하여 코호트 2의 모든 대상에서 관찰되었다. 모든 대상체는 60 mg/일에서 시작하는 프레드니손으로 면역 예방을 받았지만; 모든 대상체는 AST 및 ALT의 상승에 대한 반응으로 프레드니손의 증가된 용량을 투여받았다. 2명의 대상체(004 및 005)가 IV 메틸프레드니솔론을 투여받았다. 코호트 2에서 2명의 대상체는 CTCAE 3등급 실험실 상승을 보였다. 대상체 004는 ALT(348 U/L, ULN = 63, 5x ULN = 315)에서 한 번의 3등급 상승을 경험하였고 AST(222 U/L, ULN = 41, 5x ULN = 205)에서 한 번의 3등급 상승을 경험하였다. 3등급 상승은 둘 모두 동일한 날 발생했으며, 책임 연구자는 ACTUS-101과 확실히/아마도 관련이 있는 것으로 간주하였다. 대상체 5는 감마-글루타밀 트랜스퍼라제(GGT) (311 U/L; ULN = 50; 5x ULN = 250)에서 한 번의 3등급 상승을 경험하였으며, 이는 책임 연구자가 ACTUS-101과 관련이 있을 수 있지만 프레드니손과 관련이 있을 가능성이 더 크다고 생각했다. 모든 3등급 상승은 일시적이고 무증상이었다. 대상체 004는 3등급 실험실 상승 후 10일 이내에 간 생검을 받았으며, 그 결과 경증 및 비특이적 염증이 나타났다. 이 대상체는 또한 24주차에 지방종을 보고했으며, 이는 스테로이드 투여가 길고 오래 지속되었기 때문에, 이후 점차 감소했다. 모든 대상체는 AAV8 캡시드에 대해 양성이지만 다양한 수준의 ELISPOT T 세포 반응을 나타냈으며, 이는 언급된 간 효소 변화와 항상 일치하는 것은 아니다.

[0849] AAV8-LSP-hGAA의 실체, 강도, 품질, 순도 또는 효능에 대한 제안된 제조 변경의 효과 또는 그의 부재를 결정하기 위해 분석적 비교동등성(comparability) 연구를 수행하였다.

[0850] AAV8-LSP-hGAA의 분석적 비교동등성은 방법 표준 작동 절차(SOP) 하에 흡수 Absorption Systems Boston, LLC(ASB, Medford, MA)에서 실행되는 적격 시험관내 효능 검정을 사용하여 평가할 수 있으며, 상기 검정은 참조 표준에 대한 AAV8-LSP-hGAA 시험 샘플의 효능을 결정하기 위한 시험관내 검정이다. 간단히 말해서, 96웰 HuH-7 세포 배양 플레이트 포맷(인간 간세포 세포주, Sekisui XenoTech, 카탈로그 번호 JCRB0403)에서 6일 검정이 수행되고, 9-포인트 희석 계획을 사용하여 AAV8-LSP-hGAA로 형질도입되고, 96 ± 2시간 동안 인큐베이션된 후 4-MU 생물검정에 의해 세포 상청액에서 GAA가 측정된다. 상대 효능은 PLA 소프트웨어 3.0(Stegmann Systems)을 사용하는 평행선 검정 모델로 계산된다.

[0851] **용량 범위 연구:**

- [0852] AAV8-LSPHGAAs의 약력학적 효과를 뒷받침하기 위해 사용된 GAA 녹아웃(KO) 마우스에서의 용량 범위 연구. 수컷 및 암컷 GAA KO 품종 마우스는 꼬리 정맥(정맥내) 주사(볼루스)에 의해 8 내지 12주령(기관으로부터 피드백 9 내지 12주 전에 변형 1 제안됨)에 주사된다. 최소 6마리 수컷(M)과 6마리 암컷(F)은 문헌(Han et al (2017 and 2019))에 기반하여 2E11, 2E12 및 2E13 vg/kg으로 투여된 시험 제품(AAV8-LSPHGAAs) 중 하나를 투여받도록 무작위로 할당되고 대조군은 체형 완충제(비히클)만 투여받을 것이다. 동물을 8주 동안 추적한 다음 조직 및 혈액 수집을 위해 희생시킬 것이다. 실험 및 대조군은 오프레이터(들)에 대해 블라인드 상태로 유지될 것이다.
- [0853] 생활 모니터링은 사망물에 대한 2일 점검으로 구성될 것이다. 빈사 상태의 동물은 연구에서 제거될 것이다(조직병리학을 위한 전체 부검 조사 및 조직 샘플링으로). 혈청 GAA는 투여 후 8주째에 종료될 때까지 2주마다 모든 그룹에서 측정될 것이다.
- [0854] 부검 시 측정 및 조직 수집에는 간, 심장, 대퇴사두근, 전경골근(Tibialis A.), 횡격막, 뇌, 가자미근, 콩팥, 비장, 폐 및 다음을 포함하는 기타 임상 측정을 포함할 것이다: (1) 말기 체중, (2) 간 및 심장에 대한 기관 중량, (3) 항-GAA 항체(ADA)의 후속 측정을 위해 혈청(GAA 활성) 및 -20°C에서 보관, (4) 여러 단편으로 나뉜 간, 심장, 가자미근, 대퇴사두근, 전경골근, 허, 횡격막, 뇌. 추가 측정에는 다음을 포함한다: (4) 적격 검정에 의해 측정된, 혈청 및 조직의 GAA 활성 및 조직의 글리코겐 함량(조직은 액체 질소에서 급속 동결되고 -80°C에서 보관될 것임). 각 검정에 대해 ~50 mg의 지정된 각 조직의 표적이 수집될 것이며 (5) 부검 시 관찰된 모든 육안 병변뿐만 아니라 상기 표시된 조직의 헤모톡실린 및 에오신(H&E) 염색에 대해 10% 중성 완충 포르말린으로 고정된 조직병리학적 검사 분석. 필요에 따라 글리코겐 염색(파요오드산-Schiff 염색 - PAS), 항-GAA 항체 반응(ADA)을 측정하여 전이유전자 생성물에 대한 면역 반응의 평가, 이배체 계통 결정당 백터 계통 카피 수를 위한 총 DNA - DNA의 vg/ug(액체 질소에서 급속 동결한 다음, -80°C에서 보관)을 위해 조직의 하위섹션은 고정될 것이다.
- [0855] GAA KO 마우스에서의 용량 범위 연구는 다음 측정에서 유의한 차이가 관찰되지 않는 경우, AAV8-LSPHGAAs와 본원의 개시내용에 따라 변형된, 변형된 AAV 백터(예를 들어, 상이한 LSP, hGAA 서열의 최적화) 간의 생체내 비교동등성을 뒷받침할 것이다: 1. 수집된 모든 조직에서 입증된 수의병리학자에 의한 조직병리학(안전성); 2. AAV8-LSPHGAAs의 0.80과 1.25 사이(80-125%)의 변형된 AAV 백터의 상대 활성을 갖는 심장, 횡격막 및 대퇴사두근에서의 GAA 활성 및 글리코겐 함량(효능); 3. AAV8-LSPHGAAs의 0.80과 1.25 사이(80-125%)의 변형된 AAV 백터의 상대적 생체분포를 갖는 수집된 조직(생체분포)에서 ug DNA당 백터 계통 DNA; 4. 항-GAA ADA의 변경 없음.
- [0856] 본 발명이 다수의 특정 구현예와 관련하여 설명되고 예시되었지만, 당업자는 설명 및 청구된 바와 같이 본원에 예시된 바와 같은 본 발명의 원리를 벗어나지 않고 변동 및 변형이 이루어질 수 있음을 인식할 것이다. 본 발명은 그 취지 또는 본질적인 특성을 벗어나지 않으면서 다른 구체적인 형태로 구체화될 수 있다. 설명된 구현예에는 모든 측면에서 예시적이며 제한적이지 않은 것으로 간주된다.
- [0857] 끝으로, 본원에 나타내고 설명된 본 발명의 예시적인 구현예와 관련하여, AAV(아데노-관련 바이러스) 바이러스 비리온을 포함하는 계통 작제물이 개시되고 AAV 백터의 전달을 위해 구성된다는 것으로 인식될 것이다. 본 발명의 원리는 나타내고 설명된 것 이상의 다양한 구성으로 실행될 수 있기 때문에, 본 발명은 예시적인 구현예에 의해 어떤 식으로든 제한되지 않지만, 일반적으로 AAV(아데노-관련 바이러스) 바이러스 비리온 기관을 포함하는 계통 작제물에 관한 것이고 본 발명의 취지 및 범위를 벗어나지 않고 그렇게 하기 위해 다양한 형태를 취할 수 있다는 것으로 이해해야 한다.
- [0858] 본 발명을 수행하기 위해 발명자(들)에게 알려진 최상의 모드를 포함하여 본 발명의 특정 구현예가 본원에 설명되어 있다. 물론, 이들 설명된 구현예에 대한 변형은 전술한 설명을 읽을 때 당업자에게 명백해질 것이다. 발명자(들)는 숙련가가 그러한 변형을 적절하게 사용하기를 기대하며, 발명자(들)는 본 발명이 본원에 구체적으로 설명된 것과 다르게 실시되기를 의도한다. 따라서, 본 발명은 적용 가능한 법에 의해 허용되는 본원에 첨부된 청구범위에 인용된 주제의 모든 변형 및 등가물을 포함한다. 더욱이, 본원에서 달리 나타내지 않거나 문맥상 명백하게 모순되지 않는 한, 이의 모든 가능한 변형에서 상기 기술된 구현예의 임의의 조합은 본 발명에 포함된다.
- [0859] 본 발명의 대안적인 구현예, 요소 또는 단계의 그룹화는 제한으로 해석되어서는 안된다. 각 그룹 구성원은 개별적으로 또는 본원에 개시된 다른 그룹 구성원과 조합하여 참조 및 청구될 수 있다. 그룹의 하나 이상의 구성원이 편의 및/또는 특허성을 이유로 그룹에 포함되거나 그룹으로부터 삭제될 수 있다. 임의의 그러한 포함 또는 삭제가 발생하는 경우, 명세서는 변형된 그룹을 포함하는 것으로 간주되며 따라서 첨부된 청구범위에 사용된 모

든 마쿠쉬 그룹의 서면 설명을 충족한다.

[0860] 달리 표시되지 않는 한, 본 명세서 및 청구범위에 사용된 특성, 항목, 수량, 파라미터, 성질, 용어 등을 표현하는 모든 숫자는 모든 경우에 "약"이라는 용어에 의해 변형되는 것으로 이해되어야 한다. 본원에 사용된 바와 같이, 용어 "약"은 특성, 항목, 수량, 파라미터, 성질 또는 그렇게 한정된 용어가 명시된 특성, 항목, 수량, 파라미터, 성질 또는 용어 값의 위아래 $\pm 10\%$ 범위를 포함함을 의미한다. 따라서, 달리 표시되지 않는 한, 명세서 및 첨부된 청구범위에 기재된 수치 파라미터는 변경될 수 있는 근사치이다. 적어도, 청구범위에 대한 등가 원칙의 적용을 제한하려는 시도가 아니라, 각 숫자 표시는 적어도, 보고된 유효 자릿수에 비추어 그리고 일반적인 반올림 기술을 적용하여 해석되어야 한다. 본 발명의 넓은 범위에 기재된 수치 범위 및 값이 근사치임에도 불구하고, 구체적인 실시예에 기재된 수치 범위 및 값은 가능한 한 정확하게 보고된다. 그러나, 임의의 수치 범위 또는 값에는 본질적으로 이들 각각의 시험 측정에서 발견된 표준 편차로부터 필연적으로 발생하는 특정 오차를 포함한다. 본원에서 값의 수치 범위를 인용하는 것은 범위 내에 속하는 각각의 개별 수치를 개별적으로 언급하는 속기 방법의 역할을 하기 위한 것일 뿐이다. 본원에 달리 표시되지 않는 한, 수치 범위의 각 개별 값은 본원에 개별적으로 인용된 것처럼 본 명세서에 포함된다. 유사하게, 본원에 사용된 바와 같이, 달리 표시되지 않는 한, "실질적으로"라는 용어는 특성, 항목, 수량, 파라미터, 성질, 또는 그렇게 한정된 용어의 근사치를 나타내도록 의도된 정도의 용어로, 당업자에 의해 이해되고 구성될 수 있는 범위를 포함한다.

[0861] 구현예 또는 구현예의 양태와 관련하여 "할 수 있는(may)" 또는 "할 수 있는(can)"이라는 용어의 사용은 "할 수 없는(may not)" 또는 "할 수 없는(cannot)"의 대안적인 의미도 수반한다. 이와 같이, 본 명세서가 구현예 또는 구현예의 양태가 본 발명의 주제의 일부로서 포함될 수(may be) 있거나 포함될 수(can be) 있음을 개시하는 경우, 부정적인 제한 또는 배제 조항 또한 명시적으로 의미되며, 이는 구현예 또는 구현예의 양태가 본 발명의 주제의 일부로 포함되지 않거나 포함될 수 없음을 의미한다. 유사한 방식으로, 구현예 또는 구현예의 양태와 관련하여 "임의로"라는 용어의 사용은 그러한 구현예 또는 구현예의 양태가 본 발명의 주제의 일부로 포함될 수 있거나 본 발명의 주제의 일부로서 포함되지 않을 수 있음을 의미한다. 그러한 부정적인 제한 또는 배제 조항이 적용되는지 여부는 부정적인 제한 또는 배제 조항이 청구된 주제에 인용되었는지 여부를 기반으로 할 것이다.

[0862] 청구범위에 사용되는 경우, 보정에 따라 제출 또는 추가되었는지 여부에 관계없이 개방형 전환 용어인 "포함하는"(포함하는, "함유하는" 및 "갖는"과 같은 이의 동등한 개방형 전환 문구와 함께)은 명시적으로 언급된 모든 요소, 제한, 단계 및/또는 특징을 단독으로 또는 언급되지 않은 주제와 조합하여 포함하고; 명명된 요소, 제한 및/또는 특징은 필수적이지만 다른 명명되지 않은 요소, 제한 및/또는 특징이 추가될 수 있으며 여전히 청구범위 내에서 구성을 형성할 수 있다. 본원에 개시된 특정 구현예는 "포함하는" 대신에 또는 이에 대한 보정으로 서 폐쇄형 전환 문구 "~로 이루어진" 또는 "~로 본질적으로 이루어진"을 사용하는 청구범위에서 추가로 제한될 수 있다. 청구범위에서 사용되는 경우, 출원 또는 보정마다 추가된 폐쇄형 전환 문구 "~로 이루어진"은 청구범위에 명시적으로 언급되지 않은 요소, 제한, 단계 또는 특징을 배제한다. "~로 본질적으로 이루어진"이라는 폐쇄형 전환 문구는 명시적으로 언급된 요소, 제한, 단계 및/또는 특징 및 청구된 주제의 기본적인 새로운 특성(들)에 실질적으로 영향을 미치지 않는 임의의 기타 요소, 제한, 단계 및/또는 특징에 대한 청구 범위를 제한한다. 따라서, 개방형 전환 문구 "포함하는"의 의미는 구체적으로 인용된 모든 요소, 제한, 단계 및/또는 특징뿐만 아니라 임의의 선택적 추가의 지정되지 않은 것들을 포괄하는 것으로 정의된다. 폐쇄형 전환 문구 "~로 이루어진"의 의미는 청구범위에 구체적으로 언급된 요소, 제한, 단계 및/또는 특징만을 포함하는 것으로 정의되는 반면, 폐쇄형 전환 문구 "~로 본질적으로 이루어진"의 의미는 청구범위에 구체적으로 언급된 이들 요소, 제한, 단계 및/또는 특징 및 청구된 주제의 기본적인 새로운 특징(들)에 실질적으로 영향을 미치지 않는 이들 요소, 제한, 단계 및/또는 특징만을 포함하는 것으로 정의된다. 따라서, 개방형 전환 문구 "포함하는" (이의 동등한 개방형 전환 문구와 함께)는 그 의미 내에서 청구된 제한적인 경우로서 폐쇄형 전환 문구 "~로 이루어진" 또는 "~로 본질적으로 이루어진"에 의해 명시된 청구된 주제를 포함한다. 이와 같이, "포함하는"이라는 문구로 본원에 기술되거나 그렇게 청구된 구현예는 "~로 본질적으로 이루어진" 및 "~로 이루어진"이라는 문구에 대해 명시적으로 또는 본질적으로 모호하지 않게 설명되고, 활성화되고 뒷받침된다.

[0863] 본 발명의 양태가 적어도 하나의 예시적인 구현예를 참조하여 설명되었지만, 본 발명이 이에 제한되지 않는다는 것은 당업자에 의해 명백하게 이해되어야 한다. 오히려, 본 발명의 범위는 첨부된 청구범위와 함께만 해석되어야 하며, 여기서 발명자(들)는 청구된 주제가 본 발명이라고 믿는 것이 분명하다.

[0864] **참조문헌**

[0865] 특허 및 특허 출원 및 국제 특허 출원을 포함하지만 이에 제한되지 않는 명세서 및 실시예에 개시된 참조는 모두 그 전체가 본원에 참조로 포함된다.

[0866] 본 명세서에서 참조 및 확인된 모든 특허, 특허 간행물 및 기타 간행물은, 예를 들어, 본 발명과 관련하여 사용될 수 있는 그러한 간행물에 기재된 조성물 및 방법론을 기술하고 개시할 목적으로 그 전체가 본원에 참조로 개별적으로 명시적으로 포함된다. 이러한 간행물은 본 출원의 출원일 이전에 공개될 목적으로만 제공된다. 이와 관련하여 그 어떤 것도 발명자가 선행 발명으로 인해 또는 임의로 다른 이유로 그러한 개시보다 선행할 자격이 없다는 것을 인정하는 것으로 해석되어서는 안된다. 이들 문서의 내용에 대한 날짜 또는 표현에 대한 모든 진술은 출원인이 사용할 수 있는 정보에 기반하며 이들 문서의 날짜 또는 내용의 정확성에 대한 어떠한 승인도 구성하지 않는다.

[0867] 참조문헌:

1. L Lisowski, AP Dane, K Chu, Y Zhang, SC Cunningham, EM Wilson, et al. Selection and evaluation of clinically relevant AAV variants in a xenograft liver model *Nature*, 506 (2014), pp. 382-386 (LK03 and others LK0-19)
2. Grimm D, Lee JS, Wang L, Desai T, Akache B Storm TA, Kay MA. In vitro and in vivo gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses. *J Virol*. 2008 Jun; 82(12):5887-911. (AAV-DJ)
3. Powell SK, Khan N, Parker CL, Samulski RJ, Matsushima G, Gray SJ, McCown TJ. Characterization of a novel adeno-associated viral vector with preferential oligodendrocyte tropism. *Gene Ther*. 2016 Nov; 23(11):807-814. (Olig001)
4. Tervo DG, Hwang BY, Viswanathan S, Gaj T, Lavzin M, Ritola KD, Lindo S, Michael S, Kuleshova E, Ojala D, Huang CC, Gerfen CR, Schiller J, Dudman JT, Hantman AW, Looger LL, Schaffer DV, Karpova AY. A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons. *Neuron*. 2016 Oct 19; 92(2):372-382. (rAAV2-retro)
5. Marsic D, Govindasamy L, Currin S, Markusic DM, Tseng YS, Herzog RW, Agbandje-McKenna M, Zolotukhin S. Vector design Tour de Force: integrating combinatorial and rational approaches to derive novel adeno-associated virus variants. *Mol Ther*. 2014 Nov; 22(11):1900-9. (AAV-LiC)
6. Sallach J, Di Pasquale G, Larcher F, Niehoff N, Rubsam M, Huber A, Chiarini J, Almarza D, Eming SA, Ulus H, Nishimura S, Hacker UT, Ballek M, Niessen CM, Buning H. Tropism-modified AAV vectors overcome barriers to successful cutaneous therapy. *Mol Ther*. 2014 May; 22(5):929-39. (AAV-Keral, AAV-Kera2, and AAV-Kera3)
7. Dalkara D, Byrne LC, Klimczak RR, Visel M, Yin L, Merigan WH, Flannery JG, Schaffer DV. In vivo-directed evolution of a new adeno-associated virus for therapeutic outer retinal gene delivery from the vitreous. *Sci Transl Med*. 2013 Jun 12; 5(189):189ra76. (AAV 7m8)
8. Asuri P, Bartel MA, Vazin T, Jang JH, Wong TB, Schaffer DV. Directed evolution of adeno-associated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells. *Mol Ther*. 2012 Feb; 20(2):329-38. (AAV1.9)
9. Jang JH, Koerber JT, Kim JS, Asuri P, Vazin T, Bartel M, Keung A, Kwon I, Park KI, Schaffer DV. An evolved adeno-associated viral variant enhances gene delivery and gene targeting in neural stem cells. *Mol Ther*. 2011 Apr; 19(4):667-75. doi: 10.1038/mt.2010.287. (AAV r3.45)
10. Gray SJ, Blake BL, Criswell HE, Nicolson SC, Samulski RJ, McCown TJ, Li W. Directed evolution of a novel adeno-associated virus (AAV) vector that crosses the seizure-compromised blood-brain barrier (BBB). *Mol Ther*. 2010 Mar; 18(3):570-8 (AAV clone 32 and 83)
11. Maguire CA, Gianni D, Meijer DH, Shaket LA, Wakimoto H, Rabkin SD, Gao G, Sena-Esteves M. Directed evolution of adeno-associated virus for glioma cell transduction. *J Neurooncol*. 2010 Feb; 96(3):337-47. (AAV-U87R7-C5)
12. Koerber JT, Klimczak R, Jang JH, Dalkara D, Flannery JG, Schaffer DV. Molecular evolution of adeno-associated virus for enhanced glial gene delivery. *Mol Ther*. 2009 Dec; 17(12):2088-95. (AAV ShH13, AAV ShH19, AAV LI-12)
13. Li W, Zhang L, Johnson JS, Zhijian W, Grieger JC, Ping-Jie X, Drouin LM, Agbandje-McKenna M, Pickles RJ, Samulski RJ. Generation of novel AAV variants by directed evolution for improved CFTR delivery to human ciliated airway epithelium. *Mol Ther*. 2009 Dec; 17(12):2067-77. (AAV HAE-1, AAV HAE-2)
14. Klimczak RR, Koerber JT, Dalkara D, Flannery JG, Schaffer DV. A novel adeno-associated viral variant for efficient and selective intravitreal transduction of rat Muller cells. *PLoS One*. 2009 Oct 14; 4(10):e7467. (AAV variant ShH10)
15. Excoffon KJ, Koerber JT, Dickey DD, Murtha M, Keshavjee S, Kaspar BK, Zabner J, Schaffer DV. Directed evolution of adeno-associated virus to an infectious respiratory virus. *Proc Natl Acad Sci US A*. 2009 Mar 10; 106(10):3865-70. (AAV2.5T)
16. Sellner L, Stiefelwagen M, Kleinschmidt JA, Laufs S, Wenz F, Fruehauf S, Zeller WJ, Veldwijk MR. Generation of efficient human blood progenitor-targeted recombinant adeno-associated viral vectors (AAV) by applying an AAV random peptide library on primary human hematopoietic progenitor cells. *Exp Hematol*. 2008 Aug; 36(8):957-64. (AAV LS1-4, AAV Lsm)
17. Li W, Asokan A, Wu Z, Van Dyke T, DiPrimio N, Johnson JS, Govindaswamy L, Agbandje-McKenna M, Leichter S, Redmond DE Jr, McCown TJ, Petermann KB, Sharpless NE, Samulski RJ.

[0868]

- Engineering and selection of shuffled AAV genomes: a new strategy for producing targeted biological nanoparticles. *Mol Ther.* 2008 Jul;16(7):1252-60. (AAV1289)
18. Charbel Issa P, De Silva SR, Lipinski DM, Singh MS, Mouravlev A, You Q. Assessment of tropism and effectiveness of new primate-derived hybrid recombinant AAV serotypes in the mouse and primate retina. *PLoS ONE.* 2013;8:e60361. (AAVHSC 1-17)
 19. Huang W, McMurphy T, Liu X, Wang C, Cao L. Genetic Manipulation of Brown Fat Via Oral Administration of an Engineered Recombinant Adeno-associated Viral Serotype Vector. *Mol Ther.* 2016 Jun; 24(6):1062-9. (AAV2 Rec 1-4)
 20. Cronin T, Vandenbergh LH, Hantz P, et al. Efficient transduction and optogenetic stimulation of retinal bipolar cells by a synthetic adeno-associated virus capsid and promoter. *EMBO Mol Med* 2014;6:1175-1190 (AAV8BP2)
 21. Choudhury SR, Fitzpatrick Z, Harris AF, Maitland SA, Ferreira JS, Zhang Y, Ma S, Sharma RB, Gray-Edwards HL, Johnson JA, Johnson AK, Alonso LC, Punzo C, Wagner KR, Maguire CA, Katin RM, Martin DR, Sena-Esteves M. In Vivo Selection Yields AAV-BI Capsid for Central Nervous System and Muscle Gene Therapy. *Mol Ther.* 2016 Aug; 24(7):1247-57. (AAV-B1)
 22. Deverman BE, Prayda PL, Simpson BP, Kumar SR, Chan KY, Banerjee A, Wu WL, Yang B, Huber N, Pasca SP, Gradinaru V. Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. *Nat Biotechnol.* 2016 Feb; 34(2):204-9. doi: 10.1038/nbt.3440. (AAV-PHP.B)
 23. Pulicherla N, Shen S, Yadav S, Debbink K, Govindasamy L, Agbandje-McKenna M, Asokan A. Engineering liver-detargeted AAV9 vectors for cardiac and musculoskeletal gene transfer. *Mol Ther.* 2011 Jun; 19(6):1070-8. (AAV9 derived mutants-AAV9.45, AAV9.61, AAV9.47)
 24. Yang L, Jiang J, Drouin LM, Agbandje-McKenna M, Chen C, Qiao C, Pu D, Hu X, Wang DZ, Li J, Xiao X. A myocardium tropic adeno-associated virus (AAV) evolved by DNA shuffling and in vivo selection. *Proc Natl Acad Sci US A.* 2009 Mar 10;106(10):3946-51. (AAVM41)
 25. Korbelen J, Sieber T, Michelfelder S, Lunding L, Spies E, Hunger A, Alawi M, Rapti K, Indenbirken D, Muller OJ, Pasqualini R, Arap W, Kleinschmidt JA, Trepel M. Pulmonary Targeting of Adeno-associated Viral Vectors by Next-generation Sequencing-guided Screening of Random Capsid Displayed Peptide Libraries. *Mol Ther.* 2016 Jun; 24(6):1050-61. (AAV2 displayed peptides)
 26. Geoghegan JC, Keiser NW, Okulist A, Martins I, Wilson MS, Davidson BL. Chondroitin Sulfate is the Primary Receptor for a Peptide-Modified AAV That Targets Brain Vascular Endothelium In Vivo. *Mol Ther Nucleic Acids.* 2014 Oct 14;3:e202. (AAV2-GMN)
 27. Varadi K, Michelfelder S, Korff T, Hecker M, Trepel M, Katus HA, Kleinschmidt JA, Muller OJ. Novel random peptide libraries displayed on AAV serotype 9 for selection of endothelial cell-directed gene transfer vectors. *Gene Ther.* 2012 Aug; 19(8):800-9. (AAV9-peptide displayed)
 28. Michelfelder S, Varadi K, Raupp C, Hunger A, Korbelen J, Pahrman C, Schrepfer S, Muller OJ, Kleinschmidt JA, Trepel M. Peptide ligands incorporated into the threefold spike capsid domain to re-direct gene transduction of AAV8 and AAV9 . in vivo. *PLoS One.* 2011;6(8):e23101. (AAV8 and AAV9 peptide displayed)
 29. Yu CY, Yuan Z, Cao Z, Wang B, Qiao C, Li J, Xiao X. A muscle-targeting peptide displayed on AAV2 improves muscle tropism on systemic delivery. *Gene Ther.* 2009 Aug; 16(8):953-62.
 30. Michelfelder S, Lee MK, deLima-Hahn E, Wilmes T, Kaul F, Muller O, Kleinschmidt JA, Trepel M. Vectors selected from adeno-associated viral display peptide libraries for leukemia cell-targeted cytotoxic gene therapy. *Exp Hematol.* 2007 Dec; 35(12): 1766-76.
 31. Muller OJ, Kaul F, Weitzman MD, Pasqualini R, Arap W, Kleinschmidt JA, Trepel M. Random peptide libraries displayed on adeno-associated virus to select for targeted gene therapy vectors. *Nat Biotechnol.* 2003 Sep; 21(9):1040-6.
 32. Grifman M, Trepel M, Speece P, Gilbert LB, Arap W, Pasqualini R, Weitzman MD. Incorporation of tumor-targeting peptides into recombinant adeno-associated virus capsids. *Mol Ther.* 2001 Jun; 3(6):964-75.
 33. Anne Girod, Martin Ried, Christiane Wobus, Harald Lahm, Kristin Leike, Jurgen Kleinschmidt, Gilbert Deleage & Michael Ballek. Genetic capsid modifications allow efficient re-targeting of adeno-associated virus type 2. *Nature Medicine*, 1052 - 1056 (1999)

[0869]

34. Bello A, Chand A, Aviles J, Soule G, Auricchio A, Kobinger GP. Novel adeno-associated viruses derived from pig tissues transduce most major organs in mice. *Sci Rep*. 2014 Oct 22;4:6644. (AAVpo2.1, -po4, -poS, and -po6).
35. Gao G, Vandenberghe LH, Alvira MR, Lu Y, Calcedo R, Zhou X, Wilson JM. Clades of Adeno-associated viruses are widely disseminated in human tissues. *J Virol*. 2004 Jun; 78(12):6381-8. (AAV rh and AAV Hu)
36. Arbetman AE, Lochrie M, Zhou S, Wellman J, Scallan C, Doroudchi MM, et al. Novel caprine adeno-associated virus (AAV) capsid (AAV-Go.1) is closely related to the primate AAV-5 and has unique tropism and neutralization properties. *J Virol*. 2005;79:15238-15245. (AAV-Go.1)
37. Lochrie MA, Tatsuno GP, Arbetman AE, Jones K, Pater C, Smith PH, et al. Adeno-associated virus (AAV) capsid genes isolated from rat and mouse liver genomic DNA define two new AAV species distantly related to AAV-5. *Virology*. 2006;353:68-82. (AAV-mo.1)
38. Schmidt M, Katano H, Bossis I, Chiarini JA. Cloning and characterization of a bovine adeno-associated virus. *J Virol*. 2004;78:6509-6516. (BAAV)
39. Bossis I, Chiarini JA. Cloning of an avian adeno-associated virus (AAAV) and generation of recombinant AAAV particles. *J Virol*. 2003;77:6799-6810. (AAAV)
40. Chen CL, Jensen RL, Schnepf BC, Connell MJ, Shell R, Sferra TJ, Bartlett JS, Clark KR, Johnson PR. Molecular characterization of adeno-associated viruses infecting children. *J Virol*. 2005 Dec; 79(23):14781-92. (AAV variants)
41. Sen D, Gadkari RA, Sudha G, Gabriel N, Kumar YS, Selot R, Samuel R, Rajalingam S, Ramya V, Nair SC, Srinivasan N, Srivastava A, Jayandharan GR. Targeted modifications in adeno-associated virus serotype 8 capsid improves its hepatic gene transfer efficiency in vivo. *Hum Gene Ther Methods*. 2013 Apr; 24(2):104-16. (AAV8 K137R)
42. Li B, Ma W, Ling C, Van Vliet K, Huang LY, Agbandje-McKenna M, Srivastava A, Aslanidi GV. Site-Directed Mutagenesis of Surface-Exposed Lysine Residues Leads to Improved Transduction by AAV2, But Not AAV8, Vectors in Murine Hepatocytes In Vivo. *Hum Gene Ther Methods*. 2015 Dec; 26(6):211-20.
43. Gabriel N, Hareendran S, Sen D, Gadkari RA, Sudha G, Selot R, Hussain M, Dhaknamoorthy R, Samuel R, Srinivasan N, et al. Bioengineering of AAV2 capsid at specific serine, threonine, or lysine residues improves its transduction efficiency in vitro and in vivo. *Hum Gene Ther Methods*. 2013 Apr; 24(2):80-93.
44. Zinn E, Pacouret S, Khaychuk V, Turunen HT, Carvalho LS, Andres-Mateos E, Shah S, Shelke R, Maurer AC, Plovie E, Xiao R, Vandenberghe LH. In Silico Reconstruction of the Viral Evolutionary Lineage Yields a Potent Gene Therapy Vector. *Cell Rep*. 2015 Aug 11;12(6):1056-68. (AAV Anc80L65)
45. Shen S, Horowitz ED, Troupes AN, Brown SM, Pulicherla N, Samulski RJ, Agbandje-McKenna M, Asokan A. Engraftment of a galactose receptor footprint onto adeno-associated viral capsids improves transduction efficiency. *J Biol Chem*. 2013 Oct 4;288(40):28814-23. (AAV2G9)
46. Li C, DiPrimio N, Bowles DE, Hirsch ML, Monahan PE, Asokan A, Rabinowitz J, Agbandje-McKenna M, Samulski RJ. Single amino acid modification of adeno-associated virus capsid changes transduction and humoral immune profiles. *J Virol*. 2012 Aug; 86(15):7752-9. (AAV2 265 insertion-AAV2/265D)
47. Bowles DE, McPhee SW, Li C, Gray SJ, Samulski JJ, Camp AS, Li J, Wang B, Monahan PE, Rabinowitz JE, et al. Phase I gene therapy for Duchenne muscular dystrophy using a translational optimized AAV vector. *Mol Ther*. 2012 Feb; 20(2):443-55 (AAV2.5)
48. Messina EL, Nienaber J, Daneshmand M, Villamizar N, Samulski J, Milano C, Bowles DE. Adeno-associated viral vectors based on serotype 3b use components of the fibroblast growth factor receptor signaling complex for efficient transduction. *Hum. Gene Ther*. 2012 Oct; 23(10):1031-42. (AAV3 SASTG)
49. Asokan A, Conway JC, Phillips JL, Li C, Hegge J, Sinnott R, Yadav S, DiPrimio N, Nam HJ, Agbandje-McKenna M, McPhee S, Wolff J, Samulski RJ. Reengineering a receptor footprint of adeno-associated virus enables selective and systemic gene transfer to muscle. *Nat Biotechnol*. 2010 Jan; 28(1):79-82. (AAV2i8)

[0870]

50. Vance M, Llanga T, Bennett W, Woodard K, Murlidharan G, Chungfat N, Asokan A, Gilger B, Kurtzberg J, Samulski RJ, Hirsch ML. AAV Gene Therapy for MPSI- associated Corneal Blindness. *Sci Rep*. 2016 Feb 22;6:22131. (AAV8G9)
51. Zhong L, Li B, Mah CS, Govindasamy L, Agbandje-McKenna M, Cooper M, Herzog RW, Zolotukhin I, Warrington KH Jr, Weigel-Van Aken KA, Hobbs JA, Zolotukhin S, Muzyczka N, Srivastava A. Next generation of adeno-associated virus 2 vectors: point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proc Natl Acad Sci US A*. 2008 Jun 3;105(22):7827-32. (AAV2 tyrosine mutants AAV2 Y-F)
52. Petrs-Silva H, Dinculescu A, Li Q, Min SH, Chiodo V, Pang JJ, Zhong L, Zolotukhin S, Srivastava A, Lewin AS, Hauswirth WW. High-efficiency transduction of the mouse retina by tyrosine-mutant AAV serotype vectors. *Mol Ther*. 2009 Mar; 17(3):463-71. (AAV8 Y-F and AAV9 Y-F)
53. Qiao C, Zhang W, Yuan Z, Shin JH, Li J, Jayandharan GR, Zhong L, Srivastava A, Xiao X, Duan D. Adeno-associated virus serotype 6 capsid tyrosine-to- phenylalanine mutations improve gene transfer to skeletal muscle. *Hum Gene Ther*. 2010 Oct; 21(10):1343-8 (AAV6 Y-F)
54. Carlson M, Toelen J, Van der Perren A, Vandenberghe LH, Reumers V, Sbragia L, Gijsbers R, Backelant V, Himmelreich U, Wilson JM, Deprest J, Debysier Z. Efficient gene transfer into the mouse lung by fetal intratracheal injection of rAAV2/6.2. *Mol Ther*. 2010 Dec; 18(12):2130-8. (AAV6.2) PCT Publication No. WO2013158879A1 (lysine mutants)
55. Piacentino III, Valentino, et al. "X-linked inhibitor of apoptosis protein-mediated attenuation of apoptosis, using a novel cardiac-enhanced adeno-associated viral vector." *Human gene therapy* 23.6 (2012): 635-646.
56. Hirschhorn et al. (2001) "Glycogen Storage Disease Type II: Acid Alpha-glucosidase (Acid Maltase) Deficiency," in Scriver et al., eds., *The Metabolic and Molecular Basis of Inherited Disease*, 8th Ed., New York: McGraw-Hill, 3389-3420.
57. Hermonat, P. L., Labow, M. A., Wright, R., Berns, K. I., & Muzyczka, N. (1984). Genetics of adeno-associated virus: isolation and preliminary characterization of adeno-associated virus type 2 mutants. *Journal of Virology*, 51(2), 329-339. <https://doi.org/10.1128/jvi.51.2.329-339.1984>
58. West, M. H., Trempe, J. P., Tratschin, J. D., & Carter, B. J. (1987). Gene expression in adeno-associated virus vectors: the effects of chimeric mRNA structure, helper virus, and adenovirus VA1 RNA. *Virology*, 160(1), 38-47. [https://doi.org/10.1016/0042-6822\(87\)90041-9](https://doi.org/10.1016/0042-6822(87)90041-9)
59. Bohenzky, R. A., & Berns, K. I. (1989). Interactions between the termini of adeno-associated virus DNA. *Journal of Molecular Biology*, 206(1), 91-100. [https://doi.org/10.1016/0022-2836\(89\)90526-3](https://doi.org/10.1016/0022-2836(89)90526-3)
60. Flotte, T. R., Solow, R., Owens, R. A., Afione, S., Zeitlin, P. L., & Carter, B. J. (1992). Gene expression from adeno-associated virus vectors in airway epithelial cells. *American Journal of Respiratory Cell and Molecular Biology*, 7(3), 349-356. <https://doi.org/10.1165/ajrcmb/7.3.349>
61. Flotte, T. R., Afione, S. A., Solow, R., Drumm, M. L., Markakis, D., Guggino, W. B., Zeitlin, P. L., & Carter, B. J. (1993). Expression of the cystic fibrosis transmembrane conductance regulator from a novel adeno-associated virus promoter. *Journal of Biological Chemistry*, 268(5), 3781-3790.
62. Flotte, T., Carter, B., Conrad, C., Guggino, W., Reynolds, T., Rosenstein, B., Taylor, G., Walden, S., & Wetzel, R. (1996). A phase I study of an adeno-associated Virus-CFTR gene vector in adult CF patients with mild lung disease. In *Human Gene Therapy* (Vol. 7, Issue 9, pp. 1145-1159). Mary Ann Liebert Inc. <https://doi.org/10.1089/hum.1996.7.9-1145>
63. Rubenstein, R. C., McVeigh, U., Flotte, T. R., Guggino, W. B., & Zeitlin, P. L. (1997). CFTR gene transduction in neonatal rabbits using an adeno-associated virus (AAV) vector. *Gene Therapy*, 4(5), 384-392. <https://doi.org/10.1038/sj.gt.3300417>
64. Haberman, R. P., McCown, T. J., & Samulski, R. J. (1998). Inducible long-term gene expression in brain with adeno-associated virus gene transfer. *Gene Therapy*, 5(12), 1604-1611. <https://doi.org/10.1038/sj.gt.3300782>
65. Wang, D., Fischer, H., Zhang, L., Fan, P., Ding, R. X., & Dong, J. (1999). Efficient CFTR expression from AAV vectors packaged with promoters - The second generation. *Gene Therapy*, 6(4), 667-675. <https://doi.org/10.1038/sj.gt.3300856>
66. Hsiao, C. D., Hsieh, F. J., & Tsai, H. J. (2001). Enhanced expression and stable transmission of transgenes flanked by inverted terminal repeats from adeno-associated virus in zebrafish.

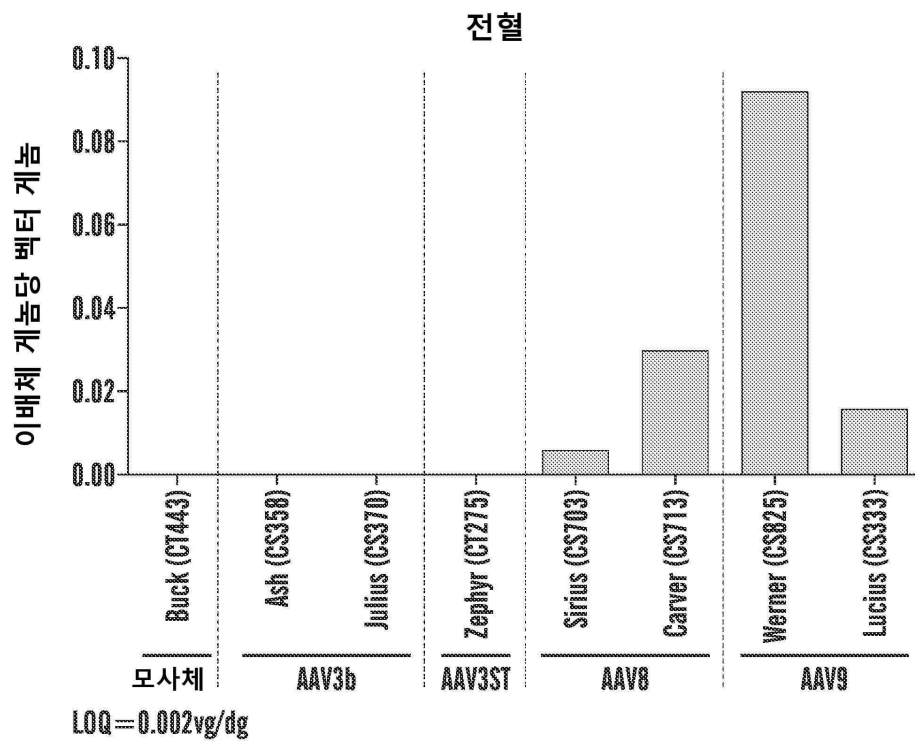
- Developmental Dynamics : An Official Publication of the American Association of Anatomists, 220(4), 323–336. <https://doi.org/10.1002/dvdy.1113>
67. Davidoff, A. M., Ng, C. Y. C., Zhou, J., Spence, Y., & Nathwani, A. C. (2003). Sex significantly influences transduction of murine liver by recombinant adeno-associated viral vectors through an androgen-dependent pathway. *Blood*, 102(2), 480–488. <https://doi.org/10.1182/blood-2002-09-2889>
68. Farris, K. D., & Pintel, D. J. (2008). Improved splicing of adeno-associated viral (AAV) capsid protein-supplying pre-mRNAs leads to increased recombinant AAV vector production. *Human Gene Therapy*, 19(12), 1421–1427. <https://doi.org/10.1089/hum.2008.118>
69. Li, C., Hirsch, M., Carter, P., Asokan, A., Zhou, X., Wu, Z., & Samulski, R. J. (2009). A small regulatory element from chromosome 19 enhances liver-specific gene expression. *Gene Therapy*, 16(1), 43–51. <https://doi.org/10.1038/gt.2008.134>
70. Dean, J., Plante, J., Huggins, G. S., Snyder, R. O., & Aikawa, R. (2009). Role of cyclic AMP-dependent kinase response element-binding protein in recombinant adeno-associated virus-mediated transduction of heart muscle cells. *Human Gene Therapy*, 20(9), 1005–1012. <https://doi.org/10.1089/hum.2009.054>
71. Ling, C., Wang, Y., Lu, Y., Wang, L., Jayandharan, G. R., Aslanidi, G. V., Li, B., Cheng, B., Ma, W., Lentz, T., Ling, C., Xiao, X., Samulski, R. J., Muzyczka, N., & Srivastava, A. (2015). Enhanced Transgene Expression from Recombinant Single-Stranded D-Sequence-Substituted Adeno-Associated Virus Vectors in Human Cell Lines In Vitro and in Murine Hepatocytes In Vivo. *Journal of Virology*, 89(2), 952–961. <https://doi.org/10.1128/jvi.02581-14>
72. Wang, L., Yin, Z., Wang, Y., Lu, Y., Zhang, D., Srivastava, A., Ling, C., Aslanidi, G. V., & Ling, C. (2015). Productive life cycle of adeno-associated virus serotype 2 in the complete absence of a conventional polyadenylation signal. *Journal of General Virology*, 96(9), 2780–2787. <https://doi.org/10.1099/jgv.0.000229>
73. Powell, S. K., Rivera-Soto, R., & Gray, S. J. (2015). Viral expression cassette elements to enhance transgene target specificity and expression in gene therapy. *Discovery Medicine*, 19(102), 49–57. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4505817/>
74. Stutika, C., Mietzsch, M., Gogol-Döring, A., Weger, S., Sohn, M., Chen, W., & Heilbronn, R. (2016). Comprehensive Small RNA-Seq of Adeno-Associated Virus (AAV)-Infected Human Cells Detects Patterns of Novel, Non-Coding AAV RNAs in the Absence of Cellular miRNA Regulation. *PLoS One*, 11(9), e0161454. <https://doi.org/10.1371/journal.pone.0161454>
75. Weger, S., Hammer, E., Gonsior, M., Stutika, C., & Heilbronn, R. (2016). A Regulatory Element Near the 3' End of the Adeno-Associated Virus rep Gene Inhibits Adenovirus Replication in cis by Means of p40 Promoter-Associated Short Transcripts. *Journal of Virology*, 90(8), 3981–3993. <https://doi.org/10.1128/jvi.03120-15>
76. Logan, G. J., Dane, A. P., Hallwirth, C. V., Smyth, C. M., Wilkie, E. E., Amaya, A. K., Zhu, E., Khandekar, N., Ginn, S. L., Liao, S. H. Y., Cunningham, S. C., Sasaki, N., Cabanes-Creus, M., Tam, P. P. L., Russell, D. W., Lisowski, L., & Alexander, I. E. (2017). Identification of liver-specific enhancer-promoter activity in the 3' untranslated region of the wild-type AAV2 genome. *Nature Genetics*, 49(8), 1267–1273. <https://doi.org/10.1038/ng.3893>
77. Lu, J., Williams, J. A., Luke, J., Zhang, F., Chu, K., & Kay, M. A. (2017). A 5' noncoding exon containing engineered intron enhances transgene expression from recombinant AAV vectors in vivo. *Human Gene Therapy*, 28(1), 125–134. <https://doi.org/10.1089/hum.2016.140>
78. Xu, H., Hao, S., Zhang, J., Chen, Z., Wang, H., & Guan, W. (2017). The formation and modification of chromatin-like structure of human parvovirus B19 regulate viral genome replication and RNA processing. *Virus Research*, 232, 134–138. <https://doi.org/10.1016/j.virusres.2017.03.001>
79. Julien, L., Chassagne, J., Peccate, C., Lorain, S., Piétri-Rouxel, F., Danos, O., & Benkhelifa-Ziyyat, S. (2018). RFX1 and RFX3 Transcription Factors Interact with the D Sequence of Adeno-Associated Virus Inverted Terminal Repeat and Regulate AAV Transduction. *Scientific Reports*, 8(1). <https://doi.org/10.1038/s41598-017-18604-3>
80. Wilmott, P., Lisowski, L., Alexander, I. E., & Logan, G. J. (2019). A User's Guide to the Inverted Terminal Repeats of Adeno-Associated Virus. *Human Gene Therapy Methods*, 30(6), 206–213. <https://doi.org/10.1089/hgtb.2019.276>
81. Lu, L., Zhang, F., Li, Y., Yang, A., Guan, C., Ding, X., Liu, Y., Liu, Y., Zhang, C.-Y., Li, L., & Zhang, Q. (2019). Dendritic targeted mRNA expression via a cis-acting RNA UTR element. *Biochemical and Biophysical Research Communications*, 509(2), 402–406. <https://doi.org/10.1016/j.bbrc.2018.12.137>
82. Earley, L. F., Conatser, L. M., Lue, V. M., Dobbins, A. L., Li, C., Hirsch, M. L., & Samulski, R. J. (2020). Adeno-Associated Virus Serotype-Specific Inverted Terminal Repeat Sequence Role in Vector Transgene Expression. *Human Gene Therapy*, 31(3–4), 151–162. <https://doi.org/10.1089/hum.2019.274>

[0872]

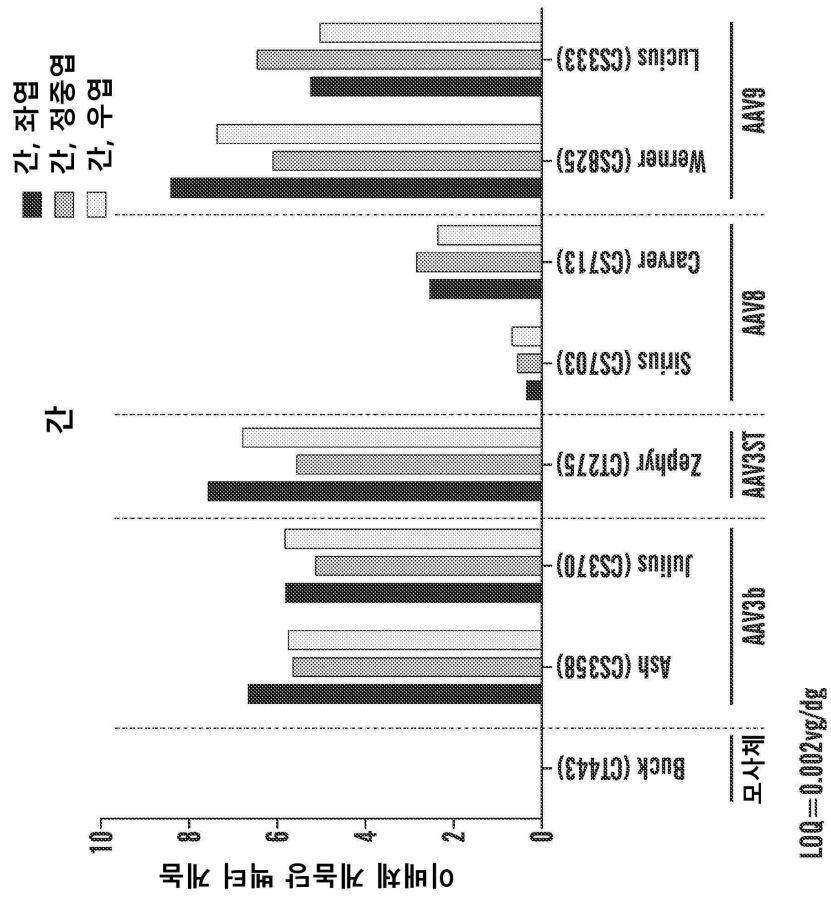
[0873]

도면

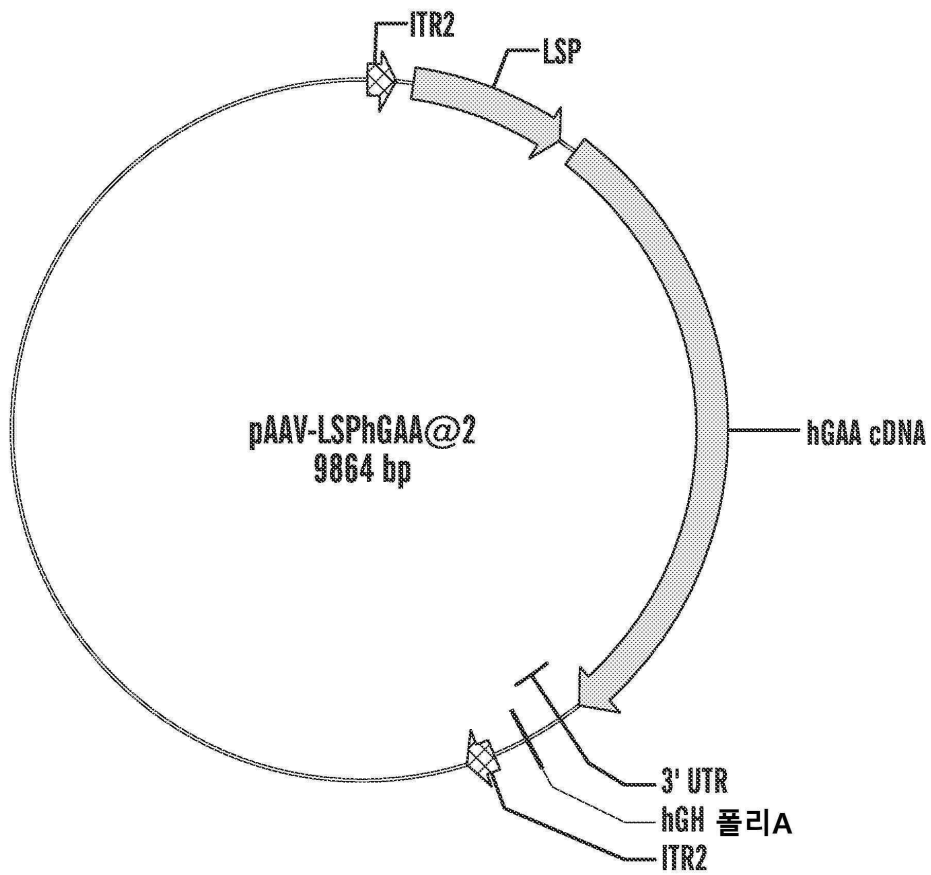
도면1



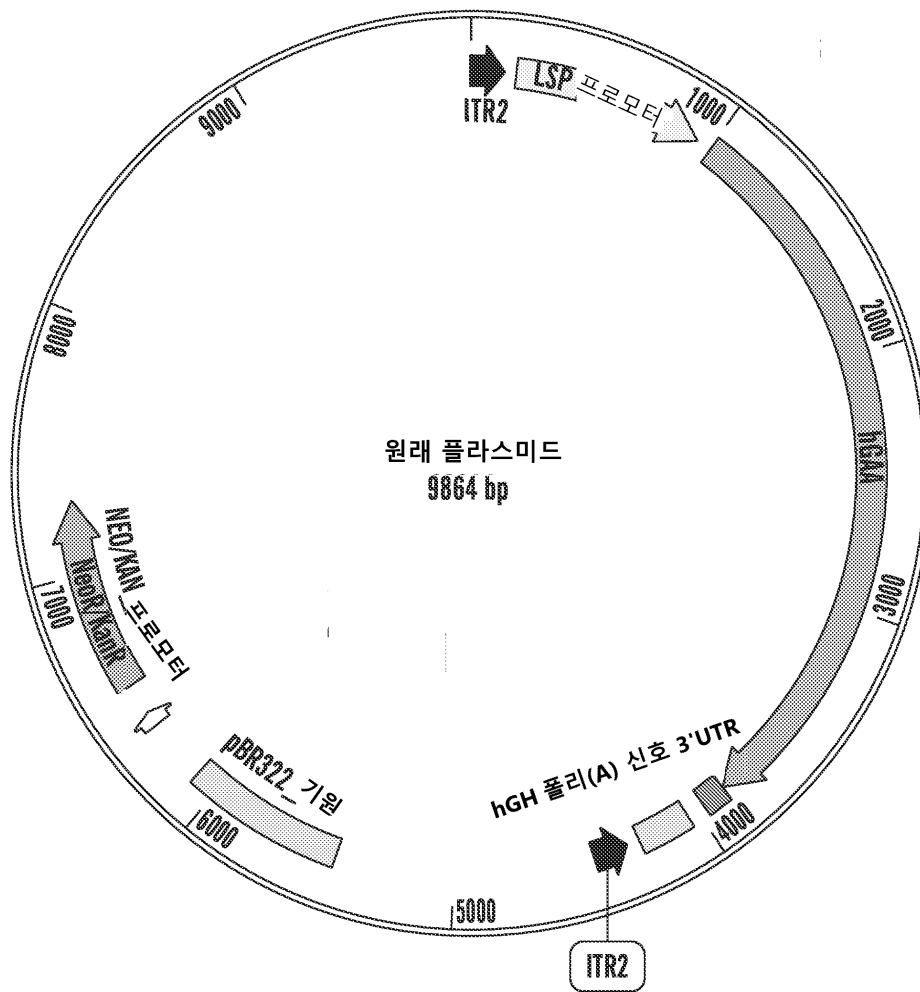
도면2



도면3a



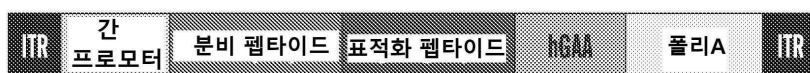
도면3b



도면4a



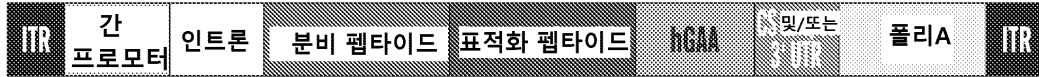
도면4b



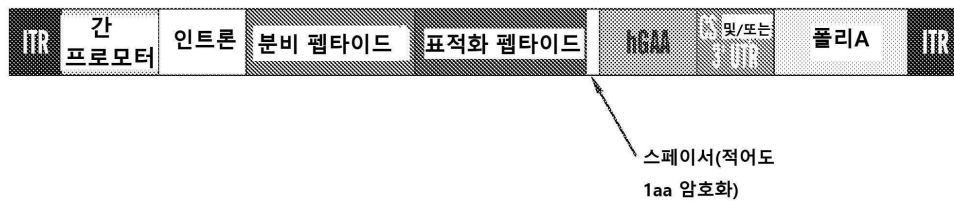
도면4c



도면4d



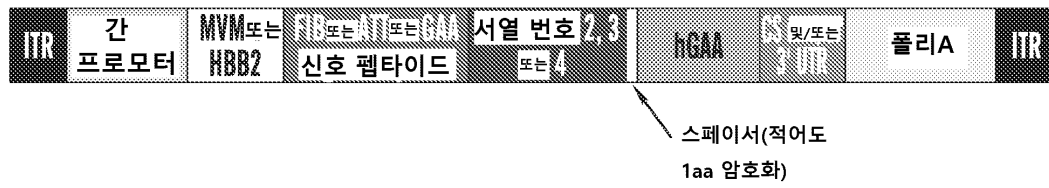
도면4e



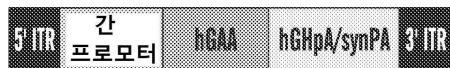
도면4f



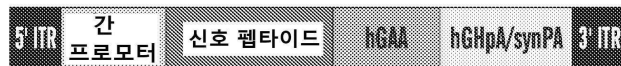
도면4g



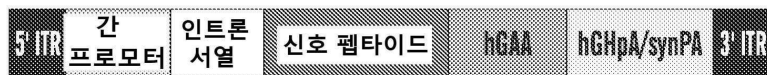
도면5a



도면5b



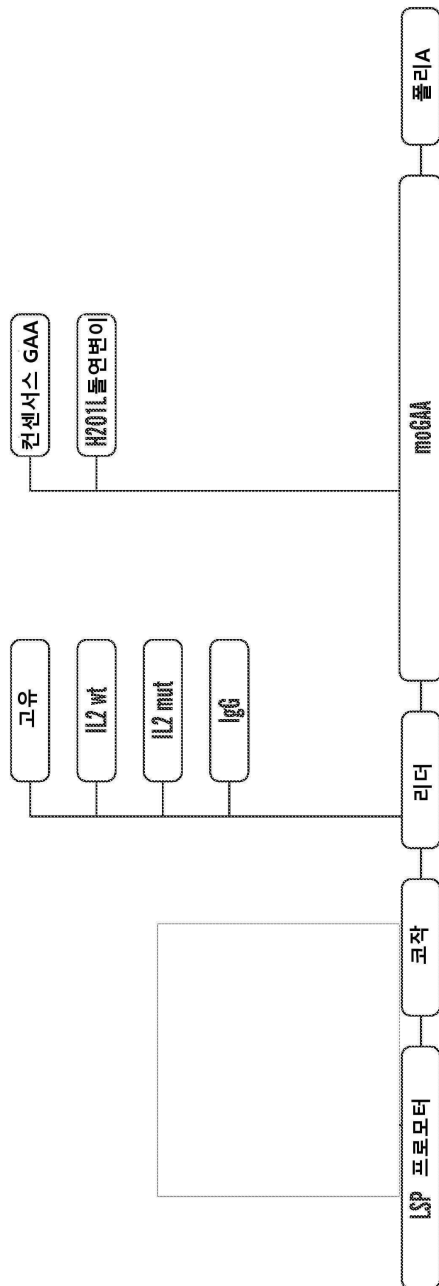
도면5c



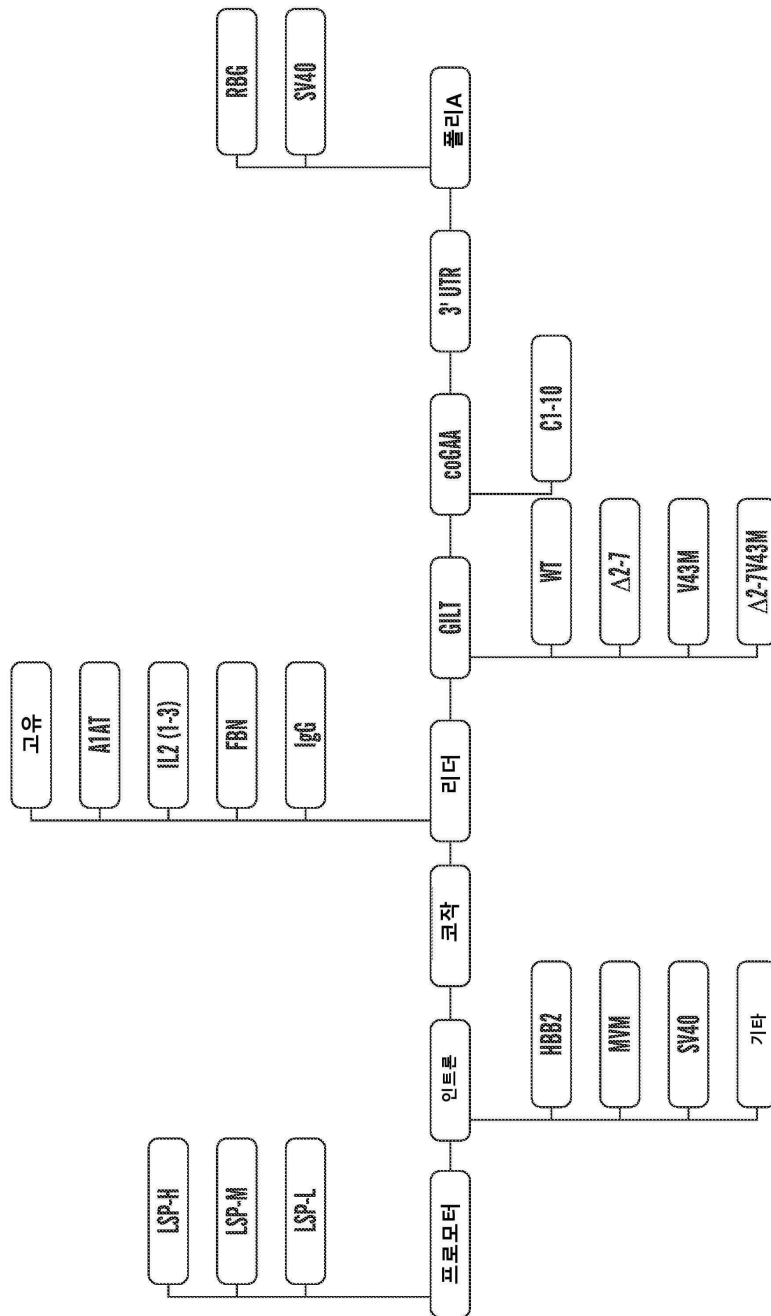
도면5d



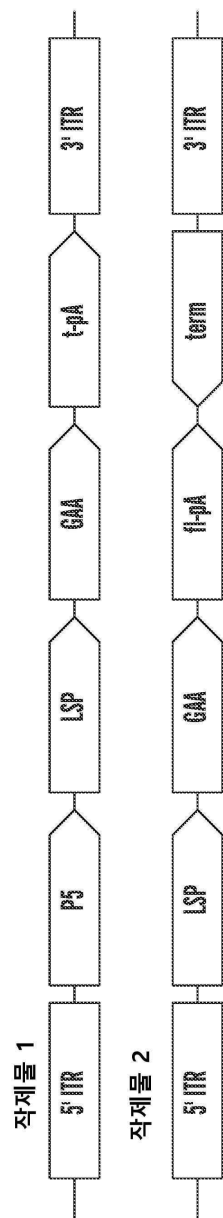
도면5e



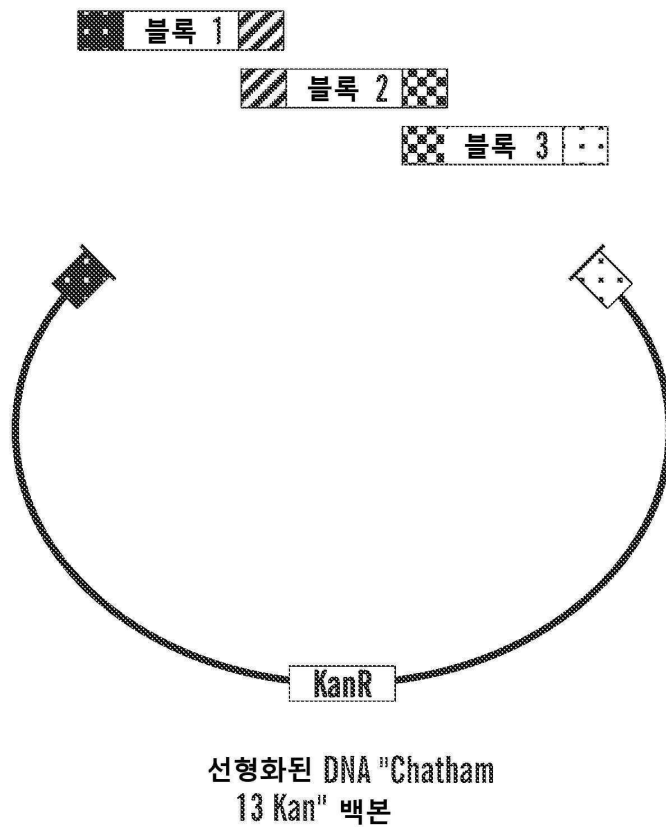
도면5f



도면5g

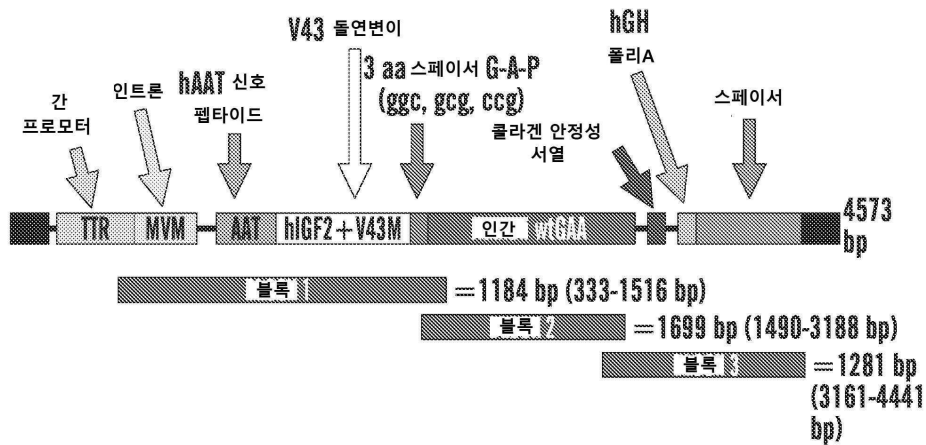


도면6



도면7

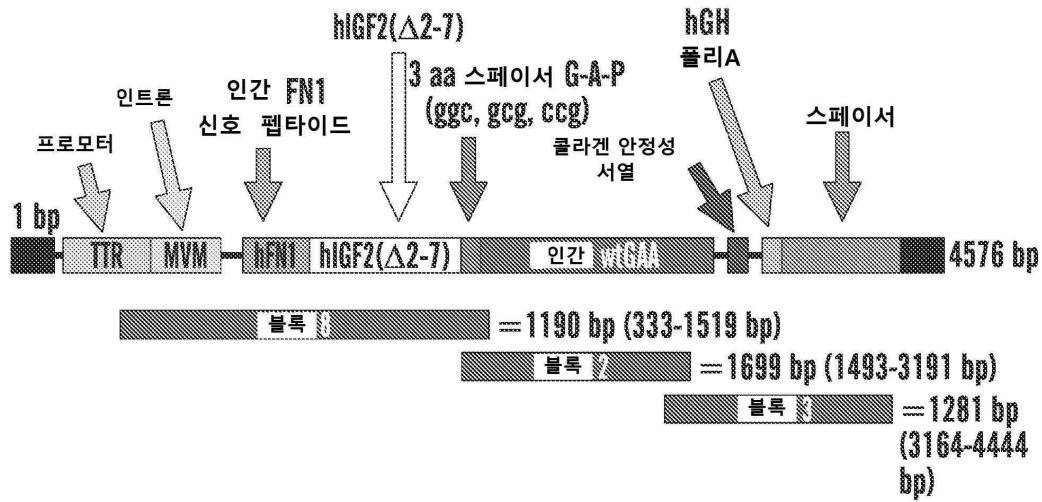
AAT-hIFG2(V43M)-wtGAA (del1-69aa)



작성된 벡터 게놈 : 1-4573 bp (플라스미드 백본에서 183-4755 bp)

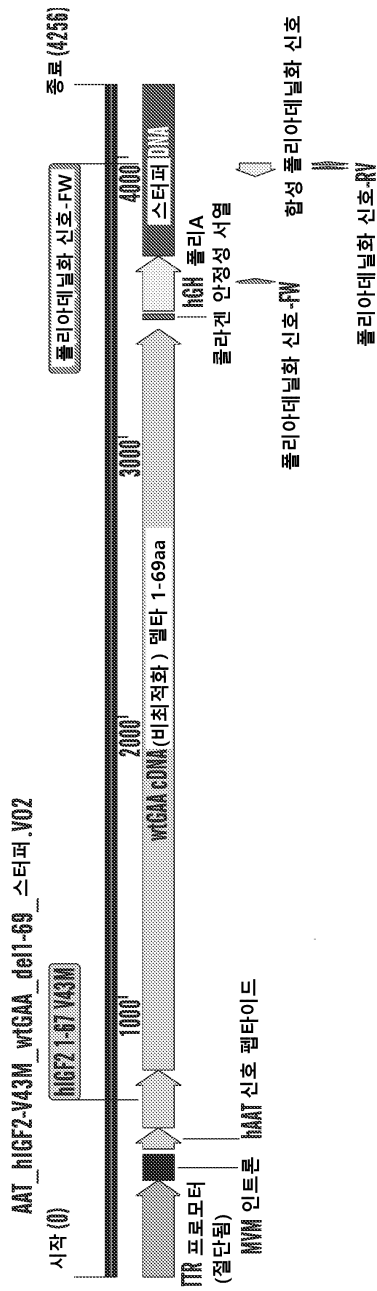
도면8

hFN1-hIGF2(Δ 2-7)-wtGAA (del1-69aa)

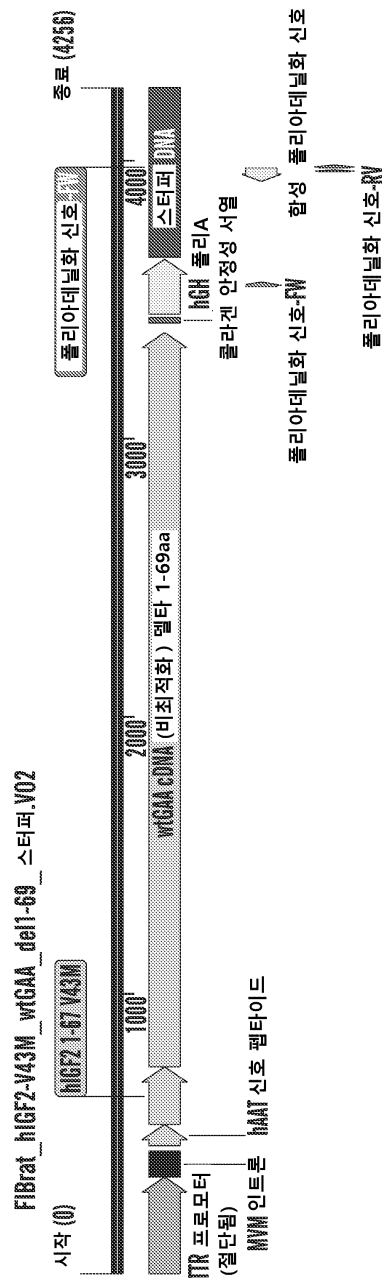


삭제된 벡터 게놈: 1-4576 bp (플라스미드 백본에서 183-4758 bp)

도면 9ab

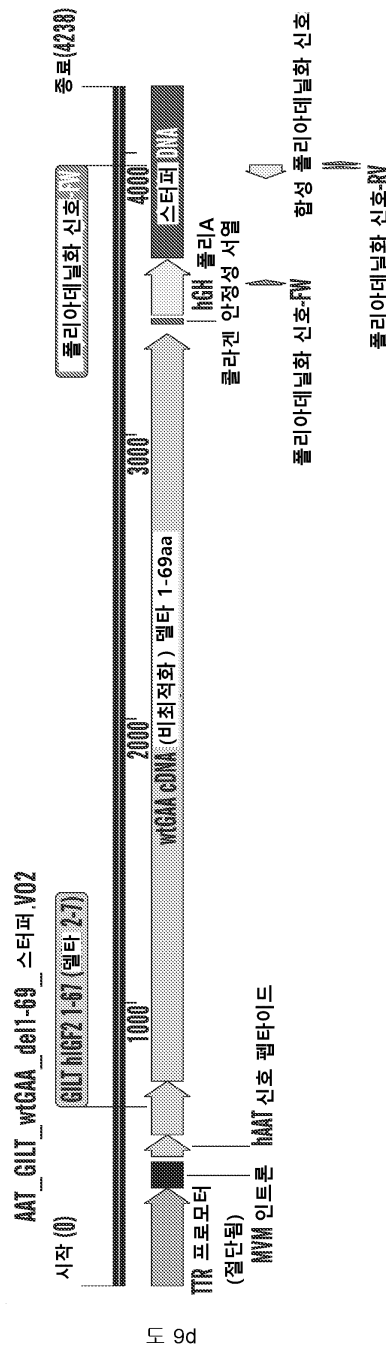
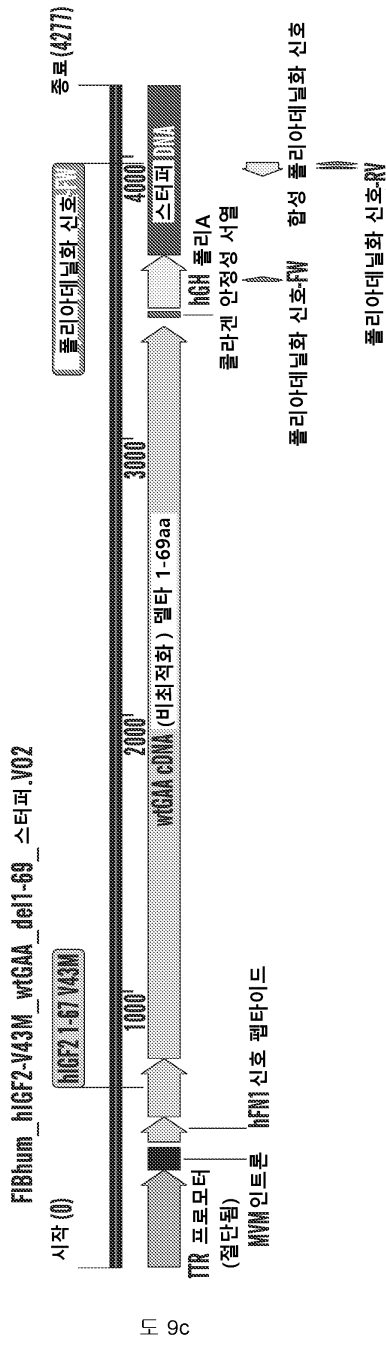


도 9a

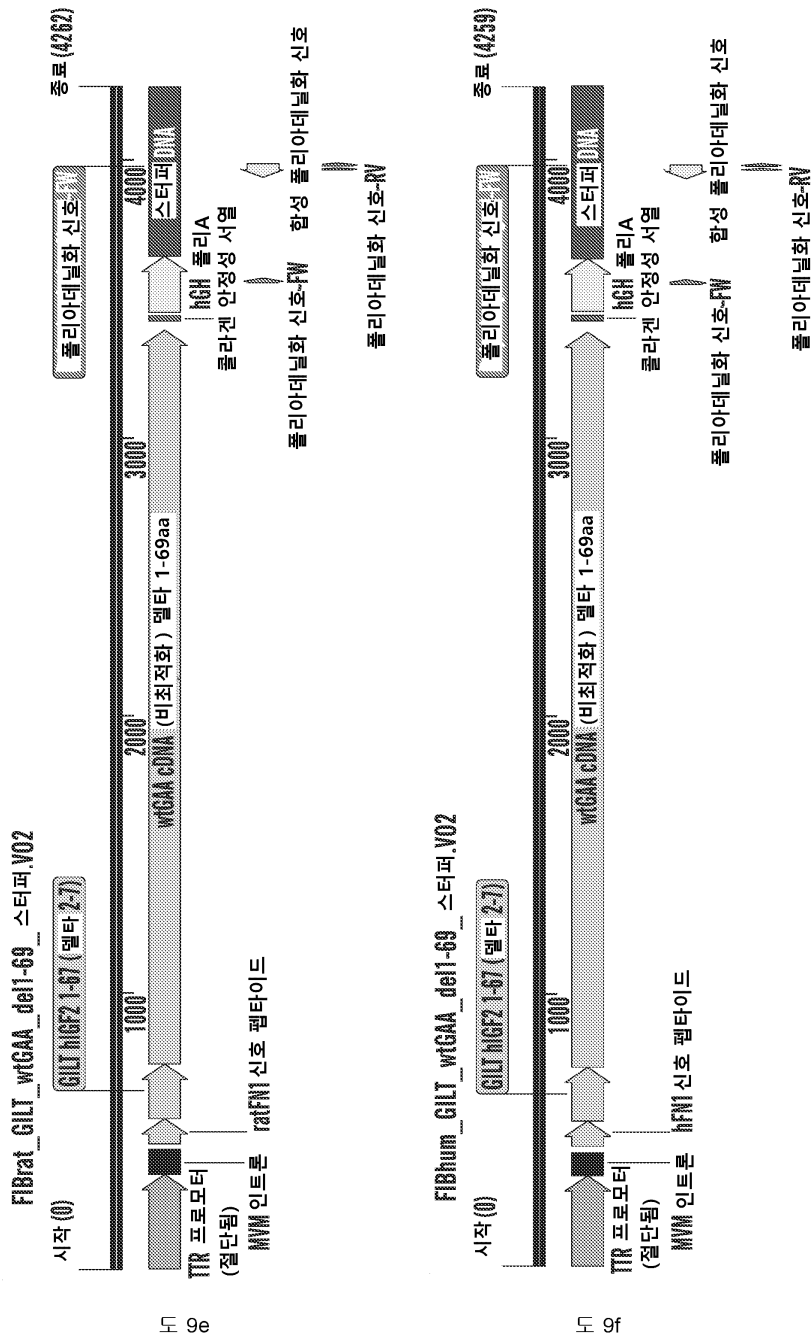


도 9b

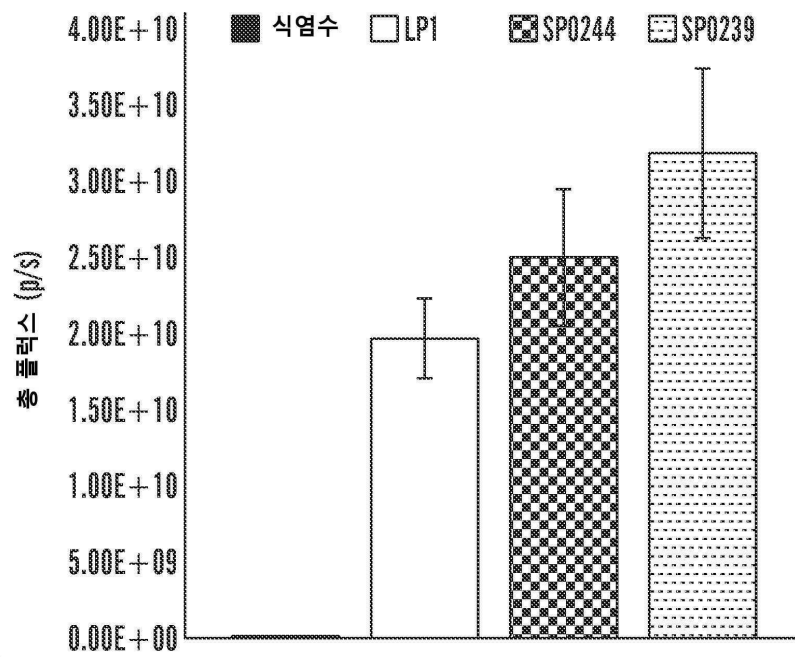
도면9cd



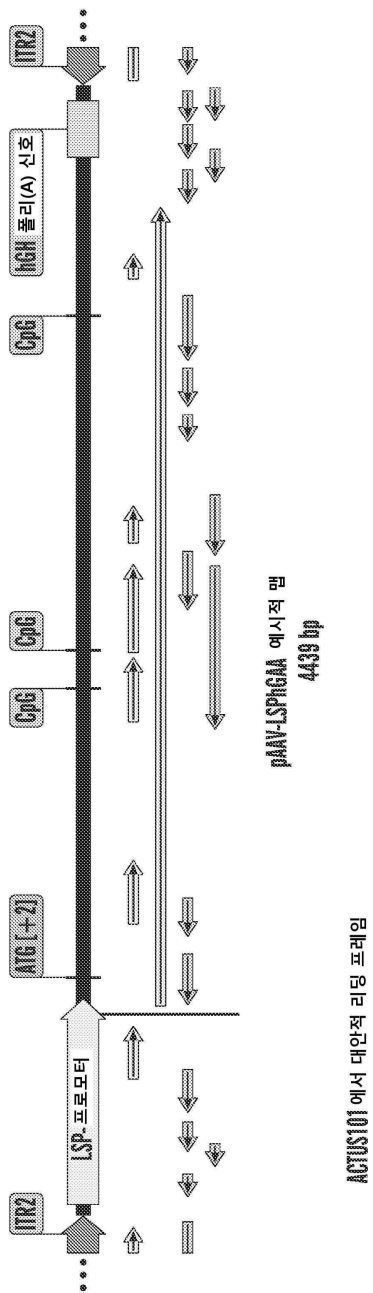
도면9ef



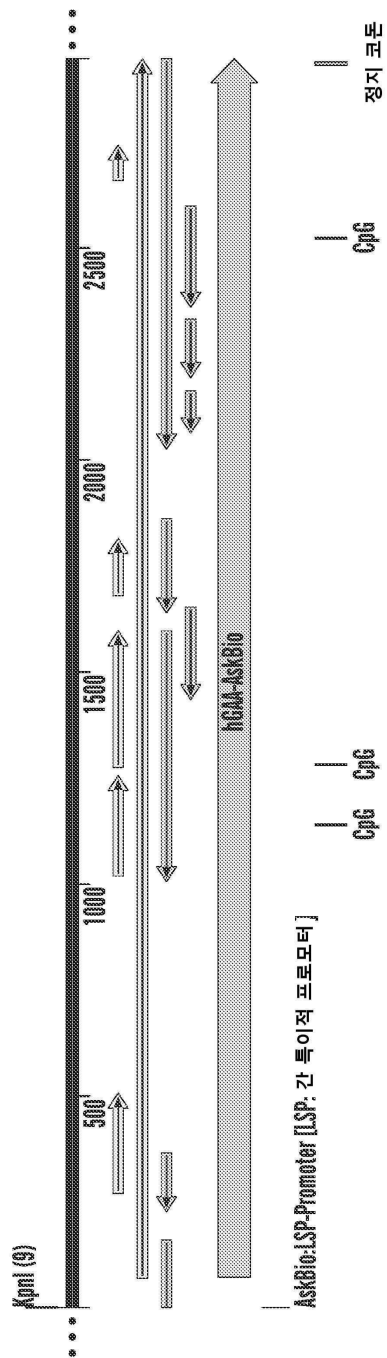
도면10



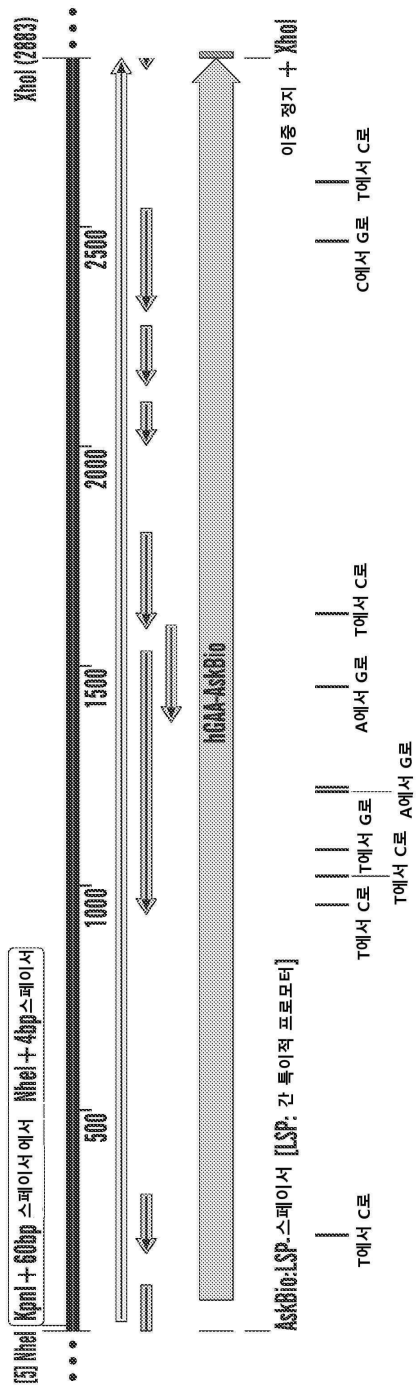
도면11a



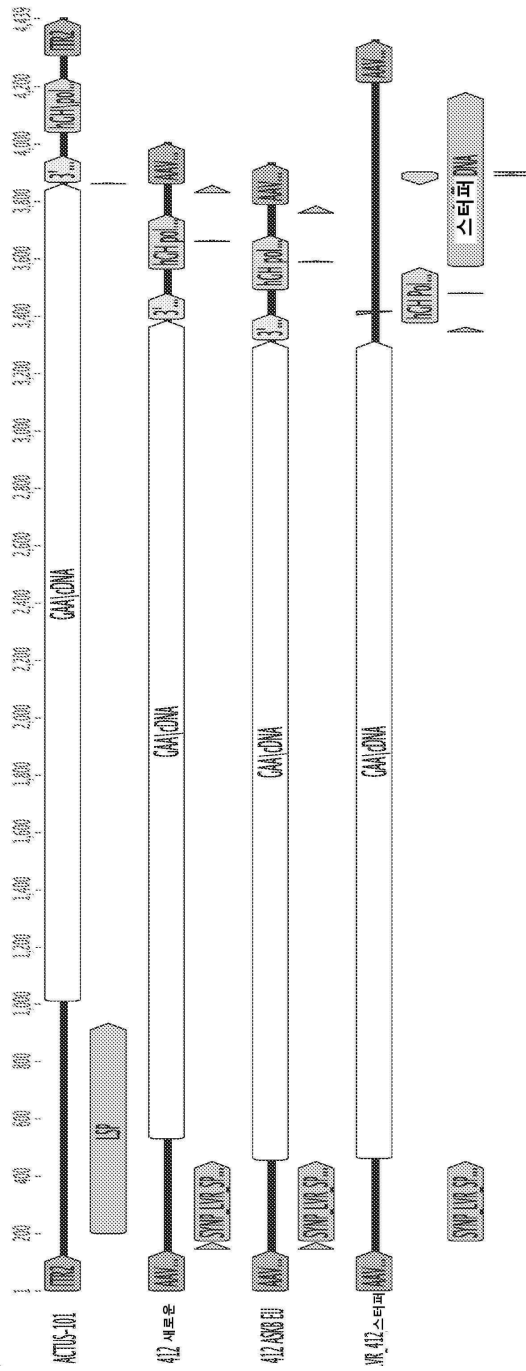
도면11c



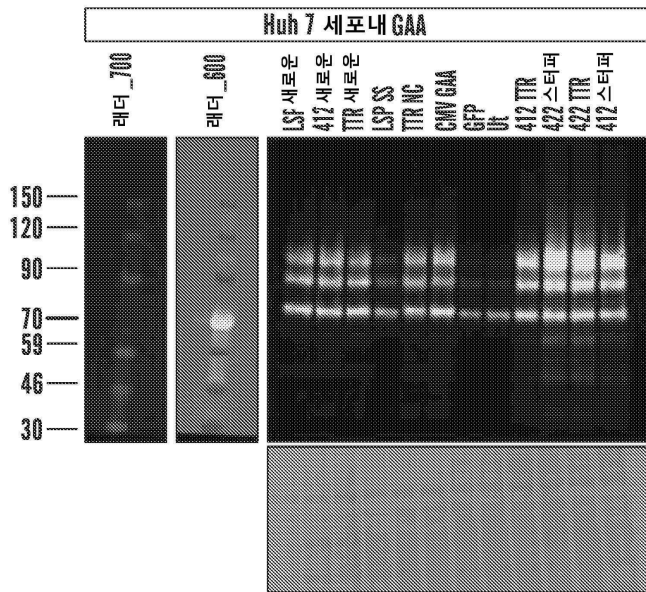
도면11d



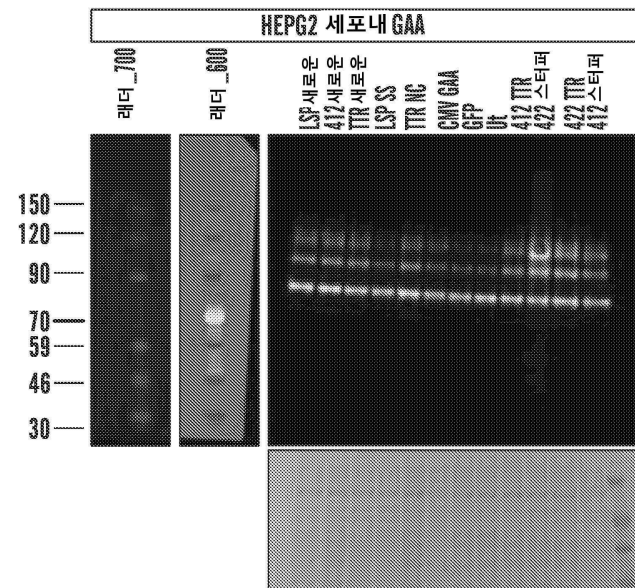
도면12



도면13a



도면13b



서열 목록

SEQUENCE LISTING

<110> ASKLEPIOS BIOPHARMACEUTICAL, INC.

<120> THERAPEUTIC ADENO-ASSOCIATED VIRUS COMPRISING LIVER-SPECIFIC
PROMOTERS FOR TREATING POMPE DISEASE AND LYSOSOMAL DISORDERS

<130> 046192-096600WOPT

<140><141><150> 63/023,570

<151> 2020-05-12

<150> 62/937,583

<151> 2019-11-19

<150> 62/937,556

<151> 2019-11-19

<160> 447

<170> PatentIn version 3.5

<210> 1

<211> 201

<212> DNA

<213> Homo sapiens

<400> 1

gcttaccgcc ccagttagac cctgtgcggc ggggagctgg tggacaccct ccagttcgctc 60

tgtggggacc gcggcttcta cttcagcagg cccgcaagcc gtgtgagccg tcgcagccgt 120

ggcatcgctt aggagtgtg tttccgcagc tgtgacctgg ccctcctgga gacgtactgt 180

gctacccccg ccaagtccga g 201

<210> 2

<211> 183

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 2

gctctgtgcg gcggggagct ggtggacacc ctccagttcg tctgtgggga ccgcggcttc 60

tacttcagca ggcccgaag ccgtgtgagc cgtcgagcc gtggcatcgt tgaggagtgc 120

tgtttccgca gctgtgacct ggccctcctg gagacgtact gtgctacccc cgccaagtcc 180

gag 183

<210> 3

<211> 180

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 3

ctgtgcggcg gggagctggt ggacacccctc cagttcgtct gtggggaccg cggcttctac 60
 ttcagcaggc ccgaagccg tgtgagccgt cgcagccgtg gcatcgttga ggagtgtgt 120
 ttccgcagct gtgacctggc cctcctggag acgtactgtg ctacccccgc caagtccgag 180

<210> 4

<211> 201

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 4

gcttaccgcc ccagttagac cctgtgcggc ggggagctgg tggacaccct ccagttcgtc 60
 tgtggggacc gcggttctta cttcagcagg cccgaagcc gtgtgagccg tcgcagccgt 120
 ggcatcatgg aggagtgtgt ttccgcagc tgtgacctgg ccctcctgga gacgtactgt 180
 gctacccccg ccaagtccga g 201

<210> 5

<211> 67

<212> PRT

<213> Homo sapiens

<400> 5

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

Lys Ser Glu

65

<210> 6

<211> 61

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polypeptide

<400> 6

Ala Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln Phe Val Cys Gly

1 5 10 15

Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala Ser Arg Val Ser Arg Arg

20 25 30

Ser Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser Cys Asp Leu Ala

35 40 45

Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala Lys Ser Glu

50 55 60

<210> 7

<211> 60

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polypeptide

<400> 7

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

1 5 10 15

Arg Gly Phe Tyr Phe Ser Arg Pro Ala Ser Arg Val Ser Arg Arg Ser

20 25 30

Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser Cys Asp Leu Ala Leu

35 40 45

Leu Glu Thr Tyr Cys Ala Thr Pro Ala Lys Ser Glu

50 55 60

<210> 8

<211> 24

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
peptide

<400> 8

Glu Glu Cys Cys Phe Arg Ser Cys Asp Leu Ala Leu Leu Glu Thr Tyr
1 5 10 15

Cys Ala Thr Pro Ala Lys Ser Glu
20

<210> 9

<211> 67

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polypeptide

<400> 9

Ala Tyr Arg Pro Ser Glu Thr Leu Cys Gly Gly Glu Leu Val Asp Thr
1 5 10 15

Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala
20 25 30

Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Met Glu Glu Cys Cys Phe

35 40 45
Arg Ser Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala

50 55 60

Lys Ser Glu

65

<210> 10

<211> 952

<212> PRT

<213> Homo sapiens

<400> 10

Met Gly Val Arg His Pro Pro Cys Ser His Arg Leu Leu Ala Val Cys
1 5 10 15

Ala Leu Val Ser Leu Ala Thr Ala Ala Leu Leu Gly His Ile Leu Leu
20 25 30

His Asp Phe Leu Leu Val Pro Arg Glu Leu Ser Gly Ser Ser Pro Val

35 40 45
 Leu Glu Glu Thr His Pro Ala His Gln Gln Gly Ala Ser Arg Pro Gly
 50 55 60
 Pro Arg Asp Ala Gln Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr
 65 70 75 80
 Gln Cys Asp Val Pro Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys
 85 90 95

 Ala Ile Thr Gln Glu Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro
 100 105 110
 Ala Lys Gln Gly Leu Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe
 115 120 125
 Phe Pro Pro Ser Tyr Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser
 130 135 140
 Glu Met Gly Tyr Thr Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe
 145 150 155 160

 Pro Lys Asp Ile Leu Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu
 165 170 175
 Asn Arg Leu His Phe Thr Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu
 180 185 190
 Val Pro Leu Glu Thr Pro His Val His Ser Arg Ala Pro Ser Pro Leu
 195 200 205
 Tyr Ser Val Glu Phe Ser Glu Glu Pro Phe Gly Val Ile Val Arg Arg
 210 215 220

 Gln Leu Asp Gly Arg Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe
 225 230 235 240
 Phe Ala Asp Gln Phe Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr
 245 250 255
 Ile Thr Gly Leu Ala Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser
 260 265 270
 Trp Thr Arg Ile Thr Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly
 275 280 285

Ala Asn Leu Tyr Gly Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly
290 295 300

Gly Ser Ala His Gly Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val
305 310 315 320

Val Leu Gln Pro Ser Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile
325 330 335

Leu Asp Val Tyr Ile Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln
340 345 350

Gln Tyr Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly
355 360 365

Leu Gly Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr
370 375 380

Arg Gln Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val
385 390 395 400

Gln Trp Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe
405 410 415

Asn Lys Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His
420 425 430

Gln Gly Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser
435 440 445

Ser Gly Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg
450 455 460

Gly Val Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val
465 470 475 480

Trp Pro Gly Ser Thr Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu
485 490 495

Ala Trp Trp Glu Asp Met Val Ala Glu Phe His Asp Gln Val Pro Phe
500 505 510

Asp Gly Met Trp Ile Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly
515 520 525

Ser Glu Asp Gly Cys Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val

530 535 540
 Pro Gly Val Val Gly Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser
 545 550 555 560
 Ser His Gln Phe Leu Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly
 565 570 575
 Leu Thr Glu Ala Ile Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly
 580 585 590
 Thr Arg Pro Phe Val Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg
 595 600 605

 Tyr Ala Gly His Trp Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu
 610 615 620
 Ala Ser Ser Val Pro Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro
 625 630 635 640
 Leu Val Gly Ala Asp Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu
 645 650 655
 Leu Cys Val Arg Trp Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg
 660 665 670

 Asn His Asn Ser Leu Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser
 675 680 685
 Glu Pro Ala Gln Gln Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala
 690 695 700
 Leu Leu Pro His Leu Tyr Thr Leu Phe His Gln Ala His Val Ala Gly
 705 710 715 720
 Glu Thr Val Ala Arg Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser
 725 730 735

 Thr Trp Thr Val Asp His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile
 740 745 750
 Thr Pro Val Leu Gln Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro
 755 760 765
 Leu Gly Thr Trp Tyr Asp Leu Gln Thr Val Pro Val Glu Ala Leu Gly
 770 775 780

Ser Leu Pro Pro Pro Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser
785 790 795 800

Glu Gly Gln Trp Val Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val
805 810 815

His Leu Arg Ala Gly Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr
820 825 830

Thr Thr Glu Ser Arg Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr
835 840 845

Lys Gly Gly Glu Ala Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser
850 855 860

Leu Glu Val Leu Glu Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala
865 870 875 880

Arg Asn Asn Thr Ile Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly
885 890 895

Ala Gly Leu Gln Leu Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala
900 905 910

Pro Gln Gln Val Leu Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr
915 920 925

Ser Pro Asp Thr Lys Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly
930 935 940

Glu Gln Phe Leu Val Ser Trp Cys
945 950

<210> 11

<211> 3837

<212> DNA

<213> Homo sapiens

<400> 11

gccccgcgac gagctcccgc cggtcacgtg acccgctctt gcgcgcccc gggcacgacc	60
ccggagtctc cgcgggcggc cagggcgcgc gtgcgcggag gtgagccggg ccggggctgc	120
ggggcttccc tgagcgcggg ccgggtcggg ggggcggtcg gctgcccgcg cggcctctca	180
gttgggaaag ctgaggttgt cgccggggcc gcgggtggag gtcggggatg aggcagcagg	240

taggacagtg acctcgggtga cgcaaggac cccggccacc tctaggttct cctcgtccgc	300
ccgtttgttca gcgagggagg ctctgcgcgt gccgcagctg acggggaaac tgaggcacgg	360
agcgggcctg taggagctgt ccagcccatc tccaacatg ggagttaggc acccgccctg	420
ctcccaccgg ctcttgccg tctgcgcct cgtgtccttg gcaaccgctg cactcctggg	480
gcacatccta ctccatgatt tctgtctggt tccccagag ctgagtggct cctccccagt	540
cctggaggag actcaccag ctcaccagca gggagccagc agaccagggc cccgggatgc	600
ccaggcacac cccggccgtc ccagagcagt gcccacacag tgcgacgtcc ccccaacag	660
ccgttcgat tgcgccctg acaaggccat caccaggaa cagtgcgagg cccgcggctg	720
ttgtacatc cctgcaaagc aggggctgca gggagcccag atggggcagc cctggtgctt	780
cttcccacc agctaccca gctacaagct ggagaacctg agtcctctg aaatgggcta	840
cacggccacc ctgaccgta ccacccccac ctctttcccc aaggacatcc tgacctgcg	900
gtcggacgtg atgatggaga ctgagaaccg cctccacttc acgatcaaag atccagctaa	960
caggcgtac gaggtgccct tggagacccc gcatgtccac agccgggcac cgtccccact	1020
ctacagcgtg gagtctccg aggagccctt cggggtgatc gtgcgccggc agctggacgg	1080
ccgcgtgctg ctgaacacga cggtagcgcc cctgttcttt gcggaccagt tccttcagct	1140
gtccacctcg ctgccctcg agtatatcac aggcctcgcc gagcacctca gtccccgat	1200
gtcagcacc agctggacca ggatcacct gtggaaccgg gacctgctg ccacgcccgg	1260
tgcgaacctc tacgggtctc accctttcta cctggcgctg gaggacggcg ggtcggcaca	1320
cggggtgttc ctgctaaaca gcaatgcat ggatgtggc ctgcagccga gccctgccct	1380
tagctggagg tcgacagtg ggatcctgga tgtctacatc ttcttgggc cagagcccaa	1440
gagcgtggtg cagcagtacc tggacgttgt gggatacccg ttcattgccg catactgggg	1500
cctgggcttc cacctgtgcc gctggggcta ctctccacc gctatcacc gccagtggt	1560
ggagaacatg accagggccc acttccccct ggacgtccag tggaacgacc tggactacat	1620
ggactcccgg agggacttca cgttcaacaa ggatggcttc cgggacttcc cggccatggt	1680
gcaggagctg caccagggcg gccggcgcta catgatgac gtggatcctg ccatcagcag	1740
ctcgggccct gccgggagct acaggcccta cgacgagggt ctgcggaggg gggttttcat	1800
caccaacgag accggccagc cgctgattgg gaaggatgg cccgggtcca ctgccttccc	1860
cgacttcacc aacccccacg ccctggcctg gtgggaggac atggtggctg agttccatga	1920
ccaggtgccc ttcgacggca tgttgattga catgaacgag ccttccaact tcatcagggg	1980
ctctgaggac ggtgcacca acaatgagct ggagaacca ccctactgac ctggggtggt	2040
tggggggacc ctccaggcgg ccaccatctg tgctccagc caccagtttc tctccacaca	2100

ctacaacctg cacaacctct acggcctgac cgaagccatc gcctcccaca gggcgctggt 2160
gaaggctcgg gggacacgcc cattttgat ctcccgtcgc acctttgctg gccacggccg 2220
atacgccggc cactggacgg gggacgtgtg gagctcctgg gagcagctcg cctcctccgt 2280
gccagaaatc ctgcagttaa acctgctggg ggtgcctctg gtcggggccg acgtctgcgg 2340

cttctggggc aacacctcag aggagctgtg tgtgcgctgg acccagctgg gggccttcta 2400
ccccctcatg cggaaccaca acagcctgct cagtctgccc caggagccgt acagcttcag 2460
cgagccggcc cagcaggcca tgaggaaagg cctcacctg cgctacgcac tctccccca 2520
cctctacaca ctgttcacc agggccacgt cgcgggggag accgtggccc ggcccccttt 2580
cctggagttc cccaaggact ctagcacctg gactgtggac caccagctcc tgtgggggga 2640
ggccctgctc atcacccag tgctccaggc cgggaaggcc gaagtgactg gctacttccc 2700
cttgggcaca tggtagacc tgcagacgtt gccagtagag gcccttggca gcctcccacc 2760

cccacctgca gctccccgtg agccagccat ccacagcgag gggcagtggtg tgacgtgcc 2820
ggccccctg gacaccatca acgtccacct cgggctggg tacatcatcc ccctgcaggg 2880
ccctggcctc acaaccacag agtcccgcca gcagcccatg gccttggtg tggcctgac 2940
caagggtggg gaggcccgag gggagctgtt ctgggacgat ggagagagcc tggaaagtgt 3000
ggagcgaggg gcctacacac aggtcatctt cctggccagg aataacacga tcgtgaatga 3060
gttggtacgt gtgaccagtg agggagctgg cctgcagctg cagaaggtga ctgtcctggg 3120
cgtggccacg gcgccccagc aggtcctctc caacggtgtc cctgtctcca acttcaccta 3180

cagccccgac accaaggctc tggacatctg tgtctcgtg ttgatgggag agcagtttct 3240
cgtcagctgg tgitagccgg gcggagtgtg ttagtctctc cagaggaggg ctggttcccc 3300
agggaagcag agcctgtgtg cgggcagcag ctgtgtgcgg gcctgggggt tgcattgttc 3360
acctggagct gggcactaac cattccaagc cgcgcgatcg cttgtttcca cctcctgggc 3420
cggggctctg gcccccaacg tgtctaggag agctttctcc ctagatcgca ctgtgggccg 3480
gggccctgga gggctgtctt gtgttaataa gattgtaagg tttgccctcc tcacctgtt 3540
ccggcatgcg ggtagtatta gccaccccc tccatctgtt cccagcacg gagaaggggg 3600

tgctcaggtg gaggtgtggg gtatgcacct gagctcctgc ttcgcgctg ctgctctgcc 3660
ccaacgcgac cgtgccccg ctgccagag ggtgggatgc ctgccgttc ccgagcaagc 3720
ctgggaactc aggaaaattc acaggacttg ggagattcta aatcttaagt gcaattattt 3780
ttaataaaaag gggcatttgg aatcagcaaa aaaaaaaaa aaaaaaaaa aaaaaaa 3837

<210> 12

<211> 349

<212> DNA

<213> Unknown

<220><223> Description of Unknown:

TTR Promoter (trunc) sequence

<400> 12

tcgagcttgg gctgcaggtc gagggcactg ggaggatgtt gagtaagatg gaaaactact 60

gatgaccctt gcagagacag agtattagga catgtttgaa caggggccgg gcgatcagca 120

ggtagctcta gaggatcccc gtctgtctgc acatttcgta gagcgagtgt tccgatactc 180

taatctccct aggcaagggt catatttttg taggttactt attctccttt tgttgactaa 240

gtcaataatc agaatcagca ggtttggagt cagcttggca gggatcagca gcctgggttg 300

gaaggagggg gtataaaagc cccttcacca ggagaagccg tcacacaga 349

<210> 13

<211> 85

<212> DNA

<213> Mouse parvovirus

<400> 13

ctagccctaa ggtaagttgg cgccgtttta gggatggttg gttggtgggg tattaatgtt 60

taattacctt tttacaggc ctgaa 85

<210> 14

<211> 441

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 14

gtacacatat tgaccaaatc agggtaattt tgcatttgta attttaaaaa atgctttctt 60

cttttaatat acttttttgt ttatcttatt tctaatactt tccctaactt ctttctttca 120

gggcaataat gatacaatgt atcatgcctc tttgcacat tctaaagaat aacagtgata 180

atttctgggt taaggcaata gcaatatttc tgcataaaa tatttctgca tataaattgt 240

aactgatgta agaggtttca tattgctaata agcagctaca atccagctac cattctgctt 300

ttattttctg gttgggataa ggctggatta ttctgagtc aagctaggcc cttttgctaa 360

tcttgttcat acctcttatt ttctccac agctcctggg caacctgctg gtctctctgc 420

tggcccatca ctttggcaaa g 441

<210> 15

<211> 196

<212> DNA

<213> Homo sapiens

<400> 15

gatctgtggc ttctagctgc ccgggtggca tcctgtgac cctccccag tgccttcct 60

ggccctggaa gttgccactc cagtgccac cagccttgc ctaataaaat taagttgcat 120

cattttgtct gactagggtg cttctataa tattatgggg tggagggggg tggataggag 180

caaggggcaa gttggg 196

<210> 16

<211> 251

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 16

aaggccgacc cgcaaatagt agatcccgcg agggcttgaa tctatcacct agagtacacc 60

ctagagaata gctagctctc aatgactaag gactaaactt ggtatttcga ctgaagcctg 120

tcccctcact gttggcgcta ggaggagagt tcgtagaaag gatagtacga ttttaagtatc 180

tctaagcctt gtgaagcact aaggttgcgt acagacgtgc ttgaattacg gataattcgg 240

gaaccttggg a 251

<210> 17

<211> 72

<212> DNA

<213> Homo sapiens

<400> 17

atgccgtctt ctgtctcgtg gggcactctc ctgctggcag gcctgtgctg cctggtcctt 60

gtctccctgg ct 72

<210> 18

<211> 32

<212> PRT

<213> Rattus norvegicus

<400> 18

Met Leu Arg Gly Pro Gly Pro Gly Arg Leu Leu Leu Leu Ala Val Leu
1 5 10 15

Cys Leu Gly Thr Ser Val Arg Cys Thr Glu Thr Gly Lys Ser Lys Arg
20 25 30

<210> 19

<211> 38

<212> PRT

<213> Rattus norvegicus

<400> 19

Met Leu Arg Gly Pro Gly Pro Gly Arg Leu Leu Leu Leu Ala Val Leu
1 5 10 15

Cys Leu Gly Thr Ser Val Arg Cys Thr Glu Thr Gly Lys Ser Lys Arg
20 25 30

Leu Ala Leu Gln Ile Val
35

<210> 20

<211> 26

<212> PRT

<213> Homo sapiens

<400> 20

Met Leu Arg Gly Pro Gly Pro Gly Leu Leu Leu Leu Ala Val Gln Cys
1 5 10 15

Leu Gly Thr Ala Val Pro Ser Thr Gly Ala
20 25

<210> 21

<211> 31

<212> PRT

<213> Xenopus laevis

<400> 21

Met Arg Arg Gly Ala Leu Thr Gly Leu Leu Leu Val Leu Cys Leu Ser
1 5 10 15

Val Val Leu Arg Ala Ala Pro Ser Ala Thr Ser Lys Lys Arg Arg

	20	25	30	
<210>	22			
<211>	72			
<212>	DNA			
<213>	Homo sapiens			
<400>	22			
	atgggaatcc caatggggaa gtcgatgctg gtgcttctca ccttcttggc cttegcctcg			60
	tgctgcattg ct			72
<210>	23			
<211>	96			
<212>	DNA			
<213>	Rattus norvegicus			
<400>	23			
	atgctcaggg gtccgggacc cgggcggctg ctgctgctag cagtcctgtg cctggggaca			60
	tcggtgcgct gcaccgaaac cgggaagagc aagagg			96
<210>				
	24			
<211>	114			
<212>	DNA			
<213>	Rattus norvegicus			
<400>	24			
	atgctcaggg gtccgggacc cgggcggctg ctgctgctag cagtcctgtg cctggggaca			60
	tcggtgcgct gcaccgaaac cgggaagagc aagaggcagg ctacagcaat cgtg			114
<210>	25			
<211>	78			
<212>	DNA			
<213>	Homo sapiens			
<400>	25			
	atgcttaggg gtccggggcc cgggctgctg ctgctggccg tccagtgcct ggggacagcg			60
	gtgccctcca cgggagcc			78
<210>	26			
<211>	93			
<212>	DNA			

<213> *Xenopus laevis*

<400> 26

atgcgcggg gggccctgac cgggctgctc ctggtcctgt gcctgagtgt tgtgctacgt 60

gcagccccct ctgcaacaag caagaagcgc agg 93

<210> 27

<211> 17

<212> PRT

<213> *Rattus norvegicus*

<400> 27

Met Thr Pro Leu Leu Leu Ala Val Leu Cys Leu Gly Thr Ala Leu

1 5 10 15

Ala

<210> 28

<211> 22

<212> PRT

<213> *Homo sapiens*

<400> 28

Met Leu Ser Phe Val Asp Thr Arg Thr Leu Leu Leu Ala Val Thr

1 5 10 15

Leu Cys Leu Ala Thr Cys

20

<210> 29

<211> 21

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 29

Met Trp Trp Arg Leu Trp Trp Leu Leu Leu Leu Leu Leu Trp

1 5 10 15

Pro Met Val Trp Ala

20

<210> 30

<211> 9

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 30

ggcgcgccg

9

<210> 31

<211> 3

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
peptide

<400> 31

Gly Ala Pro

1

<210> 32

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
peptide

<400> 32

Gly Gly Gly Gly Gly Pro

1 5

<210> 33

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
peptide

<

400> 33

Gly Gly Gly Gly Ser

1 5

<210> 34

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
peptide

<400> 34

Glu Ala Ala Ala Lys

1 5

<210> 35

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
peptide

<400> 35

Gly Gly Gly Thr Val Gly Asp Asp Asp Asp Lys

1 5 10

<210> 36

<211> 167

<212>

> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 36

aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

caatttgata aaaatcgta aattataaac aggctttgcc tgtttagcct cagtgagcga 120

gcgagcgcgc agagagggag tggccaactc catcactagg ggttcct 167

<210> 37

<211> 143

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 37

aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg	60
gataaaaatc caggctttgc ctgcctcagt gagcgagcga gcgcgcagag agggagtggc	120
caactccatc actaggggtt cct	143

<210> 38

<211> 143

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 38

aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg	60
gataaaaatc caggctttgc ctgcctcagt gagcgagcga gcgcgcagag agggagtggc	120

caactccatc actaggggtt cct	143

<210> 39

<211> 208

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 39

aggaaccct agtgatggag ttggccactc cctctctggg attgggattg gcgcctcgct	60
cgcgggattg ggattgggat tgggattggg attgggattg ataaaaatca atcccaatcc	120
caatcccaat cccaatccca atcccgcgag cgagcgcgca atcccaatcc cagagaggga	180
gtggccaact ccatcactag gggttcct	208

<210> 40

<211> 199

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 40

aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcgggattg	60
ggattgggat tgggattggg attgggattg ataaaaatca atccaatcc caatccaat	120
ccaatccca atcccgcgag cgagcgcgca ggagaggag tggccaactc catcactagg	180
ggttcctaag cttattata	199

<210> 41

<211> 154

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 41

aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctactgagg	60
gcgcctataa agataaaaat ccaggctttg cctgcctcag ttagcgagcg agcgcgcaga	120
gaggagtg ccaactccat cactaggggt tcct	154

<210> 42

<211> 127

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 42

ctagtgatgg agttggccac tccctctctg cgcgctcgct cgctcactga gggataaaaa	60
tccaggcttt gcctgcctca gtgagcgagc gagcgcgcag agagggagtg gccaaactcca	120
tcactag	127

<210> 43

<211> 2208

<212> DNA

<213> Adeno-associated virus 3B

<400> 43

atggctgctg acggttatct tccagattgg ctcgaggaca acctttctga aggcattcgt	60
gagtgtgtgg ctctgaaacc tggagtcct caacccaaag cgaaccaaca acaccaggac	120

aaccgtcggg gtcttgtgct tccgggttac aaatacctcg gacccggtaa cggactcgac	180
aaaggagagc cggtaacga ggcggacgcg gcagccctcg aacacgacaa agcttacgac	240
cagcagctca aggccggtga caacccttac ctcaagtaca accacgccga cggcagttt	300
caggagcgtc ttcaagaaga tacgtctttt gggggcaacc ttggcagagc agtcttcag	360
gccaaaaaga ggatccttga gcctcttggc ctggttgagg aagcagctaa aacggctcct	420
ggaaagaaga ggcctgtaga tcagtctcct caggaaccgg actcatcatc tgggtgtggc	480
aaatcgggca aacagcctgc cagaaaaaga ctaaatttcg gtcagactgg cgactcagag	540
tcagtccag accctcaacc tctcggagaa ccaccagcag cccccacaag ttgggatct	600
aatacaatgg cttaggcggg tggcgaccca atggcagaca ataacgaggg tgccgatgga	660
gtgggtaatt cctcaggaaa ttggcattgc gattccaat ggctgggcca cagagtcac	720
accaccagca ccagaacctg ggccctgccc acttacaaca accatctcta caagcaaate	780
tccagccaat caggagcttc aaacgacaac cactactttg gctacagcac cccttggggg	840
tattttgact ttaacagatt ccaactgccac ttctcaccac gtgactggca gcgactcatt	900
aacaacaact ggggattccg gcccaagaaa ctcagcttca agctcttcaa catccaagtt	960
aaagaggta cgcagaacga tggcacgacg actattgcca ataaccttac cagcacgggt	1020
caagtgttta cggactcgga gtatcagctc ccgtacgtgc tcgggtcggc gcaccaaggc	1080
tgtctccgc cgtttcagc ggacgtcttc atggtcctc agtatggata cctcacctg	1140
aacaacggaa gtcaagcgtt gggacgtcct tccttttact gcctggagta ctcccttcg	1200
cagatgctaa ggactggaaa taacttccaa ttcagctata ccttcgagga tgtaccttt	1260
cacagcagct acgtcacag ccagagttag gatcgctga tgaatcctct tattgatcag	1320
tatctgtact acctgaacag aacgcaagga acaacctctg gaacaacaa ccaatcacgg	1380
ctgcttttta gccaggtgg gcctcagctc atgtctttgc agccagaaa ttggctacct	1440
gggccctgct accggcaaca gagactttca aagactgcta acgacaaca caacagtaac	1500
tttcttggga cagcggccag caaatatcat ctcaatggcc gcgactcgt ggtgaatcca	1560
ggaccagcta tggccagtca caaggacgat gaagaaaaat tttccctat gcacggcaat	1620
ctaataattg gcaagaagg gacaacggca agtaacgcag aattagataa tgtaatgatt	1680
acggatgaag aagagattcg taccaccaat cctgtggcaa cagagcagta tggaaactgtg	1740
gcaaataact tgcagagctc aaatacagct cccacgacta gaactgtcaa tgatcagggg	1800
gccttacctg gcatggtgtg gcaagatcgt gacgtgtacc ttcaaggacc tatctgggca	1860

aagattcctc acacggatgg acactttcat ccttctcctc tgatgggagg ctttggactg 1920

aaacatccgc ctctcaaat catgatcaaa aatactccgg taccggcaaa tcctccgacg 1980

actttcagcc cggccaagtt tgcttcattt atcactcagt actccactgg acaggtcagc 2040

gtggaaattg agtgggagct acagaaagaa aacagcaaac gttggaatcc agagattcag 2100

tacacttcca actacaacaa gtctgttaat gtggacttta ctgtagacac taatggtgtt 2160

tatagtgaac ctgcacctat tggaacccgg tatctcacac gaaacttg 2208

<210> 44

<211> 736

<212> PRT

<213> Adeno-associated virus 3B

<400> 44

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser

1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Val Pro Gln Pro

20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Arg Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly

100 105 110

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

115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140

Pro Val Asp Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Val Gly

145 150 155 160

Lys Ser Gly Lys Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
 165 170 175
 Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
 180 185 190
 Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly
 195 200 205
 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
 210 215 220
 Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
 225 230 235 240
 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
 245 250 255
 Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
 260 265 270
 Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
 275 280 285
 Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
 290 295 300
 Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln Val
 305 310 315 320
 Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
 325 330 335
 Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
 340 345 350
 Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
 355 360 365
 Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
 370 375 380
 Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
 385 390 395 400
 Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe Glu

405 410 415
 Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
 420 425 430
 Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg Thr
 435 440 445
 Gln Gly Thr Thr Ser Gly Thr Thr Asn Gln Ser Arg Leu Leu Phe Ser

 450 455 460
 Gln Ala Gly Pro Gln Ser Met Ser Leu Gln Ala Arg Asn Trp Leu Pro
 465 470 475 480
 Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Thr Ala Asn Asp Asn
 485 490 495
 Asn Asn Ser Asn Phe Pro Trp Thr Ala Ala Ser Lys Tyr His Leu Asn
 500 505 510
 Gly Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys

 515 520 525
 Asp Asp Glu Glu Lys Phe Phe Pro Met His Gly Asn Leu Ile Phe Gly
 530 535 540
 Lys Glu Gly Thr Thr Ala Ser Asn Ala Glu Leu Asp Asn Val Met Ile
 545 550 555 560
 Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln
 565 570 575
 Tyr Gly Thr Val Ala Asn Asn Leu Gln Ser Ser Asn Thr Ala Pro Thr

 580 585 590
 Thr Arg Thr Val Asn Asp Gln Gly Ala Leu Pro Gly Met Val Trp Gln
 595 600 605
 Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
 610 615 620
 Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
 625 630 635 640
 Lys His Pro Pro Pro Gln Ile Met Ile Lys Asn Thr Pro Val Pro Ala

 645 650 655

Asn Pro Pro Thr Thr Phe Ser Pro Ala Lys Phe Ala Ser Phe Ile Thr
660 665 670
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700
Tyr Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val

705 710 715 720
Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 45

<211> 2208

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 45

atggctgctg acggttatct tccagattgg ctcgaggaca accitttctga aggcattcgt	60
gagtgggtggg ctctgaaacc tggagtcctt caaccctaaag cgaaccaaca acaccaggac	120
aaccgtcggg gtcttgtgct tccgggttac aaatacctcg gaccctgtaa cggactcgac	180
aaaggagagc cggtaacga ggcggacgcg gcagccctcg aacacgacaa agcttacgac	240
cagcagctca aggccgtga caaccgtac ctcaagtaca accacgccga cggcagttt	300
caggagcgtc ttcaagaaga tacgtctttt gggggcaacc ttggcagagc agtcttcag	360
gccaaaaaga ggatccttga gcctcttggt ctggttgagg aagcagctaa aacggctcct	420
ggaaaaaga ggcctgtaga tcagtctcct caggaaccgg actcatcctc tgggtgtggc	480
aaatcgggca aacagcctgc cagaaaaaga ctaaatttcg gtcagactgg cgactcagag	540
tcagtcccag accctcaacc tctcggagaa ccaccagcag cccccacaag ttgggatct	600
aatacaatgg cttcaggcgg tggcgacca atggcagaca ataacgaggg tgccgatgga	660
gtgggtaatt cctcaggaaa ttggcattgc gattccaat ggctgggcga cagagtcctc	720
accaccagca ccagaacctg ggcctgccc acttacaaca accatctcta caagcaaac	780
tccagccaat cagatgcttc aaacgacaac cactactttg gctacagcac cccttggggg	840
tattttgact ttaacagatt ccaactgccac ttctaccac gtgactggca gcgactcatt	900

aacaacaact ggggattccg gcccaagaaa ctcagcttca agctcttcaa catccaagtt 960

aaagaggta cgcagaacga tggcagcagc actattgccataaaccttac cagcagcgtt 1020

caagtgttta cggactcgga gtatcagctc ccgtacgtgc tcgggtcggc gcaccaaggc 1080

tgtctccgc cgttccagc ggacgtcttc atggtccctc agtatggata cctcacctg 1140

aacaacggaa gtcaagcgtt gggacgtc tcttttact gcctggagta cttcccttcg 1200

cagatgctaa ggactggaaa taacttccaa ttcagctata cttcgagga tgtacctttt 1260

cacagcagct acgtcacag ccagagtttg gatcgcttga tgaatcctct tattgatcag 1320

tatctgtact acctgaacag aacgcaagga acaacctctg gaacaaccaa ccaatcacgg 1380

ctgcttttta gccaggtcgg gcctcagctc atgtctttgc aggccagaaa ttggctacct 1440

gggccctgct accggcaaca gagactttca aagactgcta acgacaacaa caacagtaac 1500

tttcccttga cagcggccag caaatatcat ctcaatggcc gcgactcgct ggtgaatcca 1560

ggaccagcta tggccagta caaggacgat gaagaaaaat ttttcctat gcacggcaat 1620

ctaataattg gcaagaagg gacaacggca agtaacgcag aattagataa tgtaatgatt 1680

acggtatgaag aagagattcg taccaccaat cctgtggcaa cagagcagta tggaaactgtg 1740

gcaaataact tgcagagctc aaatacagct cccacgacta gaactgtcaa tgatcagggg 1800

gccttacctg gcatggtgtg gcaagatcgt gacgtgtacc ttcaaggacc tatctgggca 1860

aagattcctc acacggatgg acactttcat ccttctcctc tgatgggagg ctttggactg 1920

aaacatccgc ctctcaaat catgatcaaa aatactccgg taccggcaaa tctccgacg 1980

actttcagcc cggccaagtt tgcttcattt atcactcagt actccactgg acaggtcagc 2040

gtggaaattg agtgggagct acagaaagaa aacagcaaac gttggaatcc agagattcag 2100

tacacttcca actacaacaa gtctgttaat gtggacttta ctgtagacac taatggtgtt 2160

tatagtgaac ctgcacctat tggaacccgg tatctcacac gaaacttg 2208

<210> 46

<211> 736

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 46

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser

1

5

10

15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Val Pro Gln Pro
 20 25 30
 Lys Ala Asn Gln Gln His Gln Asp Asn Arg Arg Gly Leu Val Leu Pro
 35 40 45
 Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60
 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
 85 90 95
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
 100 105 110
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Ile Leu Glu Pro
 115 120 125
 Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
 130 135 140
 Pro Val Asp Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Val Gly
 145 150 155 160
 Lys Ser Gly Lys Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
 165 170 175
 Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
 180 185 190
 Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly
 195 200 205
 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
 210 215 220
 Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
 225 230 235 240
 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
 245 250 255
 Tyr Lys Gln Ile Ser Ser Gln Ser Asp Ala Ser Asn Asp Asn His Tyr

260 265 270
 Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
 275 280 285
 Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
 290 295 300
 Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln Val

 305 310 315 320
 Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
 325 330 335
 Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
 340 345 350
 Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
 355 360 365
 Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser

 370 375 380
 Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
 385 390 395 400
 Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe Glu
 405 410 415
 Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
 420 425 430
 Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg Thr

 435 440 445
 Gln Gly Thr Thr Ser Gly Thr Thr Asn Gln Ser Arg Leu Leu Phe Ser
 450 455 460
 Gln Ala Gly Pro Gln Ser Met Ser Leu Gln Ala Arg Asn Trp Leu Pro
 465 470 475 480
 Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Thr Ala Asn Asp Asn
 485 490 495
 Asn Asn Ser Asn Phe Pro Trp Thr Ala Ala Ser Lys Tyr His Leu Asn

 500 505 510

Gly Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Asp Asp Glu Glu Lys Phe Phe Pro Met His Gly Asn Leu Ile Phe Gly
530 535 540

Lys Glu Gly Thr Thr Ala Ser Asn Ala Glu Leu Asp Asn Val Met Ile
545 550 555 560

Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln
565 570 575

Tyr Gly Thr Val Ala Asn Asn Leu Gln Ser Ser Asn Thr Ala Pro Thr
580 585 590

Thr Arg Thr Val Asn Asp Gln Gly Ala Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
625 630 635 640

Lys His Pro Pro Pro Gln Ile Met Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asn Pro Pro Thr Thr Phe Ser Pro Ala Lys Phe Ala Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 47

<211> 2208

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 47

atggctgctg acggttatct tccagattgg ctcgaggaca acctttctga aggcattcgt	60
gagtgggtggg ctctgaaacc tggagtcctt caacccaaag cgaaccaaca acaccaggac	120
aaccgtcggg gtcttgtgct tccgggttac aaatacctcg gacccggtta cggactcgac	180
aaaggagagc cggltcaacga ggcggacgcg gcagccctcg aacacgacaa agcttacgac	240
cagcagctca aggccgtga caaccctac ctcaagtaca accacgccga cgccgagttt	300
caggagcgtc ttcaagaaga tacgtctttt gggggcaacc ttggcagagc agtcttcag	360
gccaaaaaga ggatccttga gcctcttggc ctggttgagg aagcagctaa aacggctcct	420
ggaaagaaga ggcctgtaga tcagtctcct caggaaccgg actcatcatc tgggtgtggc	480
aaatcgggca aacagcctgc cagaaaaaga ctaaatttcg gtcagactgg cgactcagag	540
tcagtcccag accctcaacc tctcggagaa ccaccagcag cccccacaag ttgggatct	600
aatacaatgg cttcaggcgg tggcgacca atggcagaca ataacgaggg tgccgatgga	660
gtgggttaatt cctcaggaat ttggcattgc gattcccaat ggctggcgca cagagtcac	720
accaccagca ccagaacctg ggccttgcct acttacaaca accatctcta caagcaaac	780
tccagccaat cagatgcttc aaacgacaac cactactttg gctacagcac cccttggggg	840
tattttgact ttaacagatt ccactgccac ttctcaccac gtgactggca gcgactcatt	900
aacaacaact ggggattccg gcccagaaa ctcagcttca agctcttcaa catccaagtt	960
aaagaggta cgcagaacga tggcacgacg actattgcca ataaccttac cagcacgggt	1020
caagtgttta cggactcgga gtatcagctc ccgtacgtgc tcgggtcggc gcaccaaggc	1080
tgtctcccgc cgtttccagc ggacgtcttc atggctccctc agtatggata cctcacctg	1140
aacaacggaa gtcaagcggg gggacgtca tccttttact gcctggagta cttcccttcg	1200
cagatgctaa ggactggaaa taacttccaa ttcagctata ccttcgagga tgtaccttt	1260
cacagcagct acgtcacag ccagagtttg gatcgcttga tgaatcctct tattgatcag	1320
tatctgtact acctgaacag aacgcaagga acaacctctg gaacaacaa ccaatcacgg	1380
ctgcttttta gccaggctgg gcctcagctc atgtctttgc aggccagaaa ttggctacct	1440
gggccctgct accggcaaca gagactttca aaggtagcta acgacaacaa caacagtaac	1500
tttccttgga cagcggccag caaatatcat ctcaatggcc gcgactcgct ggtgaatcca	1560
ggaccagcta tggccagta caaggacgat gaagaaaaat ttttccctat gcacggcaat	1620
ctaataatttg gcaagaagg gacaacggca agtaacgcag aattagataa tgtaatgatt	1680

acggatgaag aagagattcg taccaccaat cctgtggcaa cagagcagta tggaactgtg 1740

gcaaataact tgcagagctc aaatacagct cccacgacta gaactgtcaa tgatcagggg 1800

gccttacctg gcatggtgtg gcaagatcgt gacgtgtacc ttcaaggacc tatctgggca 1860

aagattcctc acacggatgg acactttcat ccttctctc tgatgggagg ctttgactg 1920

aaacatccgc ctctcaaat catgatcaaa aatactcggg taccggcaaa tcctccgacg 1980

actttcgtag cgccaagtt tgcttcattt atcactcagt actccactgg acaggtcagc 2040

gtggaaattg agtgggagct acagaaagaa aacagcaaac gttggaatcc agagattcag 2100

tacattcca actacaaca gtctgttaat gtggacttta ctgtagacac taatggtgtt 2160

tatagtgaac ctgcacctat tggaacccgg tatctcacac gaaacttg 2208

<210> 48

<211> 736

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 48

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser

1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Val Pro Gln Pro

20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Arg Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly

100 105 110

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

115 120 125
 Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
 130 135 140
 Pro Val Asp Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Val Gly
 145 150 155 160
 Lys Ser Gly Lys Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr

 165 170 175
 Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
 180 185 190
 Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly
 195 200 205
 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
 210 215 220
 Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile

 225 230 235 240
 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
 245 250 255
 Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
 260 265 270
 Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
 275 280 285
 Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp

 290 295 300
 Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln Val
 305 310 315 320
 Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
 325 330 335
 Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
 340 345 350
 Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp

 355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg Thr
435 440 445

Gln Gly Thr Thr Ser Gly Thr Thr Asn Gln Ser Arg Leu Leu Phe Ser
450 455 460

Gln Ala Gly Pro Gln Ser Met Ser Leu Gln Ala Arg Asn Trp Leu Pro
465 470 475 480

Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Val Ala Asn Asp Asn
485 490 495

Asn Asn Ser Asn Phe Pro Trp Thr Ala Ala Ser Lys Tyr His Leu Asn
500 505 510

Gly Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Asp Asp Glu Glu Lys Phe Phe Pro Met His Gly Asn Leu Ile Phe Gly
530 535 540

Lys Glu Gly Thr Thr Ala Ser Asn Ala Glu Leu Asp Asn Val Met Ile
545 550 555 560

Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln
565 570 575

Tyr Gly Thr Val Ala Asn Asn Leu Gln Ser Ser Asn Thr Ala Pro Thr
580 585 590

Thr Arg Thr Val Asn Asp Gln Gly Ala Leu Pro Gly Met Val Trp Gln
595 600 605

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

610 615 620
 Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
 625 630 635 640
 Lys His Pro Pro Pro Gln Ile Met Ile Lys Asn Thr Pro Val Pro Ala
 645 650 655
 Asn Pro Pro Thr Thr Phe Val Pro Ala Lys Phe Ala Ser Phe Ile Thr
 660 665 670
 Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln

675 680 685
 Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
 690 695 700
 Tyr Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val
 705 710 715 720
 Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
 725 730 735

<210> 49

<211> 2134

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<400> 49

gaaacctgga gtcctcaac ccaaagcgaa ccaacaacac caggacaacc gtcggggtct	60
tgtgcttcgg ggttacaat acctcggacc cggtaacgga ctgacaaag gagagccggt	120
caacgaggcg gacgcggcag ccctcgaaca cgacaaagct tacgaccagc agctcaaggc	180
cggtgacaac ccgtacctca agtacaacca cgccgacgcc gagtttcagg agcgtcttca	240
agaagatacg tcttttgggg gcaaccttgg cagagcagtc ttccaggcca aaaagaggat	300
ccttgagcct cttggtctgg ttgaggaagc agctaaaacg gtccttgga agaagaggcc	360
tgtagatcag tctcctcagg aaccggactc atcatctggt gttggcaaat cgggcaaaca	420
gcctgccaga aaaagactaa atttcggtca gactggcgac tcagagtcag tcccagacc	480
tcaacctctc ggagaaccac cagcagcccc cacaagttag ggatctaata caatggcttc	540
aggcgggtggc gcaccaatgg cagacaataa cgagggtgcc gatggagtgg gtaattcctc	600

aggaaattgg cattgcgatt cccaatggct gggcgacaga gtcacacca ccagcaccag 660
aacctgggcc ctgcccactt acaacaacca tctctacaag caaatctcca gccaatcagg 720
agcttcaaac gacaaccact actttggcta cagcacccct tgggggtatt ttgactttaa 780

cagattccac tgccacttct caccacgtga ctggcagcga ctcatlaaca acaactgggg 840
attccggccc aagaaactca gcttcaagct cttaacatc caagttaaag aggtcacgca 900
gaacgatggc acgacgacta ttgccaataa ccttaccagc acggttcaag tgtttacgga 960
ctcggagtat cagctcccgt acgtgctcgg gtcggcgcac caaggctgtc tcccgcggtt 1020
tccagcggac gtcttcattg tccctcagta tggatactc accctgaaca acggaagtca 1080
agcgggtggga cgtcctcct tttactgcct ggagtacttc ccttcgcaga tgctaaggac 1140
tggaaataac ttccaattca gctatacctt cgaggatgta ccttttcaca gcagctacgc 1200

tcacagccag agtttggatc gcttgatgaa tctctttatt gatcagtatc tgtactacct 1260
gaacagaacg caaggaacaa cctctggaac aaccaacaa tcacggctgc tttttagcca 1320
ggctgggcct cagtctatgt ctttgcaggc cagaaattgg ctacctgggc cctgctaccg 1380
gcaacagaga ctttcaaaga ctgctaacga caacaacaac agtaactttc cttggacagc 1440
ggccagcaaa tatcatctca atggccgcga ctgctggtg aatccaggac cagctatggc 1500
cagtcaaacg gacgatgaag aaaaattttt ccctatgcac ggcaatctaa tatttggcaa 1560
agaagggaca acggcaagta acgcagaatt agataatgta atgattacgg atgaagaaga 1620

gattcgtacc accaatcctg tggcaacaga gcagtatgga actgtggcaa ataacttgca 1680
gagctcaaat acagctccca cgactagaac tgtaaatgat cagggggcct tacctggcat 1740
ggtgtggcaa gatcgtgacg tgtaccttca aggacctatc tgggcaaaga ttcctcacac 1800
ggatggacac tttcatcctt ctctctgat gggaggtttt ggactgaaac atccgcctcc 1860
tcaaatcatg atcaaaaata ctccggtacc ggcaaatcct ccgacgactt tcagcccggc 1920
caagtttgct tcatttatca ctcagtactc cactggacag gtcagcgtgg aaattgagtg 1980
ggagctacag aaagaaaaca gcaaacgttg gaatccagag attcagtaca cttccaacta 2040

caacaagtct gttaatgtgg actttactgt agacactaat ggtgtttata gtgaacctcg 2100
ccctattgga acccggtatc tcacacgaaa ctgtg 2134

<210> 50

<211> 736

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 50

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser

1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Val Pro Gln Pro

20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Arg Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly

100 105 110

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

115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140

Pro Val Asp Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Val Gly

145 150 155 160

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

165 170 175

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro

180 185 190

Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly

195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser

210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile

225 230 235 240
 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
 245 250 255
 Tyr Lys Gln Ile Ser Ser Gln Ser Asp Ala Ser Asn Asp Asn His Tyr
 260 265 270
 Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
 275 280 285

 Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
 290 295 300
 Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln Val
 305 310 315 320
 Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
 325 330 335
 Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
 340 345 350

 Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
 355 360 365
 Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
 370 375 380
 Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
 385 390 395 400
 Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe Glu
 405 410 415

 Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
 420 425 430
 Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg Thr
 435 440 445
 Gln Gly Thr Thr Ser Gly Thr Thr Asn Gln Ser Arg Leu Leu Phe Ser
 450 455 460
 Gln Ala Gly Pro Gln Ser Met Ser Leu Gln Ala Arg Asn Trp Leu Pro
 465 470 475 480

Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Thr Ala Asn Asp Asn
485 490 495

Asn Asn Ser Asn Phe Pro Trp Thr Ala Ala Ser Lys Tyr His Leu Asn
500 505 510

Gly Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Asp Asp Glu Glu Lys Phe Phe Pro Met His Gly Asn Leu Ile Phe Gly
530 535 540

Lys Glu Gly Thr Ala Ala Ser Asn Ala Glu Leu Asp Asn Val Met Ile
545 550 555 560

Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln
565 570 575

Tyr Gly Thr Val Ala Asn Asn Leu Gln Ser Ser Asn Thr Ala Pro Thr
580 585 590

Thr Arg Thr Val Asn Asp Gln Gly Ala Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
625 630 635 640

Lys His Pro Pro Pro Gln Ile Met Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asn Pro Pro Thr Thr Phe Ser Pro Ala Lys Phe Ala Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val
705 710 715 720

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

725

730

735

<210> 51

<211> 2208

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 51

atggctgctg acggttatct tccagattgg ctcgaggaca acctttctga aggcattcgt	60
gagtgggtggg ctctgaaacc tggagtcctt caacccaaag cgaaccaaca acaccaggac	120
aaccgtcggg gtcttgtgct tccgggttac aaatacctcg gaccgggtta cggactcgac	180
aaaggagagc cggtaacga ggcggacgcg gcagccctcg aacacgacaa agcttacgac	240
cagcagctca aggccgggtga caaccgtac ctcaagtaca accacgccga cgccgagttt	300
caggagcgtc ttcaagaaga tacgtctttt gggggcaacc ttggcagagc agtcttcag	360
gccaaaaaga ggatccttga gcctcttggf ctggttgagg aagcagctaa aacggctcct	420
ggaaagaaga ggcctgtaga tcagtctcct caggaaccgg actcatcatc tgggtttggc	480
aatcgggca aacagcctgc cagaaaaaga ctaaatttcg gtcagactgg cgactcagag	540
tcagtcccag accctcaacc tctcggagaa ccaccagcag cccccacaag tttgggatct	600
aatacaatgg cttcaggcgg tggcgacca atggcagaca ataacgaggg tgccgatgga	660
gtgggtaatt cctcaggaaa ttggcattgc gattccaat ggctgggcga cagagtcac	720
accaccagca ccagaacctg ggccctgccc acttacaaca accatctcta caagcaaate	780
tccagccaat caggagcttc aaacgacaac cactactttg gctacagcac cccttggggg	840
tattttgact ttaacagatt ccactgccac ttctcaccac gtgactggca gcgactcatt	900
aacaacaact ggggattccg gcccaagaaa ctcagcttca agctcttcaa catccaagtt	960
aaagaggta cgcagaacga tggcagcagc actattgcc aataaccttac cagcacggtt	1020
caagtgttta cggactcgga gtatcagctc ccgtacgtgc tcgggtcggc gcaccaaggc	1080
tgtctccgcg cgtttccagc ggacgtcttc atggtcctc agtatggata cctcacctg	1140
aacaacggaa gtcaagcggg gggacgctca tccttttact gccaggagta cttcccttcg	1200
cagatgctaa ggactggaaa taacttcaa ttcagctata ccttcgagga tgtacctttt	1260
cacagcagct acgtcacag ccagagtttg gatcgcttga tgaatcctct tattgatcag	1320
tatctgtact acctgaacag aacgcaagga acaacctctg gaacaacaa ccaatcacgg	1380

ctgcttttta gccaggctgg gcctcagtct atgtctttgc aggccagaaa ttggctacct 1440
gggccctgct accggcaaca gagactttca aagactgcta acgacaacaa caacagtaac 1500
tttccttggg cagcggccag caaatatcat ctcaatggcc gcgactcgct ggtgaatcca 1560

ggaccagcta tggccagtca caaggacgat gaagaaaaat ttttcctat gcacggcaat 1620
ctaataatttg gcaagaagg gacaacggca agtaacgcag aattagataa tgtaatgatt 1680
acggatgaag aagagattcg taccaccaat cctgtggcaa cagagcagta tggaaactgtg 1740
gcaaataact tgcagagctc aaatacagct cccacgacta gaactgtcaa tgatcagggg 1800
gccttacctg gcatggtgtg gcaagatcgt gacgtgtacc ttcaaggacc tatctgggca 1860
aagattcctc acacggatgg acattttcat ccttctcctc tgatgggagg ctttggactg 1920
aaacatccgc ctctcaaat catgatcaaa aatactccgg taccggcaaa tcctccgacg 1980

actttcagcc cgccaagt tgcctcattt atcactcagt actccactgg acaggtcagc 2040
gtggaaattg agtgggagct acagaaagaa aacagcaaac gttggaatcc agagattcag 2100
tacacttcca actacaacaa gtctgttaat gtggacttta ctgtagacac taatggtgtt 2160
tatagtgaac ctgccttat tggaacccgg tatctcacac gaaacttg 2208

<210> 52

<211> 736

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 52

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser

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

65	70	75	80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Ile Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Asp Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Val Gly
145 150 155 160
Lys Ser Gly Lys Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
180 185 190
Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly
195 200 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255
Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270
Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285
Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300
Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320
Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu

325 330 335
 Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
 340 345 350
 Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
 355 360 365
 Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
 370 375 380
 Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser

 385 390 395 400
 Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe Glu
 405 410 415
 Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
 420 425 430
 Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg Thr
 435 440 445
 Gln Gly Thr Thr Ser Gly Thr Thr Asn Gln Ser Arg Leu Leu Phe Ser

 450 455 460
 Gln Ala Gly Pro Gln Ser Met Ser Leu Gln Ala Arg Asn Trp Leu Pro
 465 470 475 480
 Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Thr Ala Asn Asp Asn
 485 490 495
 Asn Asn Ser Asn Phe Pro Trp Thr Ala Ala Ser Lys Tyr His Leu Asn
 500 505 510
 Gly Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys

 515 520 525
 Asp Asp Glu Glu Lys Phe Phe Pro Met His Gly Asn Leu Ile Phe Gly
 530 535 540
 Lys Glu Gly Thr Ala Ala Ser Asn Ala Glu Leu Asp Asn Val Met Ile
 545 550 555 560
 Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln
 565 570 575

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

580 585 590

Thr Arg Thr Val Asn Asp Gln Gly Ala Leu Pro Gly Met Val Trp Gln

595 600 605

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

610 615 620

Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu

625 630 635 640

Lys His Pro Pro Pro Gln Ile Met Ile Lys Asn Thr Pro Val Pro Ala

645 650 655

Asn Pro Pro Thr Thr Phe Ser Pro Ala Lys Phe Ala Ser Phe Ile Thr

660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln

675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn

690 695 700

Tyr Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val

705 710 715 720

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

725 730 735

<210> 53

<211> 2208

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 53

atggctgctg acggttatct tccagattgg ctcgaggaca acctttctga aggcattcgt 60

gagtgggtggg ctctgaaacc tggagtcctt caaccctaaag cgaaccaaca acaccaggac 120

aaccgtcggg gtcttgtgct tccgggttac aaatacctcg gaccgcgtaa cggactcgac 180

aaaggagagc cggtaacga ggcggacgcg gcagccctcg aacacgacaa agcttacgac 240

cagcagctca aggccgggtga caaccctgtac ctcaagtaca accacgccga cgccgagttt	300
caggagcgtc ttcaagaaga tacgtctttt gggggcaacc ttggcagagc agtcttccag	360
gccaaaaaga ggatccttga gcctcttgggt ctggttgagg aagcagctaa aacggctcct	420
ggaaagaaga ggctgtaga tcagtcctcct caggaaccgg actcatcatc tgggtgtggc	480
aaatcgggca aacagcctgc cagaaaaaga ctaaatttcg gtcagactgg cgactcagag	540
tcagtcccag accctcaacc tctcggagaa ccaccagcag cccccacaag ttgggatct	600
aatacaatgg cttcaggcgg tggcgcacca atggcagaca ataacgaggg tgccgatgga	660
gtgggtaatt cctcaggaat ttggcattgc gattcccaat ggctgggcga cagagtcac	720
accaccagca ccagaacctg ggccctgccc acttacaaca accatctcta caagcaaac	780
tccagccaat caggagcttc aaacgacaac cactactttg gctacagcac cccttggggg	840
tattttgact ttaacagatt ccaatgccac ttctcaccac gtgactggca gcgactcatt	900
aacaacaact ggggattccg gcccaagaaa ctcagcttca agctcttcaa catccaagtt	960
aaagaggta cgcagaacga tggcacgacg actattgcca ataaccttac cagcacggtt	1020
caagtgttta cggactcgga gtatcagctc ccgtacgtgc tcgggtcggc gcaccaaggc	1080
tgtctccgc cgtttccagc ggacgtcttc atggctccctc agtatggata cctcacctg	1140
aacaacggaa gtcaagcggg gggacgtcga tccttttact gcctggagta cttcccttcg	1200
cagatgctaa ggactggaaa taacttccaa ttcagctata ccttcagga tgtaccttt	1260
cacagcagct acgtcacag ccagagtttg gatcgcttga tgaatcctct tattgatcag	1320
tatctgtact acctgaacag aacgcaagga acaacctctg gaacaaccaa ccaatcacgg	1380
ctgcttttta gccaggctgg gcctcagtct atgtctttgc aggccagaaa ttggctacct	1440
gggccctgct accggcaaca gagactttca aagactgcta acgacaaca caacagtaac	1500
tttcccttga cagcggccag caaatatcat ctcaatggcc ggcactcgt ggtgaatcca	1560
ggaccagcta tggccagtca caaggacgat gaagaaaaat ttttccctat gcacggcaat	1620
ctaataattt gcaagaagg gacaacggca agtaacgcag aattagataa tgtaatgatt	1680
acggatgaag aagagattcg taccaccaat cctgtggcaa cagagcagta tggaaactgtg	1740
gcaaataact tgcagagctc aaatacagct cccacgacta gaactgtcaa tgatcagggg	1800
gccttacctg gcatggtgtg gcaagatcgt gacgtgtacc ttcaaggacc tatctgggca	1860
aagattcctc acacggatgg acattttcat ccttctcctc tgatgggagg ctttggactg	1920
aaacatccgc ctctcaaat catgatcaaa aatctccgg taccggcaaa tcctccgacg	1980
actttcagcc cgccaagtt tgcttcattt atcactcagt actccactgg acaggtcagc	2040
gtggaaattg agtgggagct acagaaagaa aacagcaaac gttggaatcc agagattcag	2100

tacacttcca actacaacaa gtctgttaat gtggacttta ctgtagacac taatggtgtt 2160

tatagtgaac ctgcacctat tggaacccgg tatctcacac gaaacttg 2208

<210> 54

<211> 736

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 54

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser

1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Val Pro Gln Pro

20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Arg Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly

100 105 110

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

115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140

Pro Val Asp Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Val Gly

145 150 155 160

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

165 170 175

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro

180 185 190
Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly

195 200 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser

210 215 220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile

225 230 235 240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu

245 250 255
Tyr Lys Gln Ile Ser Ser Tyr Ser Gly Ala Ser Asn Asp Asn His Tyr

260 265 270
Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His

275 280 285
Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp

290 295 300
Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln Val

305 310 315 320
Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu

325 330 335
Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr

340 345 350
Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp

355 360 365
Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser

370 375 380
Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser

385 390 395 400
Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe Glu

405 410 415
Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg

420 425 430
Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg Thr

435 440 445
Gln Gly Thr Thr Ser Gly Thr Thr Asn Gln Ser Arg Leu Leu Phe Ser

450 455 460
Gln Ala Gly Pro Gln Ser Met Ser Leu Gln Ala Arg Asn Trp Leu Pro
465 470 475 480

Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Thr Ala Asn Asp Asn
485 490 495
Asn Asn Ser Asn Phe Pro Trp Thr Ala Ala Ser Lys Tyr His Leu Asn

500 505 510
Gly Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys

515 520 525
Asp Asp Glu Glu Lys Phe Phe Pro Met His Gly Asn Leu Ile Phe Gly

530 535 540
Lys Glu Gly Thr Thr Ala Ser Asn Ala Glu Leu Asp Asn Val Met Ile
545 550 555 560
Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln

565 570 575
Tyr Gly Thr Val Ala Asn Asn Leu Gln Ser Ser Asn Thr Ala Pro Thr

580 585 590
Thr Arg Thr Val Asn Asp Gln Gly Ala Leu Pro Gly Met Val Trp Gln

595 600 605
Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His

610 615 620
Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu

625 630 635 640
Lys His Pro Pro Pro Gln Ile Met Ile Lys Asn Thr Pro Val Pro Ala

645 650 655
Asn Pro Pro Thr Thr Phe Ser Pro Ala Lys Phe Ala Ser Phe Ile Thr

660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
 675 680 685
 Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn

690 695 700
 Tyr Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val
 705 710 715 720
 Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
 725 730 735

<210> 55

<211> 2859

<212> DNA

<213> Homo sapiens

<400> 55

atgggagtga ggcacccgcc ctgctccac cgctcctgg ccgtctgcgc cctcgtgtcc 60
 ttggcaaccg ctgcactcct ggggcacatc ctactccatg atttctgct ggttccccga 120

gagctgagt gctcctcccc agtctcggag gagactcacc cagctcacca gcagggagcc 180
 agcagaccag ggccccggga tgcccaggca cccccggcc gtcccagagc agtgcaccaca 240
 cagtgcgacg tccccccaa cagccgcttc gattgcgccc ctgacaaggc catcaccag 300
 gaacagtgcg aggcccgcggt ctgctgttac atccctgcaa agcagggggt gcagggagcc 360
 cagatggggc agccctggtg cttcttccca cccagctacc ccagctacaa gctggagaac 420
 ctgagctcct ctgaaatggg ctacacggcc accctgacct gtaccacccc caccttcttc 480
 cccaaggaca tctgaccct gcggctggac gtgatgatgg agactgagaa ccgcctccac 540

ttcacgatca aagatccagc taacaggcgc tacgaggtgc ccttggagac cccgcgtgtc 600
 cacagccggg caccgtcccc actctacagc gtggagtctt cagaggagcc cttcgggggtg 660
 atcgtgcacc ggcagctgga cgcccgctg ctgctgaaca cgacgggtggc gccctgttc 720
 tttgcggacc agttccttca gctgtccacc tcgtgcct cgcagtatat cacaggcctc 780
 gccgagcacc tcagtcctt gatgtcagc accagctgga ccaggatcac cctgtggaac 840
 cgggaccttg cgccacgcc cgggtcgaac ctctacgggt ctacccttt ctacctggcg 900
 ctggaggacg gcgggtcggc acacgggggtg ttctgtctaa acagcaatgc catggatgtg 960

gtcctgcagc cgagccctgc ccttagctgg aggtcgacag gtgggacctt ggatgtctac 1020
 atcttctgg gccagagcc caagagcgtg gtgcagcagt acctggacgt tgtgggatac 1080

ccgttcatgc cgccatactg gggcctgggc ttccacctgt gccgctgggg ctactcctcc 1140
accgctatca cccgccaggt ggtggagaac atgaccaggg cccacttccc cctggacgtc 1200
caatggaacg acctggacta catggactcc cggagggact tcacgttcaa caaggatggc 1260
ttccgggact tcccggccat ggtgcaggag ctgcaccagg gcggccggcg ctacatgatg 1320
atcgtggatc ctgccatcag cagctcgggc cctgccggga gctacaggcc ctacgacgag 1380

ggtctgcgga ggggggtttt catcaccaac gagaccggcc agccgctgat tgggaaggtg 1440
tggcccggtt cactgcctt ccccgacttc accaacccea cagccctggc ctggtgggag 1500
gacatggtgg ctgagttcca tgaccagtg cccttcgacg gcatgtggat tgacatgaac 1560
gagccttcca acttcatcag aggctctgag gacggctgcc ccaacaatga gctggagaac 1620
ccaccctacg tgcttgggtt ggttgggggg accctccagg cgccaccat ctgtgcctcc 1680
agccaccagt ttctctccac acactacaac ctgcacaacc tctacggcct gaccgaagcc 1740
atcgctccc acagggcgct ggtgaaggct cgggggacac gccatttgt gatctccgc 1800

tcgacctttg ctggccacgg ccgatacgcc ggccactgga cgggggacgt gtggagctcc 1860
tgggagcagc tcgctcctc cgtgccagaa atcctgcagt ttaacctgct gggggtgcct 1920
ctggtcgggg ccgacgtctg cggcttcttg ggcaacacct cagaggagct gtgtgtgcgc 1980
tggaccacgc tgggggcctt ctacccttc atgcggaacc acaacagcct gctcagcttg 2040
ccccaggagc cgtacagctt cagcgagccg gccagcagg ccatgaggaa ggccctcacc 2100
ctgcgtacg cactcctccc ccacctctac aactgttcc accaggccca cgtcgcgggg 2160
gagaccgtgg cccggccctt ctctctggag ttccccaagg actctagcac ctggactgtg 2220

gaccaccagc tctgtgggg ggaggccctg ctcatcacc cagtgtcca ggccgggaag 2280
gccgaagtga ctggctactt ccccttgggc acatgttacg acctgcagac ggtgccaata 2340
gaggcccttg gcagcctccc acccccacct gcagctcccc gtgagccagc catccacagc 2400
gaggggcagt ggttgacgtt gccggcccc ctggacacca tcaacgtcca cctccgggt 2460
gggtacatca tccccctgca gggccctggc ctcaacaac cagagtccg ccagcagccc 2520
atggccctgg ctgtggccct gaccaagggt ggagaggccc gaggggagct gttctgggac 2580
gatggagaga gcctggaagt gctggagcga ggggcctaca cacaggtcat ctctctggcc 2640

aggaataaca cgatcgtgaa tgagctggta cgtgtgacca gtgagggagc tggcctgcag 2700
ctgcagaagg tgactgtcct gggcgtggcc acggcgcccc agcaggctct ctccaacggt 2760
gtccctgtct ccaacttcac ctacagcccc gacaccaagg tcctggacat ctgtgtctcg 2820
ctgttgatgg gagagcagtt tctcgtcagc tgggtttag 2859

<210> 56

<211> 2652

<212> DNA

<213> Homo sapiens

<400> 56

gcacaccccg gccgtcccag agcagtgccc acacagtgcg acgtccccc caacagccgc	60
ttcgattgcg cccctgacaa ggccatcacc caggaacagt gcgaggcccg cggctgctgc	120
tacatccctg caaagcaggg gctgcaggga gcccagatgg ggagccctg gtgcttcttc	180
ccaccagct accccagcta caagctggag aacctgagct cctctgaaat gggtacacg	240
gccaccctga cccgtaccac cccaccttc tccccaaagg acatcctgac cctgcggctg	300
gacgtgatga tggagactga gaaccgcctc cacttcacga tcaaagatcc agctaacagg	360
cgctacgagg tgccttggga gaccccgct gtccacagcc gggcaccgtc cccactctac	420
agcgtggagt tctccgagga gcccttcggg gtgatcgtgc accggcagct ggacggccgc	480
gtgctgctga acacgacggt ggcgccctg tctttgcgg accagtctct tcagctgtcc	540
acctcgctgc cctgcagta taccacagc ctgcggagc acctcagtc cctgatgctc	600
agcaccagct ggaccaggat caccctgtgg aaccgggacc ttgcgccac gcccggtgcg	660
aacctctacg ggtctcacc tttctacctg gcgctggagg acggcgggtc ggcacacggg	720
gtgttctgc taaacagcaa tgccatggat gtggtcctgc agccgagccc tgcccttagc	780
tggaggtcga caggtgggat cctggatgtc tacatcttc tgggcccaga gcccagagc	840
gtggtgcagc agtacctgga cgttgtggga taccgttca tgccgccata ctggggcctg	900
ggcttcacc tgtgccgtg gggctactcc tccaccgcta tcaccgcca ggtggtggag	960
aacatgacca ggcccactt cccctggac gtccaatgga acgacctgga ctacatggac	1020
tcccggaggg acttcacgtt caacaaggat ggttccggg acttcccggc catggtgcag	1080
gagctgcacc agggcggccg gcgctacatg atgatcgtgg atcctgccat cagcagctcg	1140
ggcctgccc ggagctacag gccctacgac gagggtctgc ggaggggggt tttcatcacc	1200
aacgagaccg gccagccgct gattgggaag gtatggccc ggtccactgc cttccccgac	1260
ttaccaacc ccacagccct ggcctgggtg gaggacatgg tggctgagtt ccatgaccag	1320
gtgcccttcg acggcatgtg gattgacatg aacgagcctt ccaacttcat cagaggctct	1380
gaggacggct gcccacaaca tgagctggag aaccacccct acgtgcctgg ggtggttggg	1440
gggaccctcc aggcggccac catctgtgcc tccagccacc agtttctctc cacacactac	1500
aacctgcaca acctctacgg cctgaccgaa gccatgcct cccacagggc gctggtgaag	1560
gctcggggga cagcccat tgtgatctcc cgctcgacct ttgctggcca cgcccatac	1620

gccggccact ggacggggga cgtgtggagc tcctgggagc agctcgccctc ctccgtgcca 1680
gaaatcctgc agtttaacct gctgggggtg cctctggtcg gggccgacgt ctgcggcttc 1740
ctgggcaaca cctcagagga gctgtgtgtg cgtgggaccc agctgggggc cttctacccc 1800

ttcattcgga accacaacag cctgtctcagt ctccccagg agccgtacag cttcagcgag 1860
cgggcccagc aggccatgag gaaggccctc accctgcgt acgcactcct ccccccacctc 1920
tacacactgt tcaccaggc ccacgtcgcg ggggagaccg tggcccgcc cctcttctg 1980
gagttcccca aggactctag cacctggact gtggaccacc agctcctgtg gggggaggcc 2040
ctgtcatca cccagtgtc ccaggccggg aagccgaag tgactggcta cttccccttg 2100
ggcacatggt acgacctgca gacggtgcca atagaggccc ttggcagcct cccaccccca 2160
cctgcagctc cccgtgagcc agccatccac agcgaggggc agtgggtgac gctgccggcc 2220

ccccggaca ccatcaacgt ccacctccgg gctgggtaca tcatccccct gcagggccct 2280
ggcctcacia ccacagagtc ccgccagcag cccatggccc tggctgtggc cctgaccaag 2340
ggtggagagg cccgagggga gctgttctgg gacgatggag agagcctgga agtgctggag 2400
cgaggggcct acacacaggt catcttctg gccaggaata acacgatcgt gaatgagctg 2460
gtactgtga ccagtgagg agctggcctg cagctgcaga agtgactgt cctgggcgtg 2520
gccacggcgc cccagcaggt cctctccaac ggtgtccctg tctccaactt cacctacagc 2580
cccgacacca aggtcctgga catctgtgtc tcgtgttga tgggagagca gtttctctc 2640

agctggtgtt ag 2652

<210> 57
<211> 4573
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 57

ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc 60
cgacgcccg gctttgccc ggcggcctca gtgagcagc gagcgcgag agaggagtg 120
gccaaactca tctagggg ttctgagtt taaacttcgt cgacattcg agcttgggt 180
gcaggtcgag ggcactggga ggatgttgag taagatgaa aactactgat gacccttgca 240

gagacagagt attaggacat gtttgaacag gggccgggag atcagcaggt agctctagag 300
gatccccgtc tgtctgcaca tttcgtagag cgagtgtcc gatactctaa tctccctagg 360

caaggttcat atttgtgtag gttacttatt ctccctttgt tgactaagtc aataatcaga	420
atcagcaggt ttggagtcag cttggcaggg atcagcagcc tgggttggaa ggagggggta	480
taaaagcccc ttcaccagga gaagccgtca cacagactag ccctaaggta agttggcgcc	540
gtttaaggga tggttgggtg gtgggggtatt aatgtttaat tacctttttt acaggcctga	600
actaggcgcg ccaccgccac catgccgtct tctgtctcgt ggggcacccct cctgctggca	660
ggcctgtgct gectgggtccc tgtctccctg gctgcttacc gccccagtga gaccctgtgc	720
ggcggggagc tggtaggacac cctccagttc gtctgtgggg accgcggctt ctacttcagc	780
aggcccgcaa gccgtgtgag ccgtcgcagc cgtggcatca tggaggagtg ctgtttccgc	840
agctgtgacc tggccctcct ggagacgtac tgtgtaccc ccgccaagtc cgagggcgcg	900
ccggcacacc ccggccgtcc cagagcagtg cccacacagt gcgacgtccc ccccaacagc	960
cgcttcgatt gcgccccga caagccatc acccaggaac agtgcgaggc ccgcggctgc	1020
tgctacatcc ctgcaaagca ggggctgcag ggagcccaga tggggcagcc ctggtgcttc	1080
ttcccacca gctacccag ctacaagctg gagaacctga gctcctctga aatgggctac	1140
acggccacc tgaccegtac cacccccacc ttcttccca aggacatcct gaccctgcgg	1200
ctggacgtga tgatggagac tgagaaccgc ctccactca cgatcaaaga tccagctaac	1260
aggcgtacg aggtgccctt ggagaccccg cgtgtccaca gccgggcacc gtccccactc	1320
tacagcgtgg agttctccga ggagcccttc ggggtgatcg tgcaccggca gctggacggc	1380
cgcgtgctgc tgaacacgac ggtggcgccc ctgttctttg cggaccagtt ccttcagctg	1440
tccacctcgc tgcctcgca gtatatcaca ggcctcgccg agcacctcag tccctgatg	1500
ctcagcacca gctggaccag gatcacctg tggaaccggg accttgcgcc cacgcccgtt	1560
gcgaacctct acgggtctca cctttctac ctggcgctgg aggacggcgg gtcggcacac	1620
ggggtgttcc tgctaaacag caatgccatg gatgtgttcc tgcagccgag cctgcccctt	1680
agctggaggt cgacaggtgg gatcctggat gtctacatct tcctgggccc agagcccaag	1740
agcgtggtgc agcagtacct ggacgttgtg ggatacccg tcatgccgc atactggggc	1800
ctgggcttcc acctgtgccg ctggggctac tcctccaccg ctatcaccg ccaggtggtg	1860
gagaacatga ccagggccca cttccccctg gacgtccaat ggaacgacct ggactacatg	1920
gactcccgga gggacttcac gttcaacaag gatggcttcc gggacttccc ggccatggtg	1980
caggagctgc accagggcgg ccggcgctac atgatgatcg tggatcctgc catcagcagc	2040
tcgggcccctg ccgggagcta caggccctac gacgagggtc tgcggagggg ggttttcac	2100
accaacgaga ccggccagcc gctgatitggg aaggtatggc ccgggtccac tgccttcccc	2160
gacttcacca accccacagc cctggcctgg tgggaggaca tggtaggtga gttccatgac	2220

caggtgccct tcgacggcat gtggattgac atgaacgagc cttccaactt catcagaggc	2280
tctgaggacg gctgccccaa caatgagctg gagaaccac cctacgtgcc tggggtggtt	2340
ggggggaccc tccaggcggc caccatctgt gcctccagcc accagtttct ctccacacac	2400
tacaacctgc acaacctcta cggcctgacc gaagccatcg cctccacag ggcgctggtg	2460
aaggctcggg ggacacgccc atttgtgatc tcccgctcga cctttgtctg ccacggccga	2520
tacgccggcc actggacggg ggacgtgtgg agctcctggg agcagctcgc ctctccgtg	2580
ccagaaatcc tgcagtttaa cctgctgggg gtgcctctgg tcggggccga cgtctgcggc	2640
ttctctggca acacctcaga ggagctgtgt gtgcgctgga ccagctggg ggccttctac	2700
cccttcatgc ggaaccacaa cagcctgctc agtctgcccc aggagccgta cagcttcagc	2760
gagccggccc agcaggccat gaggaaggcc ctacacctgc gctacgact cctccccac	2820
ctctacacac tgttccacca ggcccacgtc gcgggggaga ccgtggcccg gccctcttc	2880
ctggagtacc ccaaggactc tagcacctgg actgtggacc accagctcct gtggggggag	2940
gccctgctca tcacccagc gctccaggcc gggaaggccg aagtgactgg ctacttcccc	3000
ttgggcacat ggtacgacct gcagacggtg ccaatagagg cccttggcag cctccaccc	3060
ccacctgcag ctccccgtga gccagccatc cacagcgagg ggagctgggt gacgctgccg	3120
gcccccttg acaccatcaa cgtccacctc cgggctgggt acatcatccc cctgcagggc	3180
cctggcctca caaccacaga gtcccgcag cagcccatgg ccctggctgt ggccctgacc	3240
aagggtggag agggccgagg ggagctgttc tgggacgatg gagagagcct ggaagtgtg	3300
gagcgagggg cctacacaca ggtcatcttc ctggccagga ataacacgat cgtgaatgag	3360
ctggtacgtg tgaccagtga gggagctggc ctgcagctgc agaagggtgac tgtcctgggc	3420
gtggccacgg cgtccacaga ggtcctctcc aacggtgtcc ctgtctcaa cttcacctac	3480
agccccgaca ccaaggctct ggacatctgt gtctcgtgt tgatgggaga gcagtttctc	3540
gtcagctggt gttagcgagc ggccgctctt agtagcagta tcgatccag cccacttttc	3600
cccaatacga ctacgagatc tgttgcttct agctgcccgg gtggcatccc tgtgacctt	3660
ccccagtgcc tctcctggcc ctggaagtig ccactccagt gccaccagc cttgtcctaa	3720
taaaattaag ttgcatcatt ttgtctgact aggtgtcctt ctataatatt atggggtgga	3780
ggggggtggt atggagcaag gggcaagtig ggaaggccga cccgcgaata gtagatccc	3840
cgagggttg aatctatcac ctagagtaca ccctagagaa tagctagctc tcaatgacta	3900
aggactaaac ttggtatttc gactgaagcc tgtccctca ctgttggcgc taggaggaga	3960

gttcgtagaa aggatagtac gatttaagta tctctaagcc ttgtgaagca ctaaggttgc 4020

gtacagacgt gcttgaatta cggataattc gggaaccttg ggacacacaa aaaaccaaca 4080

cacagatcta atgaaaataa agatctttta tttaggcgcc tctgacttcc tggggattga 4140

cctgagttct accttagcgt ttgctgggtc ggtgaactaa tctgtgagat ccccaactct 4200

ccgtttggga tctccactct ctgggtgctc aaccttgggtg cccactgtc tactgctagt 4260

gagaccttac gcctgagaa acgtggcggtt actctaacta agcgacgcgc acttgactc 4320

tgaatacttc taccgtaact aaccccgac ctcagaactc agacggatct acgtgtcca 4380

tcaacaccag acttagatta cctctgttaa gttaattaa gctcggaag gaaccctag 4440

tgatggagtt ggccactccc tctctgcgcg ctgctcgt cactgaggcc gggcgaccaa 4500

aggtcgccc acgccccggc tttccccggg cggcctcagt gagcgagcga gcgcgcagag 4560

agggagtggc caa 4573

<210> 58

<211> 4597

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 58

ttggccaactc cctctctgcg cgtcgtcgc ctcactgagg ccggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggccgcctca gtgagcgagc gagcgcgag agagggagtg 120

gccaactcca tctactaggg ttctgagtt taaacttcgt cgacgattcg agcttgggt 180

gcaggtcgag ggcaactgga ggatgttag taagatggaa aactactgat gacccttgca 240

gagacagagt attaggacat gtttgaacag ggccggcg atcagcaggt agctctagag 300

gatccccgtc tgtctgcaca ttctgtagag cgagtgtcc gatactctaa tctccctagg 360

caaggttcat atttgttag gttacttatt ctcttttgt tgactaagtc aataatcaga 420

atcagcaggt ttggagtcag cttggcaggg atcagcagcc tgggttgaa ggaggggta 480

taaaagcccc ttcaccagga gaagccgtca cacagactag ccctaaggta agttggcgcc 540

gtttaaggga tggttggttg gtgggtatt aatgtttaat taccttttt acaggcctga 600

actaggcgcg ccaccgccac catgctcagg ggtccgggac ccggcggt gctgctgcta 660

gcagtctgt gcctggggac atcggtgcgc tgcaccgaaa ccgggaagag caagagggt 720

taccgcccc gtgagacct gtgcggcggg gagctggtgg acacctcca gttcgtctgt 780

ggggaccgcg gcttctactt cagcaggccc gcaagccgtg tgagccgtcg cagccgtggc 840
 atcatggagg agtgcgtgtt cgcagctgtg gacctggccc tcctggagac gtactgtgct 900
 acccccgcca agtccgaggg cgcgccggca cccccggcc gtcccagagc agtgcgccaca 960

 cagtgcgacg tccccccaa cagccgcttc gattgcgccc ctgacaaggc catcacccag 1020
 gaacagtgcg aggccccggg ctgctgtctac atccctgcaa agcaggggct gcagggagcc 1080
 cagatggggc agccctggtg cttcttccca cccagctacc ccagctacaa gctggagaac 1140
 ctgagctcct ctgaaatggg ctacacggcc accctgacct gtaccacccc caccttcttc 1200
 cccaaggaca tcctgacctt gcggctggac gtgatgatgg agactgagaa ccgcctccac 1260
 ttcacgatca aagatccagc taacaggcgc tacgaggtgc ccttggagac cccgctgttc 1320
 cacagccggg caccgtcccc actctacagc gtggagtctt ccgaggagcc cttcgggggtg 1380

 atcgtgcacc ggcagctgga cggcccgctg ctgctgaaca cgacggtggc gccctgttc 1440
 ttgctggacc agttccttca gctgtccacc tcgctgcctt cgcagtatat cacaggcctc 1500
 gccgagcacc tcagtcctt gatgtcagc accagctgga ccaggatcac cctgtggaac 1560
 cgggaccttg cggccacgcc cggctgcaac ctctacgggt ctacccttt ctacctggcg 1620
 ctggaggacg gggggtcgcc acacgggggtg ttctgtctaa acagcaatgc catggatgtg 1680
 gtcttcagc cgagccctgc ccttagctgg aggtcgacag gtgggacctt ggatgtctac 1740
 atcttcttgg gccagagcc caagagcgtg gtgcagcagt acctggacgt tgtgggatac 1800

 ccgttcacgc cgccatactg gggcctgggc ttccacctgt gccgtgggg ctactcctcc 1860
 accgtataca cccgccaggt ggtggagaac atgaccaggg cccacttccc cctggacgtc 1920
 caatggaacg acctggacta catggactcc cggagggact tcacgttcaa caaggatggc 1980
 ttccgggact tccggccat ggtgcaggag ctgcaccagg gcggccggcg ctacatgatg 2040
 atcgtggatc ctgcatcag cagctcgggc cctgccggga gctacaggcc ctacgacgag 2100
 ggtctgcgga ggggggtttt catcaccaac gagaccggcc agccgctgat tgggaaggta 2160
 tggcccggtt ccaactgcctt ccccgacttc accaacccca cagccctggc ctggtgggag 2220

 gacatggtgg ctgagttcca tgaccagggt cccttcgacg gcatgtggat tgacatgaac 2280
 gacgttcca acttcacag aggtctctag gacggctgcc ccaacaatga gctggagaac 2340
 ccacctacg tgcttgggtt ggttgggggg accctccagg cgccaccat ctgtgcctcc 2400
 agccaccagt ttctctccac acactacaac ctgcacaacc tctacggcct gaccgaagcc 2460
 atgcctccc acagggcgct ggtgaaggct cgggggacac gccatttgt gatctcccgc 2520
 tcgacctttg ctggccacgg ccgatacgcc ggccactgga cgggggacgt gtggagctcc 2580

tgggagcagc tcgcctcctc cgtgccagaa atcctgcagt ttaacctgct gggggtgcct	2640
ctggtcgggg ccgacgtctg cggcttctctg ggcaacacct cagaggagct gtgtgtgcgc	2700
tggaccacgc tgggggcctt ctaccccttc atgcggaacc acaacagcct gctcagtctg	2760
ccccaggagc cgtacagctt cagcgagccg gccagcagg ccatgaggaa ggccctcacc	2820
ctgcgtacg cactctctcc ccacctctac aactgttcc accaggccca cgtcgcgggg	2880
gagaccgtgg cccggcccct ctctctggag ttccccaagg actctagcac ctggactgtg	2940
gaccaccagc tctgtgggg ggaggccctg ctcatcacc cagtgtcca ggccgggaag	3000
gccgaagtga ctggctactt ccccttgggc acatggtacg acctgcagac ggtgccaata	3060
gaggcccttg gcagcctccc acccccacct gcagctcccc gtgagccagc catccacagc	3120
gaggggcagt gggtagcgt gccggcccc ctggacacca tcaacgtcca cctccgggt	3180
gggtacatca tccccctga gggccctggc ctcaaacca cagagtcccg ccagcagccc	3240
atggcccttg ctgtggccct gaccaagggt ggagaggccc gaggggagct gttctgggac	3300
gatggagaga gcctggaagt gctggagcga ggggcctaca cacaggatcat ctctctggcc	3360
aggaataaca cgatcgtgaa tgagctggta cgtgtgacca gtgaggagc tggcctgcag	3420
ctgcagaagg tgactgtctt gggcgtggcc acggcgcccc agcaggctct ctccaacggt	3480
gtccctgtct ccaacttcac ctacagcccc gacaccaagg tcttgacat ctgtgtctcg	3540
ctgttgatgg gagagcagtt tctctcagc tgggtttagc gagcgccgc tcttagtagc	3600
agtatcgatc ccagccact tttccccaat acgactacga gatctgtggc ttctagctgc	3660
ccgggtggca tccctgtgac ccctccccag tgcctctctt ggccctggaa gttgccactc	3720
cagtgtccac cagccttctc ctaataaaat taagtgtcat catctgtct gactaggtgt	3780
ccttctataa tattatgggg tggagggggg tggatatggag caaggggcaa gttgggaagg	3840
ccgaccgcg aatagtagat cccgcgaggg ctgaatcta tcacctagag tacaccctag	3900
agaatagcta gctctcaatg actaaggact aaacttggt tttcgactga agcctgtccc	3960
ctcactgttg gcgctaggag gagagttcgt agaaaggata gtacattta agtatctcta	4020
agccttgtga agcactaagg ttgcgtacag acgtgcttga attacggata attcggaac	4080
cttgggacac acaaaaaacc aacacacaga tctaataaaa ataaagatct tttatttagg	4140
gcctctgac ttcttggga ttgacctgag ttctactcta gcgtttgctg gttcggtgaa	4200
ctaactctgt agatcccaa ctctccgttt gggatctcca ctctctggtg tcctaacctt	4260
ggtgccccac tgtctactgc tagtgagacc ttacgcgtg agaaacgtgg cgttactcta	4320

actaagcgac gcgcacttgc actctgaata cttctaccgt aactaaccce ggacctcaga 4380
 actcagacgg atctacgctg tccatcaaca ccagacttag attacctctg ttaagtttaa 4440
 ttaagctcgc gaaggaaccc ctagtgatgg agttggccac tccctctctg cgcgctcgt 4500
 cgctcactga ggccgggcca ccaaaggtcg cccgacgccc gggctttgcc cggcgggcct 4560
 cagtgcgca gcgagcgcgc agagaggagagg tggccaa 4597

<210> 59

<211> 4594

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<400> 59

ttggccactc cctctctcgc cgctcgtcgc ctactgagg ccgggcgacc aaaggtcgcc 60
 cgacgcccgg gctttgcccg ggccggcctca gtgagcagc gagcgcgcag agaggagtg 120
 gccaaactcca tctactagggg ttccctgagtt taaacttcgt cgacgattcg agcttgggct 180
 gcaggtcgag ggactggga ggatgttag taagatggaa aactactgat gacccttgca 240
 gagacagagt attaggacat gtttgaacag ggccggggcg atcagcaggt agctctagag 300
 gatccccgtc tgctgcaca tttcctagag cgagtgttcc gatactctaa tctccctagg 360
 caaggttcat atttgttag gttacttatt ctcttttgt tgactaagtc aataatcaga 420

atcagcaggt ttggagtcag cttggcaggg atcagcagcc tgggttggaa ggagggggta 480
 taaaagcccc ttaccagga gaagccgtca cacagactag ccctaaggta agttggcgcc 540
 gtttaaggga tgggtggtg gtgggtatt aatgtttaat tacctttttt acaggcctga 600
 actaggcgcg ccaccgccac catgcttagg ggtccggggc ccgggctgct gctgctggcc 660
 gtccagtgcc tggggacagc ggtgcctcc acgggagcct cgaagagcaa gagggcttac 720
 cgccccagt agacctgtg cggcggggag ctggtggaca cctccagtt cgtctgtggg 780
 gaccgcggct tctacttcag caggcccga agccgtgtga gccgtcgag ccgtggcatc 840

atggaggagt gctgtttccg cagctgtgac ctggccctcc tggagacgta ctgtgctacc 900
 cccgccaagt ccgagggcgc gccggcacac cccggccgtc ccagagcagt gcccacacag 960
 tgcgacgtcc ccccaacag ccgcttcgat tgcgcccctg acaaggccat caccaggaa 1020
 cagtgcgagg ccccggtgtg ctgctacatc cctgcaaagc aggggctgca gggagcccag 1080
 atggggcagc cctggtgctt cttccaccc agctaccca gctacaagct ggagaacctg 1140
 agctcctctg aaatgggcta cagggccacc ctgaccgta ccacccccc cttcttcccc 1200

aaggacatcc tgaccctgcg gctggacgtg atgatggaga ctgagaaccg cctccacttc 1260

acgatcaaag atccagctaa caggcgctac gaggtgccct tggagacccc gcgtgtccac 1320

agccggggcac cgtcccccact ctacagcgtg gagttctccg aggagccctt cggggtgatc 1380

gtgcaccggc agctggacgg ccgctgtctg ctgaacacga cggtaggcgc cctgttcttt 1440

gcggaccagt tcttcagct gtccacctcg ctgccctcgc agtatatcac aggcctcgcc 1500

gagcacctca gtcccctgat gctcagcacc agctggacca ggatcaccct gtggaaccgg 1560

gaccttgccg ccacgcccg tgcgaacctc tacgggtctc accctttcta cctggcgctg 1620

gaggacggcg ggtcggcaca cggggtgttc ctgctaaaca gcaatgccat ggatgtggtc 1680

ctgcagccga gccctgccct tagctggagg tcgacagggt ggatcctgga tgtctacatc 1740

tctctgggcc cagagcccaa gagcgtgggt cagcagtacc tggacgttgt gggatacccg 1800

tctatgccgc catactgggg cctgggcttc cacctgtgcc gctggggcta ctctccacc 1860

gctatcacc gccaggtggt ggagaacatg accagggccc acttccccct ggacgtccaa 1920

tggaacgacc tggaactacat ggactcccgg agggacttca cgttcaaca ggatggcttc 1980

cgggacttcc cggccatggt gcagagcgtg caccagggcg gccggcgcta catgatgatc 2040

gtggatcctg ccatcagcag ctccggccct gccgggagct acaggcccta cgacgagggt 2100

ctgcggaggg gggttttcat caccaacgag accggccagc cgctgattgg gaaggtatgg 2160

cccgggtcca ctgccttccc cgacttcacc aacccacag ccctggcctg gtgggaggac 2220

atggtggctg agttccatga ccaggtgccc ttgcacggca tgtggattga catgaacgag 2280

ccttccaact tcatcagagg ctctgaggac ggtgccccca acaatgagct ggagaacce 2340

ccctacgtgc ctggggtggt tggggggacc ctccaggcgg ccaccatctg tgcctccagc 2400

caccagtctc tctccacaca ctacaacctg cacaacctct acggcctgac cgaagccatc 2460

gcctcccaca gggcgctggt gaaggctcgg gggacacgcc catttgtgat ctcccgtcgc 2520

acctttgctg gccacggccg atacgccggc cactggacgg gggacgtgtg gagtcctgg 2580

gagcagctcg cctcctccgt gccagaaatc ctgcagtta acctgctggg ggtgcctctg 2640

gtcggggccg acgtctgcgg ctctctgggc aacacctcag aggagctgtg tgtgcgtgg 2700

accagctgg gggccttcta ccccttcatg cggaaccaca acagcctgct cagtctgccc 2760

caggagccgt acagcttcag cgagccggcc cagcaggcca tgaggaaggc cctcaccctg 2820

cgctacgcac tctccccca cctctacaca ctgttccacc agggccacgt cgcgggggag 2880

accgtggccc gggccctctt cctggagttc cccaaggact ctagcacctg gactgtggac 2940

caccagctcc tgtgggggga ggccctgctc atcaccacag tgctccaggc cggaaggcc 3000
gaagtgactg gctacttccc cttgggcaca tggtagcacc tgcagacggt gccaatagag 3060
gcccttggca gcctcccacc cccacctgca gctccccgtg agccagccat ccacagcgag 3120
gggcagtggg tgacgctgcc ggccccctg gacacatca acgtccacct cggggtggg 3180
tacatcatcc cctgcaggg ccctggcctc acaaccacag agtcccgcca gcagcccatg 3240
gccctggctg tggccctgac caagggtgga gagggccgag gggagctgtt ctgggacgat 3300
ggagagagcc tggaagtgtt ggagcgaggg gcctacacac aggtcatctt cctggccagg 3360

aataacacga tcgtgaatga gctggtacgt gtgaccagtg agggagctgg cctgcagctg 3420
cagaaggtag ctgtcctggg cgtggccacg gcgcccagc aggtcctctc caacggtgtc 3480
cctgtctcca acttcaccta cagccccgac accaaggctc tggacatctg tgtctcgctg 3540
ttgatgggag agcagtttct cgtcagctgg tgttagcgag cggccgctct tagtagcagt 3600
atcgatccca gccactttt cccaatacag actacgagat ctgtggcttc tagctgcccg 3660
ggtggcatcc ctgtgacccc tccccagtgc ctctcctggc cctggaagtt gccactccag 3720
tgcccaccag cttgtccta ataaaattaa gttgcatcat tttgtctgac taggtgtcct 3780

tctataatat tatgggggtg aggggggtg tatggagcaa ggggcaagtt gggaaggccg 3840
accgcgaat agtagatccc gcgagggtt gaattatca cctagagtac accctagaga 3900
atagctagct ctcaatgact aaggactaaa cttggtatit cgactgaagc ctgtccctc 3960
actgttggcg ctaggaggag agttcgtaga aaggatagta cgatttaagt atctctaagc 4020
cttgtgaagc actaaggttg cgtacagacg tgcttgaatt acggataatt cgggaacctt 4080
gggacacaca aaaaaccaac acacagatct aatgaaaata aagatctttt atttaggcgc 4140
ctctgacttc ctggggattg acctgagttc tactctagcg tttgctggtt cggatgaacta 4200

atctgtgaga tccccaactc tccgtttggg atctccactc tctggtgtcc taaccttgg 4260
gccccactgt ctactgctag tgagacctta cgcgctgaga aacgtggcgt tactctaact 4320
aagcgacgcg cacttgact ctgaatactt ctaccgtaac taaccccgga cctcagaact 4380
cagacggatc tacgctgtcc atcaacacca gacttagatt acctctgtta agtttaatta 4440
agctcgcgaa ggaaccctta gtgatggag tggccactcc ctctctgcgc gctcgctcgc 4500
tactgagcg cgggcgacca aaggtcgccc gacgcccggg ctttggccgg gcggcctcag 4560
tgagcgagcg agcgcgaga gagggagtgg ccaa 4594

<210> 60

<211> 4555

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 60

ttggccactc cctctctgcg cgctcgctcg ctactgagg ccggcgacc aaaggtcgcc	60
cgacgcccg gctttgccc ggccgcctca gtgagcgagc gagcgcgag agagggagt	120
gccaaactca tctactagggg ttcttgagtt taaacttcgt cgacgattcg agcttgggct	180
gcaggtcgag ggcactggga ggatgttag taagatggaa aactactgat gacccttgca	240
gagacagagt attaggacat gtttgaacag gggccgggcg atcagcaggt agctctagag	300
gatccccgtc tgctgcaca ttctgtagag cgagtgttcc gatactctaa tctccctagg	360
caaggttcat atttgttag gttacttatt ctcttttgt tgactaagtc aataatcaga	420
atcagcaggt ttggagtcag cttggcaggg atcagcagcc tgggttggaa ggagggggt	480
taaaagcccc ttaccagga gaagccgtca cacagactag ccctaaggta agttggcgcc	540
gtttaaggga tgggtgttg gtgggtatt aatgtttaat tacctttttt acaggcctga	600
actaggcgcg ccaccgccac catgccgtct tctgtctcgt ggggcatcct cctgctggca	660
ggcctgtgct gcctggtccc tgtctccctg gctgctctgt gcggcgggga gctggtggac	720
accctccagt tegtctgtgg ggaccgggc ttctacttca gcaggccgc aagccgtgtg	780
agccgtcgca gccgtggcat cgttgaggag tgctgtttcc gcagctgtga cctggccctc	840
ctggagacgt actgtgttac ccccgccaag tccgagggcg cgccggcaca ccccgccgt	900
cccagagcag tgccacaca gtgcgacgtc cccccaaca gccgtttcga ttgcgccct	960
gacaaggcca taccacagga acagtgcgag gcccgggct gctgtcatat cctgcaaag	1020
caggggctgc agggagccca gatggggcag cctggtgct tcttcccacc cagctacccc	1080
agctacaagc tggagaacct gagtcctct gaaatgggct acacggccac cctgaccct	1140
accaccccca ctttcttccc caaggacatc ctgacctgc ggctggacgt gatgatggag	1200
actgagaacc gctccactt cacgatcaaa gatccagcta acaggcgcta cgaggtgccc	1260
ttggagacc cgctgtcca cagccgggca ccgtccccc tctacagcgt ggagtcttcc	1320
gaggagccct tgggggtgat cgtgcaccgg cagctggacg gccgctgct gctgaacacg	1380
acggtggcgc cctgtttctt tgcggaccag ttcttcagc tgtccacctc gctgccctcg	1440
cagtatatca caggcctcgc cgagcacctc agtcccctga tgctcagcac cagctggacc	1500
aggatcacc tgtggaaccg ggaccttgcg cccacgccc gtgcgaacct ctacgggtct	1560
caccctttct acctggcgt ggaggacggc gggtcggcac acggggtgtt cctgctaaac	1620

agcaatgcca tggatgtggt cctgcagccg agccctgccc ttagctggag gtcgacaggt	1680
gggatcctgg atgtctacat cttcctgggc ccagagccca agagcgtggt gcagcagtac	1740
ctggacgttg tgggataccc gtatcatgcc ccatactggg gcctgggctt ccacctgtgc	1800
cgctggggct actcctccac cgctatcacc cgccaggtgg tggagaacat gaccagggcc	1860
catttcccc tggacgtcca atggaacgac ctggactaca tggactcccg gagggacttc	1920
acgttcaaca aggatggctt ccgggacttc ccggccatgg tgcaggagct gcaccagggc	1980
ggccggcgct acatgatgat cgtggatcct gccatcagca gctcgggccc tgccgggagc	2040
tacagccct acgacgagg tctgcggagg ggggttttca tcaccaacga gaccggccag	2100
ccgtgattg ggaaggtatg gcccgggtcc actgccttcc ccgacttcac caaccccaca	2160
gccctggcct ggtgggagga catggtggct gagttccatg accaggtgcc cttcgacggc	2220
atgtggattg acatgaacga gccttccaac ttcatcagag gctctgagga cggtgcccc	2280
aacaatgagc tggagaaccc accctacgtg cctggggtgg ttggggggac cctccaggcg	2340
gccaccatct gtgcctccag ccaccagttt ctctccacac actacaacct gcacaacctc	2400
tacggcctga ccgaagccat cgcctccac agggcgctgg tgaaggctcg ggggacacgc	2460
ccatttgtga tctcccgtc gacctttgct ggccacggcc gatacgccgg cactggacg	2520
ggggacgtgt ggagctcctg ggagcagctc gcctcctccg tgccagaaat cctgcagttt	2580
aacctgctgg ggtgacctt ggtcggggcc gacgtctgcg gcttctggg caacacctca	2640
gaggagctgt gtgtgcgtg gaccagctg ggggccttct accccttcat gcggaaccac	2700
aacagcctgc tcagtctgcc ccaggagccg tacagcttca gcgagccggc ccagcaggcc	2760
atgaggaagg cctcacctt gcgtacgca ctctccccc acctctacac actgttccac	2820
caggccccag tcgcggggga gaccgtggcc cgccccctct tcctggagtt cccaaggac	2880
tctagacct ggactgtgga ccaccagctc ctgtgggggg aggccttct catcaccca	2940
gtgtccagg ccgggaaggc cgaagtgact ggctacttcc ccttgggcac atggtacgac	3000
ctgcagacgg tgccaataga ggcccttggc agcctccac cccacctgc agtccccgt	3060
gagccagcca tcacagcga ggggcagtgg gtgacgtgc cgccccctt ggacaccatc	3120
aacgtccacc tccgggctgg gtacatcat cccctgcagg gccctggcct cacaaccaca	3180
gagtcgcc agcagcccat ggccctggct gtggccctga ccaagggtgg agaggccga	3240
ggggagctgt tctgggacga tggagagagc ctggaagtgc tggagcgagg ggcctacaca	3300
caggtcatct tcctggccag gaataacacg atcgtgaatg agctggtacg tgtgaccagt	3360
gaggagctg gcctgcagct gcagaagggt actgtcttgg gcgtggccac ggcgccccag	3420
caggtcctct ccaacggtgt ccctgtctcc aacttcacct acagccccga caccaaggtc	3480

ctggacatct gtgtctcgct gttgatggga gagcagtttc tcgtcagctg gtgttagcga 3540

gcggccgctc ttagtagcag tatcgatccc agcccacttt tcccaatac gactacgaga 3600

tctgtggctt ctagctgccc gggtagcatc cctgtgaccc ctccccagtg cctctcctgg 3660

ccctggaagt tgccactcca gtgcccacca gccttgctct aataaaatta agttgcatca 3720

ttttgtctga ctaggtgtcc ttctataata ttatgggggtg gaggggggtg gtatggagca 3780

aggggcaagt tgggaaggcc gaccgcgaa tagtagatcc cgcgagggt tgaatctatc 3840

acctagagta caccctagag aatagctagc tctcaatgac taaggactaa acttggtatt 3900

tcgactgaag cctgtccct cactgttggc gctaggagga gatttcgtag aaaggatagt 3960

acgatttaag tatctctaag ccttgtgaag cactaagggt gcgtacagac gtgcttgaat 4020

tacggataat tcgggaacct tgggacacac aaaaaaccaa cacacagatc taatgaaat 4080

aaagatcttt tatttaggcg cctctgactt cctggggatt gacctgagtt ctactctagc 4140

gtttgtctgt tcggtgaact aatctgtgag atccccaact ctccgtttgg gatctccact 4200

ctctggtgtc ctaaccttgg tgccccactg tctactgcta gtgagacctt acgcgctgag 4260

aaacgtggcg ttactctaac taagcgacgc gcacttgac tctgaatact tctaccgtaa 4320

ctaaccgccg acctcagaac tcagacggat ctacgtgtc catcaacacc agacttagat 4380

tacctctgtt aagtttaatt aagctcgca aggaaccct agtgatggag ttggccactc 4440

cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc cgacgcccg 4500

gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agaggagtg gccaa 4555

<210> 61

<211> 4579

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 61

ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agaggagtg 120

gccaaactcca tctactagggg ttctgagtt taaacttcgt cgacgattcg agcttgggt 180

gcaggtcgag ggcactggga ggatgttgag taagatggaa aactactgat gacccttga 240

gagacagagt attaggacat gtttgaacag gggccgggag atcagcaggt agctctagag 300

gatccccgtc tgctgcaca ttctgtagag cgagtgtcc gatactctaa tctccctagg 360

caaggttcat atttgtgtag gttacttatt ctccctttgt tgactaagtc aataatcaga	420
atcagcaggt ttggagtcag cttggcaggg atcagcagcc tgggttggaa ggagggggta	480
taaaagcccc ttcaccagga gaagccgtca cacagactag ccctaaggta agttggcgcc	540
gtttaaggga tggttggttg gtgggggtatt aatgtttaat tacctttttt acaggcctga	600
actaggcgcg ccaccgccac catgctcagg ggtccgggac ccgggcggct gctgctgcta	660
gcagtcctgt gcctggggac atcgggtgcg tgcaccgaaa ccgggaagag caagagggct	720
ctgtgcggcg gggagctggt ggacacccctc cagtctgtct gtggggaccg cggcttctac	780
ttcagcagcg ccgaagccg tgtgagccgt cgcagccgtg gcatcgttga ggagtgtgt	840
ttccgcagct gtgacctggc cctcctggag acgtactgtg ctacccccgc caagtccgag	900
ggcgcgccgg cacaccccg cgtcccaga gcagtgccca cacagtgcga cgtcccccc	960
aacagccgct tcgattgcgc cctgacaag gccatcacc aggaacagtg cgaggcccg	1020
ggctgctgct acatccctgc aaagcagggg ctgcaggag cccagatggg gcagccctgg	1080
tgcttcttcc caccagcta cccagctac aagctggaga acctgagctc ctctgaaatg	1140
ggctacacgg ccacctgac ccgtaccacc cccaccttct tcccgaagga catcctgacc	1200
ctgcggctgg acgtgatgat ggagactgag aaccgcctcc acttcacgat caaagatcca	1260
gctaacaggc gctacaggt gcccttggag acccccgctg tccacagccg ggcaccgtcc	1320
ccactctaca gcgtggagtt ctccgaggag ccttcgggg tgatcgtgca ccggcagctg	1380
gacggccgcg tgctgtgaa cagcaggtg gcgccctgt tctttgcgga ccagttcctt	1440
cagctgtcca cctcgtgcc ctgcagtat atcacaggcc tcgccgagca cctcagtc	1500
ctgatgtca gcaccagctg gaccaggatc accctgtgga accgggacct tgcgccacg	1560
cccgtgcca acctctacgg gtctcacct ttctacctgg cgtggagga cggcgggtcg	1620
gcacacgggg tgttctgct aaacagcaat gccatggatg tggctctgca gccgagccct	1680
gcccttagct ggaggtcgac aggtgggatc ctggatgtct acatcttctt gggcccagag	1740
ccaagagcg tggtcagca gtacctggac gttgtgggat accgttcat gccgccatac	1800
tggggcctgg gttccacct gtgccgctgg ggtactcct ccaccgtat caccgccag	1860
gtggtggaga acatgaccag ggcccacttc ccctggacg tccaatggaa cgacctggac	1920
tacatggact cccggaggga cttcacgttc aacaaggatg gcttccggga ctteccggcc	1980
atggtgcagg agctgcacca gggcgcccg cgctacatga tgatcgtgga tcctgccatc	2040
agcagctcgg gccctgccg gagctacagg ccctacgac agggctctgc gaggggggtt	2100
ttcatcacca acgagaccgg ccagccgctg attgggaagg tatggcccg gtccactgcc	2160
ttccccgact tcaccaaccc cacagccctg gcctggtggg aggacatggt ggctgagttc	2220

catgaccagg tgccttcga cggcatgtgg attgacatga acgagccttc caacttcac	2280
agaggctctg aggacggctg cccaacaat gagctggaga acccacccta cgtgcctggg	2340
gtggttgggg ggacctcca ggccgccacc atctgtgcct ccagccacca gtttctctec	2400
acacactaca acctgcacaa cctctacggc ctgaccgaag ccatcgctc ccacagggcg	2460
ctggtgaagg ctggggggac acgcccattt gtgatctccc gctcgacctt tcttgccac	2520
ggccgatacg cggccactg gacgggggac gtgtggagct cctgggagca gctcgctec	2580
tccgtgccag aaatcctgca gtttaacctg ctgggggtgc ctctggctcg gcccgacgtc	2640
tgcggcttc tgggcaaac ctcagaggag ctgtgtgtgc gctggacca gctgggggcc	2700
ttctaccct tcatcgga ccacaacagc ctgctcagtc tgccccagga gccgtacagc	2760
ttcagcgagc cggcccagca ggccatgagg aaggccctca ccctgcgcta cgcactctc	2820
ccccactct acacactgtt ccaccaggcc cagctcgcg ggagaccgt gggccggccc	2880
ctcttcttg agttcccaa ggactctagc acctggactg tggaccacca gctcctgtg	2940
ggggaggccc tgctcatcac ccagtgctc caggccggga aggccgaagt gactggctac	3000
ttcccttgg gcacatggt cgcactgcag acggtgcaa tagaggcct tggcagctc	3060
ccacccccac ctgcagctcc ccgtgagcca gccatccaca gcgaggggca gtgggtgacg	3120
ctgccggccc cctggacac catcaacgtc cacctccggg ctgggtacat catccccctg	3180
caggccctg gcctcacaac cacagagtcc cgcagcagc ccatggcct ggctgtggcc	3240
ctgaccaagg gtggagagg ccgaggggag ctgttctggg acgatggaga gacctggaa	3300
gtgtggagc gaggggccta cacacaggtc atcttcttg ccaggaataa cacgatcgtg	3360
aatgagctgg tacgtgtgac cagttaggga gctggcctgc agctgcagaa ggtgactgtc	3420
ctggcgctgg ccacggccc ccagcaggtc ctctcaacg gtgtccctgt ctccaactc	3480
acctacagcc ccgacaccaa ggtcttgac atctgtgtct cgctgttgat gggagagcag	3540
tttctctca gctggtgta gcgagcggcc gctcttagta gcaglatcga tcccagcca	3600
cttttccca atacgactac gagatctgtg gcttctagct gcccgggtg catccctgtg	3660
accttcccc agtgcctctc ctggccctgg aagttgccac tccagtgcc accagccttg	3720
tcctaataaa attaatgtc atcatttgt ctgactaggt gtccttctat aatattatgg	3780
ggtggagggg ggtggtatgg agcaaggggc aagttgggaa ggccgaccg cgaatagtag	3840
atcccgcgag ggcttgaatc tatcacctag agtacacct agagaatagc tagctctcaa	3900
tgactaagga ctaaacttgg tatttcgact gaagcctgtc ccctcactgt tggcgctagg	3960

aggagagttc gtagaaagga tagtacgatt taagtatctc taagccttgt gaagcactaa 4020

ggttgcgtac agacgtgctt gaattacgga taattcggga accttgggac acacaaaaaa 4080

ccaacacaca gatctaataa aaataaagat cttttattta ggcgccctctg acttctctggg 4140

gattgacctg agttctactc tagcgtttgc tggttcgggtg aactaatctg tgagatcccc 4200

aactctccgt ttgggatctc cactctctgg tgcctaacc ttggtgcccc actgtctact 4260

gctagtgaga cttacgcgc tgagaaacgt ggcgttactc taactaagcg acgcgcactt 4320

gcactctgaa tacttctacc gtaactaacc ccggacctca gaactcagac ggatctacgc 4380

tgtccatcaa caccagactt agattacctc tgttaagttt aattaagctc gcgaaggaac 4440

ccctagtgat ggagttaggc actccctctc tgcgcgctcg ctgctcact gaggccgggc 4500

gaccaaaggt cgcccgacgc ccgggctttg ccggggcggc ctgagtgcgc gagcgagcgc 4560

gcagagaggg agtggccaa 4579

<210> 62

<211> 4576

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 62

ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggccgcctca gtgagcgagc gagcgcgag agagggagtg 120

gccaactcca tctactaggg ttctgagtt taaacttcgt cgacgattcg agcttgggct 180

gcaggtcgag ggcaactgga ggatgttag taagatggaa aactactgat gacccttgca 240

gagacagagt attaggacat gtttgaacag gggccgggcg atcagcaggt agctctagag 300

gatccccgtc tgtctgcaca ttctgtagag cgagtgttcc gatactctaa tctccctagg 360

caaggttcat atttgttag gttacttatt ctccctttgt tgactaagtc aataatcaga 420

atcagcaggt ttggagtcag cttggcaggg atcagcagcc tgggttggaa ggagggggta 480

taaaagcccc ttcaccagga gaagccgtca cacagactag ccctaaggta agttggcgcc 540

gtttaaggga tggttgggtg gtgggttatt aatgtttaat tacctttttt acaggcctga 600

actaggcgcg ccaccgccac catgcttagg ggtccggggc ccgggctgct gctgctggcc 660

gtccagtgcc tggggacagc ggtgccctcc acgggagcct cgaagagcaa gagggctctg 720

tgcggcgggg agctgggtga caccctccag ttctgtctgtg gggaccgcgg cttctacttc 780

agcaggcccg caagccgtgt gagccgtcgc agccgtggca tcgttgagga gtgctgtttc	840
cgcagctgtg acctggccct cctggagacg tactgtgcta cccccgcaa gtccgagggc	900
gcgccggcac accccggccg tcccagagca gtgcccacac agtgcgacgt ccccccaac	960
agccgcttcg attgcgcccc tgacaaggcc atcaccagg aacagtgcga gggccgcggc	1020
tgctgtata tccctgcaaa gcaggggctg caggagagccc agatggggca gccctgggtc	1080
ttcttccac ccagctacce cagctacaag ctggagaacc tgagctctc tgaaatgggc	1140
tacacggcca cctgacccg taccacccc accttcttc ccaaggacat cctgaccctg	1200
cggctggacg tgatgatgga gactgagaac cgcctccact tcacgatcaa agatccagct	1260
aacaggcgct acgaggtgcc ctggagacc ccgcgtgtcc acagccgggc accgtcccca	1320
ctctacagcg tggagtctc cgaggagccc ttcgggggtga tcgtgcaccg gcagctggac	1380
ggccgcgtgc tgctgaacac gacgtggcg ccctgttct ttgcggacca gttccttcag	1440
ctgtccacct cgctgccctc gcagtatac acaggcctcg ccgagcacct cagtccctg	1500
atgctcagca ccagctggac caggatcacc ctgtggaacc gggaccttgc gcccacgcc	1560
ggtgcgaacc tctacgggtc tcaacctttc tacctggcgc tggaggacgg cgggtcgga	1620
cacggggtgt tcttgctaaa cagcaatgcc atggatgtgg tcttcagcc gagccctgcc	1680
cttagctgga ggtcgacagg tgggacctg gatgtctaca tcttctggg cccagagccc	1740
aagagcgtgg tgcagcagta cctggacgtt gtgggatacc cgttcagcc gccatactgg	1800
ggcctgggct tccacctgtg ccgctggggc tactctcca ccgtatcac ccgccaggtg	1860
gtggagaaca tgaccagggc ccacttcccc ctggacgtcc aatggaacga cctggactac	1920
atggactccc ggagggactt cacgttcaac aaggatggct tccgggactt cccggccatg	1980
gtgcaggagc tgcaccaggc cggccggcgc tacatgatga tcgtggatcc tgccatcagc	2040
agctcgggcc ctgccgggag ctacaggccc tacgacgagg gtctcggag gggggttttc	2100
atcaccaacg agaccggcca gccgtgatt gggaaggat ggcccgggtc cactgccttc	2160
cccgacttca ccaacccac agccctggcc tgggtggagg acatggtggtc tgagttccat	2220
gaccaggtgc ccttcgacgg catgttgatt gacatgaacg agccttcaa cttcatcaga	2280
ggctctgagg acggctgccc caacaatgag ctggagaacc caccctacgt gcctgggggtg	2340
gttgggggga cctccaggc ggccaccatc tgtgctcca gccaccagtt tctctccaca	2400
cactacaacc tgcacaacct ctacggcctg accgaagcca tcgcctcca cagggcgctg	2460
gtgaaggctc gggggacacg cccatttgtg atctcccgt cgacctttgc tggccacggc	2520
cgatacgccg gccactggac gggggacgtg tggagctcct gggagcagct cgcctctcc	2580

gtgccagaaa tcttgcagtt taacctgctg ggggtgcctc tggtcggggc cgacgtctgc 2640

ggcttccttg gcaacacctc agaggagctg tgtgtgcgct ggaccagct gggggccttc 2700

tacccttca tgcggaacca caacagcctg ctctagctgc cccaggagcc gtacagcttc 2760

agcgagccgg cccagcagcc catgaggaag gcctcacc tgcgtacgc actcctcccc 2820

cacctctaca cactgttcca ccaggccccc gtcgcggggg agaccgtggc ccggcccttc 2880

ttcttgaggt tccccaagga ctctagcacc tggactgtgg accaccagct cctgtggggg 2940

gaggccctgc tcatcacc agtgctccag gccgggaagg ccgaagtac tggctacttc 3000

cccttgggca catggtacga cctgcagacg gtgccaatag aggcccttgg cagcctccca 3060

ccccacctg cagctccccg tgagccagcc atccacagcg aggggcagtg ggtgacgctg 3120

ccggccccc tggacacat caacgtccac ctccgggctg ggtacatcat cccctgcag 3180

ggccctggcc tcacaaccag agagtcgcc cagcagccca tggccctggc tgtggccctg 3240

accaagggtg gagaggcccg aggggagctg ttctgggacg atggagagag cctggaagtg 3300

ctggagcgag gggcctacac acaggtcatc ttcttgcca ggaataacac gatcgtgaat 3360

gagctggtac gtgtgaccag tgaggagct ggcctgcagc tgcagaaggt gactgtcctg 3420

ggcgtggcca cggcgcccca gcaggtcctc tccaacggtg tccctgtctc caacttcacc 3480

tacagccccg acaccaaggt cctggacatc tgtgtctcgc tgttgatggg agagcagttt 3540

ctcgtcagct ggtgttagcg agcgccgct cttagtagca gtatcatcc cagccactt 3600

ttccccaata cgactacgag atctgtggct tctagctgcc cgggtggcat ccctgtgacc 3660

cctccccagt gcctctctg gccctggaag ttgccactcc agtgcacc agccttgtec 3720

taataaaatt aagtgtcatc atttgtctg actaggtgtc cttctataat attatggggt 3780

ggaggggggt ggtatggagc aaggggcaag ttgggaaggc cgaccgcga atagtagatc 3840

ccgcgagggc ttgaatctat cacctagagt acaccctaga gaatagctag ctctcaatga 3900

ctaaggacta aacttgglat ttgactgaa gcctgtcccc tcactgttgg cgctaggagg 3960

agagttcgta gaaaggatag tacgatttaa gtatctctaa gccttgtgaa gcactaaggt 4020

tgcgtacaga cgtgcttgaa ttacggataa ttcgggaacc ttgggacaca caaaaaacca 4080

acacacagat ctaatgaaaa taaagatctt ttatttaggc gcctctgact tcttggggat 4140

tgacctgagt tctactctag cgtttgctgg ttccgtgaac taatctgtga gatcccaac 4200

tctccgtttg ggatctccac tctctggtgt cctaaccttg gtgccccact gtctactgct 4260

agtgagacct tacgcctga gaaacgtggc gttactctaa ctaagcgacg cgcacttgca 4320

ctctgaatac ttctaccgta actaaccctg gacctcagaa ctcagacgga tctacgtgt 4380
ccatcaacac cagacttaga ttacctctgt taagtttaat taagctcgcg aaggaacccc 4440
tagtgatgga gttggccact cctctctctgc gcgctcgctc gctcactgag gccgggagac 4500
caaaggtcgc ccgacgcccc ggctttgccc gggcggcctc agtgagcgag cgagcgcgca 4560
gagagggagt ggccaa 4576

<210> 63

<211> 6

<212> PRT

<213> Homo sapiens

<400> 63

Tyr Arg Pro Ser Glu Thr

1 5

<210> 64

<211> 11

<212> PRT

<213> Rattus norvegicus

<400> 64

Leu Leu Leu Leu Ala Val Leu Cys Leu Gly Thr

1 5 10

<210> 65

<211> 20

<212> DNA

<213> Unknown

<220><223> Description of Unknown:

Collagen Stability (CS) sequence

<400> 65

cccagccac tttcccca 20

<210> 66

<211> 6

<212> PRT

<213> Unknown

<220><223> Description of Unknown:

Collagen Stability (CS) sequence

<400> 66

Pro Ser Pro Leu Phe Pro

1 5

<210> 67

<211> 192

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 67

gctctgtgcg gcggggagct ggtggacacc ctccagttcg tctgtgggga ccgcggcttc	60
tacttcagca ggcccgcaag ccgtgtgagc cgtcgagcc gtggcatcgt tgaggagtgc	120
tgtttccgca gctgtgacct ggccctcctg gagacgtact gtgctacccc cgccaagtcc	180
gagggcgcg cg	192

<210> 68

<211> 189

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 68

ctgtgcggcg gggagctggt ggacaccctc cagttcgtct gtggggaccg cggttctac	60
ttcagcaggc ccgaagccg tgtgagccgt cgcagccgtg gcatcgttga ggagtgcgt	120
ttccgcagct gtgacctggc cctcctggag acgtactgtg ctacccccgc caagtccgag	180
ggcgcgccg	189

<210> 69

<211> 210

<212> DNA

<213> Artificial Sequence

<220><223>

> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 69

gcttaccgcc ccagttagac cctgtgcggc ggggagctgg tggacaccct ccagttcgtc	60

tgtggggacc gcggcttcta cttcagcagg cccgcaagcc gtgtgagccg tcgcagccgt 120
ggcatcatgg aggagtgtctg ttccgcagc tgtgacctgg ccctcctgga gacgtactgt 180
gctacccccg ccaagtccga gggcgcgccg 210

<210> 70

<211> 264

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 70

atgggaatcc caatggggaa gtcatgtctg gtgcttctca cttctttggc cttegcctcg 60
tgctgcattg ctgctctgtg cggcggggag ctggtggaca ccctccagtt cgtctgtggg 120
gaccgcggct tctacttcag caggccccga agccgtgtga gccgtcgcag ccgtggcatc 180
gttgaggagt gctgtttccg cagctgtgac ctggccctcc tggagacgta ctgtgctacc 240
cccgcccaagt ccgagggcgc gccg 264

<210> 71

<211> 300

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 71

taggcgcctc tgacttctcg gggattgacc tgagttctac tctagcgttt gctggttcgg 60
tgaactaatc tgtgagatcc ccaactctcc gtttgggac tccactctct ggtgtcctaa 120
ccttggtgcc ccaactgtcta ctgctagtga gaccttacgc gctgagaaac gtggcggttac 180
tctaactaag cgacgcgcac ttgcactctg aatacttcta ccgtaactaa ccccggacct 240
cagaactcag acggatctac gctgtccatc aacaccagac ttagattacc tctgttaagt 300

<210> 72

<211> 2859

<212> DNA

<213> Homo sapiens

<400> 72

atgggagtga ggcacccgcc ctgctccac cggctcctgg ccgtctgcgc cctcgtgtcc 60

ttggcaaccg ctgcactcct ggggcacatc ctactccatg atttctgtct ggttccccga 120

gagctgagtg gtcctcccc agtctctggag gagactcacc cagctcacca gcagggagcc 180

agcagaccag ggccccggga tgcccaggca cccccggcc gtcccagagc agtgcccaca 240

cagtgcgacg tccccccaa cagccgcttc gattgcgccc ctgacaaggc catcacccag 300

gaacagtgcg aggcccgagg ctgttgctac atccctgcaa agcaggggct gcagggagcc 360

cagatggggc agccctgggt cttcttccca ccagctacc ccagctaaa gctggagaac 420

ctgagctcct ctgaaatggg ctacacggcc accctgaccc gtaccacccc caccttcttc 480

cccaaggaca tctgacctt gcggctggac gtgatgatgg agactgagaa cgcctccac 540

ttcacgatca aagatccagc taacaggcgc tacgaggtgc ccttgagac cccgcatgtc 600

cacagccggg caccgtcccc actctacagc gtggagtctt ccgaggagcc cttcgggggtg 660

atcgtgcgcc ggcagctgga cggccgcgtg ctgctgaaca cgacgggtggc gcccctgttc 720

tttgcggacc agttcttca gctgtccacc tcgtgcctt cgcagtatat cacaggcctc 780

gccgagcacc tcagtcctt gatgtcagc accagctgga ccaggatcac cctgtggaac 840

cgggaccttg cgcacacgcc cgggtgcgaac ctctacgggt ctaccccttt ctacctggcg 900

ctggaggacg gcgggtcggc acacgggggtg ttctgtctaa acagcaatgc catggatgtg 960

gtcctgcagc cgagccctgc ccttagctgg aggtcgacag gtgggacctt ggatgtctac 1020

atcttcttgg gccagagcc caagagcgtg gtgcagcagt acctggacgt tgtgggatac 1080

ccgttcatgc cgcatactg gggcctgggc ttccacctgt gccgtgggg ctactcctcc 1140

accgtatca cccgccaggt ggtggagaac atgaccaggg cccacttccc cctggacgtc 1200

cagtggaac accgtgacta catggactcc cggagggact tcacgttcaa caagatggc 1260

ttccgggact tccggccat ggtgcaggag ctgcaccagg gcggccggcg ctacatgatg 1320

atcgtggatc ctgccatcag cagctcgggc cctgccggga gctacaggcc ctacgacgag 1380

ggtctgcgga ggggggtttt catcaccaac gagaccggcc agccgtgat tgggaaggta 1440

tgccccgggt cactgcctt ccccgacttc accaaccaca cagccctggc ctggtgggag 1500

gacatggtgg ctgagttcca tgaccagggt cccttcgacg gcatgtggat tgacatgaac 1560

gagccttcca acttcatcag gggctctgag gacggctgcc ccaacaatga gctggagaac 1620

ccacctacg tgcttgggtt ggttgggggg accctccagg cggccacat ctgtgcctcc 1680

agccaccagt ttctctccac aactacaac ctgcacaacc tctacggcct gaccgaagcc 1740

atcgctccc acagggcgct ggtgaaggct cgggggacac gccatttgt gatctccgc 1800
 tcgacctttg ctggccacgg ccgatacgcc ggccactgga cgggggacgt gtggagctcc 1860
 tgggagcagc tcgcctcctc cgtgccagaa atcctgcagt ttaacctgct gggggtgcct 1920
 ctggtcgggg ccgacgtctg cggcttcctg ggcaacacct cagaggagct gtgtgtgcgc 1980
 tggaccacgc tgggggcctt ctacccttc atcggaacc acaacagcct gctcagtctg 2040
 ccccaggagc cgtacagctt cagcgagccg gccagcagg ccatgaggaa ggccctcacc 2100
 ctgcgtacg cactcctccc ccacctctac aactgttcc accaggccca cgtcgcgggg 2160

gagaccgtgg cccggccctt ctctctggag ttccccaagg actctagcac ctggactgtg 2220
 gaccaccagc tctgtgggg ggaggccctg ctcatcacc cagtgtcca ggccgggaag 2280
 gccgaagtga ctggctactt ccccttgggc acatggtacg acctgcagac ggtgccagta 2340
 gagggccctt gcagcctccc accccacct gcagctcccc gtgagccagc catccacagc 2400
 gaggggcagt ggtgtacgt gcccggcccc ctggacacca tcaacgtcca cctccgggct 2460
 ggttacatca tccccctgca ggccctggc ctcaacca cagagtccg ccagcagccc 2520
 atggccctgg ctgtggcctt gaccaagggt ggggaggccc gaggggagct gtcttgggac 2580

gatggagaga gcctggaagt gctggagcga ggggcctaca cacaggctcat ctctctggcc 2640
 aggaataaca cgatcgtgaa tgagctggta cgtgtgacca gtgaggagc tggcctgcag 2700
 ctgcagaagg tgactgtcct gggcgtggcc acggcgcccc agcaggtcct ctccaacggt 2760
 gtccctgtct ccaacttcac ctacagcccc gacaccaagg tcttggacat ctgtgtctcg 2820
 ctgttgatgg gagagcagtt tctcgtcagc tgggtttag 2859

<210> 73

<211> 2859

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<400> 73

atgggagtga ggcaccgcc ctgctccac cgctcctgg ccgtctgcgc cctcgtgtcc 60
 ttggcaaccg ctgcactcct ggggcacatc ctactccatg atttctgtct ggttccccga 120
 gagctgagt gctcctcccc agtcctggag gagactcacc cagctcacca gcaggagacc 180
 agcagaccag ggccccggga tgcccaggca cccccggcc gtcccagagc agtgcaccaca 240
 cagtgcgacg tccccccaa cagccgttc gattgcgccc ctgacaaggc catcaccag 300
 gaacagtgcg aggccccgcg ctgctgtac atccctgcaa agcaggggct gcaggagacc 360

cagatggggc agccctgggtg cttcttccca cccagctacc ccagctacaa gctggagaac 420

ctgagctcct ctgaaatggg ctacacggcc accctgaccc gtaccacccc caccttcttc 480

cccaaggaca tectgacctt gcggtctggac gtgatgatgg agactgagaa cgcctccac 540

ttcacgatca aagatccagc taacaggcgc tacgaggtgc ccttggagac cccgctgtc 600

cacagccggg caccgtcccc actctacagc gtggagtctt ccgaggagcc cttcggggtg 660

atcgtgcacc ggcagctgga cggccgctg ctgctgaaca cgacgggtggc gccctgttc 720

tttgcggacc agttccttca gctgtccacc tcgtgcctt cgcagtatat cacaggcctc 780

gccgagcacc tcagtccctt gatgtcagc accagctgga ccaggatcac cctgtggaac 840

cgggaccttg cccccacgcc cggtcgaac ctctacgggt ctacccttt ctacctggcg 900

ctggaggacg gcgggtcggc acacgggggtg ttctgctaa acagcaatgc catggatgtg 960

gtcttcgagc cgagccctgc ccttagctgg aggtcgacag gtgggacctt ggatgtctac 1020

atcttcttgg gccagagcc caagagcgtg gtgcagcagt acctggacgt tgtgggatac 1080

ccgttcacgc cgcatactg gggcctgggc ttccacctgt gccgtgggg ctactcctcc 1140

accgtatca cccgccaggt ggtggagaac atgaccaggg cccacttccc cctggacgtc 1200

caatggaacg acctggacta catggactcc cggagggact tcacgttcaa caaggatggc 1260

ttccgggact tccggccat ggtgcaggag ctgcaccagg gcggccggcg ctacatgatg 1320

atcgtggatc ctgccatcag cagctcgggc cctgccggga gctacaggcc ctacgacgag 1380

ggtctgcgga ggggggtttt catcaccaac gagaccggcc agccgtgat tgggaaggtg 1440

tggcccggtt ccaactgcctt ccccgacttc accaacccca cagccctggc ctggtgggag 1500

gacatggttg ctgatttcca tgaccagggt cccttcgacg gcatgtggat tgacatgaac 1560

gagccttcca acttcacag aggtctgtag gacggctgcc ccaacaatga gctggagaac 1620

ccacctacg tgcttgggtt ggttgggggg accctccagg cggccacat ctgtgcctcc 1680

agccaccagt ttctctccac aactacaac ctgcacaacc tctacgcctt gaccgaagcc 1740

atgcctccc acagggcgtt ggtgaaggct cgggggacac gccatttgt gatctccgc 1800

tcgaccttg ctggccacgg ccgatacgcc ggccactgga cgggggacgt gtggagctcc 1860

tgggagcagc tcgctcctc cgtgccagaa atcctgcagt ttaacctgct gggggtgcct 1920

ctggtcgggg ccgagctctg cggttctctg ggcaacacct cagaggagct gtgtgtgcgc 1980

tggaccacgc tgggggcctt ctacccttc atgcggaacc acaacagcct gctcagctctg 2040

ccccaggagc cgtacagctt cagcgagccg gccagcagg ccatgaggaa ggcctcacc 2100

ctgcgtacg cactcctccc ccacctctac aactgttcc accaggccca cgtcgcgggg 2160
gagaccgtgg cccggccctt cttcctggag ttccccaagg actctagcac ctggactgtg 2220
gaccaccagc tctgtgggg ggaggccctg ctcatcacc cagtgtcca ggccgggaag 2280
gccgaagtga ctggctactt ccccttgggc acatggtacg acctgcagac ggtgccaata 2340
gaggcccttg gcagcctccc acccccacct gcagctcccc gtgagccagc catccacagc 2400
gaggggcagt ggggtgacgt gccggccccc ctggacacca tcaacgtcca cctccgggct 2460
gggtacatca tccccctgca gggccctggc ctcaacaacca cagagtcccg ccagcagccc 2520

atggccctgg ctgtggccct gaccaagggt ggagaggccc gaggggagct gttctgggac 2580
gatggagaga gcctggaagt gctggagcga ggggcctaca cacaggtcat cttcctggcc 2640
aggaataaca cgatcgtgaa tgagctggta cgtgtgacca gtgaggagc tggcctgcag 2700
ctgcagaagg tgactgtcct gggcgtggcc acggcgcccc agcaggtcct ctccaacggt 2760
gtccctgtct ccaacttcac ctacagcccc gacaccaagg tcttgacat ctgtgtctcg 2820
ctgttgatgg gagagcagtt tctcgtcagc tgggttag 2859

<210> 74

<211> 2859

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 74

atgggagtaa ggcatccacc atgtctcat aggtctctcg ccgtatgcgc gttggtcagc 60
cttgcgaccg cagcccttct gggccacatt ctgtgcacg attttctcct cgttcgcga 120
gagttgtctg gaagctctcc agtactcgag gaaacgcac cagcgacca gcaggcgca 180
tctcgcccc gtccaagaga cgcacaagca cccccgggc ggccaagggc tgttccaact 240
cagtgcgatg ttccgcccc ttcaagattc gattgcgcac ccgataaagc tattacgaa 300
gagcagtgcg aagctcgcgg ctgctgttat atcccggcaa agcagggcct tcagggcgct 360

caaatgggac agccgtggtg tttttttcca ccgtcatatc catctacaa gcttgagaac 420
ctgtctcat ctgagatggg atacagggt accctgacac ggacaacgcc aaccttcttc 480
cccaaagaca ttcttacct gcgactggac gtgatgatgg aaacagagaa tcgactgcat 540
tttactataa aggaccggc taacaggcgg tacgaggttc ccctggagac ccctcgcgta 600
cattctcgag ctctagccc gctctactcc gtggaatfff ccgaagagcc gtttggtgta 660
atagtgcac gccaaactga cggtcgcgta ctgctcaaca ccacggtggc acctctgttc 720

tttgcggacc aattcttgca gctctctact tctctccctt cacaatatat aaccgggctc	780
gcagagcatc tgtccccgtt gatgttgtct acgtcatgga caagaatcac gctttggaac	840
cgagacctgg cccttacgcc gggagctaatt ctgtacgggt cacatccatt ctatttggt	900
ctggaggatg gcggtagtgc acacggggta tttctgctta actctaagtc gatggatgtc	960
gtgctgcagc catccccctgc gctgtcttgg cgaagcaccg gcggcattct ggacgtgtat	1020
atctttcttg gacccgagcc aaagagcgtc gtacaacagt acctcgacgt agtggggat	1080
cccttcacgc cccctattg ggggctcggg ttccatttgt gcagatgggg ctactcatcc	1140
actgcaataa cagccaggt agtagaaaat atgaccggg ctcactttcc tcttgacgtg	1200
cagtggaatg accttgacta catggatagt cgccgcgatt tcacattcaa caaggacggt	1260
ttcagagatt tccctgctat ggtgcaagag ctgcaccaag ggggccgcgc gtatatgatg	1320
atcgtagatc cgcccatatc tagctctggc cctgccgggt cttatagacc gtatgatgaa	1380
ggtctgaggc gcggtgtctt catcacaac gaaactggtc aaccactcat cggcaaggtc	1440
tggcccggt caacggcggt tcttgatttc actaatccga cggcgctggc ttggtgggaa	1500
gatatggttg ccgagtttca cgatcaagtt ccgtttgacg gaatgtggat tgatatgaat	1560
gagccatcta acittatagc cggtagtgaa gatggttgcc ctaataatga gttggaaaac	1620
ccacctatg taccgggtgt cgtcggcggc acgcttcagg ccgctaccat atgcgctagc	1680
agtcaccaat tccttagtac gcactataac ctccataatt tgtacggact tactgaggcg	1740
attgctagtc atcgagcgt cgtaaaagct cgcggtaccc ggcccttcgt gatttagtagg	1800
agtacattcg ccgggcattg tcgatatgcg ggccactgga ctggggatgt ctggtccagc	1860
tgggagcagt tggcttcctc agtccctgaa attttgcaat ttaacctgct tgggtgtccg	1920
ctgtttggcg cggacgtttg cgggtttttg ggtaatacca gtgaagaact gtgtgttagg	1980
tggacacagc ttggggcatt ctatcctttc atgcggaatc acaacagtct tctgtcattg	2040
ccgcaggaac cgtactcttt cagtgagcct gccagcagg cgatgcggaa ggccctcaca	2100
ttgcgatacg cacttctccc gcacttttat acgcttttcc accaggccca tgttgcgggc	2160
gagacagtgc ccaggccact ctctcttgag ttccgaaag acagcagcac ctggacggtt	2220
gaccatcagc tcctctgggg tgaagctttg ttgattaccc ccgtgctgca ggcgggcaag	2280
gcggaagtaa caggctactt cccgtggga acctggtatg accttcagac agtgccgatc	2340
gaagctcttg gtcactgcc accacctccg gcagccccgc gggaaccagc gattcactcc	2400
gagggccaat gggttacgt tccggtcca ctgacacca tcaatgtcca tcttcgcgcg	2460

ggatacatta tcccgttca gggacctggc ttgactacta cagagtctcg ccaacagccg 2520
atggcactcg cagtagctct gactaaagga ggtgaggcgc ggggtgaact cttttgggat 2580
gatggggaat ccctggaggt tttggagagg ggtgctaca cgcaagtaat ctttctcgt 2640
cgaaacaata cgatcgtaa tgagctcgtc agggtaacgt ctgagggggc cggtttgcaa 2700
cttcagaaag tcaactgtact tggcgtagcc actgctcctc aacaagtcct ttcaaacggg 2760
gttctctgtt caaactttac ctattcaccg gataccaagg ttttggacat atgcgtgtct 2820
ttgttgatgg gcgaacaatt tctggtttct tgggtgctag 2859

<210> 75

<211> 2859

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 75

atgggcgtta gacacctcc atgctctcat agactgctgg ccgtgtgtgc tctgggtgtt 60
cttgctacag ctgccctgct gggacatac ctgctgcacg attttctgct ggtgcccaga 120
gagctgtctg gcagctctcc agtctctggaa gaaacacacc ctgcacatca gcaggcgcc 180
tctagacctg gacctagaga tgctcaagcc catcctggca gacctagagc cgtgcctaca 240
cagtgtgacg tgccacctaa cagcagattc gactgcgcc ctgacaaggc catcacacaa 300

gagcagtgtg aagccagagg ctgctgctac attcctgcca aacaaggact gcaggcgct 360
cagatgggac agccttggtg cttcttccca ccattctacc ccagctacaa gctggaaaac 420
ctgagcagca gcgagatggg ctacaccgcc aactgacca gaaccacacc tacattcttc 480
ccaaaggaca tctgacact gcggctggac gtgatgatgg aaaccgagaa cggctgcac 540
ttcaccatca aggaccccg caatagaaga tacgaggtgc ccctggaaac ccctagagtg 600
cattctagag ccccatctcc actgtacagc gtggaattca gcgaggaacc cttcgcgctg 660
atcgtgcaca gacagctgga tggcagagtg ctgctgaata ccacagtggc ccctctgttc 720

ttcgccgacc agtttctgca gctgagcaca agcctgccta gccagtatat cacaggcctg 780
gccgaacacc tgctccact gatgctgagc accagctgga ccagaatcac cctgtggaac 840
agagatctgg ccctacacc tggcgccaat ctgtacggct ctcaccttt ttatctggcc 900
ctggaagatg gcggaagcgc ccacggtgtc tttctgctga acagcaacgc catggacgtg 960
gtgctgcaac catctctgc tctgtcttgg agaagcaccg gcggcatcct ggacgtgtac 1020
atctttctgg gccctgagcc taagagcgtg gtgcagcagt atctggatgt cgtgggctac 1080

cccttcatgc ctccittattg gggcctgggc ttccacctgt gtagatgggg atacagctcc	1140
accgccatca ccagacaggt ggtggaaaac atgacccggg ctactttccc actggatgtg	1200
cagtggaaacg acctggacta catggacagc agacgggact tcaccttcaa caaggacggc	1260
ttcagagact tccccgcat ggtgcaagaa ctgcatcaag gcggcagacg gtacatgatg	1320
atcgtggatc ctgccatctc ttctagcggc cctgccggct cttacagacc ttatgatgag	1380
ggcctgagaa gaggcgtgtt catcaccaat gagacaggcc agcctctgat cggcaaagtg	1440
tggcctggat ccaccgcctt tccagacttc accaatccta ccgctctggc ttggtgggaa	1500
gatatggtgg ccgagtttca cgatcaggtg cccttcgacg gcatgtggat cgacatgaac	1560
gagcccagca acitcatcag aggcagcagc gacggctgcc ccaacaacga actggaaaat	1620
cctccttacg tgccaggcgt tgtcggagga aactgcagg ccgccacaat ttgtgccagc	1680
agccatcagt ttctgagcac ccactacaac ctgcacaacc tgtacggcct gaccgaggcc	1740
attgtttctc acagagccct ggttaaggcc agaggcacca gacctttctg gatcagcaga	1800
agcacattcg ccggccacgg cagatatgct ggacattgga caggcgacgt gtggtctagt	1860
tgggagcagc tggttagctc tgtgcccag atcctgcagt ttaatctgct gggagtgcct	1920
ctcgtggcgc ccgatgtttg tggctttctg ggaaacacct ccgaggaact gtgtgtgcgt	1980
tggacacagc tgggcgcctt ctatcccttc atgagaaacc acaacagcct gctgagcctg	2040
cctcaagagc cttacagctt tagcgaaccc gcacagcagg ccatgagaaa ggccctgact	2100
ctgagatacg ctctgtgcc ccacctgtac accctgtttc atcaggttca tgtggccggg	2160
gaaacagtgg ctagaccctt gtctctggaa ttccccaagg atagcagcac ctggaccgtg	2220
gatcatcagc tgctttgggg cgaagctctg ctgattacac ctgtgctgca ggctggcaag	2280
gccgaagtga caggctattt tcccctcggc acttggtagc acctgcagac agtgcctatt	2340
gaggccctgg gatctcttcc tccacctcct gctgctccta gagagcctgc cattcactct	2400
gaaggccagt gggttacact gcccgtcct ctggacacca tcaatgtgca cctgagagcc	2460
ggctacatca tcctctgca aggccttgga ctgaccacaa ccgaaagcag acagcagcca	2520
atggctctgg ccgtggcact gacaaaaggc ggagaagcta gaggcgagct gttctgggat	2580
gatggcgagt ctctggaagt gctcgagaga ggcgctaca cacaagtgat ctttctcgcc	2640
cggaacaaca ccatcgtgaa cgaactctc agagtacca gtgaagggtc cggactgcag	2700
ctccagaaag tgacagtgtc tggagtggcc acagctcctc agcaggttct gtctaattggc	2760
gtgccctgt ccaacttcac atacagccct gacaccaagg tgctggatat ctgcgtgtca	2820

ctgctgatgg gcgagcagtt cctgggtgtcc tgggtgtga 2859

<210> 76

<211> 2859

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 76

atgggcgtga ggcacccacc ttgctctcac aggctgtctgg ccgtgtgcgc actggtgagc 60

ctggcaaccg ccgccctgct gggccacatc ctgctgcacg acttcctgct ggtgccaagg 120

gagctgtccg gaagctcccc agtgtctggag gagaccacc cagcacacca gcagggagca 180

tctaggccag gccccagaga tgcacaggca caccaggca gaccagagc agtgccaacc 240

cagtgcgacg tgccaccaa cagccggttt gactgtgccc ccgataaggc catcacacag 300

gagcagtgcg aggcagaggg ctgctgttat atccctgcaa agcagggact gcagggagca 360

cagatgggac agccatggtg tttctttcct ccatcttacc ccagctataa gctggagaat 420

ctgtctagct ccgagatggg ctacacagcc accctgacaa gaaccacacc aacattcttt 480

ccaagagata tctgacctt gcggctggac gtgatgatgg agacagagaa cagactgcac 540

ttcacatca aggatcccg caatcggaga tatgaggtgc ctctggagac cccaagggtg 600

cactctaggg cacctagccc actgtactcc gtggagtct ctgaggagcc atttgccgtg 660

atcgtgcacc ggcagctgga tggcagagt ctgctgaaca ccacagtggc cccctgttc 720

tttgccgacc agttcctgca gctgagcaca tcctgcctt ccagtatat caccggcctg 780

gccgagcacc tgtctctct gatgtgtct accagctgga caaggatcac cctgtggaac 840

agggacctgg caccaacccc tggagcaaat ctgtacggca gccacccctt ctatctggcc 900

ctggaggatg gaggatccg ccacggcgtg tttctgtga actctaagc catggacgtg 960

gtgtgcagc caagccccg cctgtcctgg aggtctaccg gaggcacct ggacgtgtac 1020

atcttctggt gccctgagcc aaagtccgtg gtgcagcagt acctggacgt ggtgggctat 1080

cctttcatgc ccccttactg gggactggga tttcacctgt gcagatgggg ctattctagc 1140

acagccatca cccggcaggt ggtggagaac atgaccagag cccactttcc actggatgtg 1200

cagtggaatg acctggatta catggactcc aggcgcgact tcacattcaa caaggacggc 1260

ttccgggatt ttccgccat ggtgcaggag ctgcaccagg gaggccggag atacatgatg 1320

atcgtggatc ctgcaatctc ctctagcggg cctgcaggaa gctacagacc atatgacgag 1380

ggcctgaggg gcggcgtgtt catcaciaac gagaccggcc agcctctgat cggcaaggtc 1440

tggccaggct ccaccgcctt cccagacttc accaatccaa ccgcctggc atggtgggag 1500

gacatggtgg ccgagttcca cgaccagggtg ccttttgatg gcatgtggat cgacatgaac 1560

gagccatcta atttcatcag gggaagcgag gacggatgcc caaacaatga gctggagaat 1620

ccaccatatg tgcttgaggt ggtgggaggc aactgcagg cagcaacat ctgtgcctcc 1680

tctcaccagt ttctgtctac acactataac ctgcacaatc tgtacggact gaccgaggca 1740

atcgcaagcc acagagccct ggtgaaggca aggggaacaa ggcctttcgt gatctccagg 1800

tctacctttg ccggacacgg ccgctacgca ggacactgga ccggcgacgt gtggagctcc 1860

tgggagcagc tggcctctag cgtgccagag atcctgcagt tcaacctgct gggagtcca 1920

ctggtgggag cagacgtgtg cggctttctg ggcaatacaa gcgaggagct gtgcgtgcgg 1980

tggaccacgc tgggagcctt ctatcccttt atgagaaacc acaatagcct gctgtccctg 2040

cctcaggagc catacagctt ctccgagcct gcacagcagg caatgaggaa ggccctgaca 2100

ctgagatatg cctgtctgcc acacctgtac accctgtttc accaggcaca cgtggcagga 2160

gagacagtgg ccaggccctt gttcctggag ttctctaagg attcctctac ctggacagtg 2220

gaccaccagc tgctgtgggg agaggccctg ctgatcacc ccgtgctgca ggcaggcaag 2280

gcagaggtga caggctattt ccctctgggc acatggtacg atctgcagac cgtgccaatc 2340

gaggccctgg gaagcctgcc tccaccacct gcagcaccaa gggagcctgc catccactcc 2400

gagggacagt gggtagacct gccagcacct ctggacacca tcaacgtgca cctgagggcc 2460

ggctatatca tccactgca gggacctgga ctgaccacaa ccgagagcag gcagcagcca 2520

atggcactgg cagtggccct gaccaaggga ggcgaggcca ggggcgagct gttctgggac 2580

gatggagagt ccctggaggt gctggagagg ggagcctaca cacaggtcat cttcctggcc 2640

aggaacaata caatcgtgaa tgagctggtg cgcgtgacct ctgagggagc aggactgcag 2700

ctgcagaagg tgacagtgtt gggagtggca accgcaccac agcaggtgct gtccaacggc 2760

gtgccctga gcaatttcac atactccct gataccaagg tgctggacat ctgcgtgagc 2820

ctgctgatgg gcgagcagtt tctggtgtcc tgggtgtga 2859

<210> 77

<211> 96

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 77

ccgggcggag tgtgttagtc tctccagagg gaggttggtt ccccaggga gcagagcctg	60
tgtgcgggca gcagctgtgt gcgggcctgg gggttg	96
<210> 78	
<211> 2032	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 78	
aagtgccgac ccgtctgaat agtagatccc gtcgagggt tgaatctatc accatagagt	60
acacccatag agaatagcta gctctcaatt gactaaggac taaacttggg atttcgactg	120
aagcctgtca ctctcactg ttggcgctag tgaggagagt tcgtagaaag gatagtacga	180
tttaagtatc tctaagcctt gtgaagcact aaggttgcgt acagacgtgc ttgaattacg	240
gataattcgt ggaaccttgg gatagtgcgc ctctgacttc tctgtgggta ttgacctgag	300
ttctactcta gcgtttgctg gtctgtgtga actaatctgt gagatcctca caactctccg	360
tttgggatct ccactctctg gtgtcctaac acttgggtgcc acacactgtc tactgctagt	420
gagaccttat ctgtctgctg agaaactgtg tgctgttact ctaactaagc gatctgtctg	480
cacttgcact ctgaatactt ctaccgtaac taactctcac ggactctcag aactcagacg	540
gatctactgc tgtctcaaat caacaccaga cttagattac tctctgttaa gttgaaagac	600
tctctacaaa tccacactgt ggtctagtgt aacaagcttg agagagtatc tgtctgcata	660
tcaagagtga gagtgtctct ggatttggag tagttgagaa gtggaacct agcgagtcta	720
ctgcctgtgt caacttcaat tagagtgttg aatagaggac tgactctcga attgaagtat	780
cctgccaaact gtctcaactt aattcaaagc ctattttcgt taactgtctc aactctcgtc	840
aagttgagag gaagagtaga gccatagttt gaagcctgat aagagtctgt ctgaggacgt	900
attgtgtgtc aagacgaact gaaggatttc aacaaataga gaagcttgac actgtccttc	960
aactgtctga actgtagcac tctggatagc tagctagcgt ttgtagatag tttcaaggat	1020
agcaaccaac aacacgatag tttagaaatt gttctagatc ttctgctact actcgagaga	1080
tagttgactt ctagaagact ctggatagag agcttctctgt ggtttctatt tcaacacttc	1140
tgagcgattt agatagaggc gtcaattgag aacaaaaact ggacacctgt ttctagaaca	1200
ctagtgttag tcttctcaag gcttgatact gtctgtccac tcaagagaaa tgttctggca	1260
cctgcacttg cactggggac agcctatttt gctagtttgt tttgtttcgt tttgttttga	1320

tggagagcgt atgttagtac tatcgattca cacaaaaaac caacacacag atgtaatgaa 1380
aataaagata ttttattaga gagcgagtgg atagtttaga gtcgcttgag acttatcctg 1440
tgtccaacta gccaagagta gagagagagt gactaggtgg atcgagagat cgtcgtctat 1500
ttgactgcta gcctgtgttc tcgtcgtagt ccaaatcttc aatactctgg cgactagtaa 1560
cttgaatctg tgtatctaac acgacaaggc tactcctttg agtttctagt gataagtgcg 1620

actggatcta gctagttgag tttctagaaa ctactgtcca aagctactag cactaaggtc 1680
tctctgtgaa agagtggctc acttttgagg agaagaaagt tctagcgtag tagcaattcc 1740
tttggtggaa gaggctatcg agagtcaact ctgccttcaa gtagcttcca aaagtaacaa 1800
actgtggatc gtttctaate tcaagacgct tgaagagca agagcttgta ttgatcgagg 1860
tggtattgag atagttaaga ctgtccaaga attacgacta ccaaagatct tcgttagcaa 1920
cgaaacttga gaacgtttct aagcctgtga cttaactct cctgcttggt gcctgccttt 1980
gaatctggac tcgtagagca cgaggaagag aagtcgattt cttaggcctg ct 2032

<210> 79

<211> 4256

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 79

tcgagcttgg gctgcaggctc gagggcactg ggaggatgtt gagtaagatg gaaaactact 60
gatgaccctt gcagagacag agtattagga catgtttgaa caggggccgg gcgatcagca 120
ggtagctcta gaggatcccc gtctgtctgc acatttcgta gagcgagtgt tccgatactc 180
taatctccct aggcaagggt catatttggt taggttactt attctccttt tgttgactaa 240
gtcaataatc agaatcagca ggtttggagt cagcttggca gggatcagca gcctgggttg 300

gaaggagggg gtataaaagc cccttcacca ggagaagccg tcacacagac tagccctaag 360
glaagttggc gccgtttaag ggatggttgg ttggtggggg attaatgttt aattaccttt 420
tttacaggcc tgaactagcg gcgccaccgc caccatgccg tcttctgtct cgtggggcat 480
cctctgtctg gcaggcctgt gctgccttgt cctgtctcc ctggtgtctt accgccccag 540
tgagaccctg tgcggcgggg agctgggtgga caccctccag ttcgtctgtg gggaccgcgg 600
cttctacttc agcaggcccc caagccgtgt gagccgtcgc agccgtggca tcatggagga 660
gtgctgtttc cgcagctgtg acctggccct cctggagacg tactgtgcta cccccccaa 720

gtccgagggc gcgccggcac accccggccg tcccagagca gtgcccacac agtgcgacgt 780

ccccccaac agccgcttcg attgcgcccc tgacaaggcc atcacccagg aacagtgcga 840

ggcccgcggc tgctgtctaca tccctgcaaa gcaggggctg cagggagccc agatggggca 900

gccctggtgc ttcttccac ccagctaccc cagctacaag ctggagaacc tgagctcctc 960

tgaaatgggc tacacggcca ccctgacccg taccaccccc accttcttcc ccaaggacat 1020

cctgaccctg cggttgacg tgatgatgga gactgagaac cgcctccact tcacgatcaa 1080

agatccagct aacaggcgct acgaggtgcc ctggagacc ccgctgttcc acagccgggc 1140

accgtcccca ctctacagcg tggagtctc cgaggagccc ttcggggtga tctgcaccg 1200

gcagctggac ggccgctgc tgctgaacac gacggtggcg cccctgttct ttgcggacca 1260

gttccttcag ctgtccacct cgctgccctc gcagtatctc acaggcctcg ccgagcacct 1320

cagtccctg atgtcagca ccagctggac caggatcacc ctgtggaacc gggaccttgc 1380

gcccacgccc ggtgcgaacc tctacgggtc tcacccttcc tacctggcgc tggaggacgg 1440

cgggtcggca cacggggtgt tctgtctaaa cagcaatgcc atggatgtgg tctgcagcc 1500

gagccctgcc cttagctgga ggtcgacagg tgggatcctg gatgtctaca tcttcttggg 1560

cccagagccc aagagcgtgg tgcagcagta cctggacgtt gtgggatacc cgttcatgcc 1620

gccatactgg ggccctgggt tccacctgtg ccgctggggc tactcctcca ccgctatcac 1680

ccgccaggtg gtggagaaca tgaccagggc ccacttcccc ctggacgtcc aatggaacga 1740

cctggactac atggactccc ggagggactt cacgttcaac aaggatggct tccgggactt 1800

cccggccatg gtgcaggagc tgcaccaggc cgcccgccgc tacatgatga tcttgatcc 1860

tgccatcagc agctcgggcc ctgccgggag ctacaggccc tacgacgagg gtctgcggag 1920

gggggttttc atcaccaacg agaccggcca gccgctgatt gggaaggtat ggcccgggtc 1980

cactgccttc cccgacttca ccaacccac agccctggcc tgggtgggagg acatggtggc 2040

tgagtccat gaccaggtgc ccttcgacgg catgtggatt gacatgaacg agccttccaa 2100

cttcatcaga ggctctgagg acggctgccc caacaatgag ctggagaacc caccctacgt 2160

gcctggggtg gttgggggga ccctccaggc ggccaccatc tgtgcctcca gccaccagtt 2220

tctctccaca cactacaacc tgcacaacct ctacggcctg accgaagcca tcgctccca 2280

cagggcgctg gtgaaggctc gggggacacg cccatttctg atctcccgt cgacctttgc 2340

tggccacggc cgatacgccg gccactggac gggggacgtg tggagctcct gggagcagct 2400

cgctcctcc gtgccagaaa tctgcagtt taacctgctg ggggtgcctc tggtcggggc 2460

cgacgtctgc ggcttctgg gcaacacctc agaggagctg tgtgtgcgt ggaccagct 2520

gggggccttc tacccttca tgcggaacca caacagcctg ctacgtctgc cccaggagcc 2580

gtacagcttc agcgagccgg cccagcaggc catgaggaag gccctcacc tgcgctacgc 2640
actctcccc cacctctaca cactgttcca ccaggccac gtcgcggggg agaccgtggc 2700
ccggccctc ttcttgaggt tcccaagga ctctagcacc tggactgtgg accaccagct 2760
cctgtggggg gaggcctgc tcatcacc cagtgtccag gccgggaagg ccgaagtgc 2820

tggtacttc cccttgggca catgttacga cctgcagacg gtgccaatag aggccttgg 2880
cagcctccca cccccacctg cagctccccg tgagccagcc atccacagcg aggggcagtg 2940
ggtgacgctg ccggccccc tggacacat caacgtccac ctccgggctg ggtacatcat 3000
ccccctgcag ggcctggcc tcacaaccac agagtccgc cagcagccca tggcctggc 3060
tgtggcctg accaagggtg gagaggcccg aggggagctg ttctgggacg atggagagag 3120
cctggaagtg ctggagcgag gggcctacac acaggtcatc ttcttgcca ggaataacac 3180
gatcgtgaat gagctggtag gtgtgaccag tgaggagct ggctgcagc tgcagaaggt 3240

gactgtcctg ggcgtggcca cggcgcccca gcaggtcctc tccaacggtg tccctgtctc 3300
caacttcacc tacagcccg acaccaaggt cctggacatc tgtgtctcgc tgttgatggg 3360
agagcagttt ctctcagct ggtgttagcg agcgccgct cttagtagca gtatcgatcc 3420
cagcccatt ttcccaata cgactacgag atctgtggct tctagctgcc cgggtggcat 3480
ccctgtgacc cctccccagt gcctctctg gccctggaag ttgccactcc agtgcccacc 3540
agccttgtcc taataaaatt aagttgcac attttgtctg actaggtgtc cttctataat 3600
attatggggt ggaggggggt ggtatggagc aaggggcaag ttggggtcct tcaacgtcga 3660

actgtagcac tctggatagc tagctagcgt ttgtagatag tttcaaggat agcaaccaac 3720
aacacgatag ttiagaaatt gttctagatc ttctgtact actcgagaga tagttgactt 3780
ctagaagact ctggatagag agcttctgt ggtttctatt tcaacacttc gagcgattta 3840
gatagaggcg tcaattgaga acaaaaactg gacacctgtt tctagaacac tagtgctagt 3900
cttctcaagg cttagacgt cctgccactc acacacaaa aaccaacaca cagattaatg 3960
aaaataaaga tcttttatta gagagcgagt ggatagtta ggtcgttg agacttatcc 4020
tgtgtccaac tagccaagag tagagagaga gtgactaggt ggatcgagag atcgtcgtct 4080

atttgactgc tagccttgt tctcgtcgt gtccaaatct tcaatactct ggcgactagt 4140
aacttgaatc tgtgtatcta acacgacaag gctactcctt tgagtttcta gtgataagtg 4200
cgactggatc tagctagtgt agtttctaga aactacgtcc aaagctacta gcacta 4256

<210> 80

<211> 4280

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 80

tcgagcttgg gctgcaggtc gagggcactg ggaggatgtt gagtaagatg gaaaactact	60
gatgaccctt gcagagacag agtattagga catgtttgaa caggggccgg gcgatcagca	120
ggtagctcta gaggatcccc gtctgtctgc acatttcgta gagcgagtgt tccgatactc	180
taatctccct aggcaaggtt catatttgtg taggttactt attctccttt tgttgactaa	240
gtcaataatc agaatcagca ggtttggagt cagcttggca gggatcagca gcctgggttg	300
gaaggagggg gtataaaagc cccttcacca ggagaagccg tcacacagac tagccctaag	360
gtaagtggc gccgtttaag ggatgggttg ttggtgggtt attaatgttt aattaccttt	420
tttacaggcc tgaactaggc gcgccaccgc caccatgctc aggggtccgg gacccgggcg	480
gctgctgctg ctagcagtc tgtgcctggg gacatcggtg cgctgcaccg aaaccgggaa	540
gagcaagagg gcttaccgcc ccagtgcagc cctgtgcggc ggggagctgg tggacacct	600
ccagttcgtc tgtggggacc gcggcttcta cttcagcagg cccgcaagcc gtgtgagccg	660
tcgcagccgt ggcatcatgg aggagtgtg tttccgcagc tgtgacctgg cctcctgga	720
gacgtactgt gctaccccc ccaagtccga gggcgcgcgc gcacaccccg gccgtcccag	780
agcagtgcgc acacagtgcg acgtcccccc caacagccgc ttcgattgcg ccctgacaa	840
ggccatcacc caggaacagt gcgaggcccg cggctgctgc tacatccctg caaagcaggg	900
gctgcaggga gccagatgg ggcagccctg gtgcttcttc ccaccagct accccagcta	960
caagctggag aacctgagct cctctgaaat gggtacacg gccaccctga cccgtaccac	1020
ccccaccttc ttcccccaagg acatcctgac cctgcggctg gacgtgatga tggagactga	1080
gaaccgcctc cacttcacga tcaaagatcc agctaacagg cgctacgagg tgcccttgga	1140
gaccccgct gtccacagcc gggcaccgtc cccactctac agcgtggagt tctccgagga	1200
gcccttcggg gtgategtgc accggcagct ggacggccgc gtgctgctga acacgacggt	1260
ggcgcacctg ttctttgcgg accagttcct tcagctgtcc acctcgctgc cctcgagta	1320
tatcacaggc ctgcgcgagc acctcagtc cctgatgctc agcaccagct ggaccaggat	1380
cacctgtgg aaccgggacc ttgcgccac gcccggtgcg aacctctac ggtctcacc	1440
tttctacctg gcgtggagg acggcgggtc ggcacacggg gtgttctgc taaacagcaa	1500
tgccatggat gtggtcctgc agccgagccc tgcccttagc tggaggtcga caggtgggat	1560
cctggatgtc tacatcttc tgggccaga gcccaagagc gtggtgcagc agtacctgga	1620

cgttgtggga taccggttca tgccgccata ctggggcctg ggcttccacc tgtgccgctg	1680
gggctactcc tccaccgcta tcaccgccca ggtggtggag aacatgacca gggcccactt	1740
ccccctggac gtccaatgga acgacctgga ctacatggac tcccggaggg acttcacgtt	1800
caacaaggat ggcttccggg acttcccggc catggtgcag gagctgcacc agggcggccg	1860
gcgctacatg atgatcgtgg atcctgccat cagcagctcg ggccctgccg ggagctacag	1920
gccctacgac gagggtctgc ggaggggggt ttctcatcacc aacgagaccg gccagccgct	1980
gattgggaag gtatggcccc ggtccactgc cttccccgac ttaccaacc ccacagccct	2040
ggcctggtgg gaggacatgg tggctgagtt ccatgaccag gtgcccttcg acggcatgtg	2100
gattgacatg aacgagcctt ccaacttcat cagaggctct gaggacggct gcccacaaca	2160
tgagctggag aaccacacct acgtgcctgg ggtggttggg gggaccctcc aggcggccac	2220
catctgtgcc tcagccacc agtttctctc cacacactac aacctgcaca acctctacgg	2280
cctgaccgaa gccatcgctt cccacagggc gctggtgaag gctcggggga cacgcccatt	2340
tgtgatctcc cgctcgacct ttgctggcca cggccgatac gccggccact ggacggggga	2400
cgtgtggagc tcttgggagc agctcgctc ctcctgcca gaaatcctgc agtttaacct	2460
gctgggggtg cctctggctg gggccgacgt ctgcggcttc ctgggcaaca cctcagagga	2520
gctgtgtgtg cgctggaccg agctgggggc cttctacccc ttcattgcgga accacaacag	2580
cctgctcagt ctgccccagg agccgtacag cttcagcgag ccggcccagc aggccatgag	2640
gaaggccctc accctgcgtt acgcactcct cccccactc tacacactgt tccaccaggc	2700
ccacgtcgcg ggggagaccg tggcccgcc cctcttctg gagttccca aggactctag	2760
cacctggact gtggaccacc agctcctgtg gggggaggcc ctgctcatca cccagtgtt	2820
ccaggccggg aaggccgaag tgactggcta cttcccttg ggcacatggt acgacctgca	2880
gacggtgcca atagaggccc ttggcagcct cccaccccca cctgcagctc cccgtgagcc	2940
agccatccac agcgaggggc agtgggtgac gctgccggcc cccctggaca ccatcaacgt	3000
ccacctccgg gctgggtaca tcatccccct gcagggcctt ggcctcaca ccacagagtc	3060
ccgccagcag cccatggccc tggctgtggc cctgaccaag ggtggagagg cccgagggga	3120
gctgttctgg gacgatggag agagcctgga agtgctggag cgaggggcct acacacaggt	3180
catcttctg gccaggaata acacgatcgt gaatgagctg gtacgtgtga ccagtgaggg	3240
agctggcctg cagctgcaga aggtgactgt cctgggcgtg gccacggcgc cccagcaggt	3300
cctctccaac ggtgtccctg tctccaactt cacctacagc cccgacacca aggtcctgga	3360
catctgtgtc tcgtgttga tgggagagca gtttctcgtc agctggtgtt agcgagcggc	3420

cgctcttagt agcagtatcg atcccagccc acttttcccc aatacgacta cgagatctgt 3480

ggcttctagc tgcccgggtg gcacccctgt gacccctccc cagtgcctct cctggccctg 3540

gaagttgcc aatcagtgcc caccagcctt gtcctaataa aattaagttg catcattttg 3600

tctgactagg tgiccttcta taatattatg gggtagggg gggtaggtat gagcaagggg 3660

caagttgggg tccttcaacg tcgaactgta gcactctgga tagctagcta gcgtttgtag 3720

atagtttcaa ggatagcaac caacaacacg atagtttaga aattgttcta gatcttctgc 3780

tactactcga gagatagttg acttctagaa gactctggat agagagcttc ctgtggtttc 3840

tatttcaaca cttcgagcga tttagataga ggcgtcaatt gagaacaaaa actggacacc 3900

tgtttctaga aactagtgcc tagtcttctc aaggcttgat acgtcctgcc actcacacac 3960

aaaaaaccaa cacacagatt aatgaaaata aagatctttt attagagagc gagtggatag 4020

tttagagtcg cttgagactt atcctgtgtc caactagcca agagtagaga gagagtgact 4080

aggtggatcg agagatcgtc gtctatttga ctgctagcct gtgttctcgt cgtagtccaa 4140

atcttcaata ctctggcgac tagtaacttg aatctgtgta tctaacacga caaggctact 4200

cctttgagtt tctagtata agtgcgactg gatctagcta gttgagtttc tagaaactac 4260

gtccaaagct actagcacta 4280

<210> 81

<211> 4277

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 81

tcgagcttgg gctgcaggtc gagggcactg ggaggatgtt gagtaagatg gaaaactact 60

gatgaccctt gcagagacag agtattagga catgtttgaa caggggccgg gcgatcagca 120

ggtagctcta gaggatcccc gtctgtctgc acatttcgta gagcgagtgt tccgatactc 180

taatctccct aggcaaggtt cataatttgtaggttactt attctccttt tgttgactaa 240

gtcaataatc agaatcagca ggtttggagt cagcttggca gggatcagca gcctgggttg 300

gaaggagggg gtataaaagc ccttcacca ggagaagccg tcacacagac tagccctaag 360

gtaagttggc gccgtttaag ggatgggttg ttggtggggt attaatgttt aattaccttt 420

tttacaggcc tgaactagcg gcgccaccgc caccatgctt aggggtccgg ggccccgggt 480

gctgctgctg gccgtccagt gcctggggac agcgggtgcc tccacgggag cctcgaagag 540

caagagggtt taccgcccc gtgagaccct gtgcggcggg gagctggtgg acaccctcca	600
gttcgtctgt ggggaccgcg gcttctactt cagcaggccc gcaagccgtg tgagccgtcg	660
cagccgtggc atcatggagg agtgctgttt ccgcagctgt gacctggccc tcctggagac	720
gtactgtgct acccccgcca agtccgaggg cgcgccggca cccccggcc gtcccagagc	780
agtgcaccaca cagtgcgacg tccccccaa cagccgttc gattgcgccc ctgacaaggc	840
catcaccag gaacagtgcg agggcccgcg ctgctgtac atccctgcaa agcaggggct	900
gcagggagcc cagatggggc agccctggtg ctcttccca ccagctacc ccagctacaa	960
gctggagaac ctgagctcct ctgaaatggg ctacacggcc accctgaccc gtaccacccc	1020
cacctcttc cccaaggaca tcctgaccct gcggtggac gtgatgatgg agactgagaa	1080
ccgcctccac ttcacgatca aagatccagc taacaggcgc tacgaggtgc ccttggagac	1140
ccgcgtgtc cacagccggg caccgtcccc actctacagc gtggagttct ccgaggagcc	1200
cttcgggggtg atcgtgcacc ggcagctgga cggccgctg ctgctgaaca cgacgtggc	1260
gccccgttc ttgcgacc agttccttca gctgtccacc tcgtgccct cgcagtatat	1320
cacaggcctc gccgagcacc tcagtcctt gatgtcagc accagctgga ccaggatcac	1380
cctgtggaac cgggacctg cgcacacgc cgtgcgaac ctctacgggt ctacccttt	1440
ctacctggcg ctggaggacg gcgggtcggc acacggggtg ttctgctaa acagcaatgc	1500
catggatgtg gtctgcagc cgagccctgc ccttagctgg aggtcgacag gtgggatcct	1560
ggatgtctac atcttctgg gccagagcc caagagcgtg gtgcagcagt acctggacgt	1620
tgtgggatac ccgttcagc cgccatactg gggcctgggc ttccacctgt gccgtgggg	1680
ctactctcc accgctatca cccgccaggt ggtggagaac atgaccaggg cccacttccc	1740
cctggacgtc caatggaacg acctggacta catggactcc cggagggact tcacgttcaa	1800
caaggatggc ttccgggact tccggccat ggtgcaggag ctgcaccagg gcggccggcg	1860
ctacatgatg atcgtggatc ctgccatcag cagctcgggc cctgccggga gctacaggcc	1920
ctacgacgag ggtctgcgga ggggggtttt catcaccaac gagaccggcc agccgtgat	1980
tggaaggta tggcccggtt ccactgcctt ccccgacttc accaacccca cagccctggc	2040
ctggtgggag gacatggtgg ctgagttcca tgaccaggtg ccttcgacg gcatgtggat	2100
tgacatgaac gagccttcca acttcacag aggtctctgag gacggctgcc ccaacaatga	2160
gctggagaac ccacctacg tgcttgggtt ggttgggggg accctccagg cggccaccat	2220
ctgtgcctcc agccaccagt ttctctccac aactacaac ctgcacaacc tctacgcct	2280
gaccgaagcc atcgctccc acagggcgct ggtgaaggct cgggggacac gccatttgt	2340

gatctccgc tcgacctttg ctggccacgg ccgatacgcc ggccactgga cgggggacgt	2400
gtggagctcc tgggagcagc tcgcctcttc cgtgccagaa atcctgcagt ttaacctgct	2460
gggggtgcct ctggtcgggg ccgacgtctg cggcttcctg ggcaacacct cagaggagct	2520
gtgtgtgcgc tggaccacgc tgggggcctt ctacccttc atgcggaacc acaacagcct	2580
gctcagtctg ccccaggagc cgtacagctt cagcgagccg gccagcagg ccatgaggaa	2640
ggccctcacc ctgcgtacg cactcctccc ccacctctac aactgttcc accaggccca	2700
cgtcgcgggg gagaccgtgg cccggcccct ctctctggag ttccccaagg actctagcac	2760
ctggactgtg gaccaccagc tctgtgggg ggaggccctg ctcatcacc cagtgtcca	2820
ggccgggaag gccgaagtga ctggctactt ccccttgggc acatggtacg acctgcagac	2880
ggtgccaata gaggcccttg gcagcctccc accccacact gcagctccc gtgagccagc	2940
catccacagc gaggggcagt gggtagcgt gccggccccc ctggacacca tcaacgtcca	3000
cctccgggct gggtagatca tccccctgca gggccctggc ctcaaacca cagagtccc	3060
ccagcagccc atggccctgg ctgtggccct gaccaagggt ggagaggccc gaggggagct	3120
gttctgggac gatggagaga gcctggaagt gctggagcga ggggcctaca cacaggtcat	3180
cttctggcc aggaataaca cgatcgtgaa tgagctggtg cgtgtgacca gtgaggagc	3240
tggcctgcag ctgcagaagg tgactgtctt gggcgtggcc acggcgcccc agcaggtcct	3300
ctccaacggt gtccctgtct ccaacttcac ctacagcccc gacaccaagg tcctggacat	3360
ctgtgtctcg ctgttgatgg gagagcagtt tctcgtcagc tgggtttagc gagcgccgc	3420
tcttagtagc agtatcgatc ccagcccact ttccccaat acgactacga gatctgtggc	3480
ttctagctgc ccgggtggca tccctgtgac cctccccag tgcctctctt ggccctggaa	3540
gttgccactc cagtgccac cagcctgtc ctaataaaat taagttgcat cattttgtct	3600
gactaggtgt ctttctataa tattatgggg tggagggggg tggtagggag caaggggcaa	3660
gttggggtcc ttcaacgtcg aactgtagca ctctggatag ctagctagcg ttttagata	3720
gtttcaagga tagcaaccaa caacacgata gtttagaat tgttctagat ctctgctac	3780
tactcgagag atagttgact tctagaagac tctggataga gagcttcctg tggtttctat	3840
ttcaacactt cgagcgattt agatagaggc gtcaattgag aacaaaaact ggacacctgt	3900
ttctagaaca ctagtgtagc tcttctcaag gcttgatacg tctgccact cacacacaaa	3960
aaaccaacac acagattaat gaaaataaag atcttttatt agagagcgag tggatagttt	4020
agagtcgctt gagacttacc ctgtgtccaa ctagccaaga gtagagagag agtgactagg	4080

tggatcgaga gatcgctcgtc tatttgactg ctagcctgtg ttctcgtcgt agtccaaatc	4140
ttcaatactc tggcgactag taacttgaat ctgtgtatct aacacgacaa ggctactcct	4200
ttgagtttct agtgataagt gcgactggat ctagctagtt gagtttctag aaactacgtc	4260
caaagctact agcacta	4277
<210> 82	
<211> 4238	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 82	
tcgagcttgg gctgcaggtc gagggcactg ggaggatgtt gagtaagatg gaaaactact	60
gatgaccctt gcagagacag agtattagga catgtttgaa caggggccgg gcgatcagca	120
ggtagctcta gaggatcccc gtctgtctgc acatttcgta gagcgagtgt tccgatactc	180
taatctccct aggcaaggtt catatttggt taggttactt attctccttt tgttgactaa	240
gtcaataatc agaatcagca ggtttggagt cagcttggca gggatcagca gcctgggttg	300
gaaggagggg gtataaaagc cccctcacca ggagaagccg tcacacagac tagccctaag	360
glaagttggc gccgtttaag ggatgggttg ttggtagggg attaagtgtt aattaccttt	420
tttacaggcc tgaactaggc gcgccaccgc caccatgccg tcttctgtct cgtggggcat	480
cctcctgtctg gcaggcctgt gctgcctggt cctgtctcc ctggctgctc tgtgcggcgg	540
ggagctggtg gacaccctcc agttcgtctg tggggaccgc ggcttctact tcagcaggcc	600
cgcaagccgt gtgagccgtc gcagccgtgg catcgttgag gattgctgtt tccgcagctg	660
tgacctggcc ctcttgaga cgtactgtgc taccctcgcc aagtcgagg gcgcgccggc	720
acaccccgcc cgtcccagag cagtgcacac acagtgcgac gtcccccca acagccgctt	780
cgattgcgcc cctgacaagg ccatcaccca ggaacagtgc gagggccgcg gctgctgcta	840
catccctgca aagcaggggc tgcagggagc ccagatgggg cagccctggt gcttcttccc	900
accagctac ccagctaca agctggagaa cctgagctcc tctgaaatgg gctacacggc	960
cacctgacc cgtaccaccc ccaccttctt cccaaggac atcctgacct tgcggtgga	1020
cgtgatgatg gagactgaga accgcctcca cttcacgac aaagatccag ctaacaggcg	1080
ctacagggtg ccttggaga ccccgctgt ccacagccgg gcaccgtccc cactctacag	1140
cgtggagtgc tccgaggagc ccttcggggg gatcgtgcac cggcagctgg acggccgct	1200
gctgctgaac acgacggtgg cgccctgtt ctttcgggac cagttccttc agctgtccac	1260

ctcgtgccc tcgcagtata tcacaggcct cgccgagcac ctcagtcgcc tgatgctcag 1320

caccagctgg accaggatca ccctgtggaa cggggacctt gcgcccacgc ccggtgcgaa 1380

cctctacggg tctcaccctt tctacctggc gctggaggac ggcgggctcg cacacggggt 1440

gttcctgcta aacagcaatg ccatggatgt ggtcctgcag ccgagccctg cccttagctg 1500

gaggtcgaca ggtgggatac tggatgtcta catcttctg ggcccagagc ccaagagcgt 1560

ggtgcagcag tacctggacg ttgtgggata ccggttcacg ccgccatact ggggcctggg 1620

cttcacactg tgccgctggg gctactctc caccgctatc acccgccagg tgggtggagaa 1680

catgaccagg gcccaacttc ccctggacgt ccaatggaac gacctggact acatggactc 1740

ccggaggagc ttacagtca acaaggatgg ctccgggac ttcccgcca tgggtcagga 1800

gctgcaccag ggcgcccgcc gctacatgat gatcgtggat cctgccatca gcagctcggg 1860

ccctgccggg agctacaggc cctacgacga gggtctcgcc aggggggttt tcatcacaa 1920

cgagaccggc cagccgctga ttgggaaggt atggcccgcc tccactgcct tccccgactt 1980

caccaacccc acagccctgg cctggtggga ggacatggtg gctgagttcc atgaccaggt 2040

gcccttcgac ggcatgtgga ttgacatgaa cgagccttc aactcatca gaggtctga 2100

ggacggctgc ccaacaatg agctggagaa ccacacctac gtgcctgggg tggttggggg 2160

gacctccag gcggccacca tctgtgctc cagccaccag tttctctca cactacaa 2220

cctgcacaac ctctacggc tgaccgaagc catcgctcc cacagggcgc tggtaaggc 2280

tcgggggaca cgccatttg tgatctccg ctgaccttt gctggccacg gccgatacgc 2340

cggccactgg acgggggacg tgtggagctc ctgggagcag ctgcctcct ccgtgccaga 2400

aatctgcag tttaacctgc tgggggtgcc tctggtcggg gccgacgtct gcggcttctt 2460

gggcaacacc tcagaggagc tgtgtgtgcg ctggaccacg ctgggggcct tctaccctt 2520

catgcggaac cacaacagcc tgctcagctt gcccaggag ccgtacagct tcagcgagcc 2580

ggcccagcag gccatgagga aggcctcac cctgcgtac gactcctcc cccactcta 2640

cactgttc caccagccc acgtcgccgg ggagaccgtg gcccgcccc tcttctgga 2700

gttcccaag gactctagca cctggactgt ggaccaccag ctctgtggg gggaggccct 2760

gctcatcacc ccagtgtcc aggcgggaa ggccgaagt actggctact tcccttggg 2820

catatgtac gacctgcaga ccgtgccaat agaggccctt ggagcctcc cccccacc 2880

tgcagctccc cgtgagccag ccatccacag cgaggggcag tgggtgacgc tgccggcccc 2940

cctggacacc atcaacgtc acctccgggc tgggtacatc atccccctgc agggccctgg 3000

cctcacaacc acagagtccc gccagcagcc catggccctg gctgtggccc tgaccaaggg	3060
tggagaggcc cgaggggagc tgttctggga cgtaggagag agcctggaag tgctggagcg	3120
aggggcctac acacaggcca tcttcctggc caggaataac acgacgtga atgagctggt	3180
acgtgtgacc agtgaggagc ctggcctgca gctgcagaag gtgactgtcc tgggcgtggc	3240
cacggcgccc cagcaggacc tctccaacgg tgcctctgtc tccaacttca cctacagccc	3300
cgacaccaag gtccctggaca tctgtgtctc gctgttgatg ggagagcagt ttctcgtcag	3360
ctggtgttag cgagcggccg ctcttagtag cagtatcgat cccagcccac ttttcccaa	3420
tacgactacg agatctgtgg cttctagctg cccgggtggc atccctgtga cccctccca	3480
gtgcctctcc tggccctgga agttgccact ccagtgccca ccagccttgt cctaataaaa	3540
ttaagttgca tcatTTTgtc tgactagggt tccttctata atattatggg gtggaggggg	3600
gtggtatgga gcaaggggca agttgggggt cttcaacgtc gaactgtagc actctggata	3660
gctagctagc gttttagat agtttcaagg atagcaacca acaacacgat agtttagaaa	3720
ttgttctaga tcttctgcta ctactcgaga gatagttgac ttctagaaga ctctggatag	3780
agagcttctc gtggtttcta tttcaacact tcgagcgatt tagatagagg cgtcaattga	3840
gaacaaaaac tggacacctg tttctagaac actagtgtca gtcttctcaa ggcttgatac	3900
gtcctgccac tcacacaca aaaaccaaca cacagattaa tgaaaataaa gatcttttat	3960
tagagagcga gtggatagtt tagagtcgct tgagacttat cctgtgtcca actagccaag	4020
agtagagaga gagtgttag gtggatcgag agatcgtcgt ctatttgact gctagcctgt	4080
gttctcgtcg tagtccaaat cttcaatact ctggcgacta gtaacttgaa tctgtgtatc	4140
taacacgaca aggtactacc tttgagtttc tagtgataag tgcgactgga tctagctagt	4200
tgagtttcta gaaactacgt ccaaagctac tagcacta	4238

<210> 83

<211> 4262

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 83

tcgagcttgg gctgcaggtc gagggcactg ggaggatgtt gagtaagatg gaaaactact	60
gatgaccctt gcagagacag agtattagga catgtttgaa caggggccgg gcgacagca	120
ggtagctcta gaggatcccc gtctgtctgc acatttcgta gagcgagtgt tccgatactc	180
taatctccct aggcaaggtt catatttgtg taggttactt attctctttt tgttgactaa	240

gtcaataatc agaatacagca ggtttggagt cagcttggca gggatcagca gcctgggttg 300

gaaggagggg gtataaaagc cctttcacca ggagaagccg tcacacagac tagccctaag 360

gtaagtggc gccgtttaag ggatggttgg ttggtggggg attaatgttt aattaccttt 420

tttacaggcc tgaactagc gcgccaccgc caccatgctc aggggtccgg gaccggggcg 480

gctgctgctg ctagcagtc tgtgcttggg gacatcggtg cgctgcaccg aaaccgggaa 540

gagcaagagg gctctgtgcg gcggggagct ggtggacacc ctccagttcg tctgtgggga 600

ccgcggcttc tacttcagca ggcccgaag ccgtgtgagc cgctgcagcc gtggcatcgt 660

tgaggagtgc tgtttccgca gctgtgacct ggccctcctg gagacgtact gtgtacccc 720

cgccaagtcc gagggcgcg cggcacacc cggccgtccc agagcagtc ccacacagtg 780

cgacgtccc cccaacagcc gcttcgattg cggccctgac aaggccatca cccaggaaca 840

gtgcgaggcc cgcggtgct gctacatccc tgcaaagcag gggtgcagg gagccagat 900

ggggcagccc tgggtcttct tcccaccag ctacccagc tacaagctgg agaacctgag 960

ctcctctgaa atgggtaca cggccacct gaccgtacc accccacct tcttcccaa 1020

ggacatcctg acctgcgcg tggacgtgat gatggagact gagaaccgcc tccattcac 1080

gatcaaagat ccagctaaca ggcgtacga ggtgcccttg gagacccgc gtgtccacag 1140

ccgggcaccg tccccactct acagcgtgga gttctccgag gagcccttcg gggtgatcgt 1200

gcaccggcag ctggacggc gcgtgctgct gaacacgacg gtggcgcccc tgttctttgc 1260

ggaccagttc cttcagctgt ccacctcgt gccctcgcag tatacacag gcctcgccga 1320

gcacctcagt ccctgatgc tcagaccag ctggaccagg atcacctgt ggaaccggga 1380

ccttgcgcc acgcccgtg cgaacctcta cgggtctcac ctttctacc tggcgctgga 1440

ggacggcggg tcggcacag ggtgttctt gctaaacagc aatgcatgg atgtgtcct 1500

gcagccgagc cctgcctta gctggaggtc gacaggtggg atcctggatg tctacatctt 1560

cctgggcccc gagcccaaga gcgtggtgca gcagtacctg gacgttgtgg gataccggtt 1620

catgccgcca tactggggcc tgggttcca cctgtgccg tggggctact cctccaccg 1680

tatccccgc caggtggtgg agaacatgac cagggccac tccccctgg acgtccaatg 1740

gaacgacctg gactacatgg actcccgag ggacttcacg ttcaacaagg atggcttccg 1800

ggacttccc gccatgtgc aggagctgca ccaggcggc cggcgctaca tgatgatcgt 1860

ggatcctgcc atcagcagct cgggccctgc cgggagctac aggcctacg acgagggtct 1920

gcggaggggg gttttcatca ccaacgagac cggccagccg ctgattggga aggtatggcc 1980

cgggtccact gccttccccg acttcaccaa cccacagcc ctggcctggt gggaggacat	2040
ggtggctgag ttccatgacc aggtgccctt cgacggcatg tggattgaca tgaacagacc	2100
ttccaacttc atcagaggct ctgaggacgg ctgccccaac aatgagctgg agaaccacc	2160
ctacgtgcct ggggtggttg gggggaccct ccaggcggcc accatctgtg cctccagcca	2220
ccagtttctc tccacacact acaacctgca caacctctac ggctgaccg aagccatcgc	2280
ctccacaggg gcgttggtga aggctcgggg gacacgcca tttgtgatct cccgctcgac	2340
ctttgctggc cacggccgat acgccggcca ctggacgggg gacgtgtgga gctcctggga	2400
gcagctcgcc tcttccgtgc cagaaatcct gcagtttaac ctgctggggg tgcctctggt	2460
cggggccgac gtctcgggct tcctgggcaa cacctcagag gagctgtgtg tgcgctggac	2520
ccagctgggg gccttctacc cttcatgcg gaaccacaac agcctgtca gtctgcccc	2580
ggagccgtac agcttcagcg agccgggcca gcaggccatg aggaaggccc tcacctgcg	2640
ctacgcactc ctccccacc tctacacact gttccaccag gcccacgtcg cgggggagac	2700
cgtggcccgg cccctcttcc tggagtccc caaggactct agcacctgga ctgtggacca	2760
ccagctctg tggggggagg ccctgctcat caccacagt ctccaggccg ggaaggccga	2820
agtgactggc tacttccct tgggcacatg gtacgacctg cagacggtgc caatagaggc	2880
ccttggcagc ctccccccc cacctgcagc tccccgtgag ccagccatcc acagcgaggg	2940
gcagtgggtg acgtgcggg cccccctgga caccatcaac gtccacctcc gggctgggta	3000
catcatcccc ctgcagggcc ctggcctcac aaccacagag tcccgcagc agcccatggc	3060
cctggctgtg gccctgacca aggggtggaga ggcccaggag gagctgttct gggacgatgg	3120
agagagcctg gaagtgtggt agcgaggggc ctacacacag gtcatcttcc tggccaggaa	3180
taacacgac gtgaatgagc tggtagctgt gaccagttag ggagctggcc tgcagctgca	3240
gaaggtagct gtcttgggg tggccacggc gccccagcag gtctctcca acggtgtccc	3300
tgtctccaac ttacactaca gccccgacac caaggtcctg gacatctgtg tctcgctgtt	3360
gatgggagag cagtttctcg tcagctggtg ttagcgagcg gccgtctta gtagcagtat	3420
cgatcccagc ccacttttcc ccaatacgac tacgagatct gtggcttcta gctgccggg	3480
tggcatccct gtgaccttc cccagtgcct ctctggccc tggaagtgtg cactccagtg	3540
cccaccagcc ttgtctaat aaaattaagt tgcattttt tgtctgacta ggtgtccttc	3600
tataatatta tggggtggag gggggtggta tggagcaagg ggcaagtgg ggtccttcaa	3660
cgtcgaactg tagcactctg gatagctagc tagcgtttgt agatagtctc aaggatagca	3720
accaacaaca cgatagttaa gaaattgttc tagatcttct gctactactc gagagatagt	3780
tgacttctag aagactctgg atagagagct tcctgtggtt tctatttcaa cacttcgagc	3840

gatttagata gaggcgtcaa ttgagaacaa aaactggaca cctgtttcta gaacactagt 3900

gctagtcttc tcaaggcttg atacgtcctg ccactcacac acaaaaaacc aacacacaga 3960

ttaatgaaaa taaagatctt ttattagaga gcgagtggat agtttagagt cgcttgagac 4020

ttatcctgtg tccaactagc caagagtaga gagagagtga ctaggtggat cgagagatcg 4080

tcgtctatth gactgctagc ctgtgttctc gtcgtagtcc aaatcttcaa tactctggcg 4140

actagtaact tgaatctgtg tatctaacac gacaaggcta ctctttgag tttctagtga 4200

taagtgcgac tggatctagc tagttgagtt tctagaaact acgtccaaag ctactagcac 4260

ta 4262

<210> 84

<211> 4259

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 84

tcgagcttgg gctgcaggtc gagggcactg ggaggatgtt gagtaagatg gaaaactact 60

gatgaccctt gcagagacag agtattagga catgtttgaa caggggccgg gcgatcagca 120

ggtagctcta gaggatcccc gtctgtctgc acatttcgta gagcgagtgt tccgatactc 180

taatctccct aggcaaggtt catatttgtg taggttactt attctcctt tgttgactaa 240

gtcaataatc agaatacaga ggtttggagt cagcttggca gggatcagca gcctgggttg 300

gaaggagggg gtataaaaag cccctcacca ggagaagccg tcacacagac tagccctaag 360

gtaagtggc gccgtttaag ggatggttgg ttggtggggg attaatgtt aattaccttt 420

tttacaggcc tgaactaggc gcgccaccgc caccatgctt aggggtccgg ggccccgggt 480

gctgtctgtg gccgtccagt gcctggggac agcgggtgcc tccacgggag cctcgaagag 540

caagagggct ctgtgcggcg gggagctggt ggacaccctc cagttcgtct gtggggaccg 600

cggcttctac ttcagcaggc ccgcaagccg tgtgagccgt cgcagccgtg gcatcgttga 660

ggagtgtgtt ttccgcagct gtgacctggc cctcctggag acgtactgtg ctacccccgc 720

caagtccgag ggcgcgcggg cacaccccg cgtcccaga gcagtgccca cacagtgcga 780

cgtcccccc aacagccgct tcgattgcgc cctgacaag gccatcacc aggaacagtg 840

cgaggccgcg ggctgtgtgt acatccctgc aaagcagggg ctgcagggag cccagatggg 900

gcagccctgg tgcttcttcc caccagctc cccagctac aagctggaga acctgagctc 960

ctctgaaatg ggctacacgg ccacctgac ccgtaccacc cccacattct tcccaagga	1020
catctgacc ctgcggtcgg acgtgatgat ggagactgag aaccgcctcc acttcacgat	1080
caaagatcca gctaacaggg gctacgaggt gcccttggag accccgcgtg tccacagccg	1140
ggcaccgctc ccactctaca gcgtggagtt ctccgaggag cccttcgggg tgatcgtgca	1200
ccggcagctg gacggccgcg tgctgctgaa cagcaggtg gcgcccctgt tctttgcgga	1260
ccagttcctt cagctgtcca cctcgtgcc ctgcagtat atcacaggcc tcgccgagca	1320
cctcagtccc ctgatgctca gcaccagctg gaccaggatc accctgtgga accgggacct	1380
tgcgcccacg cccggtgcga acctctacgg gtctcacctt ttctacctgg cgctggagga	1440
cggcgggtcg gcacacgggg tggttctgct aaacagcaat gccatggatg tggctcctgca	1500
gccgagccct gcccttagct ggaggtcgac aggtgggatc ctggatgtct acatcttctt	1560
gggcccagag cccaagagcg tgggtcagca gtacctggac gttgtgggat acccgttcat	1620
gccgccatac tggggcctgg gcttcacct gtcccgtgg ggctactcct ccaccgtat	1680
caccgccag gtggtggaga acatgaccag ggcccacttc cccctggacg tccaatggaa	1740
cgacctggac tacatggact cccggaggga cttcacgttc aacaaggatg gcttcggga	1800
cttcccgccc atggtgcagg agctgcacca gggcggccgg cgctacatga tgatcgtgga	1860
tcctgccatc agcagctcgg gccctgccgg gagctacagg ccctacgacg agggctctcg	1920
gaggggggtt ttcatcacca acgagaccgg ccagccgctg attgggaagg tatggcccgg	1980
gtccactgcc ttccccact tcaccaaccc cacagccctg gcctgggtggg aggacatggt	2040
ggctgagttc catgaccagg tgcccttcga cggcatgtgg attgacatga acgagccttc	2100
caacttcac agaggctctg aggacggctg cccaacaat gagctggaga acccaccta	2160
cgtgcctggg gtggttgggg ggacctcca ggcgccacc atctgtgcct ccagccacca	2220
gtttctctcc acacactaca acctgcacaa cctctacggc ctgaccgaag ccatcgctc	2280
ccacaggcgg ctggtgaagg ctggggggac acgcccattt gtgatctccc gctcgacctt	2340
tgctggccac ggccgatacg ccggccactg gacgggggac gtgtggagct cctgggagca	2400
gctcgctcc tccgtgccag aaatcctgca gtttaacctg ctgggggtgc ctctggtcgg	2460
ggccgacgtc tgcggttcc tgggcaacac ctgagaggag ctgtgtgtgc gctggacca	2520
gctgggggcc ttctaccctt tcatcgga ccacaacagc ctgctcagtc tgccccagga	2580
gccgtacagc ttcagcgagc cggcccagca ggccatgagg aaggccctca cctgcgcta	2640
cgcactctc cccacctct acacactgtt ccaccaggcc cacgtcgcgg gggagaccgt	2700
ggccggccc ctcttctgg agttcccaaa ggactctagc acctggactg tggaccacca	2760
gctcctgtgg ggggaggccc tgctcatcac ccagtgctc caggccggga aggccgaagt	2820

gactggctac ttcccccttg gcacatggta cgacctgcag acggtgccaa tagaggccct 2880
 tggcagcctc ccacccccac ctgcagctcc ccgtgagcca gccatccaca gcgaggggca 2940
 gtgggtgacg ctgccggccc ccttgacac catcaacgtc cacctccggg ctgggtacat 3000

catccccctg cagggccctg gcctcacaac cacagagtcc cgccagcagc ccatggccct 3060
 ggctgtggcc ctgaccaagg gtggagaggc ccgaggggag ctgttctggg acgatggaga 3120
 gagcctggaa gtgctggagc gaggggccta cacacaggtc atcttcttg ccaggaataa 3180
 cacgatcgtg aatgagctgg tacgtgtgac cagttaggga gctggcctgc agctgcagaa 3240
 ggtgactgtc ctgggcgtgg ccacggcgcc ccagcaggtc ctctccaacg gtgtccctgt 3300
 ctccaacttc acctacagcc ccgacaccaa ggtcctggac atctgtgtct cgctgttgat 3360
 gggagagcag ttctctgtca gctggtgtta gcgagcggcc gctcttagta gcagtatcga 3420

tcccagccca cttttccca atacgactac gagatctgtg gcttctagct gcccgggtgg 3480
 catccctgtg acccctccc agtgctctc ctggccctgg aagttgccac tccagtcccc 3540
 accagccttg tcttaataaa attaatgtgc atcatctgt ctgactaggt gtcttctat 3600
 aatattatgg ggtggagggg ggtggtatgg agcaaggggc aagttgggt cttcaacgt 3660
 cgaactgtag cactctggat agctagctag cgttttaga tagttcaag gatagcaacc 3720
 aacaacacga tagtttagaa attgttctag atcttctgct actactcgag agatagtga 3780
 ctctagaag actctggata gagagcttcc tgtggtttct attcaacac ttcgagcgat 3840

ttagatagag gcgtcaattg agaacaaaa ctggacacct gtttctagaa cactagtgt 3900
 agtcttctca aggttgata cgtctgccca ctcacacaca aaaaaccaac acacagatta 3960
 atgaaaataa agatctttta ttagagagcg agtggatagt ttagagtcgc ttgagactta 4020
 tctgtgtcc aactagccaa gagtagagag agagtgacta ggtggatcga gagatcgtcg 4080
 tctatttgac tgetagcctg tgttctgtc gtagtccaaa tcttcaatac tctggcgact 4140
 agtaacttga atctgtgtat ctaacacgac aaggctactc ctttgagttt ctagtataa 4200
 gtgcgactgg atctagctag ttgagtttct agaaactacg tccaaagcta ctgaccta 4259

<210> 85

<211> 165

<212> DNA

<213> Homo sapiens

<400> 85

ccgggcggag tgtgttagtc tctccagagg gaggtggtt cccagggaa gcagacctg 60
 tgtcgggca gcagctgtgt gcggccctgg gggtgttaa gtgcaattat ttttaataaa 120

aggggcattt ggaaaaaaa aaaaaaggta gcagtcgaca gatga 165

<210> 86

<211> 180

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 86

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60

tgtttctct ggtaataat ctcaggagca caaacattcc cctgttcaa acatgtccta 120

atactctgtc tctgcaagg tcatcagtag ttttccatct tactcaacat cctcccagtg 180

<210> 87

<211> 402

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 87

caggctttca ctttctgcc aacttacaag gcctttctgt gtaaacaata cctgaacctt 60

tacccegttg cccgcaacg gccaggtctg tgccaagtgt ttgaggttaa tttttaaaaa 120

gcagtcaaaa gtccaagtgg cccctggcag catttactct ctctgtttgc tctggttaat 180

aatctcagga gcacaaacat tccggcccgg gaggcgcct ttggaccttt tgcaatcctg 240

gcgcactgaa ccttgacct ctgccctgca gccccgcag cttgtctgtt gccactcta 300

tttgcccagc ccagccctg gagagtcctt tagcaggga aagtgaaca taggcagacc 360

ttaagggatg actcagtaac agataagctt tgtgtgcctg ca 402

<210> 88

<211> 305

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 88

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60

tgtttctct ggtaataat ctgaggagca caaacattcc tgtaccggcc cgggaggcgc 120

cctttggacc ttttgaatc ctggcgact gaaccttga cccctgccct gcagcccccgc 180

cagcttgctg ttgcccact ctatttgccc agcccagcc ctggagagtc ctttagcagg 240

gcaaagtga acataggcag accttaaggg atgactcagt aacagataag ctttgtgtgc 300

ctgca 305

<210> 89

<211> 103

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 89

caggttttca ctttctgcc aacttacaag gcctttctgt gtaaacaata cctgaacctt 60

taccctgttg cccggcaacg gccaggtctg tgccaagtgt ttg 103

<210> 90

<211> 212

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 90

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60

tgtttctct ggtaataat ctgaggagca caaacattcc ggcccgggag gcgccctttg 120

gaccttttg aatcctggcg cactgaacc ttgaccttg cctgcagcc cccgcagctt 180

gctgtttgcc cactctattt gccagcccc ag 212

<210> 91

<211> 283

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 91

aggTTAATTT tTAAAAAGCA gTCAAAAGTC CAAGTGGCCC TTGGCAGCAT TTACTCTCTC	60
TGTTTGCTCT GGTTAATAAT CTCAGGAGCA CAAACATTCC CCCTGTTCAA ACATGTCCTA	120
ATACTCTGTC TCTGCAAGGG TCAATCAGTAG TTTTCCATCT TACTCAACAT CCTCCCAGTG	180
AGTCATATGT TTGCTCACTG AAGGTTACTA GTTAACAGGC ATCCCTTAAA CAGGATATAA	240

aaggacttca gcaggactgc tcgaaacatc ccactcagcc acc	283

<210> 92

<211> 565

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 92

aggTTAATTT tTAAAAAGCA gTCAAAAGTC CAAGTGGCCC TTGGCAGCAT TTACTCTCTC	60
TGTTTGCTCT GGTTAATAAT CTCAGGAGCA CAAACATTCC CCCTGTTCAA ACATGTCCTA	120
ATACTCTGTC TCTGCAAGGG TCAATCAGTAG TTTTCCATCT TACTCAACAT CCTCCCAGTG	180
AGTCATATGT TTGCTCACTG AAGGTTACTA GTTAACAGGC ATCCCTTAAA CAGGATATAA	240

aaggacttca gcaggactgc tcgaaacatc ccacttcaga tcgcctggag acgcatcca	300
cgctgttttg acctccatag aagacaccgg gaccgatcca gcctccgcgg ccgggaacgg	360
tgcatggaa cgcggattcc ccgtgccaag agtgacgtaa gtaccgccta tagactctat	420
aggcacacc ctttggctct tatgcatgaa cggtagggagg cagtgtagtc tgagcagtac	480
tcgttctgc cgcgcgcgcc accagacata atagctgaca gactaacaga ctgttccttt	540
ccatgggtct tttctgcagg ccacc	565

<210> 93

<211> 478

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 93

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt	60
taccccgttg cccggcaacg gccaggtctg tgccaagtgt ttgagggttaa tttttaaaaa	120

gcagtcaaaa gtccaagtgg cccttggcag cttttactct ctctgtttgc tctggttaat 180
aatctcagga gcacaaacat tccggcccgg gaggcgcct ttggaccttt tgcaatcctg 240
gcgcactgaa cccttgaccc ctgccctgca gccccgcag cttgctgttt gccactcta 300
tttgcaccagc cccagccctg gagagtcctt tagcaggga aagtgaaca taggcagacc 360

ttaagggatg actcagtaac agataagctt tgtgtgcctg cagggcatat aaaacagggg 420
caaggcacag actcatagca gagcaatcac caccaagcct ggaataactg cagccacc 478

<210> 94

<211> 381

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 94

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
tgtttgctct ggtaataat ctccaggagca caaacattcc tgtaccggcc cgggaggcgc 120
cctttggacc ttttgcaatc ctggcgcact gaacccttga cccctgcct gcagcccccg 180

cagcttgctg ttgcccact ctatttgcgc agccccagcc ctggagagtc ctttagcagg 240
gcaaagtga acataggcag accttaaggg atgactcagt aacagataag ctttgtgtgc 300
ctgcagggca tataaaacag gggcaaggca cagactcata gcagagcaat caccaccaag 360
cctggaataa ctgcagccac c 381

<210> 95

<211> 379

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 95

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60

taccctgttg cccggcaacg gccaggtctg tgccaagtgt ttgccgatga cctaatgatt 120
ctgagcttgg caaaggtctt atctcccagc tcgcccaggc ccagtgttcc aggaatgtga 180
cctttgctgc agcagccgct ggagggggca gaggggatgg gctggaggtt gagcaaacag 240
agcagcagaa aaggcagttc ctcttctcca gtccctcct tccctgtctc tgccctctcc 300

tcccttcctc aggcacatcaga gcggagactt caggagagacc agagcccagc ttgccaggca 360
ctgagctaga agccctgcc 379
<210> 96
<211> 288
<212> DNA
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 96
agggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
tgtttgctct ggtaataat ctcaggagca caaacattcc ggcccgggag gcgccctttg 120
gaccttttgc aatcctggcg cactgaacct ttgaccttg ccctgcagcc cccgcagctt 180
gctgtttgcc cactctatct gccagcccc aggggcatat aaaacagggg caaggcacag 240
actcatagca gagcaatcac caccaagcct ggaataactg cagccacc 288
<210> 97
<211> 100
<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 97
agggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
tgtttgctct ggtaataat ctcaggagca caaacattcc 100
<210> 98
<211> 13
<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 98
gttaattttt aaa 13
<210>
99

<211> 11

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 99

gtggcccttg g 11

<210> 100

<211> 7

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 100

tgtttgc 7

<210> 101

<211> 15

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 101

tggttaataa tctca 15

<210> 102

<211> 7

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 102

acaaaca 7

<210> 103

<211> 124

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<220><221> modified_base

<222> (14)..(39)

<223> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (14)..(39)

<223> This region may encompass 10 to 26 nucleotides

<220><221> modified_base

<222> (51)..(72)

<223> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (51)..(72)

<223> This region may encompass 8 to 22 nucleotides

<220><221> modified_base

<222> (80)..(89)

<223> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (80)..(89)

<223> This region may encompass 1 to 10 nucleotides

<220><221> modified_base

<222> (105)..(117)

<223> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (105)..(117)

<223> This region may encompass 1 to 13 nucleotides

<220>

<223> See specification as filed for detailed description of

substitutions and preferred embodiments

<400> 103

gttaattttt aaannnnnnn nnnnnnnnnn nnnnnnnnng tggcccttgg nnnnnnnnnn

60

nnnnnnnnnn nntgtttgcn nnnnnnnnnt ggtaataat ctcannnnnn nnnnnnnaca 120

aaca 124

<210> 104

<211> 80

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 104

ccctgttcaa acatgtccta atactctgtc tctgcaaggg tcatcagtag ttttccatct 60

tactcaacat cctcccagtg 80

<210> 105

<211> 87

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220><221> modified_base

<222> (12)..(21)

<223> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (12)..(21)

<223> This region may encompass 1 to 10 nucleotides

<220><221> modified_base

<222> (33)..(42)

<223> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (33)..(42)

<223> This region may encompass 1 to 10 nucleotides

<220><221> modified_base

<222> (55)..(77)

<223> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (55)..(77)

<223> This region may encompass 8 to 23 nucleotides

<220>

<223> See specification as filed for detailed description of
substitutions and preferred embodiments

<400> 105

gttcaaacat gnnnnnnnnn nctaatactc tgnnnnnnnn nntgcaaggg tcatnnnnnn 60

nnnnnnnnnn nnnnnnntta ctcaaca 87

<210> 106

<211> 11

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 106

gttcaaacat g 11

<210> 107

<211> 11

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 107

ctaatactct g 11

<210> 108

<211> 12

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 108

tgcaagggtc at 12

<210> 109

<211> 10

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 109

ttactcaaca 10

<210> 110

<211> 95

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 110

agtcatatgt ttgtcactg aaggttacta gttaacaggc atcccttaaa caggatataa 60

aaggacttca gcaggactgc tcgaaacatc ccact 95

<210> 111

<211> 13

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 111

gttactagtt aac 13

<210> 112

<211> 15

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 112

gctcgaaaca tccca 15

<210> 113

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 113

ggccccgggag gcgccctttg gaccttttgc aatcctggcg

40

<210> 114

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 114

gcgccctttg gaccttttgc aatcctgg

28

<210> 115

<211> 13

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 115

cgccctttgg acc

13

<210> 116

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 116

gaccttttgc aatcctgg

18

<210> 117

<211> 72

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 117

cactgaaccc ttgaccctg ccctgcagcc cccgcagctt gctgtttgcc cactctattt 60

gcccgagccc ag 72

<210> 118

<211> 77

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<220><221> modified_base

<222> (19)..(48)

<223> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (19)..(48)

<223> This region may encompass 14 to 30 nucleotides

<220><221> modified_base

<222> (59)..(68)

<223> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (59)..(68)

<223> This region may encompass 1 to 10 nucleotides

<220

>

<223> See specification as filed for detailed description of

substitutions and preferred embodiments

<400> 118

actgaaccct tgaccctnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnct gtttgccnn 60

nnnnnnnnta ttgccc 77

<210> 119

<211> 23

<212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 119
 actgaaccct tgaccctgc cct 23

<210> 120
 <211> 10
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 120
 ctgtttgccc 10

<210> 121
 <211> 10
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 121
 ctatttgccc 10

<210> 122
 <211> 87
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 122
 ccctggagag tccttttagca gggcaaagtg caacataggc agaccttaag ggatgactca 60
 gtaacagata agcttttgtgt gcctgca 87

<210> 123
 <211> 53

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<220><221> modified_base

<222> (16)..(43)

<223> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (16)..(43)

<223> This region may encompass 10 to 28 nucleotides

<220>

<223> See specification as filed for detailed description of

substitutions and preferred embodiments

<400> 123

gcagggcaaa gtgcannnnn nnnnnnnnnn nnnnnnnnnn nnngatgact cag

53

<210> 124

<211> 15

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 124

gcagggcaaa gtgca

15

<210> 125

<211>

> 10

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 125

gatgactcag

10

<210> 126

<211> 76

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 126

gggcatataa aacaggggca aggcacagac tcatagcaga gcaatcacca ccaagcctgg 60

aataactgca gccacc 76

<210> 127

<211> 103

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 127

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60

taccccgttg cccggcaacg gccaggtctg tgccaagtgt ttg 103

<210> 128

<211> 129

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<220><221> modified_base

<222> (19)..(38)

<223> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (19)..(38)

<223> This region may encompass 10 to 20 nucleotides

<220><221> modified_base

<222> (52)..(61)

<223> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (52)..(61)

<223> This region may encompass 1 to 10 nucleotides

<220><221> modified_base

<222> (76)..(85)

<223> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (76)..(85)

<223> This region may encompass 1 to 10 nucleotides

<220><221> modified_base

<222> (99)..(108)

<223>

> a, c, t, g, unknown or other

<220><221> misc_feature

<222> (99)..(108)

<223> This region may encompass 1 to 10 nucleotides

<220>

<223> See specification as filed for detailed description of
substitutions and preferred embodiments

<400> 128

ctttcacttt ctgccaann nnnnnnnnnn nnnnnnnntg tgtaaacaat annnnnnnnn	60
nctgaacctt taccnnnnn nnnnngttgc cggcaacnn nnnnnnnnca ggtctgtgcc	120
aagtgtttg	129

<210> 129

<211>

11

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 129

ctttcacttt c	11

<210> 130

<211> 7

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide	
<400> 130	
tcgccaa	7
<210> 131	
<211> 10	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	
<400> 131	
tgtgtaaaca	10
<210> 132	
<211> 11	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	
<400> 132	
tgtaacaat a	11
<210> 133	
<211> 14	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	
<400> 133	
ctgaaccttt accc	14
<210> 134	
<211> 13	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	

<400> 134	
gttgcccggc aac	13
<210> 135	
<211> 15	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	
<400> 135	
caggtctgtg ccaag	15
<210> 136	
<211> 13	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	
<400> 136	
tgccaagtgt ttg	13
<210> 137	
<211> 276	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 137	
ccgatgacct aatgattctg agcttggcaa aggtcttata tcccagctcg cccaggccca	60
gtgttcagg aatgtgacct ttgtgcagc agccgctgga gggggcagag gggatgggct	120
ggaggttgag caaacagagc agcagaaaag gcagttcttc ttctccagtg ccttccttc	180
ctgtctctgc ctctccctcc ctctctcagg catcagagcg gagacttcag ggagaccaga	240
gcccagcttg ccaggcactg agctagaagc cctgcc	276
<210> 138	
<211> 13	
<212> DNA	

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 138

tggcaaaggt ctt 13

<210> 139

<211> 11

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 139

tgtgaccttt g 11

<210> 140

<211> 13

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 140

gtgacctttg ctg 13

<210> 141

<211> 14

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 141

aggttgagca aaca 14

<210> 142

<211> 9

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 142

agcaaacag

9

<210> 143

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 143

ggagaggcag agacaggga ggaggcact g

31

<210> 144

<211> 275

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 144

ggccccggag gcgcccttg gacctttgc aatcctggcg cactgaaccc ttgaccttg 60

ccctgcagcc cccgcagctt gctgtttgcc cactctatct gccagcccc agccctggag 120

agtcctttag cagggcaaag tgcaacatag gcagacctta agggatgact cagtaacaga 180

taagctttgt gtgcctgcag ggcatataaa acaggggcaa ggcacagact catagcagag 240

caatcaccac caagcctgga ataactgcag ccacc 275

<210> 145

<211> 284

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 145

tcagatgcc tggagagccc atccacgtg ttttgacctc catagaagac accgggaccg 60

atccagctc cgcgcccgga aacggtgcat tggaaacggg attccccgtg ccaagagtga 120

cgtaagtacc gcctatagac tctataggca caccctttg gctcttatgc atgaacggtg 180
gagggcagtg tagtctgagc agtactcgtt gctgccgcgc gcgccaccag acataatagc 240
tgacagacta acagactgtt cctttccatg ggtcttttct gcag 284

<210> 146

<211> 636

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 146

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
tgtttgctct ggtaataat ctgaggagca caaacattcc tgtaccggcc cgggaggcgc 120
cctttggacc ttttgcaatc ctggcgact gaacccttga cccctgcct gcagcccccg 180
cagcttgctg tttgccact ctatttgcgc agcccagcc ctggagagtc ctttagcagg 240
gcaaagtga acataggcag accttaaggg atgactcagt aacagataag ctttgtgtgc 300

ctgcagggca tataaaacag gggcaaggca cagactcata gcagagtcag atgccttga 360
gacgccatcc acgtgtttt gacctcata gaagacaccg ggaccgatcc agcctccgcg 420
gccgggaacg gtgcattgga acgcggattc cccgtgcaa gagtgacgta agtaccgct 480
atagactcta taggcacacc cttttggctc ttatgcatga acggtggagg gcagtgtagt 540
ctgagcagta ctggttctg ccgcgcgcgc caccagacat aatagctgac agactaacag 600
actgttcctt tccatgggtc ttttctgcag gccacc 636

<210> 147

<211> 733

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 147

caggctttca ctttctgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60
tacccggtt cccggcaacg gccaggtctg tgccaagtgt ttgaggttaa tttttaaaaa 120
gcagtcaaaa gtccaagtgg cctttggcag catttactct ctctgtttgc tctggttaat 180
aatctcagga gcacaaacat tccgccccg gaggcgcct ttggacctt tgcaatcctg 240

gcgactgaa ccttgacc ctgccctgca gccccgcag cttgtgttt gccactcta 300
 ttgcccagc ccagccctg gagagtcctt tagcaggga aagtgaaca taggcagacc 360

ttaagggatg actcagtaac agataagctt tgtgtgctg caggcatat aaaacagggg 420
 caaggcacag acatagaca gaggcagat gcctggagac gccatccag ctgttttgac 480
 ctccatagaa gacaccgga ccgattcagc ctccgcggcc gggaacggtg cattggaacg 540
 cggattcccc gtgccaagag tgacgtaagt accgcctata gactctatag gcacaccct 600
 ttggctctta tgcataaac gtggaggga gtgtagtctg agcagtactc gttgtgccg 660
 cgcgccac cagacataat agctgacaga ctaacagact gttcctttcc atgggtcttt 720
 tctgcaggcc acc 733

<210> 148

<211> 653

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 148

caggcttca cttctcgcc aactacaag gcctttctgt gtaacaata cctgaacctt 60
 taccctgtg ccggcaacg gccaggtctg tgccaagtgt ttgccgatga cctaatgatt 120
 ctgagcttg caaaggtctt atctccagc tgcaccaggc ccagtgttc aggaatgtga 180
 cctttgtgc agcagccgt ggagggggca gagggatgg gctggaggt gagcaaacag 240
 agcagcagaa aaggcagtc ctctctcca gtgcccctct tccctgtctc tgcctctccc 300

tccctctctc aggcacaga gcggagactt caggagacc agagcccagc ttgccaggca 360
 ctgtcagatc gcctggagac gccatccag ctgttttgac ctccatagaa gacaccgga 420
 ccgatccagc ctccgcggcc gggaacggtg cattggaacg cggattcccc gtgccaagag 480
 tgacgtaagt accgcctata gactctatag gcacaccct ttggctctta tgcataaacg 540
 gtggaggga gtgtagtctg agcagtactc gttgtgccg cgcgccac cagacataat 600
 agctgacaga ctaacagact gttcctttcc atgggtcttt tctgcaggcc acc 653

<210> 149

<211> 543

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 149

aggTTAattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgttttctct ggTTaataat ctCaggagca caaacattcc ggcccgggag gcgccctttg	120
gaccttttgc aatcctggcg cactgaaccc ttgaccctg ccctgcagcc cccgcagctt	180
gctgtttgcc cactctatTT gcccagcccc aggggcatat aaaacagggg caaggcacag	240
actcatagca gattcagatc gcctggagac gccatccacg ctgttttgac ctccatagaa	300
gacaccggga ccgatccagc ctccgcggcc gggaacgggtg cattggaacg cggattcccc	360

gtgccaagag tgacgtaagt accgcctata gactctatag gcacaccctt ttggctctta	420
tgcattgaacg gtggagggca gtgtagtctg agcagtactc gttgctgccg cgcgcgccac	480
cagacataat agctgacaga ctaacagact gttcctttcc atgggtcttt tctgcaggcc	540
acc	543

<210> 150

<211> 530

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 150

ggcccgggag gcgccctttg gaccttttgc aatcctggcg cactgaaccc ttgaccctg	60
ccctgcagcc cccgcagctt gctgtttgcc cactctatTT gcccagcccc agccctggag	120
agtcctttag cagggcaaag tgcaacatag gcagacctta agggatgact cagtaacaga	180
taagctttgt gtgcctgcag ggcatataaa acaggggcaa ggacagact catagcagag	240
tcagatgcc ttgagacgcc atccacgtg ttttgacctc catagaagac accgggaccg	300
atccagcctc cgcggccggg aacggtgcat tggaacgcgg attccccgtg ccaagagtga	360
cgtaagtacc gcctatagac tctataggca caccctttg gctcttatgc atgaacgggtg	420
gagggcagtg tagtctgagc agtactcgtt gctgccgcgc gcgccaccag acataatagc	480

tgacagacta acagactgtt cctttccatg ggtcttttct gcaggccacc	530

<210> 151

<211> 103

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 151

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60

tacccegttg cccggcaacg gccaggtctg tgccaagtgt ttg 103

<210> 152

<211> 80

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 152

ccctgttcaa acatgtccta atactctgtc tctgcaaggg tcatcagtag ttttccatct 60

tactcaacat cctcccagtg 80

<210> 153

<211> 6

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 153

gccacc 6

<210> 154

<211> 6557

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 154

gccattcgcc attcaggtcg cgcaactgtt gggaagggcg atcggtgcgg gcctcttcgc 60

tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag 120

ggttttccca gtcacgacgt tgtaaaacga cggccagtga attcctcgag atttaaatga 180

cgttggccac tccctctctg cgcgctcgt cgtcactga ggccgggcga ccaaaggtcg 240

cccgacgccc gggctttgcc cgggcggcct cagtgagcga gcgagcgcgc agagaggag 300

tggccaactc catcactagg ggttctgag tttaaacttc gtcgaggccg cacgcgtagg 360

ttaattttta aaaagcagtc aaaagtccaa gtggcccttg gcagcattta ctctctctgt 420

ttgctctggt taataatctc aggagcacia acattccccc tgttcaaaca tgtcctaata 480

ctctgtctct gcaagggtca tcagtagttt tccatcttac tcaacatcct cccagtgagt 540

catatgtttg ctactgaag gttactagtt aacaggcatc ccttaaacag gatataaaag 600

gacttcagca ggactgctcg aaacatccca ctacagccacc atgggagtga ggcacccgcc 660

ctgtccccc cggctcctgg ccgtctgcgc cctcgtgtcc ttggcaaccg ctgactcct 720

ggggcacatc ctactccatg atttctgct ggttccccga gagctgagt gtcctcccc 780

agtctggag gagactcacc cagctacca gcaggagacc agcagaccag ggccccggga 840

tgcccaggca caccgccgcc gtcccagagc agtgcccaca cagtgcgacg tccccccaa 900

cagccgcttc gattgcgccc ctgacaaggc catcaccag gaacagtgcg agggccgcgg 960

ctgtctctac atccctgcaa agcaggggct gcaggagacc cagatggggc agccctggtg 1020

cttcttccca cccagctacc ccagctacaa gctggagaac ctgagctcct ctgaaatggg 1080

ctacacggcc accctgacc gtaccacccc caccttcttc cccaaggaca tctgaccct 1140

gcggctggac gtgatgatg agactgagaa ccgcctccac ttcacgatca aagatccagc 1200

taacaggcgc tacgaggtgc ccttgagac ccgcgtgtc cacagccggg caccgtcccc 1260

actctacagc gtggagtct ccgaggagcc ctccgggtg atcgtgcacc ggcagctgga 1320

cggcccgctg ctgtgaaca cgacggtggc gccctgttc ttgctggacc agttccttca 1380

gctgtccacc tcgtgccct cgcagtatat cacaggctc gccagcacc tcagtccct 1440

gatgtcagc accagctgga ccaggatcac cctgtggaac cgggacctg cggccacgcc 1500

cgggtgcgaac ctctacgggt ctacaccttt ctacctggcg ctggaggacg gcgggtcggc 1560

acacgggggtg ttctgtctaa acagcaatgc catggatgtg gtcctgcagc cgagccctgc 1620

ccttagctgg aggtcgacag gtgggatect ggatgtctac atcttcttg gccagagcc 1680

caagagcgtg gtgcagcagt acctggacgt tgtgggatac ccgttcatgc cgccatactg 1740

gggcctgggc ttccacctgt gccgtgggg ctactctcc accgtatca cccgccaggt 1800

ggtggagaac atgaccagg cccacttccc cctggacgtc caatggaacg acctggacta 1860

catggactcc cggagggact tcacgttcaa caaggatggc ttccgggact tcccggccat 1920

ggtgcaggag ctgcaccagg gcggccggcg ctacatgatg atcgtggatc ctgcatcag 1980

cagctcgggc cctgccggga gctacaggcc ctacgacgag ggtctgcgga ggggggtttt 2040

catcaccaac gagaccggcc agccgctgat tgggaaggta tggccccggt cactgcctt	2100
ccccgacttc accaacccca cagccctggc ctggtgggag gacatggtgg ctgagttcca	2160
tgaccagggtg cccttcgacg gcatgtggat tgacatgaac gagccttcca acttcacag	2220
aggctctgag gacggctgcc ccaacaatga gctggagaac ccacctacg tgcctggggt	2280
ggttgggggg accctccagg cgccacat ctgtgcctcc agccaccagt ttctctccac	2340
acactacaac ctgcacaacc tctacggcct gaccgaagcc atgcctccc acaggcgct	2400
ggtgaaggct cgggggacac gccatttgi gatctccgc tcgaccttg ctggccacgg	2460
ccgatacgcc ggccactgga cgggggacgt gtggagctcc tgggagcagc tcgctcctc	2520
cggtgccagaa atcctgcagt ttaacctgct gggggtgcct ctggtcgggg ccgacgtctg	2580
cggcttcttg ggcaacacct cagaggagct gtgtgtgcgc tggaccagc tgggggcctt	2640
ctaccccttc atcggaacc acaacagcct gctcagctg cccagagc cgtacagctt	2700
cagcgagccg gccagcagg ccatgaggaa ggccctcacc ctgcctacg cactcctccc	2760
ccacctctac acactgttc accaggccca cgtcgcgggg gagaccgtg cccggccct	2820
cttctggag ttcccaagg actctagcac ctggactgtg gaccaccagc tctgtgggg	2880
ggaggccctg ctcatcccc cagtgtcca ggccgggaag gccgaagtga ctggctactt	2940
ccccctgggc acatggtacg acctgcagac ggtgccata gaggccttg gcagcctccc	3000
acccccacct gcagctcccc gtgagccagc catccacagc gaggggcagt ggtgacgt	3060
gccggcccc ctggacacca tcaactcca cctccgggt ggtacatca tccccctgca	3120
gggcctggc ctcaacaac cagagtccg ccagcagccc atggccctgg ctgtggcct	3180
gaccaagggt ggagaggccc gaggggagct gtctgggac gatggagaga gcctggaagt	3240
gctggagcga ggggcctaca cacaggtcat ctctctggcc aggaataaca cgatcgtgaa	3300
tgagctggta cgtgtgacca gtgaggagc tggcctgcag ctgcagaagg tgactgtcct	3360
gggcgtggcc acggcgcccc agcaggtcct ctccaacggt gtcctgtct ccaactcac	3420
ctacagcccc gacaccaagg tctggacat ctgtgtctcg ctgttgatgg gagagcagtt	3480
tctcgtcagc tgggtgttagc cgggcggagt gtgttagtct ctccagagg aggctggttc	3540
cccaggaag cagagcctgt gtgcgggcag cagctgtgtg cgggcctggg ggttgtaag	3600
tgcaattatt ttaataaaa ggggcatttg gaaaaaaaa aaaaaggtag cagtcgacag	3660
atgaattctg cagatctgtg gcttctagct gcccggtgg catcctgtg acccctccc	3720
agtgcctctc ctggccctgg aagttgccac tccagtgcc accagccttg tcctaataaa	3780

attaagtgc atcattttgt ctgactaggt gtcctttctat aatattatgg ggtggagggg	3840
ggtggtatgg agcaaggggc aagtgggaa gacaacctga gttgttgga ttccagcat	3900
cgagtagata agtagcatgg cgggttaatc attaaactaca cacgtgcgga ccgagcggcc	3960
taagctcgcg aaggaacccc tagtgatgga gttggccact cctctctgc gcgctcgctc	4020
gctcactgag gccgggcgac caaaggtcgc ccgacgcccg ggctttgcc gggcggcctc	4080
agtgagcgag cgagcgcga gagaggaggt ggccaagacg atttaaatga caagcttggc	4140
gtaatcatgg tcatagctgt ttctgtgtg aaattgttat ccgctcaca ttccacaca	4200
catacgagcc ggaagcataa agtgtaaagc ctggggtgcc taatgagtga gtaactcac	4260
attaattgcg ttgcgctcac tgcccgttt ccagtcggga aacctgtcgt gccagctgca	4320
ttaatgaatc ggccaacgag cggggagagg cggtttgcgt attgggcgt cttccgcttc	4380
ctcgctcact gactcgctgc gctcggctgt tcggctgcgg cgagcggat cagctcactc	4440
aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc	4500
aaaaggccag caaaaggcca ggaaccgtta aaaggccgag ttgctggcgt tttccatag	4560
gctccgcccc cctgacgagc atcacaaaaa tcgacgtca agtcagaggt ggcgaaaccc	4620
gacaggacta taaagatacc aggcgtttcc ccttgaagc tccctcgtgc gctctcctgt	4680
tccgacctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgt	4740
ttctcatagc tcacgtgta ggtatctcag ttcggtgtag gtcgttcgt ccaagctggg	4800
ctgtgtgcac gaacccccg ttacagccga ccgctgcgcc ttatccgta actatcgtct	4860
tgagtccaac ccgtaagac acgacttctc gccactggca gcagccactg gtaacaggat	4920
tagcagagcg aggtatgtag gcggtgtctc agagtcttg aagtgggtgc ctaactacgg	4980
ctacactaga agaacagtat ttggtatctg cgctctgctg aagccagta ccttcggaaa	5040
aagagtgggt agctcttgat ccggcaaaaca aaccaccgt ggtagcgggtg gttttttgt	5100
ttgcaagcag cagattacgc gcagaaaaa aggatctcaa gaagatcctt tgatctttc	5160
tacggggtct gacgtcagt ggaacgaaa ctcacgttaa gggattttgg tcatgagatt	5220
atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta	5280
aagtatatat gagtaacctt ggtctgacag ttagaaaaa tcatcgagca tcaaatgaaa	5340
ctgcaattta ttcatatcag gattatcaat accatatctt tgaaaaagcc gtttctgtaa	5400
tgaaggagaa aactcacga ggcagttcca taggatggca agatcctggt atcggtctgc	5460
gattccgact cgtccaacat caatacaacc tattaatttc cctcgtcaa aaataaggtt	5520
atcaagttag aaatcaccat gactgacgac tgaatccgtt gagaatggca aaagtttatg	5580
catttctttc cagacttgtt caacaggcca gccattacgc tcgtcatcaa aatcactcgc	5640

atcaacaaaa ccgttattca ttctgtattg cgcctgagcg agacgaaata cgcgatcgct 5700
gttaaaagga caattacaaa caggaatcga atgcaaccgg cgcaggaaca ctgccagcgc 5760
atcaacaata ttttcacctg aatcaggata ttctttctaat acctggaatg ctgtttttcc 5820
ggggatcgca gtggtgagta accatgcac atcaggagta cggataaaat gcttgatggt 5880

cggaagaggc ataaattccg tcagccagtt tagtctgacc atctcatctg taacatcatt 5940
ggcaacgcta cctttgccat gtttcagaaa caactctggc gcatcgggct tcccatacaa 6000
gcatagatt gtgcacactg attgcccac attatcgca gccatttat accatataa 6060
atcagcatcc atgttggat ttaatcgcg cctcgacgtt tcccgttgaa tatggctcat 6120
actcttctt tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata 6180
catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa 6240
agtgccacct gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg 6300

tatcacgagg ccctttcgtc tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat 6360
gcagctccc gagacggta cagcttgtct gtaagcggat gccgggagca gacaagccc 6420
tcaggcgcg tcagcgggtg ttggcgggtg tcggggctgg cttaactatg cggcatcaga 6480
gcagattgta ctgagagtc accatatgcg gtgtgaaata ccgcacagat gcgtaaggag 6540
aaaataccgc atcaggc 6557

<210> 155

<211> 6694

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 155

gccattcgcc attcaggctg cgcaactgtt gggaagggcg atcgggtcgg gcctcttcgc 60
tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag 120
ggttttccca gtcacgacgt tgtaaacga cggccagtga attcctcgag atttaaatga 180
cgttggccac tccctctctg cgcgtctgct cgtcactga ggccggcgga ccaaaggteg 240
cccagcggc gggctttgcc cgggcggcct cagtgagcga gcgagcgcgc agagaggag 300
tggccaactc catcactagg ggttctctgag tttaaaactt gtcaggccg cacgcgtcga 360
gcttgggctg caggtcgagg gcaactgggag gatgttgagt aagatggaaa actactgatg 420

acccttgag agacagagta ttaggacatg ttgaaacagg ggccgggcga tcagcaggta 480

gctctagagg atccccgtct gtctgcacat ttcgtagagc gagtgttccg atactctaata 540
 ctccctaggc aaggttcata tttgtgtagg ttacttattc tctttttgtt gactaagtca 600
 ataatacagaa tcagcaggtt tggagtcagc ttggcaggga tcagcagcct ggggttggaag 660
 gaggggggtat aaaagccctt tcaccaggag aagccgtcac acagaggtac cgagctcgga 720
 tccactagta acggccgcca gtgtgctgga attcggtttt ccaggccatc tccaaccatg 780
 ggagttaggc acccgccctg ctcccaccgg ctctggccg tctgcgcctt cgtgtccttg 840

 gcaaccgctg cactcctggg gcacatccta ctccatgatt tcctgctggt tccccgagag 900
 ctgagtggct cctccccagt cctggaggag actcaccag ctcaccagca gggagccagc 960
 agaccagggc cccgggatgc ccaggcacac cccggccgtc ccagagcagt gcccacacag 1020
 tgcgacgtcc ccccaacag ccgcttcgat tgcgcccctg acaaggccat caccaggaa 1080
 cagtgcgagg cccgcggtg ctgctacatc cctgcaaagc aggggctgca gggagcccag 1140
 atggggcagc cctgggtgctt ctccaccac agctaccca gctacaagct ggagaacctg 1200
 agctcctctg aaatgggcta cacggccacc ctgaccgta ccacccccc cttcttcccc 1260

 aaggacatcc tgacctgctg gctggacgtg atgatggaga ctgagaaccg cctccacttc 1320
 acgatcaaag atccagctaa caggcgctac gaggtgcctt tggagacccc gcgtgtccac 1380
 agccgggcac cgtccccact ctacagcgtg gagttctccg aggagccctt cggggtgatc 1440
 gtgcaccggc agctggacgg ccgctgtctg ctgaacacga cggtggcgcc cctgttcttt 1500
 gcggaccagt tcttcagct gtccacctcg ctgccctcgc agtatatcac aggcctcgcc 1560
 gagcacctca gtccctgat gctcagcacc agctggacca ggatcacctt gtggaaccgg 1620
 gaccttgccg ccacgcccg tgcgaacctc tacgggtctc accctttcta cctggcgctg 1680

 gaggacggcg ggtcggcaca cggggtgttc ctgctaaca gcaatgccat ggatgtggtc 1740
 ctgcagccga gccctgcctt tagctggagg tcgacagggt ggatcctgga tgtctacatc 1800
 ttcctgggcc cagagcccaa gagcgtggtg cagcagtacc tggacgttgt gggatacccg 1860
 ttcctgccc cactactggg cctgggcttc cactgtgcc gctggggcta ctctccacc 1920
 gctatcacc gccaggtggt ggagaacatg accaggggcc acttccccct ggacgtccaa 1980
 tggaacgacc tggactacat ggactcccgg agggacttca cgttcaaca ggatggcttc 2040
 cgggacttcc cggccatggt gcaggagctg caccaggcg gccggcgcta catgatgatc 2100

 gtggatcctg ccatcagcag ctccggccct gccgggagct acaggcccta cgacagggtt 2160
 ctgcggaggg gggttttcat caccaacgag accggccagc cgctgattgg gaaggtatgg 2220
 cccgggtcca ctgccttccc cgacttcacc aacccacag cctggcctg gtgggaggac 2280
 atggtggctg agttccatga ccaggtgccc ttcgacggca tgtggattga catgaacgag 2340

ccctccaact tcatcagagg ctctgaggac ggctgcccc acaatgagct ggagaaccca	2400
ccctactgtc ctgggggtgt tggggggacc ctccaggcgg ccaccatctg tgcctccagc	2460
caccagtttc tctccacaca ctacaacctg cacaacctct acggcctgac cgaagccatc	2520
gcctcccaca gggcgcttgt gaaggctcgg gggacacgcc catttgtgat ctcccgtctg	2580
acctttgtct gccacggccg atacgccggc cactggacgg gggacgtgtg gagctcctgg	2640
gagcagctcg cctcctccgt gccagaaatc ctgcagtta acctgctggg ggtgcctctg	2700
gtcggggccg acgtctgcgg ctctctgggc aacacctcag aggagctgtg tgtgcgtgtg	2760
accagctgg gggccttcta ccccttcctg cggaaccaca acagcctgct cagtctgccc	2820
caggagccgt acagcttcag cgagccggcc cagcaggcca tgaggaaggc cctcacctg	2880
cgctacgcac tcttccccca cctctacaca ctgttcacc aggccacgt cgcgggggag	2940
accgtggccc ggccctctt cctggagtgc cccaaggact ctagcacctg gactgtggac	3000
caccagctcc tgtgggggga ggccctgtc atcacccag tgctccaggc cgggaaggcc	3060
gaagtactg gctacttccc ctggggcaca tggtagacc tgcagacggt gccaatagag	3120
ggccttggca gctcccacc cccacctgca gctccccgtg agccagccat ccacagcgag	3180
gggcagtggg tgacgtgccc gggccccctg gacaccatca acgtccacct cgggctggg	3240
tacatcatcc cctgcaggg cctggcctc acaaccacag agtcccga gcagccatg	3300
ggcctggctg tggccctgac caagggtgga gaggcccgag gggagctgtt ctgggacgat	3360
ggagagagcc tggaagtgt ggagcgagg gcctacacac aggtcatctt cctggccagg	3420
aataacacga tctgaatga gctgttacgt gtgaccagtg agggagctgg cctgcagctg	3480
cagaaggatg ctgtcctggg cgtggccacg gcgccccagc aggtcctctc caacggtgtc	3540
cctgtctcca acttcacct cagccccgac accaaggctc tggacatctg tgtctcgctg	3600
ttgatgggag agcagtttct cgtcagctgg tgttagccgg gcggagtgtg ttagtctctc	3660
cagagggagg ctggttcccc aggggaagcag agcctgtgtg cgggcagcag ctgtgtcgg	3720
gcctgggggt tgtaagtgc aattatttt aataaaagg gcatttgaa aaaaaaaaa	3780
aaggtagcag tcgacagatg aattctgcag atctgtggct tctagctgcc cgggtggcat	3840
ccctgtgacc cctccccagt gcctctctg gcctggaag ttgccactcc agtgcacc	3900
agcctgttc taataaaatt aagtgtcatc attttgtctg actaggtgtc ctctataat	3960
attatgggtt ggaggggggt ggtatggagc aaggggcaag ttgggaagac aacctgagtt	4020
gttgggattc caggcatcga gtagataagt agcatggcgg gttaatcatt aactacacac	4080
gtcgggaccg agcgccctaa gctcgcaag gaaccctag tgatggagtt ggccactccc	4140

tctctgcgcg ctgcctcgct cactgaggcc gggcgaccaa aggtcgcccg acgcccgggc	4200
tttgcgggg cgccctcagt gagcgagcga gcgcgcagag agggagtggc caagacgatt	4260
taaatgacaa gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg	4320
ctcacaattc cacacaacat acgagccgga agcataaagt gtaaagcctg gggcgcctaa	4380
tgagtgagct aactcacatt aattgcgttg cgtcactgc ccgctttcca gtcgggaaac	4440
ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg ttgctgatt	4500
gggcgctctt ccgcttcctc gctcactgac tcgtcgcgt cggctgttcg gctgcggcga	4560
gcggtatcag ctactcaaa ggcggttaata cggttatcca cagaatcagg ggataacgca	4620
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg	4680
ctggcgtttt tccataggtt ccgccccct gacgagcatc acaaaaatcg acgtcaagt	4740
cagaggtggc gaaacccgac aggactataa agataaccagg cgtttcccc tggaaagctcc	4800
ctcgtcgcgt ctctgttcc gacctgccc cttaccgat accgtccgc ctttctcct	4860
tcgggaagcg tggcgctttc tcatagctca cgtgttaggt atctcagttc ggtgtaggtc	4920
gttcgtcca agctgggctg tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta	4980
tccggttaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca	5040
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gtctctgaag	5100
tggcgccta actacggcta cactagaaga acagtatttg gtatctgcgc tctgctgaag	5160
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgtggt	5220
agcgggtggtt ttttgtttg caagcagcag attacgcgca gaaaaaagg atctcaagaa	5280
gatactttga tctttctac ggggtctgac gctcagtga acgaaaactc acgttaaggg	5340
atattggtca tgagattatc aaaaaggatc ttacactaga tccttttaaa ttaaaaatga	5400
agttttaaat caatctaaag tatatatgag taaacttggc ctgacagtta gaaaaactca	5460
tcgagcatca aatgaaactg caatttattc atatcaggat tatcaatacc atatttttga	5520
aaaagccgtt tctgtaatga aggagaaaac tcaccgaggc agttccatag gatggcaaga	5580
tcctggtatc ggctgcgat tccgactcgt ccaacatcaa tacaacctat taatttcccc	5640
tcgtcaaaaa taaggttatc aagtgagaaa tcacatgag tgacgactga atccggtgag	5700
aatggcaaaa gtttatgcat ttctttccag acttgttcaa caggccagcc attacgctcg	5760
tcatcaaaat cactcgcatc aaccaaacccg ttattcattc gtgattgcgc ctgagcgaga	5820
cgaaatacgc gatcgtgtt aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc	5880

aggaacactg ccagcgcatc aacaatatTT tcacctgaat caggatatTC ttctaatacc 5940
 tggaatgctg tttttccggg gatcgcatg gtgagtaacc atgcatcatc aggagtacgg 6000
 ataaaatgct tgatggtcgg aagaggcata aattccgtca gccagtttag tctgaccatc 6060
 tcatctgtaa catcattggc aacgttacct ttgccatgtt tcagaaacaa ctctggcgca 6120
 tcgggcttcc catacaagcg atagattgtc gcacctgatt gcccacatt atcgcgagcc 6180
 catttatacc catataaatc agcatccatg ttggaattta atcgcggcct cgacgtttcc 6240
 cgttgaatat ggctcatact ctcccttttt caatatattt gaagcattta tcagggttat 6300

tgctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6360
 cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6420
 acctataaaa ataggcgtat cagcaggccc ttctgtctcg cgcgtttcgg tgatgacggt 6480
 gaaaacctct gacacatgca gctcccgag acggtcacag cttgtctgta agcggatgcc 6540
 gggagcagac aagcccgta gggcgctca gcgggtgttg gcgggtgtcg gggctggctt 6600
 aactatgcgg catcagagca gattgtactg agagtgcacc atatgcggtg tgaatatccg 6660
 cacagatgcg taaggagaaa ataccgcatc aggc 6694

<210> 156

<211> 6984

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 156

gccattcgcc attcaggctg cgcaactgtt gggaaggcg atcgggtcgg gcctcttcgc 60
 tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag 120
 ggttttccca gtcacagct tgtaaacga cggccagtga attcctcgag atttaaatga 180
 cgttggccac tcctctctg cgcgtcgtc cgtcactga ggccgggcga ccaaaggctc 240
 cccgacgccc gggctttgcc cgggcggcct cagttagcga gcgagcgcgc agagaggag 300

tggccaactc catcactagg ggttctgag tttaaacttc gtcaggccg cagcgtagg 360
 ttaattttta aaaagcagtc aaaagtccaa gtggcccttg gcagcattta ctctctctgt 420
 ttgtctgtgt taataatctc aggagcacia acattcccc tgttcaaaca tgtcctaata 480
 ctctgtctct gcaagggtca tcagtagttt tccatcttac tcaacatcct cccagttagt 540
 catatgtttg ctactgaag gttactagtt aacaggcatc ccttaaacag gatataaaag 600
 gacttcagca ggactgctcg aaacatccca ctcagccacc atgggagtga ggcacccgcc 660

ctgctccac cggtctctgg ccgtctgcgc cctcgtgtcc ttggcaaccg ctgcactcct	720
ggggcacatc ctactccatg atttctgtct ggttccccga gagctgagtg gtcctctccc	780
agtctctggag gagactcacc cagctcacca gcagggagcc agcagaccag ggccccggga	840
tgccaggca caccctggcc gtcccagagc agtgcccaca cagtgcgacg tccccccaa	900
cagccgttc gattgcgcc ctgacaaggc catcaccag gaacagtgcg agggccgcgg	960
ctgtctctac atccctgcaa agcaggggct gcagggagcc cagatggggc agccctggtg	1020
cttcttccca ccagctacc ccagctacaa gctggagaac ctgagctcct ctgaaatggg	1080
ctacacggcc accctgacc gtaccacccc caccttcttc cccaaggaca tctgaccct	1140
gcggctggac gtgatgatgg agactgagaa ccgcctccac ttcacgatca aagatccagc	1200
taacaggcgc tacgagggtc ccttgagac ccgcgtgtc cacagccggg caccgtcccc	1260
actctacagc gtggagtct cagaggagcc ctctgggggtg atcgtgcacc ggcagctgga	1320
cggcccgctg ctgtgaaca cgacggtggc gccctgttc tttgcggacc agttccttca	1380
gctgtccacc tcgtgccct cgcagtatat cacaggcctc gccgagcacc tcagtccct	1440
gatgtcagc accagctgga ccaggtacac cctgtggaac cgggaccttg cggccacgc	1500
cgggtcgaac ctctacgggt ctacaccttt ctacctggcg ctggaggacg gcgggtcggc	1560
acacggggtg ttctgtctaa acagcaatgc catggatgtg gtctgcagc cgagccctgc	1620
ccttagctgg aggtcgacag gtgggacct ggatgtctac atcttctgg gccagagcc	1680
caagagcgtg gtgcagcagt acctggacgt tgtgggatac ccgttcatgc cgccatactg	1740
gggcctgggc ttccacctgt gccctgggg ctactctcc accgtatca cccgccaggt	1800
ggtggagaac atgaccaggg ccacttccc cctggacgtc caatggaacg acctggacta	1860
catggactcc cggagggact tcacgttcaa caagatggc ttccgggact tcccggccat	1920
ggtgcaggag ctgcaccagg gcggccggcg ctacatgatg atcgtggatc ctgcatcag	1980
cagctcgggc ctgcggga gctacaggcc ctacgacgag ggtctgcgga ggggggtttt	2040
catcaccaac gagaccggcc agccgtgat tgggaaggta tggccgggt cactgcctt	2100
ccccgacttc accaaccceca cagccctggc ctggtgggag gacatggtgg ctgagttcca	2160
tgaccaggtg cccttcgacg gcatgtgat tgacatgaac gagccttcca acttcatcag	2220
aggctctgag gacggctgcc ccaacaatga gctggagaac ccacctacg tgctgggggt	2280
ggttgggggg accctccagg cggccacat ctgtgcctcc agccaccagt ttctctccac	2340
acactacaac ctgcacaacc tctacggcct gaccgaagcc atgcctccc acaggcgct	2400

ggtgaaggct cgggggacac gcccatTTgt gatctcccgc tcgacctttg ctggccacgg	2460
ccgatacgcc ggccactgga cgggggacgt gtggagctcc tgggagcagc tcgcctcctc	2520
cgtgccagaa atcctgcagt ttaacctgct ggggggtgcct ctggtcgggg ccgacgtctg	2580
cggcttcctg ggcaaacctt cagaggagct gtgtgtgcgc tggaccacgc tgggggcctt	2640
ctacccttc atgcggaacc acaacagcct gctcagtctg ccccaggagc cgtacagctt	2700
cagcgagccg gcccgagcagg ccatgaggaa ggccctcacc ctgcgctacg cactcctccc	2760
ccacctctac aactgttcc accaggccca cgtcgcgggg gagaccgtgg cccggccctt	2820
cttctggag tttcccaagg actctagcac ctggactgtg gaccaccagc tctgtgggg	2880
ggaggccctg ctcatcacc cagtgtcca ggccgggaag gccgaagtga ctggctactt	2940
ccccttgggc acatggtacg acctgcagac ggtgccaata gaggcccttg gcagcctccc	3000
acccccact gcagctcccc gtgagccagc catccacagc gaggggcagt gggtagcgt	3060
gccggccccc ctggacacca tcaacgtcca cctccgggt gggtagatca tccccctgca	3120
gggccttggc ctcaaacca cagagtcccg ccagcagccc atggccctgg ctgtggcctt	3180
gaccaaggtt ggagaggccc gaggggagct gttctgggac gatggagaga gcctggaagt	3240
gctggagcga ggggcctaca cacaggtcat cttcctggcc aggaataaca cgatcgtgaa	3300
tgagctggta cgtgtgacca gtgaggagc tggcctgcag ctgcagaagg tgactgtcct	3360
gggcgtggcc acggcgcccc agcaggctct ctccaacggt gtcctgtct ccaacttcac	3420
ctacagcccc gacaccaagg tcttgacat ctgtgtctcg ctgttgatgg gagagcagtt	3480
tctcgtcagc tgggtgttagc gagcgccgc tcttagtagc agtatcgatc ccagccact	3540
tttcccaat acgactacga gatctgtggc ttttagctgc ccgggtggca tccctgtgac	3600
ccctccccag tgcctctcct ggccctggaa gttgccactc cagtgccac cagccttgtc	3660
ctaataaaat taagttgcat cttttgtct gactaggtgt cttctataa tattatgggg	3720
tggagggggg tggtaggag caaggggcaa gttgggtcc ttcaacgtcg aactgtagca	3780
ctctggatag ctagctagcg ttttagata gtttcaagga tagcaacaa caacacgata	3840
gtttagaaat tgttctagat cttctgtctac tactcgagag atagttgact tctagaagac	3900
tctggataga gagcttcctg tggtttctat ttcaacctt cgagcgattt agatagaggc	3960
gtcaattgag aaaaaaact ggacacctgt ttttagaaca ctagtctag tcttctcaag	4020
gcttgatagc tctgccact cacacacaaa aaaccaacac acagattaat gaaaataaag	4080
atcttttatt agagagcgag tggatagttt agagtcgctt gagacttacc ctgtgtccaa	4140
ctagccaaga gtagagagag agtgactagg tggatcgaga gatcgtcgtc tatttgactg	4200
ctagcctgtg ttctcgtcgt agtccaaatc ttcaatactc tggcgactag taacttgaat	4260

ctgtgtatct aacacgacaa ggctactcct ttgagtttct agtgataagt gcgactggat 4320
ctagctagtt gagtttctag aaactacgtc caaagctact agcactacac gtgcggaccg 4380
agcggcctaa gctcgcgaag gaacccttag tgatggagtt ggccactccc tctctgcgcg 4440
ctcgtctgct cactgaggcc gggcgaccaa aggtcgcccc acgcccgggc ttgcccggg 4500

cggcctcagt gagcgagcga gcgcgcagag agggagtggc caagacgatt taaatgacaa 4560
gcttggcgta atcatggta tagctgtttc ctgtgtgaaa ttgttatccg ctcacaattc 4620
cacacaacat acgagccgga agcataaagt gtaaagcctg gggtgcctaa tgagtgaagt 4680
aactcacatt aattgcgttg cgctcactgc ccgtttcca gtcgggaaac ctgtcgtgcc 4740
agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt gggcgctctt 4800
ccgttctctc gctcactgac tcgtcgcgt cggtcgttcg gctgcggcga gcggtatcag 4860
ctcactcaaa ggcggaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 4920

tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt 4980
tccataggct ccgccccct gacgagcatc aaaaaaatcg acgtcaagt cagaggtggc 5040
gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgt 5100
ctcctgttcc gacctgccg cttaccgat acctgtccgc ctttctccct tcgggaagcg 5160
tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gtctcgtcca 5220
agctgggctg tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta tccggttaact 5280
atcgtcttga gtccaaccg gtaagacacg acttatcgcc actggcagca gccactggta 5340

acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tgggtgccta 5400
actacggcta cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct 5460
tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgtggt agcgggtggtt 5520
tttttgttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatccttga 5580
tcttttctac ggggtctgac gctcagtga acgaaaactc acgttaaggg attttggta 5640
tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agtttttaaat 5700
caatctaaag tatatatgag taaacttggt ctgacagtta gaaaaactca tcgagcatca 5760

aatgaaactg caatttatc atatcaggat tatcaatacc atatttttga aaaagccgtt 5820
tctgtaatag aggagaaaaa tcaccgaggc agttccatag gatggcaaga tcctggtatc 5880
ggctctcgat tccgactcgt ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa 5940
taaggttatc aagtigagaaa tcacatgag tgacgactga atccggtgag aatggcaaaa 6000
gtttatgcat ttctttccag actgtttcaa caggccagcc attacgtcg tcatcaaaat 6060
cactgcacac aaccaaacgg ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc 6120

gatcgctgtt aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc aggaacactg 6180

ccagcgcac aacaatat tccactgaat caggatattc ttctaatacc tggaatgctg 6240

ttttccggg gatcgagtg gtgagtaacc atgcatcacc aggagtacgg ataaaatgct 6300

tgatggtcgg aagaggcata aattccgtca gccagtttag tctgaccatc tcatctgtaa 6360

catcattggc aacgctacct ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc 6420

catacaagcg atagattgtc gcacctgatt gcccagacatt atcgcgagcc catttatacc 6480

catataaatc agcatccatg ttggaattta atcgcggcct cgacgtttcc cgttgaatat 6540

ggctcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga 6600

gcgatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 6660

cccgaagaat gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa 6720

ataggcgat cagcaggccc ttctgtctcg cgcgtttcgg tgatgacgtt gaaaacctct 6780

gacacatgca gtcctcggag acggtcacag ctgtgtctgta agcggatgcc gggagcagac 6840

aagcccgta gggcgctga gcgggtgttg gcgggtgtcg gggctggctt aactatgcgg 6900

catcagagca gattgtactg agagtgcacc atatgcggtg tgaaataccg cacagatgcg 6960

taaggagaaa ataccgcac aggc 6984

<210> 157

<211> 6839

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 157

gccattcgcc attcaggctg cgcaactgtt gggaaggcgc atcggatgcg gcctcttcgc 60

tattacgcca gctggcgaaa ggggatgtg ctgcaaggcg attaagttgg gtaacgccag 120

ggttttccca gtcacgacgt tgtaaaacga cgccagtgat attcctcgag atttaaatga 180

cgttggccac tccctctctg cgcgctcgct cgctcactga ggccgggcga ccaaaggctc 240

cccagcgcgc gggctttgcc cgggcggcct cagtgcgcga gcgagcgcgc agagaggag 300

tggccaactc catcactagg ggttctgag tttaaacttc gtcgaggccg cacgcgtagg 360

ttaattttta aaaagcagtc aaaagtccaa gtggcccttg gcagcattta ctctctctgt 420

ttgtctggt taataatctc aggagcaca acattccccc tgttcaaaca tgtcctaata 480

ctctgtctct gcaagggtca tcagtagttt tccatcttac tcaacatcct ccagtgagt 540

catatgtttg ctactgaag gttactagtt aacaggcatc ccttaaacag gatataaaag	600
gacttcagca ggactgctcg aaacatccca cttcagatcg cctggagacg ccatccacgc	660
tgttttgacc tccatagaag acaccgggac cgatccagcc tccgcgcccg ggaacgggtgc	720
attggaacgc ggattccccg tgccaagagt gacgtaagta cgcctatag actctatagg	780
cacacccctt tggtctttat gcatgaacgg tggagggcag tgtagtctga gcagtactcg	840
ttgctgccgc gcgcgccacc agacataata gctgacagac taacagactg ttcctttcca	900
tgggtctttt ctgcaggcca ccatgggagt gaggcacccg ccctgctccc accggctcct	960
ggcctctgc gccctcgtgt ccttggcaac cgtgcactc ctggggcaca tcctactcca	1020
tgatttcctg ctggttcccc gagagctgag tggctcctcc ccagtcctgg aggagactca	1080
cccagctcac cagcaggagg ccagcagacc agggcccccg gatgccagg cacaccccg	1140
ccgtcccaga gcagtgcaca cacagtgcga cgtcccccc aacagccgt tgcattgcgc	1200
ccctgacaag gccatcacc aggaacagt cgaggccgc ggctgctgct acatccctgc	1260
aaagcagggg ctgcaggagg cccagatggg gcagccctgg tgcctcttcc caccagcta	1320
ccccagctac aagctggaga acctgagctc ctctgaaatg ggctacacgg ccacctgac	1380
ccgtaccacc cccaccttct tcccgaagga catcctgacc ctgcggctgg acgtgatgat	1440
ggagactgag aaccgcctcc acttcacgat caaagatcca gctaacaggc gctacgaggt	1500
gcccttggag accccgcgtg tccacagccg ggcaccgtcc ccactctaca gcgtggagtt	1560
ctccgaggag cccttcgggg tgatcgtgca cggcagctg gacggccgcg tgctgctgaa	1620
cacgacggtg gcgcccctgt tctttcgga ccagttcctt cagctgtcca cctcgtgcc	1680
ctcgcagtat atcacaggcc tcgccgagca cctcagtcct ctgatgtca gcaccagctg	1740
gaccaggatc accctgtgga accgggacct tgcgccacg cccggtgcga acctctacgg	1800
gtctcacctt ttctacctgg cgctggagga cggcgggtcg gcacacgggg tgttctgct	1860
aaacagcaat gccatggatg tggctctgca gccgagccct gcccttagct ggaggctcac	1920
aggtgggac ctggatgtct acatcttctt gggcccagag cccaagagcg tggatgcagca	1980
gtacctggac gtgtgggat acccgttcat gccgccatac tggggcctgg gcttccacct	2040
gtgccgctgg ggctactcct ccaccgtat caccgccag gtgtggaga acatgaccag	2100
ggccacttc cccctggacg tccaatggaa cgacctggac tacatggact cccggaggga	2160
cttcacgttc aacaaggatg gcttccggga ctccccgcc atggtgcagg agctgcacca	2220
gggcggcccg cgctacatga tgatcgtgga tcctgccatc agcagctcgg gccctgccgg	2280
gagctacagg ccctacgacg agggctctgcg gaggggggtt ttcatcacca acgagaccgg	2340

ccagccgctg attgggaagg tatggcccgg gtccactgcc ttccccgact tcaccaaccc 2400

cacagccctg gcctgggtggg aggacatggt ggctgagttc catgaccagg tgcctttcga 2460

cggcatgtgg attgacatga acgagccctc caacttcac agaggctctg aggacggctg 2520

ccccacaat gagctggaga acccacccia cgtgcctggg gtggttgggg ggaccctcca 2580

ggcggccacc atctgtgcct ccagccacca gtttctctcc acacactaca acctgcacaa 2640

cctctacggc ctgaccgaag ccatcgctc ccacagggcg ctggtgaagg ctcgggggac 2700

acgcccattt gtgatctccc gctcgacctt tgctggccac ggccgatacg ccggccactg 2760

gacggggggac gtgtggagct cctgggagca gctcgctcc tccgtgccag aaatcctgca 2820

gtttaacctg ctgggggtgc ctctggtcgg ggccgacgtc tgcggcttcc tgggcaacac 2880

ctcagaggag ctgtgtgtgc gctggacca gctggggggc ttctaccct tcatgcggaa 2940

ccacaacagc ctgctcagtc tgcgccagga gccgtacagc ttcagcgagc cggcccagca 3000

ggccatgagg aaggccctca ccctgcgta cgcactcctc cccacctct acacactgtt 3060

ccaccaggcc cagtcgagg gggagaccgt ggcccggccc ctcttcttg agttcccaa 3120

ggactctagc acctggactg tggaccacca gctcctgtgg ggggaggccc tgctcatcac 3180

cccagtgtc caggccggga aggcgaagt gactggctac ttccccttg gcacatggt 3240

cgacctgcag acggtgcaa tagaggccct tggcagcctc ccacccccac ctgcagctcc 3300

ccgtgagcca gccatccaca gcgaggggca gtgggtgacg ctgccggccc ccctggacac 3360

catcaacgtc cacctccggg ctgggtacat catccccctg cagggccctg gcctcacaac 3420

cacagagtc cgcagcagc ccatggccct ggctgtggcc ctgaccaagg gtggagaggc 3480

ccgaggggag ctgttctggg acgatggaga gagcctggaa gtgctggagc gaggggccta 3540

cacacagtc atcttcttg ccaggaataa cagatcgtg aatgagctg tacgtgtgac 3600

cagtggagg gctggcctgc agctgcagaa ggtgactgtc ctgggcgtgg ccacggcgcc 3660

ccagcaggtc ctctccaacg gtgtccctgt ctccaactc acctacagcc ccgacaccaa 3720

ggtcctggac atctgtgtct cgctgttgat gggagagcag tttctctca gctggtgtta 3780

gccgggcgga gtgtgttagt ctctccagag ggaggctggt tccccaggga agcagagcct 3840

gtgtgcgggc agcagctgtg tgcgggcctg ggggttgta agtgcaatta ttttaataa 3900

aaggggcatt tggaaaaaa aaaaaaagg agcagtcgac agatgaattc tgcagatctg 3960

tggcttctag ctgcccgggt ggcatccctg tgaccctcc ccagtgcctc tcctggccct 4020

ggaagttgcc actccagtgc ccaccagcct tgcctaata aaattaagtt gcatcathtt 4080

gtctgactag gtgtccttct ataatattat ggggtggagg ggggtggat ggagcaaggg 4140
gcaagttggg aagacaacct gaggtttg gattccaggc atcgagtaga taagtagcat 4200
ggcgggttaa tcattaacta cacacgtgcg gaccgagcgg cctaagctcg cgaaggaacc 4260
cctagtgatg gaggttggcca ctccctctct gcgcgtctgc tcgctcactg aggccgggcg 4320
accaaagtc gcccagacc cgggctttgc cggggcgcc tcagttagcg agcgagcgcg 4380
cagagaggga gtggccaaga cgatttaa gacaagcttg gcgtaatcat ggtcatagct 4440
gtttctgtg tgaattgtt atccgtcac aattccacac aacatacgag ccggaagcat 4500

aaagtgtaaa gctgggggtg cctaagagt gagtaactc acattaattg cgttgcgtc 4560
actgcccgt ttccagtcgg gaaacctgtc gtccagctg cattaatgaa tcggccaacg 4620
cgcggggaga ggcggtttgc gtattggcg ctcttcgct tcctcgctca ctgactcgt 4680
gcgctcggtc gttcggtcgc ggcgagcgg atcagctcac tcaaaggcgg taatacggtt 4740
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 4800
caggaacctg aaaaaggcgc gtttctggc gttttccat aggtcgcgc cccctgacga 4860
gcatcacaaa aatcgacgt caagtcagag gtggcgaaac ccgacaggac tataaagata 4920

ccaggcgttt cccctggaa gctccctcgt gcgtctcct gttccgacc tgcgcttac 4980
cggatactg tcgccttcc tccctcggg aagcgtggcg ctttctcata gtcacgctg 5040
taggtatctc agttcgggt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 5100
cgttcagccc gaccgtcgc cttatccgg taactatcgt cttgagtcca acccgtaag 5160
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt 5220
aggcgggtc acagagtct tgaagtgtg gcctaactac ggctaccta gaagaacagt 5280
atttggtatc tgcgtctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 5340

atccggcaaa caaaccacc ctggtagcgg tggtttttt gtttgaagc agcagattac 5400
gcgcagaaaa aaaggatctc aagaagatcc ttgatcttt tctacgggt ctgacgtca 5460
gtggaacgaa aactcacgtt aagggtttt ggtcatgaga ttatcaaaaa ggtcttcac 5520
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 5580
ttggtctgac agttagaaaa actcatcgag catcaaatga aactgcaatt tattcatatc 5640
aggattatca ataccatatt ttgaaaaag ccgtttctgt aatgaaggag aaaactcacc 5700
gaggcagttc cataggatgg caagatcctg gtatcggtct gcgattccga ctgctccaac 5760

atcaatacaa cctattaatt tccctcgtc aaaaataagg ttatcaagt agaaatcacc 5820
atgagtgcg actgaatccg gtgagaatgg caaaagtta tgcatttctt tccagacttg 5880
ttcaacagcg cagccattac gctcgtcatc aaaactcctc gcatcaacca aaccgttatt 5940

cattcgtgat tgcgcctgag cgagacgaaa tacgcgatcg ctgttaaaag gacaattaca 6000
aacaggaatc gaatgcaacc ggcgcaggaa cactgccagc gcatcaacaa tattttcacc 6060
tgaatcagga tattcttcta atacctggaa tgcgtttttt ccgggggatcg cagtggtag 6120
taaccatgca tcatcaggag tacggataaa atgcttgatg gtcggaagag gcataaattc 6180

cgtcagccag tttagtctga ccatctcatc tgtaacatca ttggcaacgc tacctttgcc 6240
atgtttcaga aacaactctg gcgcacggg cttcccatc aagcataga ttgtcgcacc 6300
tgattgcccg acattatcgc gagccattt ataccatata aatcagcat ccatgttga 6360
atttaatcgc ggctcgcag tttcccggtt aatatggctc atactcttc tttttcaata 6420
ttattgaagc atttatcagg gttattgtct catgagcgga tacatattg aatgtattta 6480
gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgcac ctgacgtcta 6540
agaaaccatt attatcatga cattaacctt taaaaatagg cgtatcacga ggccctttcg 6600

tctcgcgcgt ttcggtgatg acggtgaaaa cctctgacac atgcagctcc cggagacggt 6660
cacagcttgt ctgtaagcgg atgccgggag cagacaagcc cgtcaggcg cgtcagcggg 6720
tgttggcggg tgcggggct ggcttaacta tgcggcatca gagcagattg tactgagagt 6780
gcaccatag cgggtgtgaaa taccgcacag atgcgtaagg agaaaatacc gcatcaggc 6839

<210> 158
<211> 7266
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 158

gccattcgcc attcaggctg cgcaactgtt gggaaggcg atcggtagcg gcctcttcgc 60

tattaccca gctggcgaaa ggggatgtg ctgcaaggcg attaagttgg gtaacgccag 120
ggttttccca gtcacgacgt tgtaaaacga cgccagtga attctcagc atttaaatga 180
cgttggccac tccctctctg cgcgctcgt cgtcactga ggccgggcga ccaaaggctc 240
cccgacgcc gggttttgc cgggcggcct cagtgagcga gcgagcgcgc agagaggag 300
tgccaactc catcactagg ggttctgag tttaaacttc gtcgaggcgc cacgcgtagg 360
ttaattttta aaaagcagtc aaaagtcaa gtggcccttg gcagcattta ctctctctgt 420
ttgtctgtgt taataatctc aggagcacia acattcccc tgttcaaaca tgtcctaata 480

ctctgtctct gcaagggtca tcagtagttt tccatcttac tcaacatcct cccagttagt 540

catatgtttg ctactgaag gttactagtt aacaggcatc ctttaaacag gatataaaag	600
gacttcagca ggactgctcg aaacatccca cttcagatcg cctggagacg ccatccacgc	660
tgttttgacc tccatagaag acaccgggac cgatccagcc tccgcggccg ggaacgggtc	720
attggaacgc ggattcccc tgccaagagt gacgtaagta ccgcctatag actctatagg	780
cacaccctt tggtctttat gcatgaacgg tggagggcag tgtagtctga gcagtactcg	840
ttgctgccgc gcgcgccacc agacataata gctgacagac taacagactg ttcctttcca	900
tggtctttt ctgcaggcca ccatgggagt gaggcacccg cctgctccc accggtcct	960
ggcctctgc gccctcgtgt ccttggaac cgtgcactc ctggggcaca tctactcca	1020
tgatttcctg ctggttcccc gagagctgag tggctcctcc ccagtcctgg aggagactca	1080
cccagctcac cagcaggag ccagcagacc agggccccgg gatgccagg cacaccgg	1140
ccgtcccaga gcagtgcga cacagtgcga cgtcccccc aacagccgt tcgattgcgc	1200
ccctgacaag gccatcacc aggaacagt cgaggccgc ggctgctgt acatccctgc	1260
aaagcagggg ctgcaggag cccagatggg gcagccctgg tgcctcttc caccagcta	1320
ccccagctac aagctggaga acctgagctc ctctgaaatg ggctacacgg ccacctgac	1380
ccgtaccacc cccacctct tcccaagga catcctgacc ctgcggctgg acgtgatgat	1440
ggagactgag aaccgcctcc acttcacgat caaagatcca gctaacagc gctacaggt	1500
gcccttggag accccgcgtg tccacagccg ggccacctcc cactctaca gcgtggagtt	1560
ctccgaggag ccttcgggg tgatcgtgca ccggcagctg gacggccgcg tgctgtgaa	1620
cacgacggtg gcgccccgt tctttgcgga ccagttcct cagctgtcca ctcgtgcc	1680
ctcgcagtat atcacaggcc tcgccgagca cctcagtcct ctgatgtca gcaccagctg	1740
gaccaggatc accctgtgga accgggacct tgcgccacg cccggtgcga acctctacgg	1800
gtctcacctt ttctacctgg cgctggagga cggcgggtcg gcacacgggg tttctctgt	1860
aaacagcaat gccatggatg tggctctgca gccgagccct gcccttagct ggaggtcgac	1920
aggtgggac ctggatgtct acatcttctt gggcccagag cccaagagcg tggtcagca	1980
gtacctggac gttgtgggat acccgttcat gccgccatac tggggcctgg gcttcacct	2040
gtgccgtgg ggctactcct ccaccgtat caccgccag gtggtggaga acatgaccag	2100
ggccacttc cccctggacg tccaatggaa cgacctggac tacatggact cccggaggga	2160
cttcacgttc aacaaggatg gcttccggga cttccggcc atggtgcagg agctgcacca	2220
gggcggccgg cgctacatga tgatcgtgga tctgccatc agcagctcgg gccctgccgg	2280
gagctacagg cctacgacg agggctctcg gaggggggtt ttcacacca acgagaccgg	2340
ccagccgtg attgggaagg tatggcccg gtccactgcc tttcccgact tcaccaaccc	2400

cacagccctg gcctggtggg aggacatggt ggctgagttc catgaccagg tgcccttcga	2460
cggcatgtgg attgacatga acgagccttc caacttcac agaggctctg aggacggctg	2520
ccccacaat gagctggaga acccacccta cgtgcctggg gtggttgggg ggaccctcca	2580
ggcggccacc atctgtgcct ccagccacca gtttctctcc acacactaca acctgcacaa	2640
cctctacggc ctgaccgaag ccatgcctc ccacagggcg ctggtgaagg ctggggggac	2700
acgccattt gtgatctccc gctcgacctt tcttgccac ggccgatacg cggccactg	2760
gacgggggac gtgtggagct cctgggagca gctgcctcc tccgtgccag aaatcctgca	2820
gtttaacctg ctgggggtgc ctctggtcgg ggccgacgtc tgcggcttcc tgggcaacac	2880
ctcagaggag ctgtgtgtgc gctggacca gctgggggcc ttctaccct tcatgcggaa	2940
ccacaacagc ctgctcagtc tgccccagga gccgtacagc ttcagcgagc cggcccagca	3000
ggccatgagg aaggccctca cctgcgcta cgcactctc cccacctct acacactgtt	3060
ccaccaggcc cacgtcgcg ggagaccgt ggcccggccc ctcttctgg agttcccaa	3120
ggactctagc acctggactg tggaccacca gctcctgtgg ggggaggccc tgctcatcac	3180
cccagtgtc caggccggga aggccgaagt gactggtac ttccttgg gcacatgta	3240
cgacctgcag acggtgccaa tagaggcctt tggcagctc ccacccccc ctgcagctcc	3300
ccgtgagcca gccatccaca gcgaggggca gtgggtgacg ctgccggccc ccttgacac	3360
catcaacgtc cacctccggg ctgggtacat catccccctg cagggcctg gcctcacaac	3420
cacagagtc cgccagcagc ccatggcctt ggctgtggcc ctgaccaagg gtggagaggc	3480
ccgaggggag ctgttctggg acgatggaga gagcctggaa gtgctggagc gaggggccta	3540
cacacaggtc atcttctgg ccaggaataa cagatcgtg aatgagctgg tacgtgtgac	3600
cagtgaggga gctggcctgc agctgcagaa ggtgactgtc ctgggcgtgg ccacggcgcc	3660
ccagcaggtc ctctccaacg gtgtccctgt ctccaattc acctacagcc ccgacaccaa	3720
ggtcctggac atctgtgtct cgctgttgat gggagagcag tttctcgtca gctggtgtta	3780
gcgagcggcc gctcttagta gcagtatcga tcccagccca cttttccca atacgactac	3840
gagatctgtg gcttctagct gcccgggtgg catccctgtg accctcccc agtgcctctc	3900
ctggccctgg aagttgccac tccagtgcc accagccttg tcctaataaa attaatgtc	3960
atcattttgt ctgactaggt gtcttctat aatattatgg ggtggagggg ggtggtatgg	4020
agcaaggggc aagttgggt ctttaacgt cgaactgtag cactctggat agctagctag	4080
cgttttaga tagtttcaag gatagcaacc aacaacacga tagtttagaa attgttctag	4140
atcttctgct actactcgag agatagttga ctctagaag actctggata gagagcttcc	4200

tgtggtttct atttcaacac ttcgagcgat ttagatagag gcgtaattg agaacaaaaa	4260
ctggacacct gtttctagaa cactagtgc tgccttctca aggccttgata cgtcctgcca	4320
ctcacacaca aaaaaccaac acacagatta atgaaaataa agatctttta ttagagagcg	4380
agtggatagt ttagagtcgc ttgagactta tctgtgtcc aactagccaa gagtagagag	4440
agagtgacta ggtggatcga gagatcgctg tctatttgac tgctagcctg tgttctcgtc	4500
gtagtccaaa tcttcaatac tctggcgact agtaacttga atctgtgtat ctaacacgac	4560
aaggctactc ctttgagitt ctagtgataa gtgcgactgg atctagctag ttgagtttct	4620
agaaactacg tccaaagcta ctagcactac acgtgcggac cgagcggcct aagctcgca	4680
aggaaccctt agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg	4740
ccgggcgacc aaaggtcgcc cgacgcccg gctttgccc ggccgctca gtgagcgagc	4800
gagcgcgcag agagggagtg gccaaagcga tttaaatgac aagcttggcg taatcatggt	4860
catagctgtt tctgtgtga aattgttatt cgctcacaat tccacacaac atacgagccg	4920
gaagcataaa gtgtaaagcc tggggtgcct aatgagttag ctaactcaca ttaattgcgt	4980
tgcgctcact gcccgcttcc cagtgcggaa acctgtcgtg ccagctgcat taatgaatcg	5040
gccaacgcgc ggggagagcg ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg	5100
actcgctgcg ctcggctggt cggtgcggc gagcggatc agctcactca aaggcggtaa	5160
tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc	5220
aaaaggccag gaaccgtaaa aaggccgcgt tgctggcggt tttccatagg ctccgcccc	5280
ctgacgagca tcacaaaaat cgacgtcaa gtgagagggt gcgaaaccg acaggactat	5340
aaagatacca ggcgtttccc cctggaagct cctcgtgcg ctctcctgtt ccgacctgc	5400
cgcttaccgg atacctgtcc gcctttctcc ctccgggaag cgtggcgctt tctcatagct	5460
cacgtgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg	5520
aacccccgt tcagcccgac cgctgcgcct tatccgtaa ctatcgtctt gattccaacc	5580
cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga	5640
ggtatgtagg cgggtgctaca gatttcttga agtgggtggc taactacggc tacactagaa	5700
gaacagtatt tggatctgct gctctgctga agccagttac cttcgaaaa agagttggta	5760
gctcttgatc cggcaaaaa accaccgtg gtagcggtgg ttttttgtt tgcaagcagc	5820
agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg	5880
acgtcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga	5940

tcttcaccta gatcctttta aattaaaaat gaagttttta atcaatctaa agtatatatg	6000
agtaaacttg gtctgacagt tagaaaaact catcgagcat caaatgaaac tgcaatttat	6060
tcataatcagg attatcaata ccatatTTTT gaaaaagccg tttctgtaat gaaggagaaa	6120
actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg attccgactc	6180
gtccaacatc aatacaacct attaatTTCC cctcgtcaaa aataaggTTa tcaagtgaga	6240
aatcaccatg agtgacgact gaatccggtg agaatggcaa aagtttatgc atttctttcc	6300
agacttgTtc aacaggccag ccattacgct cgtcatcaaa atcactcgca tcaaccaaac	6360
cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg ttaaaaggac	6420
aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca tcaacaatat	6480
tttcacctga atcaggatat tcttctaata cctggaatgc tgtttttccg gggatcgcag	6540
tgggtagtaa ccatgcata tcaggagtac ggataaaatg cttgatggtc ggaagaggca	6600
taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg gcaacgctac	6660
ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaag cgatagattg	6720
tcgcacctga ttgcccgaca ttatcgcgag cccatttata cccatataaa tcagcatcca	6780
tgttggaatt taatcgcggc ctgcagttt cccgttgaat atggctcata ctcttccttt	6840
ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat	6900
gtatttagaa aaataaaca ataggggttc cgcgcacatt tccccgaaaa gtgccacctg	6960
acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc	7020
cctttcgtct cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg cagctcccgg	7080
agacggtcac agcttgctg taagcggatg ccgggagcag acaagcccgt cagggcgcgt	7140
cagcgggtgt tggcgggtgt cggggctggc ttaactatgc ggcatcagag cagattgtac	7200
tgagagtgca ccatatgcgg tgtgaaatac cgcacagatg cgtaaggaga aaataccgca	7260
tcaggc	7266
<210> 159	
<211> 6629	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 159	
gccattcgcc attcaggtcg cgcaactgtt gggaaggcg atcgggtcgg gcctcttcgc	60
tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag	120

ggttttccca gtcacgacgt tgtaaacga cgccagtgga attcctcgag atttaaatga 180

cgttggccac tccctctctg cgcgtctgct cgctcactga ggccgggcga ccaaaggtcg 240

cccgaagccc gggttttgcc cgggcggcct cagtgcgcga gcgagcgcgc agagaggag 300

tggccaactc catcactagg ggttctctgag tttaaaacttc gtagaggccg cacgcgtagg 360

ttaattttta aaaagcagtc aaaagtccaa gtggcccttg gcagcattta ctctctctgt 420

ttgtctctgt taataatctc aggagcacia acattccccc tgttcaaaca tgtcctaata 480

ctctgtctct gcaagggtca tcagtgttt tccatcttac tcaacatcct ccagtgagt 540

catatgtttg ctactgaag gttactagt aacaggcatc ccttaaacag gatataaaag 600

gacttcagca ggactgctcg aaacatccca ctgagccacc ggtaccgagc tcgcatccac 660

tagtaacggc cgccagtgtg ctggaattcg gctttccagg ccatctcaa ccatgggagt 720

gaggcacccg cctgtctccc accggtcctt ggccgtctgc gccctcgtgt ccttggcaac 780

cgctgcactc ctggggcaca tctactcca tgatttcttg ctggttcccc gagagctgag 840

tggctcctcc ccagtctgg aggagactca ccagctcac cagcaggag ccagcagacc 900

agggccccgg gatgccagg cacaccccg ccgtcccaga gcagtgccca cacagtgcga 960

ctccccccc aacagccgct tcgattgcgc cctgacaag gccatcacc aggaacagtg 1020

cgaggccgc ggctgctgct acatccctgc aaagcagggg ctgcaggag cccagatggg 1080

gcagccctgg tgcttttcc caccagctca cccagctac aagctggaga acctgagctc 1140

ctctgaaatg ggctacacgg ccacctgac ccgtaccacc cccaccttct tccccagga 1200

catctgacc ctgcggtgg acgtgatgat ggagactgag aaccgcctcc acttcacgat 1260

caaagatcca gctaacaggg gctacagggt gcccttggag acccgcgtg tccacagccg 1320

ggcaccgtcc cactctaca gcgtggagt ctccgaggag ccttcgggg tgatctgca 1380

ccggcagctg gacggccggt tgctgctgaa cagcagggtg gcgccccgt tctttcgga 1440

ccagttcctt cagctgtcca cctcgtgcc ctgcagtat atcacaggcc tcgccgagca 1500

cctcagtcct ctgatgtca gcaccagtg gaccaggatc acctgtgga accgggacct 1560

tgcgccacg cccggtgcga acctctacgg gtctaccct ttctacctgg cgctggagga 1620

cggcgggtcg gcacacgggg tgttctgct aaacagcaat gccatggatg tggctctgca 1680

gccgagccct gcccttagct ggaggtcgac aggtgggac ctggatgtct acatcttct 1740

gggccagag cccaagagcg tggtcagca gtacctggac gttgtggat accggttcat 1800

gccgccatac tggggcctgg gcttcacct gtgccgtgg ggctactcct ccaccgtat 1860

cacccgccag gtggtggaga acatgaccag ggcccacttc cccctggacg tccaatggaa	1920
cgacctggac tacatggact cccggaggga cttcacgttc aacaaggatg gcttcaggga	1980
cttcccgcc atggtgcagg agctgcacca gggcgccgg cgctacatga tgatcgtgga	2040
tcctgccatc agcagctcgg gccctgccgg gagctacagg ccctacgacg agggctctcg	2100
gaggggggtt ttcatcacca acgagaccgg ccagccgtg attgggaagg tatggccgg	2160
gtccactgcc ttccccgact tcaccaaccc cacagccctg gcctggtggg aggacatggt	2220
ggctgagttc catgaccagg tgcccttcga cggcatgtgg attgacatga acgagccttc	2280
caacttcac agaggctctg aggacggctg cccaacaat gagctggaga acccaccta	2340
cgtgcctggg gtggttgggg ggacctcca ggccggccacc atctgtgcct ccagccacca	2400
gtttctctcc acacactaca acctgcacaa cctctacggc ctgaccgaag ccatcgcctc	2460
ccacaggcgg ctggtgaagg ctggggggac acgcccattt gtgatctccc gctcgacctt	2520
tgctggccac ggccgatagc ccggccactg gacgggggac gtgtggagct cctgggagca	2580
gctcgctcc tccgtgccag aaatcctgca gttaacctg ctgggggtgc ctctggtcgg	2640
ggccgacgtc tgcggcttcc tgggcaacac ctgagaggag ctgtgtgtgc gctggacca	2700
gctgggggcc ttctaccct tcatgcggaa ccacaacagc ctgctcagtc tgccccagga	2760
gccgtacagc ttcagcgagc cggcccagca ggccatgagg aaggccctca cctgcgcta	2820
cgcactctc cccacctct acacactgtt ccaccaggcc cactcgcgg gggagaccgt	2880
ggccccggcc ctcttctgg agttcccaa ggactctagc acctggactg tggaccacca	2940
gctcctgtgg ggggaggccc tgcctacac ccagtgctc caggccggga aggccgaagt	3000
gactggctac ttccccctgg gcacatggta cgacctgcag acggtgcaa tagaggccct	3060
tggcagctc ccacccccac ctgcagctcc ccgtgagcca gccatccaca gcgaggggca	3120
gtgggtgacg ctcccgccc ccctggacac catcaacgtc cacctccggg ctgggtacat	3180
catccccctg cagggccctg gcctcacaac cacagagtcc cgccagcagc ccatggccct	3240
ggctgtggcc ctgaccaagg gtggagaggc ccgaggggag ctgttctggg acgatggaga	3300
gagcctggaa gtgtggagc gaggggccta cacacaggtc atcttctgg ccaggaataa	3360
cacgatcgtg aatgagctgg tacgtgtgac cagttaggga gctggcctgc agctgcagaa	3420
ggtgactgtc ctgggctgg ccacggcgcc ccagcaggtc ctctccaacg gtgtccctgt	3480
ctccaacttc acctacagcc ccgacaccaa ggtcctggac atctgtgtct cgctgttgat	3540
gggagagcag tttctctga gctggtgtta gccgggcgga gtgtgttagt ctctccagag	3600
ggaggtggt tccccaggga agcagagcct gtgtgcgggc agcagctgtg tgcgggcctg	3660
ggggttgtta agtgcaatta ttttaataa aaggggcatt tggaaaaaa aaaaaaggt	3720

agcagtcgac agatgaattc tgcagatctg tggcttctag ctgcccgggt ggcatccctg 3780
 tgaccctcc ccagtgcctc tcttggcctt ggaagttgcc actccagtgc ccaccagcct 3840
 tgtcctaata aaattaagtt gcatcatctt gtctgactag gtgtccttct ataataattat 3900
 ggggtggagg ggggtgggat ggagcaaggg gcaagttggg aagacaacct gagttgttgg 3960

 gattccaggc atcgagtaga taagtagcat ggcggtttaa tcattaacta cacacgtgcg 4020
 gaccgagcgg cctaagctcg cgaaggaacc ctagtgatg gagttggcca ctccctctct 4080
 gcgcgtctcg tcgtcactg aggcggggcg accaaaggct gcccagacgc cgggctttgc 4140
 cggggcgccc tcagtgcgcg agcgagcgcg cagagaggga gtggccaaga cgatttaaat 4200
 gacaagcttg gcgtaatcat ggtcatagct gtttctgtg tgaaattgtt atccgtcac 4260
 aattccacac aacatacgag ccggaagcat aaagtgtaaa gcctggggtg cctaagagt 4320
 gagctaactc acattaattg cgttgcgtc actgcccgtt tccagtcgg gaaacctgtc 4380

 gtgccagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc gtattgggcg 4440
 ctcttcgct tctcgtca ctgactcgt gcgtcggtc gttcggctgc ggcgagcgg 4500
 atcagctcac tcaaaggcgg taatacgtt atccacagaa tcaggggata acgcaggaaa 4560
 gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggcgg cgttgctggc 4620
 gtttttccat aggtccgcc cccctgacga gcatcacaaa aatcgacgt caagtcagag 4680
 gtggcgaaac ccgacagcac tataaagata ccaggcgtt cccctggaa gtcctcgt 4740
 gcgtctcct gtccgaccc tgccgcttac cggatactg tccgccttc tccctcggg 4800

 aagcgtggcg ctttctcata gtcacgctg taggtatctc agttcgtgt aggtcgttcg 4860
 ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg cttatccgg 4920
 taactatcgt cttgagtcca acccggttaag acacactta tcgccactgg cagcagccac 4980
 tggtaacagg attagcagag cgaggtatgt aggcgggtct acagagttct tgaagtgggtg 5040
 gcctaactac ggctacacta gaagaacagt atttggatc tgcgtctgc tgaagccagt 5100
 taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg 5160
 tggttttttt gtttgaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc 5220

 tttgatctt tctacggggt ctgacgtca gtggaacgaa aactcacgtt aagggtttt 5280
 ggtcatgaga ttatcaaaaa ggatcttcac ctgatacctt ttaaattaaa aatgaagttt 5340
 taaatcaatc taaagtatat atgagtaaac ttggtctgac agttagaaaa actcatcgag 5400
 catcaaatga aactgcaatt tattcatatc aggattatca ataccatatt ttgaaaaag 5460
 ccgtttctgt aatgaaggag aaaactcacc gaggcagttc cataggatgg caagatcctg 5520
 glatcggctt gcgattccga ctctccaac atcaatacaa cctattaatt tcccctcgtc 5580

aaaaataagg ttatcaagtg agaaatcacc atgagtgcg actgaatccg gtgagaatgg 5640

 caaaagttaa tgcatttctt tccagacttg ttcaacaggc cagccattac gctcgtcatc 5700
 aaaatcactc gcatcaacca aaccgttatt cattcgtgat tgcgcctgag cgagacgaaa 5760
 tacgcgatcg ctgttaaaag gacaattaca aacaggaatc gaatgcaacc ggccgaggaa 5820
 cactgccagc gcatcaacaa tattttcacc tgaatcagga tattcttcta atacctggaa 5880
 tgctgttttt ccggggatcg cagtgggtgag taacctatga tcatcaggag tacggataaa 5940
 atgcttgatg gtcggaagag gcataaattc cgtcagccag tttagtctga ccatctcatc 6000
 tgtaacatca ttggcaacgc tacctttgcc atgtttcaga aacaactctg gcgcacggg 6060

 ctcccatcac aagcgaatga ttgtgcgacc tgattgccc acattatcgc gagcccatct 6120
 ataccatata aaatcagcat ccatgttga attaatcgc ggccctgcagc tttccgcttg 6180
 aatatggctc atactcttc tttttcaata ttattgaagc atttatcagg gttattgtct 6240
 catgagcgga tacatatctg aatgtattta gaaaaataaa caaatagggg ttccgcgcac 6300
 atttccccga aaagtgcac ctgacgtcta agaaaccatt attatcatga cattaaccta 6360
 taaaaatagg cgtatcacga ggccctttcg tctgcgcgt ttcggtgatg acggtgaaaa 6420
 cctctgacac atgcagctcc cggagacggt cacagcttgt ctgtaagcgg atgccgggag 6480

 cagacaagcc cgtcagggcg cgtcagcggg tgttggcggg tgcggggct ggcttaacta 6540
 tgcggcatca gagcagattg tactgagagt gcaccatag cggtgtgaaa taccgcacag 6600
 atgcgtaagg agaaaatacc gcatcagcg 6629

 <210> 160
 <211> 7132
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 polynucleotide
 <400> 160

 gccattcgcc attcaggctg cgcaactgtt gggaaggcg atcgggtcgg gcctcttcgc 60
 tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag 120

 ggttttccca gtcacgacgt tgtaaaacga cgccagtgga attcctcgag atttaaatga 180
 cgttggccac tccctctctg cgcgtcgtc cgtcactga ggccgggcga ccaaaggctc 240
 cccgacgccc gggttttgc cgggcgccct cagtgcgga gcgagcgcgc agagaggag 300
 tggccaactc catcactagg ggttctctg tttaaacttc gtcgaggccg cacgcgtgaa 360

ttacgtcata gggttaggga ggtcggccgc tctaggagtt aatttttaaa aagcagtcaa	420
aagtccaagt gcccttgca gcatctactc tctctgtttg ctctggttaa taatctcagg	480
agcacaaca ttccttacta gttctaggag ttaattttta aaaagcagtc aaaagtccaa	540
gtgcccttgc gagcatttac tctctctgtt tgctctggtt aataatctca ggagcacaaa	600
cattccttac tagttctaga gcggccgcca gtgtgctgga attcggtttt tttagggtg	660
gaagctacct ttgacatcat ttctctgcg aatgcatgta taatttctac agaacctatt	720
agaaaggatc acccagcctc tgcttttgta caactttccc ttaaaaaact gccaatccca	780
ctgctgtttg gcccaatagt gagaactttt tcctgctgcc tcttggtgct tttgcctatg	840
gccccattc tgctgctga agacactctt gccagcatgg acttaaaccc ctccagctct	900
gacaatcctc tttctctttt gttttacatg aagggtctgg cagccaaagc aatcactcaa	960
agttcaaacc ttatcatttt ttgctttgtt cctcttggcc ttggttttgt acatcagctt	1020
tgaaaatacc atcccagggt taatgctggg gttaatattat aactgagagt gctctagttc	1080
tgcaatacag gacatgctat aaaaatggaa agatgttgct ttctgagaga tcagcttaca	1140
tgtggtaccg agctcgatc cactagtaac ggccgccagt gtgctggaat tcggctttcc	1200
aggccatctc caaccatggg agtgaggcac ccgccctgct cccaccggct cctggccgtc	1260
tgcgccctcg tgccttggc aaccgctgca ctcttggggc acatcctact ccatgatttc	1320
ctgctggttc cccgagagct gagtggctcc tccccagtc ttgaggagac tcaccagct	1380
caccagcagg gagccagcag accagggccc cgggatgccc aggcacaccc cggccgtccc	1440
agagcagtgc ccacacagtg cgacgtcccc cccaacagcc gcttcgattg cggccctgac	1500
aaggccatca ccaggaaca gtgcgaggcc cgcggctgct gctacatccc tgcaaagcag	1560
gggctgcagg gagcccatg ggggcagccc tggctgttct tccccaccag ctacccagc	1620
tacaagctgg agaacctgag ctctctgaa atgggctaca cggccaccct gacccgtacc	1680
acccccacct tcttcccaa ggacatcctg accctgcggc tggacgtgat gatggagact	1740
gagaaccgcc tccattcac gatcaaagat ccagctaaca ggcgtacga ggtgcccttg	1800
gagacccgc gtgtccacag ccgggcaccg tccccactct acagcgtgga gttctccgag	1860
gagcccttcg gggtagctgt gcaccggcag ctggacggcc gcgtgctgct gaacacgacg	1920
gtggcgcgcc tgttcttgc ggaccagttc ctctagctgt ccacctcgt gacctcgag	1980
tatatcacag gcctcgcca gcacctcagt cccctgatgc tcagcaccag ctggaccagg	2040
atcacctgt ggaaccggga ccttgcgcc acgcccgggtg cgaacctcta cgggtctcac	2100
cctttctacc tggcgttga ggacggcggg tcggcacacg ggggtttct gctaaacagc	2160

aatgccatgg atgtggctct gcagccgagc cctgccctta gctggaggtc gacaggtggg	2220
atcctggatg tctacatctt cctgggcca gagccaaga gcgtggtgca gcagtacctg	2280
gacgttgtgg gatacccggt catgccgcca tactggggcc tgggcttcca cctgtgccgc	2340
tggggctact cctccaccgc tatcaccgc caggtggtgg agaacaatgac cagggccac	2400
ttccccctgg acgtccaatg gaacgacctg gactacatgg actcccgag ggacttcacg	2460
ttcaacaagg atggcttcg ggacttcccg gccatggtgc aggagctgca ccaggcggc	2520
cggcgctaca tgatgatcgt ggatcctgcc atcagcagct cgggccctgc cgggagctac	2580
aggccctacg acgagggtct gcggaggggg gttttcatca ccaacgagac cggccagccg	2640
ctgattggga aggtatggcc cgggtccact gccttccccg acttaccaa cccacagcc	2700
ctggcctggt gggaggacat ggtggtgag ttccatgacc aggtgccctt cgacggcatg	2760
tggattgaca tgaacgacc ttccaacttc atcagaggct ctgaggacgg ctgccccaac	2820
aatgagctgg agaaccacc ctacgtgcct ggggtggttg gggggaccct ccaggcggc	2880
accatctgtg cctccagcca ccagtttctc tccacacact acaacctgca caacctctac	2940
ggcctgaccg aagccatcgc ctcccacagg gcgctggtga aggtcgggg gacacgcca	3000
tttgtgatct cccgctcgac ctttctggc caggccgat acgccggcca ctggacgggg	3060
gacgtgtgga gtcctggga gcagctgcc tcctccgtgc cagaaatcct gcagtttaac	3120
ctgctggggg tgcctctggt cggggccgac gtctgcggt tcctgggcaa cacctcagag	3180
gagctgtgtg tgcgctggac ccagctgggg gccttctacc cttcatgcg gaaccacaac	3240
agcctgctca gtctgcccc ggagccgtac agcttcagcg agccggccca gcaggccatg	3300
aggaaggccc tcacctgcg ctacgcactc ctccccacc tctacacact gtccaccag	3360
gcccacgtcg cgggggagac cgtggccccg cccctcttcc tggagtccc caaggactct	3420
agcacctgga ctgtggacca ccagctcctg tggggggagg cctgctcat caccacagt	3480
ctccaggccg ggaaggccga agtgactggc tacttccct tgggcacatg gtacgacctg	3540
cagacggtgc caatagagc ccttggcagc ctcccaccc cacctgcagc tccccgtgag	3600
ccagccatcc acagcgagg gcagtgggtg acgtgccgg cccccctgga caccatcaac	3660
gtccacctcc gggctgggt catcatccc ctgcagggcc ctggcctcac aaccacagag	3720
tcccgcagc agcccatggc cctggctgtg gccctgacca aggtggaga ggcccagggg	3780
gagctgttct gggacgatgg agagagcctg gaagtgtg agcgaggggc ctacacacag	3840
gtcatcttcc tggccaggaa taacacgatc gtgaatgagc tggtagtgt gaccagttag	3900

ggagctggcc tgcagctgca gaaggtgact gtcctgggcg tggccacggc gccccagcag 3960
 gtcctctcca acggtgtccc tgtctccaac ttcacctaca gccccgacac caaggtcctg 4020
 gacatctgtg tctcgtgttt gatgggagag cagtttctcg tcagctggtg ttagccgggc 4080
 ggagtgtgtt agtctctcca gagggaggct ggttccccag ggaagcagag cctgtgtgcg 4140
 ggcagcagct gtgtgcgggc ctgggggttg ttaagtgcaa ttatttttaa taaaaggggc 4200
 atttgaaaa aaaaaaaaaa ggtagcagtc gacagatgaa ttctgcagat ctgtggcttc 4260
 tagctgcccc ggtagcatcc ctgtgacccc tccccagtgc ctctcctggc cctggaagtt 4320

 gccactccag tgcccaccag ccttgtctta ataaaattaa gttgcatcat tttgtctgac 4380
 taggtgtcct tctataatat tatgggggtg aggggggttg tatggagcaa ggggcaagtt 4440
 gggaagacaa cctgagttgt tgggattcca ggcacagagt agataagtag catggcgggt 4500
 taatcattaa ctacacagct gcggaccgag cgccctaagc tcgcgaagga acccctagtg 4560
 atggagttag ccactccctc tctgcgcgtc cgtcgcctca ctgaggccgg gcgaccaaag 4620
 gtcgcccagc gcccgggctt tgcccgggcg gcctcagtga gcgagcgagc gcgcagagag 4680
 ggagtggcca agacgattta aatgacaagc ttggcgtaat catggtcata gctgtttcct 4740

 gtgtgaaatt gttatccgt cacaattcca cacaacatac gagccggaag cataaagtt 4800
 aaagcctggg gtgcctaata agtgagctaa ctacattaa ttgcgttgcg ctactgccc 4860
 gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcgcca acgcgcgggg 4920
 agaggcgggt tgcgtattgg gcgctcttcc gcttctcgc tcactgactc gctgcgtcgc 4980
 gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cgtaataacg gttatccaca 5040
 gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac 5100
 cgtaaaaagg ccgcgttgct ggcgtttttc cataggtccc gccccctga cgagcatcac 5160

 aaaaatcgac gctcaagtca gaggtggcga aaccgacag gactataaag ataccaggcg 5220
 tttcccctg gaagctccct cgtgcgtctc cctgttccga cctgccgct taccggatac 5280
 ctgtccgct tttcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat 5340
 ctgattcgg ttaggtcgt tcgctccaag ctgggctgtg tgcacgaacc cccggttcag 5400
 cccgaccgct gcgccttacc cggttaactat cgtcttgagt ccaaccgggt aagacacgac 5460
 ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta ttaggcgggt 5520
 gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac agtatttggt 5580

 atctgcgtc tgcgaagcc agttacctc ggaaaaagag ttggtagctc ttgatccggc 5640
 aaacaaacca ccgtggtag cggtaggtttt ttgtttgca agcagcagat tacgcgcaga 5700
 aaaaaaggat ctcaagaaga tctttgatc tttctacgg ggtctgacgc tcagtggaac 5760

gaaaactcac gtttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc 5820
cttttaaatt aaaaatgaag ttttaaatca atctaaagta tataatgagta aacttggctc 5880
gacagttaga aaaactcatc gagcatcaaa tgaactgca atttattcat atcaggatta 5940
tcaataccat atttttgaaa aagccgtttc tgaatgaag gagaaaactc accgaggcag 6000

ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata 6060
caacctatta atttccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 6120
acgactgaat ccggtgagaa tggcaaaagt ttatgcattt ctttccagac ttgttcaaca 6180
ggccagccat tacgtctgtc atcaaaatca ctgcgatcaa ccaaaccgtt attcattcgt 6240
gattgcgcct gagcgagacg aaatacgcg aatcgctgtta aaggacaatt acaaacagga 6300
atcgaatgca accggcgagc gaacactgcc agcgcatcaa caatattttc acctgaatca 6360
ggatatcttt ctaatacctg gaatgctgtt tttccgggga tcgcagtggg gagtaacat 6420

gcatcatcag gatlacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 6480
cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc 6540
agaaacaact ctggcgcatc gggcttccca tacaagcgat agattgtcgc acctgattgc 6600
ccgacattat cgcgagccca tttatacca tataaatcag catccatgtt ggaatttaat 6660
cgcggcctcg acgtttcccg ttgaatatgg ctcatactct tcctttttca atattattga 6720
agcatttate agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat 6780
aaacaaatag gggttccgag cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc 6840

attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtctcgcg 6900
cgtttcgggtg atgacgggtga aaacctctga cacatgcagc tcccggagac ggtcacagct 6960
tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg gcgcgtcagc ggggtgttggc 7020
gggtgtcggg gctggtctaa ctatcgggca tcagagcaga ttgtactgag agtgcacat 7080
atgcgggtgtg aaataccgca cagatgcgta aggagaaaat accgcatcag gc 7132

<210> 161
<211> 145
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 161
ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc 60

cgacgccccg gctttgcccc ggcggcctca gtgagcgagc gagcgcgag agaggggagtg 120

gccaaactcca tcactagggg ttctt 145

<210> 162

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 162

gagttttaaac ttcgtcgagg ccgcacgcgt 30

<210> 163

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 163

ggtaccgagc tcgcatccac tagtaacggc cgccagtgtg ctggaat 47

<210> 164

<211> 193

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 164

ctgtggcttc tagctgcccc ggtggcatcc ctgtgacccc tccccagtgc ctctcctggc 60

cctggaagtt gccactccag tgcccaccag ccttgtccta ataaaattaa gttgcatcat 120

tttgtctgac taggtgtcct tctataatat tatgggggtgg aggggggtgg tatggagcaa 180

ggggcaagtt ggg 193

<210> 165

<211> 145

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 165

aggaaccct agtatggag ttggccactc cctctctgcg cgctcgctcg ctactgagg	60
ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggccgacctca gtgagcgagc	120
gagcgcgag agagggagtg gccaa	145

<210> 166

<211> 193

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 166

ctgtggcttc tagtgcccc ggtggcatcc ctgtgacccc tccccagtgc ctctctggc	60
cctggaagtt gccactccag tgcccaccag cctgtccta ataaaattaa gttgcatcat	120
tttgtctgac tagtgtcct tctataatat tatgggggtgg aggggggtgg tatggagcaa	180
ggggcaagtt ggg	193

<210> 167

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 167

cacgtcgga ccgagcgcc taagctcgcg a	31

<210> 168

<211> 611

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 168

gtccttcaac gtcgaactgt agcactctgg atagctagct agcgtttgta gatagtttca	60

aggatagcaa ccaacaacac gatagtttag aaattgttct agatcttctg ctactactcg 120
 agagatagtt gacttctaga agactctgga tagagagctt cctgtgggtt ctatttcaac 180
 acttcgagcg atttagatag aggcgtcaat tgagaacaaa aactggacac ctgtttctag 240

aacactagtg ctagtcttct caaggcttga tacgtcctgc cactcacaca caaaaaacca 300
 acacacagat taatgaaaat aaagatcttt tattagagag cgagtggata gtttagagtc 360
 gcttgagact tatcctgtgt ccaactagcc aagagtagag agagagtgac taggtggatc 420
 gagagatcgt cgctctattg actgctagcc tgtgttctcg tcgtagtcca aatcttcaat 480
 actctggcga ctagtaactt gaatctgtgt atctaacacg acaaggctac tcctttgagt 540
 ttctagtgat aagtgcgact ggatctagct agttgagttt ctagaaacta cgtccaaagc 600
 tactagcact a 611

<210> 169

<400> 169

000

<210> 170

<211> 952

<212> PRT

<213> Homo sapiens

<400> 170

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

1 5 10 15

Ala Leu Val Ser Leu Ala Thr Ala Ala Leu Leu Gly His Ile Leu Leu

20 25 30

His Asp Phe Leu Leu Val Pro Arg Glu Leu Ser Gly Ser Ser Pro Val

35 40 45

Leu Glu Glu Thr His Pro Ala His Gln Gln Gly Ala Ser Arg Pro Gly

50 55 60

Pro Arg Asp Ala Gln Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr

65 70 75 80

Gln Cys Asp Val Pro Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys

85 90 95

Ala Ile Thr Gln Glu Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro

100 105 110

Ala Lys Gln Gly Leu Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe

115 120 125
Phe Pro Pro Ser Tyr Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser

130 135 140
Glu Met Gly Tyr Thr Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe

145 150 155 160
Pro Lys Asp Ile Leu Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu

165 170 175
Asn Arg Leu His Phe Thr Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu

180 185 190
Val Pro Leu Glu Thr Pro Arg Val His Ser Arg Ala Pro Ser Pro Leu

195 200 205
Tyr Ser Val Glu Phe Ser Glu Glu Pro Phe Gly Val Ile Val His Arg

210 215 220
Gln Leu Asp Gly Arg Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe

225 230 235 240
Phe Ala Asp Gln Phe Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr

245 250 255
Ile Thr Gly Leu Ala Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser

260 265 270
Trp Thr Arg Ile Thr Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly

275 280 285
Ala Asn Leu Tyr Gly Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly

290 295 300
Gly Ser Ala His Gly Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val

305 310 315 320
Val Leu Gln Pro Ser Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile

325 330 335
Leu Asp Val Tyr Ile Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln

340 345 350
Gln Tyr Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly

355 360 365
 Leu Gly Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr

 370 375 380
 Arg Gln Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val
 385 390 395 400
 Gln Trp Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe
 405 410 415
 Asn Lys Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His
 420 425 430
 Gln Gly Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser

 435 440 445
 Ser Gly Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg
 450 455 460
 Gly Val Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val
 465 470 475 480
 Trp Pro Gly Ser Thr Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu
 485 490 495
 Ala Trp Trp Glu Asp Met Val Ala Glu Phe His Asp Gln Val Pro Phe

 500 505 510
 Asp Gly Met Trp Ile Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly
 515 520 525
 Ser Glu Asp Gly Cys Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val
 530 535 540
 Pro Gly Val Val Gly Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser
 545 550 555 560
 Ser His Gln Phe Leu Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly

 565 570 575
 Leu Thr Glu Ala Ile Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly
 580 585 590
 Thr Arg Pro Phe Val Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg
 595 600 605

Tyr Ala Gly His Trp Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu
 610 615 620
 Ala Ser Ser Val Pro Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro

 625 630 635 640
 Leu Val Gly Ala Asp Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu
 645 650 655
 Leu Cys Val Arg Trp Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg
 660 665 670
 Asn His Asn Ser Leu Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser
 675 680 685
 Glu Pro Ala Gln Gln Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala

 690 695 700
 Leu Leu Pro His Leu Tyr Thr Leu Phe His Gln Ala His Val Ala Gly
 705 710 715 720
 Glu Thr Val Ala Arg Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser
 725 730 735
 Thr Trp Thr Val Asp His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile
 740 745 750
 Thr Pro Val Leu Gln Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro

 755 760 765
 Leu Gly Thr Trp Tyr Asp Leu Gln Thr Val Pro Ile Glu Ala Leu Gly
 770 775 780
 Ser Leu Pro Pro Pro Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser
 785 790 795 800
 Glu Gly Gln Trp Val Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val
 805 810 815
 His Leu Arg Ala Gly Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr

 820 825 830
 Thr Thr Glu Ser Arg Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr
 835 840 845
 Lys Gly Gly Glu Ala Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser

850 855 860
 Leu Glu Val Leu Glu Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala
 865 870 875 880
 Arg Asn Asn Thr Ile Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly

885 890 895
 Ala Gly Leu Gln Leu Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala
 900 905 910
 Pro Gln Gln Val Leu Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr
 915 920 925
 Ser Pro Asp Thr Lys Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly
 930 935 940
 Glu Gln Phe Leu Val Ser Trp Cys
 945 950

<210> 171

<211> 952

<212> PRT

<213> Homo sapiens

<400> 171

Met Gly Val Arg His Pro Pro Cys Ser His Arg Leu Leu Ala Val Cys
 1 5 10 15
 Ala Leu Val Ser Leu Ala Thr Ala Ala Leu Leu Gly His Ile Leu Leu
 20 25 30
 His Asp Phe Leu Leu Val Pro Arg Glu Leu Ser Gly Ser Ser Pro Val
 35 40 45
 Leu Glu Glu Thr His Pro Ala His Gln Gln Gly Ala Ser Arg Pro Gly

50 55 60
 Pro Arg Asp Ala Gln Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr
 65 70 75 80
 Gln Cys Asp Val Pro Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys
 85 90 95
 Ala Ile Thr Gln Glu Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro
 100 105 110

Ala Lys Gln Gly Leu Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe

115 120 125
Phe Pro Pro Ser Tyr Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser

130 135 140
Glu Met Gly Tyr Thr Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe

145 150 155 160
Pro Lys Asp Ile Leu Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu

165 170 175
Asn Arg Leu His Phe Thr Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu

180 185 190
Val Pro Leu Glu Thr Pro Arg Val Leu Ser Arg Ala Pro Ser Pro Leu

195 200 205
Tyr Ser Val Glu Phe Ser Glu Glu Pro Phe Gly Val Ile Val His Arg

210 215 220
Gln Leu Asp Gly Arg Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe

225 230 235 240
Phe Ala Asp Gln Phe Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr

245 250 255
Ile Thr Gly Leu Ala Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser

260 265 270
Trp Thr Arg Ile Thr Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly

275 280 285
Ala Asn Leu Tyr Gly Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly

290 295 300
Gly Ser Ala His Gly Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val

305 310 315 320
Val Leu Gln Pro Ser Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile

325 330 335
Leu Asp Val Tyr Ile Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln

340 345 350
Gln Tyr Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly

355 360 365
Leu Gly Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr

370 375 380
Arg Gln Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val

385 390 395 400
Gln Trp Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe

405 410 415
Asn Lys Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His

420 425 430
Gln Gly Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser

435 440 445
Ser Gly Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg

450 455 460
Gly Val Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val

465 470 475 480
Trp Pro Gly Ser Thr Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu

485 490 495
Ala Trp Trp Glu Asp Met Val Ala Glu Phe His Asp Gln Val Pro Phe

500 505 510
Asp Gly Met Trp Ile Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly

515 520 525
Ser Glu Asp Gly Cys Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val

530 535 540
Pro Gly Val Val Gly Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser

545 550 555 560
Ser His Gln Phe Leu Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly

565 570 575
Leu Thr Glu Ala Ile Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly

580 585 590
Thr Arg Pro Phe Val Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg

595 600 605

Tyr Ala Gly His Trp Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu
 610 615 620
 Ala Ser Ser Val Pro Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro

 625 630 635 640
 Leu Val Gly Ala Asp Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu
 645 650 655
 Leu Cys Val Arg Trp Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg
 660 665 670
 Asn His Asn Ser Leu Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser
 675 680 685
 Glu Pro Ala Gln Gln Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala

 690 695 700
 Leu Leu Pro His Leu Tyr Thr Leu Phe His Gln Ala His Val Ala Gly
 705 710 715 720
 Glu Thr Val Ala Arg Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser
 725 730 735
 Thr Trp Thr Val Asp His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile
 740 745 750
 Thr Pro Val Leu Gln Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro

 755 760 765
 Leu Gly Thr Trp Tyr Asp Leu Gln Thr Val Pro Ile Glu Ala Leu Gly
 770 775 780
 Ser Leu Pro Pro Pro Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser
 785 790 795 800
 Glu Gly Gln Trp Val Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val
 805 810 815
 His Leu Arg Ala Gly Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr

 820 825 830
 Thr Thr Glu Ser Arg Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr
 835 840 845
 Lys Gly Gly Glu Ala Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser

850 855 860
 Leu Glu Val Leu Glu Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala
 865 870 875 880
 Arg Asn Asn Thr Ile Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly

885 890 895
 Ala Gly Leu Gln Leu Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala
 900 905 910
 Pro Gln Gln Val Leu Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr
 915 920 925
 Ser Pro Asp Thr Lys Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly
 930 935 940
 Glu Gln Phe Leu Val Ser Trp Cys
 945 950

<210> 172

<211> 945

<212> PRT

<213> Homo sapiens

<400> 172

Met Glu Phe Gly Leu Ser Trp Val Phe Leu Val Ala Leu Leu Lys Gly
 1 5 10 15
 Val Gln Cys Glu Gly His Ile Leu Leu His Asp Phe Leu Leu Val Pro
 20 25 30
 Arg Glu Leu Ser Gly Ser Ser Pro Val Leu Glu Glu Thr His Pro Ala
 35 40 45
 His Gln Gln Gly Ala Ser Arg Pro Gly Pro Arg Asp Ala Gln Ala His

50 55 60
 Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys Asp Val Pro Pro Asn
 65 70 75 80
 Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile Thr Gln Glu Gln Cys
 85 90 95
 Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys Gln Gly Leu Gln Gly
 100 105 110

Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro Pro Ser Tyr Pro Ser

115 120 125
Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met Gly Tyr Thr Ala Thr

130 135 140
Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys Asp Ile Leu Thr Leu

145 150 155 160
Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg Leu His Phe Thr Ile

165 170 175
Lys Asp Pro Ala Asn Arg Arg Tyr Glu Val Pro Leu Glu Thr Pro Arg

180 185 190
Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser Val Glu Phe Ser Glu

195 200 205
Glu Pro Phe Gly Val Ile Val His Arg Gln Leu Asp Gly Arg Val Leu

210 215 220
Leu Asn Thr Thr Val Ala Pro Leu Phe Phe Ala Asp Gln Phe Leu Gln

225 230 235 240
Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr Gly Leu Ala Glu His

245 250 255
Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr Arg Ile Thr Leu Trp

260 265 270
Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn Leu Tyr Gly Ser His

275 280 285
Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser Ala His Gly Val Phe

290 295 300
Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu Gln Pro Ser Pro Ala

305 310 315 320
Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp Val Tyr Ile Phe Leu

325 330 335
Gly Pro Glu Pro Lys Ser Val Val Gln Gln Tyr Leu Asp Val Val Gly

340 345 350
Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly Phe His Leu Cys Arg

355 360 365
Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln Val Val Glu Asn Met

370 375 380
Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp Asn Asp Leu Asp Tyr
385 390 395 400
Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys Asp Gly Phe Arg Asp
405 410 415
Phe Pro Ala Met Val Gln Glu Leu His Gln Gly Gly Arg Arg Tyr Met
420 425 430
Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly Pro Ala Gly Ser Tyr

435 440 445
Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val Phe Ile Thr Asn Glu
450 455 460
Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro Gly Ser Thr Ala Phe
465 470 475 480
Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp Trp Glu Asp Met Val
485 490 495
Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly Met Trp Ile Asp Met

500 505 510
Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu Asp Gly Cys Pro Asn
515 520 525
Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly Val Val Gly Gly Thr
530 535 540
Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His Gln Phe Leu Ser Thr
545 550 555 560
His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr Glu Ala Ile Ala Ser

565 570 575
His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg Pro Phe Val Ile Ser
580 585 590
Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala Gly His Trp Thr Gly
595 600 605

Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser Ser Val Pro Glu Ile
 610 615 620
 Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val Gly Ala Asp Val Cys

 625 630 635 640
 Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys Val Arg Trp Thr Gln
 645 650 655
 Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His Asn Ser Leu Leu Ser
 660 665 670
 Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln Ala Met
 675 680 685
 Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu Pro His Leu Tyr Thr

 690 695 700
 Leu Phe His Gln Ala His Val Ala Gly Glu Thr Val Ala Arg Pro Leu
 705 710 715 720
 Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp Thr Val Asp His Gln
 725 730 735
 Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro Val Leu Gln Ala Gly
 740 745 750
 Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly Thr Trp Tyr Asp Leu

 755 760 765
 Gln Thr Val Pro Ile Glu Ala Leu Gly Ser Leu Pro Pro Pro Pro Ala
 770 775 780
 Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly Gln Trp Val Thr Leu
 785 790 795 800
 Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu Arg Ala Gly Tyr Ile
 805 810 815
 Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg Gln Gln

 820 825 830
 Pro Met Ala Leu Ala Val Ala Leu Thr Lys Gly Gly Glu Ala Arg Gly
 835 840 845
 Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu Val Leu Glu Arg Gly

850 855 860
Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn Asn Thr Ile Val Asn
865 870 875 880
Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly Leu Gln Leu Gln Lys

885 890 895
Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln Gln Val Leu Ser Asn
900 905 910
Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro Asp Thr Lys Val Leu
915 920 925
Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln Phe Leu Val Ser Trp
930 935 940

Cys

945

<210> 173

<211> 945

<212> PRT

<213> Homo sapiens

<400> 173

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

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

Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro Pro Ser Tyr Pro Ser
115 120 125

Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met Gly Tyr Thr Ala Thr
130 135 140

Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys Asp Ile Leu Thr Leu
145 150 155 160

Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg Leu His Phe Thr Ile
165 170 175

Lys Asp Pro Ala Asn Arg Arg Tyr Glu Val Pro Leu Glu Thr Pro Arg
180 185 190

Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser Val Glu Phe Ser Glu
195 200 205

Glu Pro Phe Gly Val Ile Val His Arg Gln Leu Asp Gly Arg Val Leu
210 215 220

Leu Asn Thr Thr Val Ala Pro Leu Phe Phe Ala Asp Gln Phe Leu Gln
225 230 235 240

Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr Gly Leu Ala Glu His
245 250 255

Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr Arg Ile Thr Leu Trp
260 265 270

Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn Leu Tyr Gly Ser His
275 280 285

Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser Ala His Gly Val Phe
290 295 300

Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu Gln Pro Ser Pro Ala
305 310 315 320

Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp Val Tyr Ile Phe Leu
325 330 335

Gly Pro Glu Pro Lys Ser Val Val Gln Gln Tyr Leu Asp Val Val Gly
340 345 350

Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly Phe His Leu Cys Arg

355 360 365
 Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln Val Val Glu Asn Met
 370 375 380

 Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp Asn Asp Leu Asp Tyr
 385 390 395 400
 Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys Asp Gly Phe Arg Asp
 405 410 415
 Phe Pro Ala Met Val Gln Glu Leu His Gln Gly Gly Arg Arg Tyr Met
 420 425 430
 Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly Pro Ala Gly Ser Tyr
 435 440 445

 Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val Phe Ile Thr Asn Glu
 450 455 460
 Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro Gly Ser Thr Ala Phe
 465 470 475 480
 Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp Trp Glu Asp Met Val
 485 490 495
 Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly Met Trp Ile Asp Met
 500 505 510

 Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu Asp Gly Cys Pro Asn
 515 520 525
 Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly Val Val Gly Gly Thr
 530 535 540
 Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His Gln Phe Leu Ser Thr
 545 550 555 560
 His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr Glu Ala Ile Ala Ser
 565 570 575

 His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg Pro Phe Val Ile Ser
 580 585 590
 Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala Gly His Trp Thr Gly
 595 600 605

Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser Ser Val Pro Glu Ile
 610 615 620
 Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val Gly Ala Asp Val Cys
 625 630 635 640

 Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys Val Arg Trp Thr Gln
 645 650 655
 Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His Asn Ser Leu Leu Ser
 660 665 670
 Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln Ala Met
 675 680 685
 Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu Pro His Leu Tyr Thr
 690 695 700

 Leu Phe His Gln Ala His Val Ala Gly Glu Thr Val Ala Arg Pro Leu
 705 710 715 720
 Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp Thr Val Asp His Gln
 725 730 735
 Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro Val Leu Gln Ala Gly
 740 745 750
 Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly Thr Trp Tyr Asp Leu
 755 760 765

 Gln Thr Val Pro Ile Glu Ala Leu Gly Ser Leu Pro Pro Pro Pro Ala
 770 775 780
 Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly Gln Trp Val Thr Leu
 785 790 795 800
 Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu Arg Ala Gly Tyr Ile
 805 810 815
 Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg Gln Gln
 820 825 830

 Pro Met Ala Leu Ala Val Ala Leu Thr Lys Gly Gly Glu Ala Arg Gly
 835 840 845
 Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu Val Leu Glu Arg Gly

850 855 860
Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn Asn Thr Ile Val Asn
865 870 875 880
Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly Leu Gln Leu Gln Lys
885 890 895

Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln Gln Val Leu Ser Asn
900 905 910
Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro Asp Thr Lys Val Leu
915 920 925
Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln Phe Leu Val Ser Trp
930 935 940

Cys

945

<210> 174

<211> 945

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 174

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

Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys Asp Val Pro Pro Asn
65 70 75 80
Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile Thr Gln Glu Gln Cys
85 90 95

Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys Gln Gly Leu Gln Gly
 100 105 110
 Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro Pro Ser Tyr Pro Ser
 115 120 125

 Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met Gly Tyr Thr Ala Thr
 130 135 140
 Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys Asp Ile Leu Thr Leu
 145 150 155 160
 Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg Leu His Phe Thr Ile
 165 170 175
 Lys Asp Pro Ala Asn Arg Arg Tyr Glu Val Pro Leu Glu Thr Pro Arg
 180 185 190

 Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser Val Glu Phe Ser Glu
 195 200 205
 Glu Pro Phe Gly Val Ile Val His Arg Gln Leu Asp Gly Arg Val Leu
 210 215 220
 Leu Asn Thr Thr Val Ala Pro Leu Phe Phe Ala Asp Gln Phe Leu Gln
 225 230 235 240
 Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr Gly Leu Ala Glu His
 245 250 255

 Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr Arg Ile Thr Leu Trp
 260 265 270
 Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn Leu Tyr Gly Ser His
 275 280 285
 Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser Ala His Gly Val Phe
 290 295 300
 Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu Gln Pro Ser Pro Ala
 305 310 315 320

 Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp Val Tyr Ile Phe Leu
 325 330 335
 Gly Pro Glu Pro Lys Ser Val Val Gln Gln Tyr Leu Asp Val Val Gly

340 345 350
 Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly Phe His Leu Cys Arg
 355 360 365
 Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln Val Val Glu Asn Met
 370 375 380

 Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp Asn Asp Leu Asp Tyr
 385 390 395 400
 Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys Asp Gly Phe Arg Asp
 405 410 415
 Phe Pro Ala Met Val Gln Glu Leu His Gln Gly Gly Arg Arg Tyr Met
 420 425 430
 Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly Pro Ala Gly Ser Tyr
 435 440 445

 Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val Phe Ile Thr Asn Glu
 450 455 460
 Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro Gly Ser Thr Ala Phe
 465 470 475 480
 Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp Trp Glu Asp Met Val
 485 490 495
 Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly Met Trp Ile Asp Met
 500 505 510

 Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu Asp Gly Cys Pro Asn
 515 520 525
 Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly Val Val Gly Gly Thr
 530 535 540
 Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His Gln Phe Leu Ser Thr
 545 550 555 560
 His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr Glu Ala Ile Ala Ser
 565 570 575

 His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg Pro Phe Val Ile Ser
 580 585 590

Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala Gly His Trp Thr Gly
595 600 605

Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser Ser Val Pro Glu Ile
610 615 620

Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val Gly Ala Asp Val Cys
625 630 635 640

Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys Val Arg Trp Thr Gln
645 650 655

Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His Asn Ser Leu Leu Ser
660 665 670

Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln Ala Met
675 680 685

Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu Pro His Leu Tyr Thr
690 695 700

Leu Phe His Gln Ala His Val Ala Gly Glu Thr Val Ala Arg Pro Leu
705 710 715 720

Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp Thr Val Asp His Gln
725 730 735

Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro Val Leu Gln Ala Gly
740 745 750

Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly Thr Trp Tyr Asp Leu
755 760 765

Gln Thr Val Pro Ile Glu Ala Leu Gly Ser Leu Pro Pro Pro Pro Ala
770 775 780

Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly Gln Trp Val Thr Leu
785 790 795 800

Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu Arg Ala Gly Tyr Ile
805 810 815

Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg Gln Gln
820 825 830

Pro Met Ala Leu Ala Val Ala Leu Thr Lys Gly Gly Glu Ala Arg Gly

835 840 845
Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu Val Leu Glu Arg Gly
850 855 860
Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn Asn Thr Ile Val Asn
865 870 875 880
Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly Leu Gln Leu Gln Lys
885 890 895
Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln Gln Val Leu Ser Asn
900 905 910
Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro Asp Thr Lys Val Leu
915 920 925
Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln Phe Leu Val Ser Trp
930 935 940
Cys
945
<210> 175
<211> 27
<212> PRT
<213> Homo sapiens
<400> 175
Met Gly Val Arg His Pro Pro Cys Ser His Arg Leu Leu Ala Val Cys

1 5 10 15
Ala Leu Val Ser Leu Ala Thr Ala Ala Leu Leu
20 25
<210> 176
<211> 20
<212> PRT
<213> Homo sapiens
<400> 176
Met Glu Phe Gly Leu Ser Trp Val Phe Leu Val Ala Leu Leu Lys Gly
1 5 10 15
Val Gln Cys Glu
20

<210> 177

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 177

atggagttcg ggctgagctg ggtctttctg gtggccctgc tgaaggaggag ccagtgcgag 60

<210> 178

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 178

Met Tyr Arg Met Gln Leu Leu Ser Cys Ile Ala Leu Ser Leu Ala Leu

1 5 10 15

Val Thr Asn Ser

20

<210> 179

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 179

atgtaccgta tgcagctgct gtcctgcac gcactgtcac tggcactggt taccaactca 60

<210> 180

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 180

Met Tyr Arg Met Gln Leu Leu Leu Leu Ile Ala Leu Ser Leu Ala Leu

1 5 10 15

Val Thr Asn Ser

20

<210> 181

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 181

atgtaccgta tgcagctgct gctgctgac gcactgtcac tggcactggt taccaactca 60

<210> 182

<211> 2868

<212> DNA

<213> Homo sapiens

<400> 182

atgggagtga ggcacccgcc ctgctccac cgctcctgg ccgtctgcgc cctcgtgtcc 60

ttggcaaccg ctgcactcct ggggcacatc ctactccatg atttctgtct ggttccccga 120

gagctgagt gctcctcccc agtctctggag gagactcacc cagctcacca gcaggagacc 180

agcagaccag ggccccggga cgcccaggca cccccggcc gtcccagagc agtgcccaca 240

cagtgcgacg tccccccaa cagecgcttc gattgcgccc ctgacaaggc catcacccag 300

gaacagtgcg aggccccggg ctgctgctac atccctgcaa agcaggggct gcaggagacc 360

cagatggggc agccctggtg cttcttccca ccagctacc ccagctacaa gctggagaac 420

ctgagctcct ctgaaatggg ctacacggcc accctgacc gtaccacccc caccttcttc 480

cccaaggaca tctgaccct gcggctggac gtgatgatgg agactgagaa ccgcctccac 540

ttcacgatca aagatccagc taacaggcgc tacgaggtgc ccttgagac cccgctgtc 600

cacagccggg caccgtcccc actctacagc gtggagtct cagaggagcc cttcggggtg 660

atcgtgcacc ggcagctgga cggcccgctg ctgctgaaca cgacggtggc gccctgttc 720

tttgcggacc agttccttca gctgtccacc tcgctgccct cgcagtatat cacaggcctc 780

gccgagcacc tcagtccct gatgtcagc accagctgga ccaggatcac cctgtggaac 840

cgggaccttg cgccacgcc cggcgcgaac ctctacgggt ctcacccctt ctacctggcg	900
ctggaggacg gcgggtcgcc acacgggggtg ttcttgctaa acagcaacgc catggacgtg	960
gtctgcagc cgagccctgc ccttagctgg aggtcgacag gtgggacct ggacgtctac	1020
atcttcttg gccagagcc caagagcgtg gtgcagcagt acctggacgt ggtgggatac	1080
ccgttcatgc cgccatactg gggcctgggc ttccacctgt gccgtgggg ctactcctcc	1140
accgctatca cccgccaggt ggtggagaac atgaccaggg cccacttccc cctggacgtc	1200
cagtggaaac atctggacta catggactcc cggagggact tcacgttcaa caaggatggc	1260
ttccgggact tcccgccat ggtgcaggag ctgcaccagg gcggccggcg ctacatgatg	1320
atcgtggatc ctgccatcag cagctcgggc cctgccggga gctacaggcc ctacgacgag	1380
ggtctgcgga ggggggtttt catcaccaac gagaccggcc agccgctgat tgggaagggtg	1440
tggcccggt ccactgcctt ccccgacttc accaacccca cagccctggc ctggtgggag	1500
gacatggtgg ctgagttcca tgaccagggt cccttcgacg gcatgtggat tgacatgaac	1560
gagccttcca acttcatcag aggcctctgag gacggctgcc ccaacaacga gctggagaac	1620
ccacctacg tgcttgggt ggttgggggg accctccagg cgccaccat ctgtgcctcc	1680
agccaccagt ttctctccac acactacaac ctgcacaacc tctacggcct gaccgaagcc	1740
atgcctccc acagggcgct ggtgaaggct cgggggacac gccatttgt gatctccgc	1800
tcgacctttg ctggccacgg ccgatacgcc ggccactgga cgggggacgt gtggagctcc	1860
tgggagcagc tcgcctctc cgtgccagaa atcctgcagt ttaacctgct gggggtgcct	1920
ctggtcgggg ccgacgtctg cggcttctg ggcaacacct cagaggagct gtgtgtgcgc	1980
tggaccacgc tgggggcctt ctacccttc atcggaacc acaacagcct gtcagtctg	2040
ccccaggagc cgtacagctt cagcgagccg gccagcagg ccatgaggaa ggcctcacc	2100
ctgcgtacg cactcctccc ccacctctac aactgttcc accaggccca cgtcgcgggg	2160
gagaccgtgg cccggccctt ctctctggag ttcccaagg actctagcac ctggactgtg	2220
gaccaccagc tctgtgggg ggagccctg ctcatcacc cagtgtcca ggccgggaag	2280
gccgaagtga ctggctactt ccccttgggc acatggtacg acctgcagac ggtgccaata	2340
gaggcccttg gcagcctccc accccacct gcagctccc gtgagccagc catccacagc	2400
gaggggcagt ggtgacgt gccggcccc ctggacacca tcaactgca cctccgggt	2460
gggtacatca tccccctgca gggccctggc ctcaacca cagagtccg ccagcagccc	2520
atggccctgg ctgtggcct gaccaagggt ggagaggccc gaggggagct gttctgggac	2580
gacggagaga gcctggaagt gctggagcga ggggcctaca cacaggtcat ctctctggcc	2640
aggaataaca cgatcgtgaa tgagctggt cgtgtgacca gtgaggagc tggcctgcag	2700

ctgcagaagg tgactgtcct gggcgtggcc acggcgcccc agcaggtcct ctccaacggt 2760
gtccctgtct ccaacttcac ctacagcccc gacaccaagg tcctggacat ctgtgtctcg 2820

ctgttgatgg gagagcagtt tctcgtcagc tgggtttgat aactcgag 2868

<210> 183

<400> 183

000

<210> 184

<400> 184

000

<210> 185

<211> 677

<212> PRT

<213> Homo sapiens

<400> 185

Met Pro Gly Phe Leu Val Arg Ile Leu Pro Leu Leu Leu Val Leu Leu

1 5 10 15

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

20 25 30

Glu Ile Asp Tyr Ser Arg Asp Ser Phe Leu Lys Asp Gly Gln Pro Phe

35 40 45

Arg Tyr Ile Ser Gly Ser Ile His Tyr Ser Arg Val Pro Arg Phe Tyr

50 55 60

Trp Lys Asp Arg Leu Leu Lys Met Lys Met Ala Gly Leu Asn Ala Ile

65 70 75 80

Gln Thr Tyr Val Pro Trp Asn Phe His Glu Pro Trp Pro Gly Gln Tyr

85 90 95

Gln Phe Ser Glu Asp His Asp Val Glu Tyr Phe Leu Arg Leu Ala His

100 105 110

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

115 120 125

Glu Trp Glu Met Gly Gly Leu Pro Ala Trp Leu Leu Glu Lys Glu Ser

130 135 140

Ile Leu Leu Arg Ser Ser Asp Pro Asp Tyr Leu Ala Ala Val Asp Lys
145 150 155 160
Trp Leu Gly Val Leu Leu Pro Lys Met Lys Pro Leu Leu Tyr Gln Asn

 165 170 175
Gly Gly Pro Val Ile Thr Val Gln Val Glu Asn Glu Tyr Gly Ser Tyr
 180 185 190
Phe Ala Cys Asp Phe Asp Tyr Leu Arg Phe Leu Gln Lys Arg Phe Arg
 195 200 205
His His Leu Gly Asp Asp Val Val Leu Phe Thr Thr Asp Gly Ala His
 210 215 220
Lys Thr Phe Leu Lys Cys Gly Ala Leu Gln Gly Leu Tyr Thr Thr Val

225 230 235 240
Asp Phe Gly Thr Gly Ser Asn Ile Thr Asp Ala Phe Leu Ser Gln Arg
 245 250 255
Lys Cys Glu Pro Lys Gly Pro Leu Ile Asn Ser Glu Phe Tyr Thr Gly
 260 265 270
Trp Leu Asp His Trp Gly Gln Pro His Ser Thr Ile Lys Thr Glu Ala
 275 280 285
Val Ala Ser Ser Leu Tyr Asp Ile Leu Ala Arg Gly Ala Ser Val Asn

290 295 300
Leu Tyr Met Phe Ile Gly Gly Thr Asn Phe Ala Tyr Trp Asn Gly Ala
305 310 315 320
Asn Ser Pro Tyr Ala Ala Gln Pro Thr Ser Tyr Asp Tyr Asp Ala Pro
 325 330 335
Leu Ser Glu Ala Gly Asp Leu Thr Glu Lys Tyr Phe Ala Leu Arg Asn
 340 345 350
Ile Ile Gln Lys Phe Glu Lys Val Pro Glu Gly Pro Ile Pro Pro Ser

355 360 365
Thr Pro Lys Phe Ala Tyr Gly Lys Val Thr Leu Glu Lys Leu Lys Thr
370 375 380
Val Gly Ala Ala Leu Asp Ile Leu Cys Pro Ser Gly Pro Ile Lys Ser

385 390 395 400
 Leu Tyr Pro Leu Thr Phe Ile Gln Val Lys Gln His Tyr Gly Phe Val
 405 410 415
 Leu Tyr Arg Thr Thr Leu Pro Gln Asp Cys Ser Asn Pro Ala Pro Leu

 420 425 430
 Ser Ser Pro Leu Asn Gly Val His Asp Arg Ala Tyr Val Ala Val Asp
 435 440 445
 Gly Ile Pro Gln Gly Val Leu Glu Arg Asn Asn Val Ile Thr Leu Asn
 450 455 460
 Ile Thr Gly Lys Ala Gly Ala Thr Leu Asp Leu Leu Val Glu Asn Met
 465 470 475 480
 Gly Arg Val Asn Tyr Gly Ala Tyr Ile Asn Asp Phe Lys Gly Leu Val

 485 490 495
 Ser Asn Leu Thr Leu Ser Ser Asn Ile Leu Thr Asp Trp Thr Ile Phe
 500 505 510
 Pro Leu Asp Thr Glu Asp Ala Val Cys Ser His Leu Gly Gly Trp Gly
 515 520 525
 His Arg Asp Ser Gly His His Asp Glu Ala Trp Ala His Asn Ser Ser
 530 535 540
 Asn Tyr Thr Leu Pro Ala Phe Tyr Met Gly Asn Phe Ser Ile Pro Ser

 545 550 555 560
 Gly Ile Pro Asp Leu Pro Gln Asp Thr Phe Ile Gln Phe Pro Gly Trp
 565 570 575
 Thr Lys Gly Gln Val Trp Ile Asn Gly Phe Asn Leu Gly Arg Tyr Trp
 580 585 590
 Pro Ala Arg Gly Pro Gln Leu Thr Leu Phe Val Pro Gln His Ile Leu
 595 600 605
 Met Thr Ser Ala Pro Asn Thr Ile Thr Val Leu Glu Leu Glu Trp Ala

 610 615 620
 Pro Cys Ser Ser Asp Asp Pro Glu Leu Cys Ala Val Thr Phe Val Asp
 625 630 635 640

Arg Pro Val Ile Gly Ser Ser Val Thr Tyr Asp His Pro Ser Lys Pro
645 650 655
Val Glu Lys Arg Leu Met Pro Pro Pro Pro Gln Lys Asn Lys Asp Ser
660 665 670
Trp Leu Asp His Val
675
<210
> 186
<211> 529
<212> PRT
<213> Homo sapiens
<400> 186
Met Thr Ser Ser Arg Leu Trp Phe Ser Leu Leu Leu Ala Ala Ala Phe
1 5 10 15
Ala Gly Arg Ala Thr Ala Leu Trp Pro Trp Pro Gln Asn Phe Gln Thr
20 25 30
Ser Asp Gln Arg Tyr Val Leu Tyr Pro Asn Asn Phe Gln Phe Gln Tyr
35 40 45
Asp Val Ser Ser Ala Ala Gln Pro Gly Cys Ser Val Leu Asp Glu Ala
50 55 60
Phe Gln Arg Tyr Arg Asp Leu Leu Phe Gly Ser Gly Ser Trp Pro Arg
65 70 75 80
Pro Tyr Leu Thr Gly Lys Arg His Thr Leu Glu Lys Asn Val Leu Val
85 90 95
Val Ser Val Val Thr Pro Gly Cys Asn Gln Leu Pro Thr Leu Glu Ser
100 105 110
Val Glu Asn Tyr Thr Leu Thr Ile Asn Asp Asp Gln Cys Leu Leu Leu
115 120 125
Ser Glu Thr Val Trp Gly Ala Leu Arg Gly Leu Glu Thr Phe Ser Gln
130 135 140
Leu Val Trp Lys Ser Ala Glu Gly Thr Phe Phe Ile Asn Lys Thr Glu
145 150 155 160
Ile Glu Asp Phe Pro Arg Phe Pro His Arg Gly Leu Leu Leu Asp Thr

165 170 175
Ser Arg His Tyr Leu Pro Leu Ser Ser Ile Leu Asp Thr Leu Asp Val

180 185 190
Met Ala Tyr Asn Lys Leu Asn Val Phe His Trp His Leu Val Asp Asp

195 200 205
Pro Ser Phe Pro Tyr Glu Ser Phe Thr Phe Pro Glu Leu Met Arg Lys

210 215 220
Gly Ser Tyr Asn Pro Val Thr His Ile Tyr Thr Ala Gln Asp Val Lys

225 230 235 240
Glu Val Ile Glu Tyr Ala Arg Leu Arg Gly Ile Arg Val Leu Ala Glu

245 250 255
Phe Asp Thr Pro Gly His Thr Leu Ser Trp Gly Pro Gly Ile Pro Gly

260 265 270
Leu Leu Thr Pro Cys Tyr Ser Gly Ser Glu Pro Ser Gly Thr Phe Gly

275 280 285
Pro Val Asn Pro Ser Leu Asn Asn Thr Tyr Glu Phe Met Ser Thr Phe

290 295 300
Phe Leu Glu Val Ser Ser Val Phe Pro Asp Phe Tyr Leu His Leu Gly

305 310 315 320
Gly Asp Glu Val Asp Phe Thr Cys Trp Lys Ser Asn Pro Glu Ile Gln

325 330 335
Asp Phe Met Arg Lys Lys Gly Phe Gly Glu Asp Phe Lys Gln Leu Glu

340 345 350
Ser Phe Tyr Ile Gln Thr Leu Leu Asp Ile Val Ser Ser Tyr Gly Lys

355 360 365
Gly Tyr Val Val Trp Gln Glu Val Phe Asp Asn Lys Val Lys Ile Gln

370 375 380
Pro Asp Thr Ile Ile Gln Val Trp Arg Glu Asp Ile Pro Val Asn Tyr

385 390 395 400
Met Lys Glu Leu Glu Leu Val Thr Lys Ala Gly Phe Arg Ala Leu Leu

405 410 415

Ser Ala Pro Trp Tyr Leu Asn Arg Ile Ser Tyr Gly Pro Asp Trp Lys
420 425 430

Asp Phe Tyr Ile Val Glu Pro Leu Ala Phe Glu Gly Thr Pro Glu Gln
435 440 445

Lys Ala Leu Val Ile Gly Gly Glu Ala Cys Met Trp Gly Glu Tyr Val
450 455 460

Asp Asn Thr Asn Leu Val Pro Arg Leu Trp Pro Arg Ala Gly Ala Val
465 470 475 480

Ala Glu Arg Leu Trp Ser Asn Lys Leu Thr Ser Asp Leu Thr Phe Ala
485 490 495

Tyr Glu Arg Leu Ser His Phe Arg Cys Glu Leu Leu Arg Arg Gly Val
500 505 510

Gln Ala Gln Pro Leu Asn Val Gly Phe Cys Glu Gln Glu Phe Glu Gln
515 520 525

Thr

<210> 187
<211> 556
<212> PRT
<213> Homo sapiens
<400> 187

Met Glu Leu Cys Gly Leu Gly Leu Pro Arg Pro Pro Met Leu Leu Ala
1 5 10 15

Leu Leu Leu Ala Thr Leu Leu Ala Ala Met Leu Ala Leu Leu Thr Gln
20 25 30

Val Ala Leu Val Val Gln Val Ala Glu Ala Ala Arg Ala Pro Ser Val
35 40 45

Ser Ala Lys Pro Gly Pro Ala Leu Trp Pro Leu Pro Leu Leu Val Lys
50 55 60

Met Thr Pro Asn Leu Leu His Leu Ala Pro Glu Asn Phe Tyr Ile Ser
65 70 75 80

His Ser Pro Asn Ser Thr Ala Gly Pro Ser Cys Thr Leu Leu Glu Glu

85

90

95

Ala Phe Arg Arg Tyr His Gly Tyr Ile Phe Gly Phe Tyr Lys Trp His

100

105

110

His Glu Pro Ala Glu Phe Gln Ala Lys Thr Gln Val Gln Gln Leu Leu

115

120

125

Val Ser Ile Thr Leu Gln Ser Glu Cys Asp Ala Phe Pro Asn Ile Ser

130

135

140

Ser Asp Glu Ser Tyr Thr Leu Leu Val Lys Glu Pro Val Ala Val Leu

145

150

155

160

Lys Ala Asn Arg Val Trp Gly Ala Leu Arg Gly Leu Glu Thr Phe Ser

165

170

175

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

180

185

190

Thr Ile Ile Asp Ser Pro Arg Phe Ser His Arg Gly Ile Leu Ile Asp

195

200

205

Thr Ser Arg His Tyr Leu Pro Val Lys Ile Ile Leu Lys Thr Leu Asp

210

215

220

Ala Met Ala Phe Asn Lys Phe Asn Val Leu His Trp His Ile Val Asp

225

230

235

240

Asp Gln Ser Phe Pro Tyr Gln Ser Ile Thr Phe Pro Glu Leu Ser Asn

245

250

255

Lys Gly Ser Tyr Ser Leu Ser His Val Tyr Thr Pro Asn Asp Val Arg

260

265

270

Met Val Ile Glu Tyr Ala Arg Leu Arg Gly Ile Arg Val Leu Pro Glu

275

280

285

Phe Asp Thr Pro Gly His Thr Leu Ser Trp Gly Lys Gly Gln Lys Asp

290

295

300

Leu Leu Thr Pro Cys Tyr Ser Arg Gln Asn Lys Leu Asp Ser Phe Gly

305

310

315

320

Pro Ile Asn Pro Thr Leu Asn Thr Thr Tyr Ser Phe Leu Thr Thr Phe

325

330

335

Phe Lys Glu Ile Ser Glu Val Phe Pro Asp Gln Phe Ile His Leu Gly
340 345 350

Gly Asp Glu Val Glu Phe Lys Cys Trp Glu Ser Asn Pro Lys Ile Gln
355 360 365

Asp Phe Met Arg Gln Lys Gly Phe Gly Thr Asp Phe Lys Lys Leu Glu
370 375 380

Ser Phe Tyr Ile Gln Lys Val Leu Asp Ile Ile Ala Thr Ile Asn Lys
385 390 395 400

Gly Ser Ile Val Trp Gln Glu Val Phe Asp Asp Lys Ala Lys Leu Ala
405 410 415

Pro Gly Thr Ile Val Glu Val Trp Lys Asp Ser Ala Tyr Pro Glu Glu
420 425 430

Leu Ser Arg Val Thr Ala Ser Gly Phe Pro Val Ile Leu Ser Ala Pro
435 440 445

Trp Tyr Leu Asp Leu Ile Ser Tyr Gly Gln Asp Trp Arg Lys Tyr Tyr
450 455 460

Lys Val Glu Pro Leu Asp Phe Gly Gly Thr Gln Lys Gln Lys Gln Leu
465 470 475 480

Phe Ile Gly Gly Glu Ala Cys Leu Trp Gly Glu Tyr Val Asp Ala Thr
485 490 495

Asn Leu Thr Pro Arg Leu Trp Pro Arg Ala Ser Ala Val Gly Glu Arg
500 505 510

Leu Trp Ser Ser Lys Asp Val Arg Asp Met Asp Asp Ala Tyr Asp Arg
515 520 525

Leu Thr Arg His Arg Cys Arg Met Val Glu Arg Gly Ile Ala Ala Gln
530 535 540

Pro Leu Tyr Ala Gly Tyr Cys Asn His Glu Asn Met
545 550 555

<210> 188

<211> 429

<212> PRT

<213> Homo sapiens

<400> 188

Met Gln Leu Arg Asn Pro Glu Leu His Leu Gly Cys Ala Leu Ala Leu

1 5 10 15

Arg Phe Leu Ala Leu Val Ser Trp Asp Ile Pro Gly Ala Arg Ala Leu

20 25 30

Asp Asn Gly Leu Ala Arg Thr Pro Thr Met Gly Trp Leu His Trp Glu

35 40 45

Arg Phe Met Cys Asn Leu Asp Cys Gln Glu Glu Pro Asp Ser Cys Ile

50 55 60

Ser Glu Lys Leu Phe Met Glu Met Ala Glu Leu Met Val Ser Glu Gly

65 70 75 80

Trp Lys Asp Ala Gly Tyr Glu Tyr Leu Cys Ile Asp Asp Cys Trp Met

85 90 95

Ala Pro Gln Arg Asp Ser Glu Gly Arg Leu Gln Ala Asp Pro Gln Arg

100 105 110

Phe Pro His Gly Ile Arg Gln Leu Ala Asn Tyr Val His Ser Lys Gly

115 120 125

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

130 135 140

Phe Pro Gly Ser Phe Gly Tyr Tyr Asp Ile Asp Ala Gln Thr Phe Ala

145 150 155 160

Asp Trp Gly Val Asp Leu Leu Lys Phe Asp Gly Cys Tyr Cys Asp Ser

165 170 175

Leu Glu Asn Leu Ala Asp Gly Tyr Lys His Met Ser Leu Ala Leu Asn

180 185 190

Arg Thr Gly Arg Ser Ile Val Tyr Ser Cys Glu Trp Pro Leu Tyr Met

195 200 205

Trp Pro Phe Gln Lys Pro Asn Tyr Thr Glu Ile Arg Gln Tyr Cys Asn

210 215 220

His Trp Arg Asn Phe Ala Asp Ile Asp Asp Ser Trp Lys Ser Ile Lys

225	230	235	240
Ser Ile Leu Asp Trp Thr Ser Phe Asn Gln Glu Arg Ile Val Asp Val			
245	250	255	
Ala Gly Pro Gly Gly Trp Asn Asp Pro Asp Met Leu Val Ile Gly Asn			
260	265	270	
Phe Gly Leu Ser Trp Asn Gln Gln Val Thr Gln Met Ala Leu Trp Ala			
275	280	285	
Ile Met Ala Ala Pro Leu Phe Met Ser Asn Asp Leu Arg His Ile Ser			
290	295	300	

Pro	Gln	Ala	Lys	Ala	Leu	Leu	Gln	Asp	Lys	Asp	Val	Ile	Ala	Ile	Asn
305					310					315					320
Gln	Asp	Pro	Leu	Gly	Lys	Gln	Gly	Tyr	Gln	Leu	Arg	Gln	Gly	Asp	Asn
				325					330					335	
Phe	Glu	Val	Trp	Glu	Arg	Pro	Leu	Ser	Gly	Leu	Ala	Trp	Ala	Val	Ala
			340						345				350		
Met	Ile	Asn	Arg	Gln	Glu	Ile	Gly	Gly	Pro	Arg	Ser	Tyr	Thr	Ile	Ala
		355					360					365			

Val	Ala	Ser	Leu	Gly	Lys	Gly	Val	Ala	Cys	Asn	Pro	Ala	Cys	Phe	Ile
370						375					380				
Thr	Gln	Leu	Leu	Pro	Val	Lys	Arg	Lys	Leu	Gly	Phe	Tyr	Glu	Trp	Thr
385					390					395					400
Ser	Arg	Leu	Arg	Ser	His	Ile	Asn	Pro	Thr	Gly	Thr	Val	Leu	Leu	Gln
				405					410					415	
Leu	Glu	Asn	Thr	Met	Gln	Met	Ser	Leu	Lys	Asp	Leu	Leu			
		420						425							

<210> 189

<211> 536

<212> PRT

<213> Homo sapiens

<400> 189

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> 190
 <211> 509
 <212> PRT
 <213> Homo sapiens
 <400> 190
 Met Ser Met Gly Ala Pro Arg Ser Leu Leu Leu Ala Leu Ala Ala Gly
 1 5 10 15
 Leu Ala Val Ala Arg Pro Pro Asn Ile Val Leu Ile Phe Ala Asp Asp
 20 25 30
 Leu Gly Tyr Gly Asp Leu Gly Cys Tyr Gly His Pro Ser Ser Thr Thr
 35 40 45
 Pro Asn Leu Asp Gln Leu Ala Ala Gly Gly Leu Arg Phe Thr Asp Phe
 50 55 60
 Tyr Val Pro Val Ser Leu Cys Thr Pro Ser Arg Ala Ala Leu Leu Thr
 65 70 75 80
 Gly Arg Leu Pro Val Arg Met Gly Met Tyr Pro Gly Val Leu Val Pro
 85 90 95
 Ser Ser Arg Gly Gly Leu Pro Leu Glu Glu Val Thr Val Ala Glu Val
 100 105 110
 Leu Ala Ala Arg Gly Tyr Leu Thr Gly Met Ala Gly Lys Trp His Leu
 115 120 125
 Gly Val Gly Pro Glu Gly Ala Phe Leu Pro Pro His Gln Gly Phe His
 130 135 140
 Arg Phe Leu Gly Ile Pro Tyr Ser His Asp Gln Gly Pro Cys Gln Asn
 145 150 155 160
 Leu Thr Cys Phe Pro Pro Ala Thr Pro Cys Asp Gly Gly Cys Asp Gln
 165 170 175

Gly Leu Val Pro Ile Pro Leu Leu Ala Asn Leu Ser Val Glu Ala Gln
 180 185 190
 Pro Pro Trp Leu Pro Gly Leu Glu Ala Arg Tyr Met Ala Phe Ala His
 195 200 205
 Asp Leu Met Ala Asp Ala Gln Arg Gln Asp Arg Pro Phe Phe Leu Tyr
 210 215 220
 Tyr Ala Ser His His Thr His Tyr Pro Gln Phe Ser Gly Gln Ser Phe
 225 230 235 240
 Ala Glu Arg Ser Gly Arg Gly Pro Phe Gly Asp Ser Leu Met Glu Leu
 245 250 255
 Asp Ala Ala Val Gly Thr Leu Met Thr Ala Ile Gly Asp Leu Gly Leu
 260 265 270
 Leu Glu Glu Thr Leu Val Ile Phe Thr Ala Asp Asn Gly Pro Glu Thr
 275 280 285
 Met Arg Met Ser Arg Gly Gly Cys Ser Gly Leu Leu Arg Cys Gly Lys
 290 295 300
 Gly Thr Thr Tyr Glu Gly Gly Val Arg Glu Pro Ala Leu Ala Phe Trp
 305 310 315 320
 Pro Gly His Ile Ala Pro Gly Val Thr His Glu Leu Ala Ser Ser Leu
 325 330 335
 Asp Leu Leu Pro Thr Leu Ala Ala Leu Ala Gly Ala Pro Leu Pro Asn
 340 345 350
 Val Thr Leu Asp Gly Phe Asp Leu Ser Pro Leu Leu Leu Gly Thr Gly
 355 360 365
 Lys Ser Pro Arg Gln Ser Leu Phe Phe Tyr Pro Ser Tyr Pro Asp Glu
 370 375 380
 Val Arg Gly Val Phe Ala Val Arg Thr Gly Lys Tyr Lys Ala His Phe
 385 390 395 400
 Phe Thr Gln Gly Ser Ala His Ser Asp Thr Thr Ala Asp Pro Ala Cys
 405 410 415
 His Ala Ser Ser Ser Leu Thr Ala His Glu Pro Pro Leu Leu Tyr Asp

420 425 430
 Leu Ser Lys Asp Pro Gly Glu Asn Tyr Asn Leu Leu Gly Gly Val Ala
 435 440 445
 Gly Ala Thr Pro Glu Val Leu Gln Ala Leu Lys Gln Leu Gln Leu Leu
 450 455 460
 Lys Ala Gln Leu Asp Ala Ala Val Thr Phe Gly Pro Ser Gln Val Ala

 465 470 475 480
 Arg Gly Glu Asp Pro Ala Leu Gln Ile Cys Cys His Pro Gly Cys Thr
 485 490 495
 Pro Arg Pro Ala Cys Cys His Cys Pro Asp Pro His Ala
 500 505
 <210> 191
 <211> 685
 <212> PRT
 <213> Homo sapiens
 <400> 191
 Met Ala Glu Trp Leu Leu Ser Ala Ser Trp Gln Arg Arg Ala Lys Ala
 1 5 10 15
 Met Thr Ala Ala Ala Gly Ser Ala Gly Arg Ala Ala Val Pro Leu Leu

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

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

Tyr Glu Trp Trp Leu Met Lys Glu Ala Lys Lys Arg Asn Pro Asn Ile
 130 135 140
 Thr Leu Ile Gly Leu Pro Trp Ser Phe Pro Gly Trp Leu Gly Lys Gly

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

 210 215 220
 Ala Ser Asp Asn Leu Trp Glu Ser Ile Ser Ala Ser Met Leu Leu Asp
 225 230 235 240
 Ala Glu Leu Phe Lys Val Val Asp Val Ile Gly Ala His Tyr Pro Gly
 245 250 255
 Thr His Ser Ala Lys Asp Ala Lys Leu Thr Gly Lys Lys Leu Trp Ser
 260 265 270
 Ser Glu Asp Phe Ser Thr Leu Asn Ser Asp Met Gly Ala Gly Cys Trp

 275 280 285
 Gly Arg Ile Leu Asn Gln Asn Tyr Ile Asn Gly Tyr Met Thr Ser Thr
 290 295 300
 Ile Ala Trp Asn Leu Val Ala Ser Tyr Tyr Glu Gln Leu Pro Tyr Gly
 305 310 315 320
 Arg Cys Gly Leu Met Thr Ala Gln Glu Pro Trp Ser Gly His Tyr Val
 325 330 335
 Val Glu Ser Pro Val Trp Val Ser Ala His Thr Thr Gln Phe Thr Gln

 340 345 350
 Pro Gly Trp Tyr Tyr Leu Lys Thr Val Gly His Leu Glu Lys Gly Gly
 355 360 365
 Ser Tyr Val Ala Leu Thr Asp Gly Leu Gly Asn Leu Thr Ile Ile Ile

370 375 380
 Glu Thr Met Ser His Lys His Ser Lys Cys Ile Arg Pro Phe Leu Pro
 385 390 395 400
 Tyr Phe Asn Val Ser Gln Gln Phe Ala Thr Phe Val Leu Lys Gly Ser

 405 410 415
 Phe Ser Glu Ile Pro Glu Leu Gln Val Trp Tyr Thr Lys Leu Gly Lys
 420 425 430
 Thr Ser Glu Arg Phe Leu Phe Lys Gln Leu Asp Ser Leu Trp Leu Leu
 435 440 445
 Asp Ser Asp Gly Ser Phe Thr Leu Ser Leu His Glu Asp Glu Leu Phe
 450 455 460
 Thr Leu Thr Thr Leu Thr Thr Gly Arg Lys Gly Ser Tyr Pro Leu Pro

 465 470 475 480
 Pro Lys Ser Gln Pro Phe Pro Ser Thr Tyr Lys Asp Asp Phe Asn Val
 485 490 495
 Asp Tyr Pro Phe Phe Ser Glu Ala Pro Asn Phe Ala Asp Gln Thr Gly
 500 505 510
 Val Phe Glu Tyr Phe Thr Asn Ile Glu Asp Pro Gly Glu His His Phe
 515 520 525
 Thr Leu Arg Gln Val Leu Asn Gln Arg Pro Ile Thr Trp Ala Ala Asp

 530 535 540
 Ala Ser Asn Thr Ile Ser Ile Ile Gly Asp Tyr Asn Trp Thr Asn Leu
 545 550 555 560
 Thr Ile Lys Cys Asp Val Tyr Ile Glu Thr Pro Asp Thr Gly Gly Val
 565 570 575
 Phe Ile Ala Gly Arg Val Asn Lys Gly Gly Ile Leu Ile Arg Ser Ala
 580 585 590
 Arg Gly Ile Phe Phe Trp Ile Phe Ala Asn Gly Ser Tyr Arg Val Thr

 595 600 605
 Gly Asp Leu Ala Gly Trp Ile Ile Tyr Ala Leu Gly Arg Val Glu Val
 610 615 620

Thr Ala Lys Lys Trp Tyr Thr Leu Thr Leu Thr Ile Lys Gly His Phe
625 630 635 640
Thr Ser Gly Met Leu Asn Asp Lys Ser Leu Trp Thr Asp Ile Pro Val
645 650 655
Asn Phe Pro Lys Asn Gly Trp Ala Ala Ile Gly Thr His Ser Phe Glu
660 665 670
Phe Ala Gln Phe Asp Asn Phe Leu Val Glu Ala Thr Arg
675 680 685
<210> 192
<211> 631
<212> PRT
<213> Homo sapiens
<400> 192
Met Pro Arg Tyr Gly Ala Ser Leu Arg Gln Ser Cys Pro Arg Ser Gly
1 5 10 15
Arg Glu Gln Gly Gln Asp Gly Thr Ala Gly Ala Pro Gly Leu Leu Trp
20 25 30
Met Gly Leu Val Leu Ala Leu Ala Leu Ala Leu Ala Leu Ala
35 40 45
Leu Ser Asp Ser Arg Val Leu Trp Ala Pro Ala Glu Ala His Pro Leu
50 55 60
Ser Pro Gln Gly His Pro Ala Arg Leu His Arg Ile Val Pro Arg Leu
65 70 75 80
Arg Asp Val Phe Gly Trp Gly Asn Leu Thr Cys Pro Ile Cys Lys Gly
85 90 95
Leu Phe Thr Ala Ile Asn Leu Gly Leu Lys Lys Glu Pro Asn Val Ala
100 105 110
Arg Val Gly Ser Val Ala Ile Lys Leu Cys Asn Leu Leu Lys Ile Ala
115 120 125
Pro Pro Ala Val Cys Gln Ser Ile Val His Leu Phe Glu Asp Asp Met
130 135 140
Val Glu Val Trp Arg Arg Ser Val Leu Ser Pro Ser Glu Ala Cys Gly

145 150 155 160
 Leu Leu Leu Gly Ser Thr Cys Gly His Trp Asp Ile Phe Ser Ser Trp

 165 170 175
 Asn Ile Ser Leu Pro Thr Val Pro Lys Pro Pro Pro Lys Pro Pro Ser
 180 185 190
 Pro Pro Ala Pro Gly Ala Pro Val Ser Arg Ile Leu Phe Leu Thr Asp
 195 200 205
 Leu His Trp Asp His Asp Tyr Leu Glu Gly Thr Asp Pro Asp Cys Ala
 210 215 220
 Asp Pro Leu Cys Cys Arg Arg Gly Ser Gly Leu Pro Pro Ala Ser Arg

 225 230 235 240
 Pro Gly Ala Gly Tyr Trp Gly Glu Tyr Ser Lys Cys Asp Leu Pro Leu
 245 250 255
 Arg Thr Leu Glu Ser Leu Leu Ser Gly Leu Gly Pro Ala Gly Pro Phe
 260 265 270
 Asp Met Val Tyr Trp Thr Gly Asp Ile Pro Ala His Asp Val Trp His
 275 280 285
 Gln Thr Arg Gln Asp Gln Leu Arg Ala Leu Thr Thr Val Thr Ala Leu

 290 295 300
 Val Arg Lys Phe Leu Gly Pro Val Pro Val Tyr Pro Ala Val Gly Asn
 305 310 315 320
 His Glu Ser Thr Pro Val Asn Ser Phe Pro Pro Pro Phe Ile Glu Gly
 325 330 335
 Asn His Ser Ser Arg Trp Leu Tyr Glu Ala Met Ala Lys Ala Trp Glu
 340 345 350
 Pro Trp Leu Pro Ala Glu Ala Leu Arg Thr Leu Arg Ile Gly Gly Phe

 355 360 365
 Tyr Ala Leu Ser Pro Tyr Pro Gly Leu Arg Leu Ile Ser Leu Asn Met
 370 375 380
 Asn Phe Cys Ser Arg Glu Asn Phe Trp Leu Leu Ile Asn Ser Thr Asp
 385 390 395 400

Pro Ala Gly Gln Leu Gln Trp Leu Val Gly Glu Leu Gln Ala Ala Glu

405 410 415

Asp Arg Gly Asp Lys Val His Ile Ile Gly His Ile Pro Pro Gly His

420 425 430

Cys Leu Lys Ser Trp Ser Trp Asn Tyr Tyr Arg Ile Val Ala Arg Tyr

435 440 445

Glu Asn Thr Leu Ala Ala Gln Phe Phe Gly His Thr His Val Asp Glu

450 455 460

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

465 470 475 480

Ala Phe Leu Ala Pro Ser Ala Thr Thr Tyr Ile Gly Leu Asn Pro Gly

485 490 495

Tyr Arg Val Tyr Gln Ile Asp Gly Asn Tyr Ser Gly Ser Ser His Val

500 505 510

Val Leu Asp His Glu Thr Tyr Ile Leu Asn Leu Thr Gln Ala Asn Ile

515 520 525

Pro Gly Ala Ile Pro His Trp Gln Leu Leu Tyr Arg Ala Arg Glu Thr

530 535 540

Tyr Gly Leu Pro Asn Thr Leu Pro Thr Ala Trp His Asn Leu Val Tyr

545 550 555 560

Arg Met Arg Gly Asp Met Gln Leu Phe Gln Thr Phe Trp Phe Leu Tyr

565 570 575

His Lys Gly His Pro Pro Ser Glu Pro Cys Gly Thr Pro Cys Arg Leu

580 585 590

Ala Thr Leu Cys Ala Gln Leu Ser Ala Arg Ala Asp Ser Pro Ala Leu

595 600 605

Cys Arg His Leu Met Pro Asp Gly Ser Leu Pro Glu Ala Gln Ser Leu

610 615 620

Trp Pro Arg Pro Leu Phe Cys

625 630

<210> 193

<211> 1278

<212> PRT

<213> Homo sapiens

<400> 193

Met Thr Ala Arg Gly Leu Ala Leu Gly Leu Leu Leu Leu Leu Cys

1 5 10 15

Pro Ala Gln Val Phe Ser Gln Ser Cys Val Trp Tyr Gly Glu Cys Gly

20 25 30

Ile Ala Tyr Gly Asp Lys Arg Tyr Asn Cys Glu Tyr Ser Gly Pro Pro

35 40 45

Lys Pro Leu Pro Lys Asp Gly Tyr Asp Leu Val Gln Glu Leu Cys Pro

50 55 60

Gly Phe Phe Phe Gly Asn Val Ser Leu Cys Cys Asp Val Arg Gln Leu

65 70 75 80

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

85 90 95

Cys Pro Ser Cys Phe Tyr Asn Leu Leu Asn Leu Phe Cys Glu Leu Thr

100 105 110

Cys Ser Pro Arg Gln Ser Gln Phe Leu Asn Val Thr Ala Thr Glu Asp

115 120 125

Tyr Val Asp Pro Val Thr Asn Gln Thr Lys Thr Asn Val Lys Glu Leu

130 135 140

Gln Tyr Tyr Val Gly Gln Ser Phe Ala Asn Ala Met Tyr Asn Ala Cys

145 150 155 160

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

165 170 175

Cys Gly Lys Asp Ala Asp Ala Cys Asn Ala Thr Asn Trp Ile Glu Tyr

180 185 190

Met Phe Asn Lys Asp Asn Gly Gln Ala Pro Phe Thr Ile Thr Pro Val

195 200 205

Phe Ser Asp Phe Pro Val His Gly Met Glu Pro Met Asn Asn Ala Thr

210 215 220

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

225 230 235 240

Gln Asp Cys Ser Ile Val Cys Gly Pro Lys Pro Gln Pro Pro Pro Pro

245 250 255

Pro Ala Pro Trp Thr Ile Leu Gly Leu Asp Ala Met Tyr Val Ile Met

260 265 270

Trp Ile Thr Tyr Met Ala Phe Leu Leu Val Phe Phe Gly Ala Phe Phe

275 280 285

Ala Val Trp Cys Tyr Arg Lys Arg Tyr Phe Val Ser Glu Tyr Thr Pro

290 295 300

Ile Asp Ser Asn Ile Ala Phe Ser Val Asn Ala Ser Asp Lys Gly Glu

305 310 315 320

Ala Ser Cys Cys Asp Pro Val Ser Ala Ala Phe Glu Gly Cys Leu Arg

325 330 335

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

340 345 350

Val Ile Phe Phe Ser Leu Val Phe Ile Thr Ala Cys Ser Ser Gly Leu

355 360 365

Val Phe Val Arg Val Thr Thr Asn Pro Val Asp Leu Trp Ser Ala Pro

370 375 380

Ser Ser Gln Ala Arg Leu Glu Lys Glu Tyr Phe Asp Gln His Phe Gly

385 390 395 400

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

405 410 415

Lys His Ile Tyr Gln Pro Tyr Pro Ser Gly Ala Asp Val Pro Phe Gly

420 425 430

Pro Pro Leu Asp Ile Gln Ile Leu His Gln Val Leu Asp Leu Gln Ile

435 440 445

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

450 455 460

Gln Asp Ile Cys Leu Ala Pro Leu Ser Pro Tyr Asn Thr Asn Cys Thr

465 470 475 480
 Ile Leu Ser Val Leu Asn Tyr Phe Gln Asn Ser His Ser Val Leu Asp

 485 490 495
 His Lys Lys Gly Asp Asp Phe Phe Val Tyr Ala Asp Tyr His Thr His
 500 505 510
 Phe Leu Tyr Cys Val Arg Ala Pro Ala Ser Leu Asn Asp Thr Ser Leu
 515 520 525
 Leu His Asp Pro Cys Leu Gly Thr Phe Gly Gly Pro Val Phe Pro Trp
 530 535 540
 Leu Val Leu Gly Gly Tyr Asp Asp Gln Asn Tyr Asn Asn Ala Thr Ala

 545 550 555 560
 Leu Val Ile Thr Phe Pro Val Asn Asn Tyr Tyr Asn Asp Thr Glu Lys
 565 570 575
 Leu Gln Arg Ala Gln Ala Trp Glu Lys Glu Phe Ile Asn Phe Val Lys
 580 585 590
 Asn Tyr Lys Asn Pro Asn Leu Thr Ile Ser Phe Thr Ala Glu Arg Ser
 595 600 605
 Ile Glu Asp Glu Leu Asn Arg Glu Ser Asp Ser Asp Val Phe Thr Val

 610 615 620
 Val Ile Ser Tyr Ala Ile Met Phe Leu Tyr Ile Ser Leu Ala Leu Gly
 625 630 635 640
 His Met Lys Ser Cys Arg Arg Leu Leu Val Asp Ser Lys Val Ser Leu
 645 650 655
 Gly Ile Ala Gly Ile Leu Ile Val Leu Ser Ser Val Ala Cys Ser Leu
 660 665 670
 Gly Val Phe Ser Tyr Ile Gly Leu Pro Leu Thr Leu Ile Val Ile Glu

 675 680 685
 Val Ile Pro Phe Leu Val Leu Ala Val Gly Val Asp Asn Ile Phe Ile
 690 695 700
 Leu Val Gln Ala Tyr Gln Arg Asp Glu Arg Leu Gln Gly Glu Thr Leu
 705 710 715 720

Asp Gln Gln Leu Gly Arg Val Leu Gly Glu Val Ala Pro Ser Met Phe
 725 730 735
 Leu Ser Ser Phe Ser Glu Thr Val Ala Phe Phe Leu Gly Ala Leu Ser
 740 745 750
 Val Met Pro Ala Val His Thr Phe Ser Leu Phe Ala Gly Leu Ala Val
 755 760 765
 Phe Ile Asp Phe Leu Leu Gln Ile Thr Cys Phe Val Ser Leu Leu Gly
 770 775 780
 Leu Asp Ile Lys Arg Gln Glu Lys Asn Arg Leu Asp Ile Phe Cys Cys
 785 790 795 800
 Val Arg Gly Ala Glu Asp Gly Thr Ser Val Gln Ala Ser Glu Ser Cys
 805 810 815
 Leu Phe Arg Phe Phe Lys Asn Ser Tyr Ser Pro Leu Leu Leu Lys Asp
 820 825 830
 Trp Met Arg Pro Ile Val Ile Ala Ile Phe Val Gly Val Leu Ser Phe
 835 840 845
 Ser Ile Ala Val Leu Asn Lys Val Asp Ile Gly Leu Asp Gln Ser Leu
 850 855 860
 Ser Met Pro Asp Asp Ser Tyr Met Val Asp Tyr Phe Lys Ser Ile Ser
 865 870 875 880
 Gln Tyr Leu His Ala Gly Pro Pro Val Tyr Phe Val Leu Glu Glu Gly
 885 890 895
 His Asp Tyr Thr Ser Ser Lys Gly Gln Asn Met Val Cys Gly Gly Met
 900 905 910
 Gly Cys Asn Asn Asp Ser Leu Val Gln Gln Ile Phe Asn Ala Ala Gln
 915 920 925
 Leu Asp Asn Tyr Thr Arg Ile Gly Phe Ala Pro Ser Ser Trp Ile Asp
 930 935 940
 Asp Tyr Phe Asp Trp Val Lys Pro Gln Ser Ser Cys Cys Arg Val Asp
 945 950 955 960
 Asn Ile Thr Asp Gln Phe Cys Asn Ala Ser Val Val Asp Pro Ala Cys

965	970	975
Val Arg Cys Arg Pro Leu Thr Pro Glu Gly Lys Gln Arg Pro Gln Gly		
980	985	990
Gly Asp Phe Met Arg Phe Leu Pro Met Phe Leu Ser Asp Asn Pro Asn		
995	1000	1005
Pro Lys Cys Gly Lys Gly Gly His Ala Ala Tyr Ser Ser Ala Val		
1010	1015	1020
Asn Ile Leu Leu Gly His Gly Thr Arg Val Gly Ala Thr Tyr Phe		
1025	1030	1035
Met Thr Tyr His Thr Val Leu Gln Thr Ser Ala Asp Phe Ile Asp		
1040	1045	1050
Ala Leu Lys Lys Ala Arg Leu Ile Ala Ser Asn Val Thr Glu Thr		
1055	1060	1065
Met Gly Ile Asn Gly Ser Ala Tyr Arg Val Phe Pro Tyr Ser Val		
1070	1075	1080
Phe Tyr Val Phe Tyr Glu Gln Tyr Leu Thr Ile Ile Asp Asp Thr		
1085	1090	1095
Ile Phe Asn Leu Gly Val Ser Leu Gly Ala Ile Phe Leu Val Thr		
1100	1105	1110
Met Val Leu Leu Gly Cys Glu Leu Trp Ser Ala Val Ile Met Cys		
1115	1120	1125
Ala Thr Ile Ala Met Val Leu Val Asn Met Phe Gly Val Met Trp		
1130	1135	1140
Leu Trp Gly Ile Ser Leu Asn Ala Val Ser Leu Val Asn Leu Val		
1145	1150	1155
Met Ser Cys Gly Ile Ser Val Glu Phe Cys Ser His Ile Thr Arg		
1160	1165	1170
Ala Phe Thr Val Ser Met Lys Gly Ser Arg Val Glu Arg Ala Glu		
1175	1180	1185
Glu Ala Leu Ala His Met Gly Ser Ser Val Phe Ser Gly Ile Thr		
1190	1195	1200

Leu Thr Lys Phe Gly Gly Ile Val Val Leu Ala Phe Ala Lys Ser
1205 1210 1215

Gln Ile Phe Gln Ile Phe Tyr Phe Arg Met Tyr Leu Ala Met Val
1220 1225 1230

Leu Leu Gly Ala Thr His Gly Leu Ile Phe Leu Pro Val Leu Leu
1235 1240 1245

Ser Tyr Ile Gly Pro Ser Val Asn Lys Ala Lys Ser Cys Ala Thr
1250 1255 1260

Glu Glu Arg Tyr Lys Gly Thr Glu Arg Glu Arg Leu Leu Asn Phe
1265 1270 1275

<210> 194

<211> 151

<212> PRT

<213> Homo sapiens

<400> 194

Met Arg Phe Leu Ala Ala Thr Phe Leu Leu Leu Ala Leu Ser Thr Ala
1 5 10 15

Ala Gln Ala Glu Pro Val Gln Phe Lys Asp Cys Gly Ser Val Asp Gly
20 25 30

Val Ile Lys Glu Val Asn Val Ser Pro Cys Pro Thr Gln Pro Cys Gln
35 40 45

Leu Ser Lys Gly Gln Ser Tyr Ser Val Asn Val Thr Phe Thr Ser Asn
50 55 60

Ile Gln Ser Lys Ser Ser Lys Ala Val Val His Gly Ile Leu Met Gly
65 70 75 80

Val Pro Val Pro Phe Pro Ile Pro Glu Pro Asp Gly Cys Lys Ser Gly
85 90 95

Ile Asn Cys Pro Ile Gln Lys Asp Lys Thr Tyr Ser Tyr Leu Asn Lys
100 105 110

Leu Pro Val Lys Ser Glu Tyr Pro Ser Ile Lys Leu Val Val Glu Trp
115 120 125

Gln Leu Gln Asp Asp Lys Asn Gln Ser Leu Phe Cys Trp Glu Ile Pro

130 135 140

Val Gln Ile Val Ser His Leu

145 150

<210> 195

<211> 392

<212> PRT

<213> Homo sapiens

<400> 195

Met Asn Cys Cys Ile Gly Leu Gly Glu Lys Ala Arg Gly Ser His Arg

1 5 10 15

Ala Ser Tyr Pro Ser Leu Ser Ala Leu Phe Thr Glu Ala Ser Ile Leu

20 25 30

Gly Phe Gly Ser Phe Ala Val Lys Ala Gln Trp Thr Glu Asp Cys Arg

35 40 45

Lys Ser Thr Tyr Pro Pro Ser Gly Pro Thr Tyr Arg Gly Ala Val Pro

50 55 60

Trp Tyr Thr Ile Asn Leu Asp Leu Pro Pro Tyr Lys Arg Trp His Glu

65 70 75 80

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

85 90 95

Lys Asn Met Ile Asn Thr Phe Val Pro Ser Gly Lys Ile Met Gln Val

100 105 110

Val Asp Glu Lys Leu Pro Gly Leu Leu Gly Asn Phe Pro Gly Pro Phe

115 120 125

Glu Glu Glu Met Lys Gly Ile Ala Ala Val Thr Asp Ile Pro Leu Gly

130 135 140

Glu Ile Ile Ser Phe Asn Ile Phe Tyr Glu Leu Phe Thr Ile Cys Thr

145 150 155 160

Ser Ile Val Ala Glu Asp Lys Lys Gly His Leu Ile His Gly Arg Asn

165 170 175

Met Asp Phe Gly Val Phe Leu Gly Trp Asn Ile Asn Asn Asp Thr Trp

180 185 190
Val Ile Thr Glu Gln Leu Lys Pro Leu Thr Val Asn Leu Asp Phe Gln
195 200 205
Arg Asn Asn Lys Thr Val Phe Lys Ala Ser Ser Phe Ala Gly Tyr Val
210 215 220
Gly Met Leu Thr Gly Phe Lys Pro Gly Leu Phe Ser Leu Thr Leu Asn
225 230 235 240

Glu Arg Phe Ser Ile Asn Gly Gly Tyr Leu Gly Ile Leu Glu Trp Ile
245 250 255
Leu Gly Lys Lys Asp Val Met Trp Ile Gly Phe Leu Thr Arg Thr Val
260 265 270
Leu Glu Asn Ser Thr Ser Tyr Glu Glu Ala Lys Asn Leu Leu Thr Lys
275 280 285
Thr Lys Ile Leu Ala Pro Ala Tyr Phe Ile Leu Gly Gly Asn Gln Ser
290 295 300

Gly Glu Gly Cys Val Ile Thr Arg Asp Arg Lys Glu Ser Leu Asp Val
305 310 315 320
Tyr Glu Leu Asp Ala Lys Gln Gly Arg Trp Tyr Val Val Gln Thr Asn
325 330 335
Tyr Asp Arg Trp Lys His Pro Phe Phe Leu Asp Asp Arg Arg Thr Pro
340 345 350
Ala Lys Met Cys Leu Asn Arg Thr Ser Gln Glu Asn Ile Ser Phe Glu
355 360 365

Thr Met Tyr Asp Val Leu Ser Thr Lys Pro Val Leu Asn Lys Leu Thr
370 375 380
Val Tyr Thr Thr Leu Ile Asp Val
385 390

<210> 196

<211> 399

<212> PRT

<213> Homo sapiens

<400> 196

Met Lys Met Arg Phe Leu Gly Leu Val Val Cys Leu Val Leu Trp Thr
1 5 10 15
Leu His Ser Glu Gly Ser Gly Gly Lys Leu Thr Ala Val Asp Pro Glu
20 25 30
Thr Asn Met Asn Val Ser Glu Ile Ile Ser Tyr Trp Gly Phe Pro Ser
35 40 45
Glu Glu Tyr Leu Val Glu Thr Glu Asp Gly Tyr Ile Leu Cys Leu Asn
50 55 60
Arg Ile Pro His Gly Arg Lys Asn His Ser Asp Lys Gly Pro Lys Pro
65 70 75 80
Val Val Phe Leu Gln His Gly Leu Leu Ala Asp Ser Ser Asn Trp Val
85 90 95
Thr Asn Leu Ala Asn Ser Ser Leu Gly Phe Ile Leu Ala Asp Ala Gly
100 105 110
Phe Asp Val Trp Met Gly Asn Ser Arg Gly Asn Thr Trp Ser Arg Lys
115 120 125
His Lys Thr Leu Ser Val Ser Gln Asp Glu Phe Trp Ala Phe Ser Tyr
130 135 140
Asp Glu Met Ala Lys Tyr Asp Leu Pro Ala Ser Ile Asn Phe Ile Leu
145 150 155 160
Asn Lys Thr Gly Gln Glu Gln Val Tyr Tyr Val Gly His Ser Gln Gly
165 170 175
Thr Thr Ile Gly Phe Ile Ala Phe Ser Gln Ile Pro Glu Leu Ala Lys
180 185 190
Arg Ile Lys Met Phe Phe Ala Leu Gly Pro Val Ala Ser Val Ala Phe
195 200 205
Cys Thr Ser Pro Met Ala Lys Leu Gly Arg Leu Pro Asp His Leu Ile
210 215 220
Lys Asp Leu Phe Gly Asp Lys Glu Phe Leu Pro Gln Ser Ala Phe Leu
225 230 235 240
Lys Trp Leu Gly Thr His Val Cys Thr His Val Ile Leu Lys Glu Leu

245 250 255
 Cys Gly Asn Leu Cys Phe Leu Leu Cys Gly Phe Asn Glu Arg Asn Leu
 260 265 270
 Asn Met Ser Arg Val Asp Val Tyr Thr Thr His Ser Pro Ala Gly Thr
 275 280 285

 Ser Val Gln Asn Met Leu His Trp Ser Gln Ala Val Lys Phe Gln Lys
 290 295 300
 Phe Gln Ala Phe Asp Trp Gly Ser Ser Ala Lys Asn Tyr Phe His Tyr
 305 310 315 320
 Asn Gln Ser Tyr Pro Pro Thr Tyr Asn Val Lys Asp Met Leu Val Pro
 325 330 335
 Thr Ala Val Trp Ser Gly Gly His Asp Trp Leu Ala Asp Val Tyr Asp
 340 345 350

 Val Asn Ile Leu Leu Thr Gln Ile Thr Asn Leu Val Phe His Glu Ser
 355 360 365
 Ile Pro Glu Trp Glu His Leu Asp Phe Ile Trp Gly Leu Asp Ala Pro
 370 375 380
 Trp Arg Leu Tyr Asn Lys Ile Ile Asn Leu Met Arg Lys Tyr Gln
 385 390 395
 <210> 197
 <211> 653
 <212> PRT
 <213> Homo sapiens
 <400> 197
 Met Arg Pro Leu Arg Pro Arg Ala Ala Leu Leu Ala Leu Leu Ala Ser

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

Val Leu Ser Trp Asp Gln Gln Leu Asn Leu Ala Tyr Val Gly Ala Val

65 70 75 80

Pro His Arg Gly Ile Lys Gln Val Arg Thr His Trp Leu Leu Glu Leu

85 90 95

Val Thr Thr Arg Gly Ser Thr Gly Arg Gly Leu Ser Tyr Asn Phe Thr

100 105 110

His Leu Asp Gly Tyr Leu Asp Leu Leu Arg Glu Asn Gln Leu Leu Pro

115 120 125

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

130 135 140

Asp Lys Gln Gln Val Phe Glu Trp Lys Asp Leu Val Ser Ser Leu Ala

145 150 155 160

Arg Arg Tyr Ile Gly Arg Tyr Gly Leu Ala His Val Ser Lys Trp Asn

165 170 175

Phe Glu Thr Trp Asn Glu Pro Asp His His Asp Phe Asp Asn Val Ser

180 185 190

Met Thr Met Gln Gly Phe Leu Asn Tyr Tyr Asp Ala Cys Ser Glu Gly

195 200 205

Leu Arg Ala Ala Ser Pro Ala Leu Arg Leu Gly Gly Pro Gly Asp Ser

210 215 220

Phe His Thr Pro Pro Arg Ser Pro Leu Ser Trp Gly Leu Leu Arg His

225 230 235 240

Cys His Asp Gly Thr Asn Phe Phe Thr Gly Glu Ala Gly Val Arg Leu

245 250 255

Asp Tyr Ile Ser Leu His Arg Lys Gly Ala Arg Ser Ser Ile Ser Ile

260 265 270

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

275 280 285

Lys Phe Ala Asp Thr Pro Ile Tyr Asn Asp Glu Ala Asp Pro Leu Val

290 295 300

Gly Trp Ser Leu Pro Gln Pro Trp Arg Ala Asp Val Thr Tyr Ala Ala

305 310 315 320
 Met Val Val Lys Val Ile Ala Gln His Gln Asn Leu Leu Leu Ala Asn

 325 330 335
 Thr Thr Ser Ala Phe Pro Tyr Ala Leu Leu Ser Asn Asp Asn Ala Phe
 340 345 350
 Leu Ser Tyr His Pro His Pro Phe Ala Gln Arg Thr Leu Thr Ala Arg
 355 360 365
 Phe Gln Val Asn Asn Thr Arg Pro Pro His Val Gln Leu Leu Arg Lys
 370 375 380
 Pro Val Leu Thr Ala Met Gly Leu Leu Ala Leu Leu Asp Glu Glu Gln

 385 390 395 400
 Leu Trp Ala Glu Val Ser Gln Ala Gly Thr Val Leu Asp Ser Asn His
 405 410 415
 Thr Val Gly Val Leu Ala Ser Ala His Arg Pro Gln Gly Pro Ala Asp
 420 425 430
 Ala Trp Arg Ala Ala Val Leu Ile Tyr Ala Ser Asp Asp Thr Arg Ala
 435 440 445
 His Pro Asn Arg Ser Val Ala Val Thr Leu Arg Leu Arg Gly Val Pro

 450 455 460
 Pro Gly Pro Gly Leu Val Tyr Val Thr Arg Tyr Leu Asp Asn Gly Leu
 465 470 475 480
 Cys Ser Pro Asp Gly Glu Trp Arg Arg Leu Gly Arg Pro Val Phe Pro
 485 490 495
 Thr Ala Glu Gln Phe Arg Arg Met Arg Ala Ala Glu Asp Pro Val Ala
 500 505 510
 Ala Ala Pro Arg Pro Leu Pro Ala Gly Gly Arg Leu Thr Leu Arg Pro

 515 520 525
 Ala Leu Arg Leu Pro Ser Leu Leu Leu Val His Val Cys Ala Arg Pro
 530 535 540
 Glu Lys Pro Pro Gly Gln Val Thr Arg Leu Arg Ala Leu Pro Leu Thr
 545 550 555 560

Gln Gly Gln Leu Val Leu Val Trp Ser Asp Glu His Val Gly Ser Lys
565 570 575

Cys Leu Trp Thr Tyr Glu Ile Gln Phe Ser Gln Asp Gly Lys Ala Tyr
580 585 590

Thr Pro Val Ser Arg Lys Pro Ser Thr Phe Asn Leu Phe Val Phe Ser
595 600 605

Pro Asp Thr Gly Ala Val Ser Gly Ser Tyr Arg Val Arg Ala Leu Asp
610 615 620

Tyr Trp Ala Arg Pro Gly Pro Phe Ser Asp Pro Val Pro Tyr Leu Glu
625 630 635 640

Val Pro Val Pro Arg Gly Pro Pro Ser Pro Gly Asn Pro
645 650

<210> 198
<211> 550
<212> PRT
<213> Homo sapiens
<400> 198

Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu Val
1 5 10 15

Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr Gln Ala Asn Ser
20 25 30

Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg
35 40 45

Pro Ser Leu Gly Cys Tyr Gly Asp Lys Leu Val Arg Ser Pro Asn Ile
50 55 60

Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln
65 70 75 80

Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg
85 90 95

Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val His
100 105 110

Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr
115 120 125

Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly Ile Ser Ser Asn
130 135 140

His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro Pro Tyr His Pro
145 150 155 160

Ser Ser Glu Lys Tyr Glu Asn Thr Lys Thr Cys Arg Gly Pro Asp Gly
165 170 175

Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val Leu Asp Val Pro
180 185 190

Glu Gly Thr Leu Pro Asp Lys Gln Ser Thr Glu Gln Ala Ile Gln Leu
195 200 205

Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe Leu Ala Val Gly
210 215 220

Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys Glu Phe Gln Lys
225 230 235 240

Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val Pro
245 250 255

Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln
260 265 270

Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile
275 280 285

Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val
290 295 300

Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser Ala Leu Asp Asp
305 310 315 320

Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His Gly
325 330 335

Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp
340 345 350

Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala

355 360 365
 Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe
 370 375 380
 Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu
 385 390 395 400
 Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu
 405 410 415
 Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys
 420 425 430

 Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu
 435 440 445
 Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser
 450 455 460
 Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro
 465 470 475 480
 Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp
 485 490 495

 Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala
 500 505 510
 Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp
 515 520 525
 Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu
 530 535 540
 Phe Gln Leu Leu Met Pro
 545 550
 <210> 199
 <211> 502
 <212> PRT
 <213> Homo sapiens
 <400>
 199
 Met Ser Cys Pro Val Pro Ala Cys Cys Ala Leu Leu Leu Val Leu Gly

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

 Thr Ser Val Ser Ser Cys Ser Pro Ser Arg Ala Ser Leu Leu Thr Gly
 65 70 75 80
 Leu Pro Gln His Gln Asn Gly Met Tyr Gly Leu His Gln Asp Val His
 85 90 95
 His Phe Asn Ser Phe Asp Lys Val Arg Ser Leu Pro Leu Leu Leu Ser
 100 105 110
 Gln Ala Gly Val Arg Thr Gly Ile Ile Gly Lys Lys His Val Gly Pro
 115 120 125

 Glu Thr Val Tyr Pro Phe Asp Phe Ala Tyr Thr Glu Glu Asn Gly Ser
 130 135 140
 Val Leu Gln Val Gly Arg Asn Ile Thr Arg Ile Lys Leu Leu Val Arg
 145 150 155 160
 Lys Phe Leu Gln Thr Gln Asp Asp Arg Pro Phe Phe Leu Tyr Val Ala
 165 170 175
 Phe His Asp Pro His Arg Cys Gly His Ser Gln Pro Gln Tyr Gly Thr
 180 185 190

 Phe Cys Glu Lys Phe Gly Asn Gly Glu Ser Gly Met Gly Arg Ile Pro
 195 200 205
 Asp Trp Thr Pro Gln Ala Tyr Asp Pro Leu Asp Val Leu Val Pro Tyr
 210 215 220
 Phe Val Pro Asn Thr Pro Ala Ala Arg Ala Asp Leu Ala Ala Gln Tyr
 225 230 235 240
 Thr Thr Val Gly Arg Met Asp Gln Gly Val Gly Leu Val Leu Gln Glu
 245 250 255

Leu Arg Asp Ala Gly Val Leu Asn Asp Thr Leu Val Ile Phe Thr Ser
 260 265 270
 Asp Asn Gly Ile Pro Phe Pro Ser Gly Arg Thr Asn Leu Tyr Trp Pro
 275 280 285
 Gly Thr Ala Glu Pro Leu Leu Val Ser Ser Pro Glu His Pro Lys Arg
 290 295 300
 Trp Gly Gln Val Ser Glu Ala Tyr Val Ser Leu Leu Asp Leu Thr Pro
 305 310 315 320

 Thr Ile Leu Asp Trp Phe Ser Ile Pro Tyr Pro Ser Tyr Ala Ile Phe
 325 330 335
 Gly Ser Lys Thr Ile His Leu Thr Gly Arg Ser Leu Leu Pro Ala Leu
 340 345 350
 Glu Ala Glu Pro Leu Trp Ala Thr Val Phe Gly Ser Gln Ser His His
 355 360 365
 Glu Val Thr Met Ser Tyr Pro Met Arg Ser Val Gln His Arg His Phe
 370 375 380

 Arg Leu Val His Asn Leu Asn Phe Lys Met Pro Phe Pro Ile Asp Gln
 385 390 395 400
 Asp Phe Tyr Val Ser Pro Thr Phe Gln Asp Leu Leu Asn Arg Thr Thr
 405 410 415
 Ala Gly Gln Pro Thr Gly Trp Tyr Lys Asp Leu Arg His Tyr Tyr Tyr
 420 425 430
 Arg Ala Arg Trp Glu Leu Tyr Asp Arg Ser Arg Asp Pro His Glu Thr
 435 440 445

 Gln Asn Leu Ala Thr Asp Pro Arg Phe Ala Gln Leu Leu Glu Met Leu
 450 455 460
 Arg Asp Gln Leu Ala Lys Trp Gln Trp Glu Thr His Asp Pro Trp Val
 465 470 475 480
 Cys Ala Pro Asp Gly Val Leu Glu Glu Lys Leu Ser Pro Gln Cys Gln
 485 490 495
 Pro Leu His Asn Glu Leu
 500

<210> 200

<211> 743

<212> PRT

<213> Homo sapiens

<400>

```

200
Met Glu Ala Val Ala Val Ala Ala Ala Val Gly Val Leu Leu Leu Ala
1           5           10          15
Gly Ala Gly Gly Ala Ala Gly Asp Glu Ala Arg Glu Ala Ala Ala Val
          20          25          30
Arg Ala Leu Val Ala Arg Leu Leu Gly Pro Gly Pro Ala Ala Asp Phe
        35          40          45
Ser Val Ser Val Glu Arg Ala Leu Ala Ala Lys Pro Gly Leu Asp Thr
       50          55          60

Tyr Ser Leu Gly Gly Gly Gly Ala Ala Arg Val Arg Val Arg Gly Ser
65           70           75           80
Thr Gly Val Ala Ala Ala Ala Gly Leu His Arg Tyr Leu Arg Asp Phe
          85          90          95
Cys Gly Cys His Val Ala Trp Ser Gly Ser Gln Leu Arg Leu Pro Arg
        100         105         110
Pro Leu Pro Ala Val Pro Gly Glu Leu Thr Glu Ala Thr Pro Asn Arg
       115         120         125

Tyr Arg Tyr Tyr Gln Asn Val Cys Thr Gln Ser Tyr Ser Phe Val Trp
130          135          140
Trp Asp Trp Ala Arg Trp Glu Arg Glu Ile Asp Trp Met Ala Leu Asn
145          150          155          160
Gly Ile Asn Leu Ala Leu Ala Trp Ser Gly Gln Glu Ala Ile Trp Gln
        165         170         175
Arg Val Tyr Leu Ala Leu Gly Leu Thr Gln Ala Glu Ile Asn Glu Phe
       180         185         190

Phe Thr Gly Pro Ala Phe Leu Ala Trp Gly Arg Met Gly Asn Leu His
195          200          205

```

Thr Trp Asp Gly Pro Leu Pro Pro Ser Trp His Ile Lys Gln Leu Tyr
210 215 220

Leu Gln His Arg Val Leu Asp Gln Met Arg Ser Phe Gly Met Thr Pro
225 230 235 240

Val Leu Pro Ala Phe Ala Gly His Val Pro Glu Ala Val Thr Arg Val
245 250 255

Phe Pro Gln Val Asn Val Thr Lys Met Gly Ser Trp Gly His Phe Asn
260 265 270

Cys Ser Tyr Ser Cys Ser Phe Leu Leu Ala Pro Glu Asp Pro Ile Phe
275 280 285

Pro Ile Ile Gly Ser Leu Phe Leu Arg Glu Leu Ile Lys Glu Phe Gly
290 295 300

Thr Asp His Ile Tyr Gly Ala Asp Thr Phe Asn Glu Met Gln Pro Pro
305 310 315 320

Ser Ser Glu Pro Ser Tyr Leu Ala Ala Ala Thr Thr Ala Val Tyr Glu
325 330 335

Ala Met Thr Ala Val Asp Thr Glu Ala Val Trp Leu Leu Gln Gly Trp
340 345 350

Leu Phe Gln His Gln Pro Gln Phe Trp Gly Pro Ala Gln Ile Arg Ala
355 360 365

Val Leu Gly Ala Val Pro Arg Gly Arg Leu Leu Val Leu Asp Leu Phe
370 375 380

Ala Glu Ser Gln Pro Val Tyr Thr Arg Thr Ala Ser Phe Gln Gly Gln
385 390 395 400

Pro Phe Ile Trp Cys Met Leu His Asn Phe Gly Gly Asn His Gly Leu
405 410 415

Phe Gly Ala Leu Glu Ala Val Asn Gly Gly Pro Glu Ala Ala Arg Leu
420 425 430

Phe Pro Asn Ser Thr Met Val Gly Thr Gly Met Ala Pro Glu Gly Ile
435 440 445

Ser Gln Asn Glu Val Val Tyr Ser Leu Met Ala Glu Leu Gly Trp Arg

450 455 460
 Lys Asp Pro Val Pro Asp Leu Ala Ala Trp Val Thr Ser Phe Ala Ala
 465 470 475 480
 Arg Arg Tyr Gly Val Ser His Pro Asp Ala Gly Ala Ala Trp Arg Leu
 485 490 495
 Leu Leu Arg Ser Val Tyr Asn Cys Ser Gly Glu Ala Cys Arg Gly His
 500 505 510

 Asn Arg Ser Pro Leu Val Arg Arg Pro Ser Leu Gln Met Asn Thr Ser
 515 520 525
 Ile Trp Tyr Asn Arg Ser Asp Val Phe Glu Ala Trp Arg Leu Leu Leu
 530 535 540
 Thr Ser Ala Pro Ser Leu Ala Thr Ser Pro Ala Phe Arg Tyr Asp Leu
 545 550 555 560
 Leu Asp Leu Thr Arg Gln Ala Val Gln Glu Leu Val Ser Leu Tyr Tyr
 565 570 575

 Glu Glu Ala Arg Ser Ala Tyr Leu Ser Lys Glu Leu Ala Ser Leu Leu
 580 585 590
 Arg Ala Gly Gly Val Leu Ala Tyr Glu Leu Leu Pro Ala Leu Asp Glu
 595 600 605
 Val Leu Ala Ser Asp Ser Arg Phe Leu Leu Gly Ser Trp Leu Glu Gln
 610 615 620
 Ala Arg Ala Ala Ala Val Ser Glu Ala Glu Ala Asp Phe Tyr Glu Gln
 625 630 635 640

 Asn Ser Arg Tyr Gln Leu Thr Leu Trp Gly Pro Glu Gly Asn Ile Leu
 645 650 655
 Asp Tyr Ala Asn Lys Gln Leu Ala Gly Leu Val Ala Asn Tyr Tyr Thr
 660 665 670
 Pro Arg Trp Arg Leu Phe Leu Glu Ala Leu Val Asp Ser Val Ala Gln
 675 680 685
 Gly Ile Pro Phe Gln Gln His Gln Phe Asp Lys Asn Val Phe Gln Leu
 690 695 700

Glu Gln Ala Phe Val Leu Ser Lys Gln Arg Tyr Pro Ser Gln Pro Arg
705 710 715 720
Gly Asp Thr Val Asp Leu Ala Lys Lys Ile Phe Leu Lys Tyr Tyr Pro
725 730 735
Arg Trp Val Ala Gly Ser Trp
740
<210> 201
<211> 635
<212> PRT
<213> Homo sapiens
<400> 201
Met Ser Gly Ala Gly Arg Ala Leu Ala Ala Leu Leu Leu Ala Ala Ser
1 5 10 15
Val Leu Ser Ala Ala Leu Leu Ala Pro Gly Gly Ser Ser Gly Arg Asp
20 25 30
Ala Gln Ala Ala Pro Pro Arg Asp Leu Asp Lys Lys Arg His Ala Glu
35 40 45
Leu Lys Met Asp Gln Ala Leu Leu Leu Ile His Asn Glu Leu Leu Trp
50 55 60
Thr Asn Leu Thr Val Tyr Trp Lys Ser Glu Cys Cys Tyr His Cys Leu
65 70 75 80
Phe Gln Val Leu Val Asn Val Pro Gln Ser Pro Lys Ala Gly Lys Pro
85 90 95
Ser Ala Ala Ala Ala Ser Val Ser Thr Gln His Gly Ser Ile Leu Gln
100 105 110
Leu Asn Asp Thr Leu Glu Glu Lys Glu Val Cys Arg Leu Glu Tyr Arg
115 120 125
Phe Gly Glu Phe Gly Asn Tyr Ser Leu Leu Val Lys Asn Ile His Asn
130 135 140
Gly Val Ser Glu Ile Ala Cys Asp Leu Ala Val Asn Glu Asp Pro Val
145 150 155 160
Asp Ser Asn Leu Pro Val Ser Ile Ala Phe Leu Ile Gly Leu Ala Val

165 170 175
 Ile Ile Val Ile Ser Phe Leu Arg Leu Leu Leu Ser Leu Asp Asp Phe
 180 185 190
 Asn Asn Trp Ile Ser Lys Ala Ile Ser Ser Arg Glu Thr Asp Arg Leu
 195 200 205

 Ile Asn Ser Glu Leu Gly Ser Pro Ser Arg Thr Asp Pro Leu Asp Gly
 210 215 220
 Asp Val Gln Pro Ala Thr Trp Arg Leu Ser Ala Leu Pro Pro Arg Leu
 225 230 235 240
 Arg Ser Val Asp Thr Phe Arg Gly Ile Ala Leu Ile Leu Met Val Phe
 245 250 255
 Val Asn Tyr Gly Gly Gly Lys Tyr Trp Tyr Phe Lys His Ala Ser Trp
 260 265 270

 Asn Gly Leu Thr Val Ala Asp Leu Val Phe Pro Trp Phe Val Phe Ile
 275 280 285
 Met Gly Ser Ser Ile Phe Leu Ser Met Thr Ser Ile Leu Gln Arg Gly
 290 295 300
 Cys Ser Lys Phe Arg Leu Leu Gly Lys Ile Ala Trp Arg Ser Phe Leu
 305 310 315 320
 Leu Ile Cys Ile Gly Ile Ile Ile Val Asn Pro Asn Tyr Cys Leu Gly
 325 330 335

 Pro Leu Ser Trp Asp Lys Val Arg Ile Pro Gly Val Leu Gln Arg Leu
 340 345 350
 Gly Val Thr Tyr Phe Val Val Ala Val Leu Glu Leu Leu Phe Ala Lys
 355 360 365
 Pro Val Pro Glu His Cys Ala Ser Glu Arg Ser Cys Leu Ser Leu Arg
 370 375 380
 Asp Ile Thr Ser Ser Trp Pro Gln Trp Leu Leu Ile Leu Val Leu Glu
 385 390 395 400

 Gly Leu Trp Leu Gly Leu Thr Phe Leu Leu Pro Val Pro Gly Cys Pro
 405 410 415

Thr Gly Tyr Leu Gly Pro Gly Gly Ile Gly Asp Phe Gly Lys Tyr Pro
420 425 430

Asn Cys Thr Gly Gly Ala Ala Gly Tyr Ile Asp Arg Leu Leu Leu Gly
435 440 445

Asp Asp His Leu Tyr Gln His Pro Ser Ser Ala Val Leu Tyr His Thr
450 455 460

Glu Val Ala Tyr Asp Pro Glu Gly Ile Leu Gly Thr Ile Asn Ser Ile
465 470 475 480

Val Met Ala Phe Leu Gly Val Gln Ala Gly Lys Ile Leu Leu Tyr Tyr
485 490 495

Lys Ala Arg Thr Lys Asp Ile Leu Ile Arg Phe Thr Ala Trp Cys Cys
500 505 510

Ile Leu Gly Leu Ile Ser Val Ala Leu Thr Lys Val Ser Glu Asn Glu
515 520 525

Gly Phe Ile Pro Val Asn Lys Asn Leu Trp Ser Leu Ser Tyr Val Thr
530 535 540

Thr Leu Ser Ser Phe Ala Phe Phe Ile Leu Leu Val Leu Tyr Pro Val
545 550 555 560

Val Asp Val Lys Gly Leu Trp Thr Gly Thr Pro Phe Phe Tyr Pro Gly
565 570 575

Met Asn Ser Ile Leu Val Tyr Val Gly His Glu Val Phe Glu Asn Tyr
580 585 590

Phe Pro Phe Gln Trp Lys Leu Lys Asp Asn Gln Ser His Lys Glu His
595 600 605

Leu Thr Gln Asn Ile Val Ala Thr Ala Leu Trp Val Leu Ile Ala Tyr
610 615 620

Ile Leu Tyr Arg Lys Lys Ile Phe Trp Lys Ile
625 630 635

<210> 202

<211> 552

<212> PRT

<213> Homo sapiens

<400> 202

Met Arg Leu Leu Pro Leu Ala Pro Gly Arg Leu Arg Arg Gly Ser Pro
1 5 10 15

Arg His Leu Pro Ser Cys Ser Pro Ala Leu Leu Leu Val Leu Gly
20 25 30

Gly Cys Leu Gly Val Phe Gly Val Ala Ala Gly Thr Arg Arg Pro Asn
35 40 45

Val Val Leu Leu Leu Thr Asp Asp Gln Asp Glu Val Leu Gly Gly Met
50 55 60

Thr Pro Leu Lys Lys Thr Lys Ala Leu Ile Gly Glu Met Gly Met Thr
65 70 75 80

Phe Ser Ser Ala Tyr Val Pro Ser Ala Leu Cys Cys Pro Ser Arg Ala
85 90 95

Ser Ile Leu Thr Gly Lys Tyr Pro His Asn His His Val Val Asn Asn
100 105 110

Thr Leu Glu Gly Asn Cys Ser Ser Lys Ser Trp Gln Lys Ile Gln Glu
115 120 125

Pro Asn Thr Phe Pro Ala Ile Leu Arg Ser Met Cys Gly Tyr Gln Thr
130 135 140

Phe Phe Ala Gly Lys Tyr Leu Asn Glu Tyr Gly Ala Pro Asp Ala Gly
145 150 155 160

Gly Leu Glu His Val Pro Leu Gly Trp Ser Tyr Trp Tyr Ala Leu Glu
165 170 175

Lys Asn Ser Lys Tyr Tyr Asn Tyr Thr Leu Ser Ile Asn Gly Lys Ala
180 185 190

Arg Lys His Gly Glu Asn Tyr Ser Val Asp Tyr Leu Thr Asp Val Leu
195 200 205

Ala Asn Val Ser Leu Asp Phe Leu Asp Tyr Lys Ser Asn Phe Glu Pro
210 215 220

Phe Phe Met Met Ile Ala Thr Pro Ala Pro His Ser Pro Trp Thr Ala
225 230 235 240

Ala Pro Gln Tyr Gln Lys Ala Phe Gln Asn Val Phe Ala Pro Arg Asn
245 250 255

Lys Asn Phe Asn Ile His Gly Thr Asn Lys His Trp Leu Ile Arg Gln
260 265 270

Ala Lys Thr Pro Met Thr Asn Ser Ser Ile Gln Phe Leu Asp Asn Ala
275 280 285

Phe Arg Lys Arg Trp Gln Thr Leu Leu Ser Val Asp Asp Leu Val Glu
290 295 300

Lys Leu Val Lys Arg Leu Glu Phe Thr Gly Glu Leu Asn Asn Thr Tyr
305 310 315 320

Ile Phe Tyr Thr Ser Asp Asn Gly Tyr His Thr Gly Gln Phe Ser Leu
325 330 335

Pro Ile Asp Lys Arg Gln Leu Tyr Glu Phe Asp Ile Lys Val Pro Leu
340 345 350

Leu Val Arg Gly Pro Gly Ile Lys Pro Asn Gln Thr Ser Lys Met Leu
355 360 365

Val Ala Asn Ile Asp Leu Gly Pro Thr Ile Leu Asp Ile Ala Gly Tyr
370 375 380

Asp Leu Asn Lys Thr Gln Met Asp Gly Met Ser Leu Leu Pro Ile Leu
385 390 395 400

Arg Gly Ala Ser Asn Leu Thr Trp Arg Ser Asp Val Leu Val Glu Tyr
405 410 415

Gln Gly Glu Gly Arg Asn Val Thr Asp Pro Thr Cys Pro Ser Leu Ser
420 425 430

Pro Gly Val Ser Gln Cys Phe Pro Asp Cys Val Cys Glu Asp Ala Tyr
435 440 445

Asn Asn Thr Tyr Ala Cys Val Arg Thr Met Ser Ala Leu Trp Asn Leu
450 455 460

Gln Tyr Cys Glu Phe Asp Asp Gln Glu Val Phe Val Glu Val Tyr Asn
465 470 475 480

Leu Thr Ala Asp Pro Asp Gln Ile Thr Asn Ile Ala Lys Thr Ile Asp

485 490 495
 Pro Glu Leu Leu Gly Lys Met Asn Tyr Arg Leu Met Met Leu Gln Ser
 500 505 510
 Cys Ser Gly Pro Thr Cys Arg Thr Pro Gly Val Phe Asp Pro Gly Tyr
 515 520 525

 Arg Phe Asp Pro Arg Leu Met Phe Ser Asn Arg Gly Ser Val Arg Thr
 530 535 540
 Arg Arg Phe Ser Lys His Leu Leu
 545 550
 <210> 203
 <211> 522
 <212> PRT
 <213> Homo sapiens
 <400> 203
 Met Ala Ala Val Val Ala Ala Thr Arg Trp Trp Gln Leu Leu Leu Val
 1 5 10 15
 Leu Ser Ala Ala Gly Met Gly Ala Ser Gly Ala Pro Gln Pro Pro Asn
 20 25 30

 Ile Leu Leu Leu Leu Met Asp Asp Met Gly Trp Gly Asp Leu Gly Val
 35 40 45
 Tyr Gly Glu Pro Ser Arg Glu Thr Pro Asn Leu Asp Arg Met Ala Ala
 50 55 60
 Glu Gly Leu Leu Phe Pro Asn Phe Tyr Ser Ala Asn Pro Leu Cys Ser
 65 70 75 80
 Pro Ser Arg Ala Ala Leu Leu Thr Gly Arg Leu Pro Ile Arg Asn Gly
 85 90 95

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

His Arg Pro Gln Phe His Pro Leu Lys His Gly Phe Asp Glu Trp Phe
145 150 155 160

Gly Ser Pro Asn Cys His Phe Gly Pro Tyr Asp Asn Lys Ala Arg Pro
165 170 175

Asn Ile Pro Val Tyr Arg Asp Trp Glu Met Val Gly Arg Tyr Tyr Glu
180 185 190

Glu Phe Pro Ile Asn Leu Lys Thr Gly Glu Ala Asn Leu Thr Gln Ile
195 200 205

Tyr Leu Gln Glu Ala Leu Asp Phe Ile Lys Arg Gln Ala Arg His His
210 215 220

Pro Phe Phe Leu Tyr Trp Ala Val Asp Ala Thr His Ala Pro Val Tyr
225 230 235 240

Ala Ser Lys Pro Phe Leu Gly Thr Ser Gln Arg Gly Arg Tyr Gly Asp
245 250 255

Ala Val Arg Glu Ile Asp Asp Ser Ile Gly Lys Ile Leu Glu Leu Leu
260 265 270

Gln Asp Leu His Val Ala Asp Asn Thr Phe Val Phe Phe Thr Ser Asp
275 280 285

Asn Gly Ala Ala Leu Ile Ser Ala Pro Glu Gln Gly Gly Ser Asn Gly
290 295 300

Pro Phe Leu Cys Gly Lys Gln Thr Thr Phe Glu Gly Gly Met Arg Glu
305 310 315 320

Pro Ala Leu Ala Trp Trp Pro Gly His Val Thr Ala Gly Gln Val Ser
325 330 335

His Gln Leu Gly Ser Ile Met Asp Leu Phe Thr Thr Ser Leu Ala Leu
340 345 350

Ala Gly Leu Thr Pro Pro Ser Asp Arg Ala Ile Asp Gly Leu Asn Leu
355 360 365

Leu Pro Thr Leu Leu Gln Gly Arg Leu Met Asp Arg Pro Ile Phe Tyr
370 375 380

Tyr Arg Gly Asp Thr Leu Met Ala Ala Thr Leu Gly Gln His Lys Ala

385 390 395 400
 His Phe Trp Thr Trp Thr Asn Ser Trp Glu Asn Phe Arg Gln Gly Ile
 405 410 415

 Asp Phe Cys Pro Gly Gln Asn Val Ser Gly Val Thr Thr His Asn Leu
 420 425 430
 Glu Asp His Thr Lys Leu Pro Leu Ile Phe His Leu Gly Arg Asp Pro
 435 440 445
 Gly Glu Arg Phe Pro Leu Ser Phe Ala Ser Ala Glu Tyr Gln Glu Ala
 450 455 460
 Leu Ser Arg Ile Thr Ser Val Val Gln Gln His Gln Glu Ala Leu Val
 465 470 475 480

 Pro Ala Gln Pro Gln Leu Asn Val Cys Asn Trp Ala Val Met Asn Trp
 485 490 495
 Ala Pro Pro Gly Cys Glu Lys Leu Gly Lys Cys Leu Thr Pro Pro Glu
 500 505 510
 Ser Ile Pro Lys Lys Cys Leu Trp Ser His
 515 520

 <210> 204
 <211> 533
 <212> PRT
 <213> Homo sapiens
 <400> 204
 Met Gly Pro Arg Gly Ala Ala Ser Leu Pro Arg Gly Pro Gly Pro Arg
 1 5 10 15

 Arg Leu Leu Leu Pro Val Val Leu Pro Leu Leu Leu Leu Leu Leu
 20 25 30
 Ala Pro Pro Gly Ser Gly Ala Gly Ala Ser Arg Pro Pro His Leu Val
 35 40 45
 Phe Leu Leu Ala Asp Asp Leu Gly Trp Asn Asp Val Gly Phe His Gly
 50 55 60
 Ser Arg Ile Arg Thr Pro His Leu Asp Ala Leu Ala Ala Gly Gly Val

65				70				75				80			
Leu	Leu	Asp	Asn	Tyr	Tyr	Thr	Gln	Pro	Leu	Cys	Thr	Pro	Ser	Arg	Ser
				85				90				95			
Gln	Leu	Leu	Thr	Gly	Arg	Tyr	Gln	Ile	Arg	Thr	Gly	Leu	Gln	His	Gln
				100				105				110			
Ile	Ile	Trp	Pro	Cys	Gln	Pro	Ser	Cys	Val	Pro	Leu	Asp	Glu	Lys	Leu
				115				120				125			
Leu	Pro	Gln	Leu	Leu	Lys	Glu	Ala	Gly	Tyr	Thr	Thr	His	Met	Val	Gly
				130				135				140			
Lys	Trp	His	Leu	Gly	Met	Tyr	Arg	Lys	Glu	Cys	Leu	Pro	Thr	Arg	Arg
145				150				155				160			
Gly	Phe	Asp	Thr	Tyr	Phe	Gly	Tyr	Leu	Leu	Gly	Ser	Glu	Asp	Tyr	Tyr
				165				170				175			
Ser	His	Glu	Arg	Cys	Thr	Leu	Ile	Asp	Ala	Leu	Asn	Val	Thr	Arg	Cys
				180				185				190			
Ala	Leu	Asp	Phe	Arg	Asp	Gly	Glu	Glu	Val	Ala	Thr	Gly	Tyr	Lys	Asn
				195				200				205			
Met	Tyr	Ser	Thr	Asn	Ile	Phe	Thr	Lys	Arg	Ala	Ile	Ala	Leu	Ile	Thr
210				215				220							
Asn	His	Pro	Pro	Glu	Lys	Pro	Leu	Phe	Leu	Tyr	Leu	Ala	Leu	Gln	Ser
225				230				235				240			
Val	His	Glu	Pro	Leu	Gln	Val	Pro	Glu	Glu	Tyr	Leu	Lys	Pro	Tyr	Asp
				245				250				255			
Phe	Ile	Gln	Asp	Lys	Asn	Arg	His	His	Tyr	Ala	Gly	Met	Val	Ser	Leu
				260				265				270			
Met	Asp	Glu	Ala	Val	Gly	Asn	Val	Thr	Ala	Ala	Leu	Lys	Ser	Ser	Gly
275				280				285							
Leu	Trp	Asn	Asn	Thr	Val	Phe	Ile	Phe	Ser	Thr	Asp	Asn	Gly	Gly	Gln
290				295				300							
Thr	Leu	Ala	Gly	Gly	Asn	Asn	Trp	Pro	Leu	Arg	Gly	Arg	Lys	Trp	Ser
305				310				315				320			

Leu Trp Glu Gly Gly Val Arg Gly Val Gly Phe Val Ala Ser Pro Leu
325 330 335

Leu Lys Gln Lys Gly Val Lys Asn Arg Glu Leu Ile His Ile Ser Asp
340 345 350

Trp Leu Pro Thr Leu Val Lys Leu Ala Arg Gly His Thr Asn Gly Thr
355 360 365

Lys Pro Leu Asp Gly Phe Asp Val Trp Lys Thr Ile Ser Glu Gly Ser
370 375 380

Pro Ser Pro Arg Ile Glu Leu Leu His Asn Ile Asp Pro Asn Phe Val
385 390 395 400

Asp Ser Ser Pro Cys Pro Arg Asn Ser Met Ala Pro Ala Lys Asp Asp
405 410 415

Ser Ser Leu Pro Glu Tyr Ser Ala Phe Asn Thr Ser Val His Ala Ala
420 425 430

Ile Arg His Gly Asn Trp Lys Leu Leu Thr Gly Tyr Pro Gly Cys Gly
435 440 445

Tyr Trp Phe Pro Pro Pro Ser Gln Tyr Asn Val Ser Glu Ile Pro Ser
450 455 460

Ser Asp Pro Pro Thr Lys Thr Leu Trp Leu Phe Asp Ile Asp Arg Asp
465 470 475 480

Pro Glu Glu Arg His Asp Leu Ser Arg Glu Tyr Pro His Ile Val Thr
485 490 495

Lys Leu Leu Ser Arg Leu Gln Phe Tyr His Lys His Ser Val Pro Val
500 505 510

Tyr Phe Pro Ala Gln Asp Pro Arg Cys Asp Pro Lys Ala Thr Gly Val
515 520 525

Trp Gly Pro Trp Met
530

<210> 205

<211> 651

<212> PRT

<213> Homo sapiens

<400> 205

Met Ala Arg Gly Ser Ala Val Ala Trp Ala Ala Leu Gly Pro Leu Leu

1 5 10 15

Trp Gly Cys Ala Leu Gly Leu Gln Gly Gly Met Leu Tyr Pro Gln Glu

20 25 30

Ser Pro Ser Arg Glu Cys Lys Glu Leu Asp Gly Leu Trp Ser Phe Arg

35 40 45

Ala Asp Phe Ser Asp Asn Arg Arg Arg Gly Phe Glu Glu Gln Trp Tyr

50 55 60

Arg Arg Pro Leu Trp Glu Ser Gly Pro Thr Val Asp Met Pro Val Pro

65 70 75 80

Ser Ser Phe Asn Asp Ile Ser Gln Asp Trp Arg Leu Arg His Phe Val

85 90 95

Gly Trp Val Trp Tyr Glu Arg Glu Val Ile Leu Pro Glu Arg Trp Thr

100 105 110

Gln Asp Leu Arg Thr Arg Val Val Leu Arg Ile Gly Ser Ala His Ser

115 120 125

Tyr Ala Ile Val Trp Val Asn Gly Val Asp Thr Leu Glu His Glu Gly

130 135 140

Gly Tyr Leu Pro Phe Glu Ala Asp Ile Ser Asn Leu Val Gln Val Gly

145 150 155 160

Pro Leu Pro Ser Arg Leu Arg Ile Thr Ile Ala Ile Asn Asn Thr Leu

165 170 175

Thr Pro Thr Thr Leu Pro Pro Gly Thr Ile Gln Tyr Leu Thr Asp Thr

180 185 190

Ser Lys Tyr Pro Lys Gly Tyr Phe Val Gln Asn Thr Tyr Phe Asp Phe

195 200 205

Phe Asn Tyr Ala Gly Leu Gln Arg Ser Val Leu Leu Tyr Thr Thr Pro

210 215 220

Thr Thr Tyr Ile Asp Asp Ile Thr Val Thr Thr Ser Val Glu Gln Asp

225	230							235							240						
Ser	Gly	Leu	Val	Asn	Tyr	Gln	Ile	Ser	Val	Lys	Gly	Ser	Asn	Leu	Phe						
				245				250				255									
Lys	Leu	Glu	Val	Arg	Leu	Leu	Asp	Ala	Glu	Asn	Lys	Val	Val	Ala	Asn						
				260				265				270									
Gly	Thr	Gly	Thr	Gln	Gly	Gln	Leu	Lys	Val	Pro	Gly	Val	Ser	Leu	Trp						
				275				280				285									
Trp	Pro	Tyr	Leu	Met	His	Glu	Arg	Pro	Ala	Tyr	Leu	Tyr	Ser	Leu	Glu						
				290				295				300									
Val	Gln	Leu	Thr	Ala	Gln	Thr	Ser	Leu	Gly	Pro	Val	Ser	Asp	Phe	Tyr						
305				310				315				320									
Thr	Leu	Pro	Val	Gly	Ile	Arg	Thr	Val	Ala	Val	Thr	Lys	Ser	Gln	Phe						
				325				330				335									
Leu	Ile	Asn	Gly	Lys	Pro	Phe	Tyr	Phe	His	Gly	Val	Asn	Lys	His	Glu						
				340				345				350									
Asp	Ala	Asp	Ile	Arg	Gly	Lys	Gly	Phe	Asp	Trp	Pro	Leu	Leu	Val	Lys						
				355				360				365									
Asp	Phe	Asn	Leu	Leu	Arg	Trp	Leu	Gly	Ala	Asn	Ala	Phe	Arg	Thr	Ser						
370				375				380													
His	Tyr	Pro	Tyr	Ala	Glu	Glu	Val	Met	Gln	Met	Cys	Asp	Arg	Tyr	Gly						
385				390				395				400									
Ile	Val	Val	Ile	Asp	Glu	Cys	Pro	Gly	Val	Gly	Leu	Ala	Leu	Pro	Gln						
				405				410				415									
Phe	Phe	Asn	Asn	Val	Ser	Leu	His	His	His	Met	Gln	Val	Met	Glu	Glu						
				420				425				430									
Val	Val	Arg	Arg	Asp	Lys	Asn	His	Pro	Ala	Val	Val	Met	Trp	Ser	Val						
435				440				445													
Ala	Asn	Glu	Pro	Ala	Ser	His	Leu	Glu	Ser	Ala	Gly	Tyr	Tyr	Leu	Lys						
450				455				460													
Met	Val	Ile	Ala	His	Thr	Lys	Ser	Leu	Asp	Pro	Ser	Arg	Pro	Val	Thr						
465				470				475				480									

Phe Val Ser Asn Ser Asn Tyr Ala Ala Asp Lys Gly Ala Pro Tyr Val
485 490 495

Asp Val Ile Cys Leu Asn Ser Tyr Tyr Ser Trp Tyr His Asp Tyr Gly
500 505 510

His Leu Glu Leu Ile Gln Leu Gln Leu Ala Thr Gln Phe Glu Asn Trp
515 520 525

Tyr Lys Lys Tyr Gln Lys Pro Ile Ile Gln Ser Glu Tyr Gly Ala Glu
530 535 540

Thr Ile Ala Gly Phe His Gln Asp Pro Pro Leu Met Phe Thr Glu Glu
545 550 555 560

Tyr Gln Lys Ser Leu Leu Glu Gln Tyr His Leu Gly Leu Asp Gln Lys
565 570 575

Arg Arg Lys Tyr Val Val Gly Glu Leu Ile Trp Asn Phe Ala Asp Phe
580 585 590

Met Thr Glu Gln Ser Pro Thr Arg Val Leu Gly Asn Lys Lys Gly Ile
595 600 605

Phe Thr Arg Gln Arg Gln Pro Lys Ser Ala Ala Phe Leu Leu Arg Glu
610 615 620

Arg Tyr Trp Lys Ile Ala Asn Glu Thr Arg Tyr Pro His Ser Val Ala
625 630 635 640

Lys Ser Gln Cys Leu Glu Asn Ser Leu Phe Thr
645 650

<210> 206

<211> 1011

<212> PRT

<213> Homo sapiens

<400> 206

Met Gly Ala Tyr Ala Arg Ala Ser Gly Val Cys Ala Arg Gly Cys Leu
1 5 10 15

Asp Ser Ala Gly Pro Trp Thr Met Ser Arg Ala Leu Arg Pro Pro Leu
20 25 30

Pro Pro Leu Cys Phe Phe Leu Leu Leu Leu Ala Ala Ala Gly Ala Arg
35 40 45

Ala Gly Gly Tyr Glu Thr Cys Pro Thr Val Gln Pro Asn Met Leu Asn
50 55 60

Val His Leu Leu Pro His Thr His Asp Asp Val Gly Trp Leu Lys Thr
65 70 75 80

Val Asp Gln Tyr Phe Tyr Gly Ile Lys Asn Asp Ile Gln His Ala Gly
85 90 95

Val Gln Tyr Ile Leu Asp Ser Val Ile Ser Ala Leu Leu Ala Asp Pro
100 105 110

Thr Arg Arg Phe Ile Tyr Val Glu Ile Ala Phe Phe Ser Arg Trp Trp
115 120 125

His Gln Gln Thr Asn Ala Thr Gln Glu Val Val Arg Asp Leu Val Arg
130 135 140

Gln Gly Arg Leu Glu Phe Ala Asn Gly Gly Trp Val Met Asn Asp Glu
145 150 155 160

Ala Ala Thr His Tyr Gly Ala Ile Val Asp Gln Met Thr Leu Gly Leu
165 170 175

Arg Phe Leu Glu Asp Thr Phe Gly Asn Asp Gly Arg Pro Arg Val Ala
180 185 190

Trp His Ile Asp Pro Phe Gly His Ser Arg Glu Gln Ala Ser Leu Phe
195 200 205

Ala Gln Met Gly Phe Asp Gly Phe Phe Phe Gly Arg Leu Asp Tyr Gln
210 215 220

Asp Lys Trp Val Arg Met Gln Lys Leu Glu Met Glu Gln Val Trp Arg
225 230 235 240

Ala Ser Thr Ser Leu Lys Pro Pro Thr Ala Asp Leu Phe Thr Gly Val
245 250 255

Leu Pro Asn Gly Tyr Asn Pro Pro Arg Asn Leu Cys Trp Asp Val Leu
260 265 270

Cys Val Asp Gln Pro Leu Val Glu Asp Pro Arg Ser Pro Glu Tyr Asn

275 Ala Lys Glu Leu Val Asp Tyr Phe Leu Asn Val Ala Thr Ala Gln Gly
 290 295 300
 Arg Tyr Tyr Arg Thr Asn His Thr Val Met Thr Met Gly Ser Asp Phe
 305 310 315 320
 Gln Tyr Glu Asn Ala Asn Met Trp Phe Lys Asn Leu Asp Lys Leu Ile
 325 330 335
 Arg Leu Val Asn Ala Gln Gln Ala Lys Gly Ser Ser Val His Val Leu
 340 345 350

 Tyr Ser Thr Pro Ala Cys Tyr Leu Trp Glu Leu Asn Lys Ala Asn Leu
 355 360 365
 Thr Trp Ser Val Lys His Asp Asp Phe Phe Pro Tyr Ala Asp Gly Pro
 370 375 380
 His Gln Phe Trp Thr Gly Tyr Phe Ser Ser Arg Pro Ala Leu Lys Arg
 385 390 395 400
 Tyr Glu Arg Leu Ser Tyr Asn Phe Leu Gln Val Cys Asn Gln Leu Glu
 405 410 415

 Ala Leu Val Gly Leu Ala Ala Asn Val Gly Pro Tyr Gly Ser Gly Asp
 420 425 430
 Ser Ala Pro Leu Asn Glu Ala Met Ala Val Leu Gln His His Asp Ala
 435 440 445
 Val Ser Gly Thr Ser Arg Gln His Val Ala Asn Asp Tyr Ala Arg Gln
 450 455 460
 Leu Ala Ala Gly Trp Gly Pro Cys Glu Val Leu Leu Ser Asn Ala Leu
 465 470 475 480

 Ala Arg Leu Arg Gly Phe Lys Asp His Phe Thr Phe Cys Gln Gln Leu
 485 490 495
 Asn Ile Ser Ile Cys Pro Leu Ser Gln Thr Ala Ala Arg Phe Gln Val
 500 505 510
 Ile Val Tyr Asn Pro Leu Gly Arg Lys Val Asn Trp Met Val Arg Leu
 515 520 525

Pro Val Ser Glu Gly Val Phe Val Val Lys Asp Pro Asn Gly Arg Thr
530 535 540

Val Pro Ser Asp Val Val Ile Phe Pro Ser Ser Asp Ser Gln Ala His
545 550 555 560

Pro Pro Glu Leu Leu Phe Ser Ala Ser Leu Pro Ala Leu Gly Phe Ser
565 570 575

Thr Tyr Ser Val Ala Gln Val Pro Arg Trp Lys Pro Gln Ala Arg Ala
580 585 590

Pro Gln Pro Ile Pro Arg Arg Ser Trp Ser Pro Ala Leu Thr Ile Glu
595 600 605

Asn Glu His Ile Arg Ala Thr Phe Asp Pro Asp Thr Gly Leu Leu Met
610 615 620

Glu Ile Met Asn Met Asn Gln Gln Leu Leu Leu Pro Val Arg Gln Thr
625 630 635 640

Phe Phe Trp Tyr Asn Ala Ser Ile Gly Asp Asn Glu Ser Asp Gln Ala
645 650 655

Ser Gly Ala Tyr Ile Phe Arg Pro Asn Gln Gln Lys Pro Leu Pro Val
660 665 670

Ser Arg Trp Ala Gln Ile His Leu Val Lys Thr Pro Leu Val Gln Glu
675 680 685

Val His Gln Asn Phe Ser Ala Trp Cys Ser Gln Val Val Arg Leu Tyr
690 695 700

Pro Gly Gln Arg His Leu Glu Leu Glu Trp Ser Val Gly Pro Ile Pro
705 710 715 720

Val Gly Asp Thr Trp Gly Lys Glu Val Ile Ser Arg Phe Asp Thr Pro
725 730 735

Leu Glu Thr Lys Gly Arg Phe Tyr Thr Asp Ser Asn Gly Arg Glu Ile
740 745 750

Leu Glu Arg Arg Arg Asp Tyr Arg Pro Thr Trp Lys Leu Asn Gln Thr
755 760 765

Glu Pro Val Ala Gly Asn Tyr Tyr Pro Val Asn Thr Arg Ile Tyr Ile

770 775 780
 Thr Asp Gly Asn Met Gln Leu Thr Val Leu Thr Asp Arg Ser Gln Gly
 785 790 795 800

Gly Ser Ser Leu Arg Asp Gly Ser Leu Glu Leu Met Val His Arg Arg
 805 810 815
 Leu Leu Lys Asp Asp Gly Arg Gly Val Ser Glu Pro Leu Met Glu Asn
 820 825 830
 Gly Ser Gly Ala Trp Val Arg Gly Arg His Leu Val Leu Leu Asp Thr
 835 840 845
 Ala Gln Ala Ala Ala Ala Gly His Arg Leu Leu Ala Glu Gln Glu Val
 850 855 860

Leu Ala Pro Gln Val Val Leu Ala Pro Gly Gly Gly Ala Ala Tyr Asn
 865 870 875 880
 Leu Gly Ala Pro Pro Arg Thr Gln Phe Ser Gly Leu Arg Arg Asp Leu
 885 890 895
 Pro Pro Ser Val His Leu Leu Thr Leu Ala Ser Trp Gly Pro Glu Met
 900 905 910
 Val Leu Leu Arg Leu Glu His Gln Phe Ala Val Gly Glu Asp Ser Gly
 915 920 925

Arg Asn Leu Ser Ala Pro Val Thr Leu Asn Leu Arg Asp Leu Phe Ser
 930 935 940
 Thr Phe Thr Ile Thr Arg Leu Gln Glu Thr Thr Leu Val Ala Asn Gln
 945 950 955 960
 Leu Arg Glu Ala Ala Ser Arg Leu Lys Trp Thr Thr Asn Thr Gly Pro
 965 970 975
 Thr Pro His Gln Thr Pro Tyr Gln Leu Asp Pro Ala Asn Ile Thr Leu
 980 985 990

Glu Pro Met Glu Ile Arg Thr Phe Leu Ala Ser Val Gln Trp Lys Glu
 995 1000 1005
 Val Asp Gly
 1010

<210> 207

<211> 879

<212> PRT

<213> Homo sapiens

<400> 207

Met Arg Leu His Leu Leu Leu Leu Leu Ala Leu Cys Gly Ala Gly Thr

1 5 10 15

Thr Ala Ala Glu Leu Ser Tyr Ser Leu Arg Gly Asn Trp Ser Ile Cys

20 25 30

Asn Gly Asn Gly Ser Leu Glu Leu Pro Gly Ala Val Pro Gly Cys Val

35 40 45

His Ser Ala Leu Phe Gln Gln Gly Leu Ile Gln Asp Ser Tyr Tyr Arg

50 55 60

Phe Asn Asp Leu Asn Tyr Arg Trp Val Ser Leu Asp Asn Trp Thr Tyr

65 70 75 80

Ser Lys Glu Phe Lys Ile Pro Phe Glu Ile Ser Lys Trp Gln Lys Val

85 90 95

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

100 105 110

Glu Val Thr Ile Gly Glu Thr Asp Asn Met Phe Asn Arg Tyr Ser Phe

115 120 125

Asp Ile Thr Asn Val Val Arg Asp Val Asn Ser Ile Glu Leu Arg Phe

130 135 140

Gln Ser Ala Val Leu Tyr Ala Ala Gln Gln Ser Lys Ala His Thr Arg

145 150 155 160

Tyr Gln Val Pro Pro Asp Cys Pro Pro Leu Val Gln Lys Gly Glu Cys

165 170 175

His Val Asn Phe Val Arg Lys Glu Gln Cys Ser Phe Ser Trp Asp Trp

180 185 190

Gly Pro Ser Phe Pro Thr Gln Gly Ile Trp Lys Asp Val Arg Ile Glu

195 200 205

Ala Tyr Asn Ile Cys His Leu Asn Tyr Phe Thr Phe Ser Pro Ile Tyr

210 215 220
Asp Lys Ser Ala Gln Glu Trp Asn Leu Glu Ile Glu Ser Thr Phe Asp

Val Val Ser Ser Lys Pro Val Gly Gly Gln Val Ile Val Ala Ile Pro

Lys Leu Gln Thr Gln Gln Thr Tyr Ser Ile Glu Leu Gln Pro Gly Lys

260 265 270

Arg Ile Val Glu Leu Phe Val Asn Ile Ser Lys Asn Ile Thr Val Glu

275 280 285

Thr Trp Trp Pro His Gly His Gly Asn Gln Thr Gly Tyr Asn Met Thr

290 295 300

Val Leu Phe Glu Leu Asp Gly Gly Leu Asn Ile Glu Lys Ser Ala Lys

Val Tyr Phe Arg Thr Val Glu Leu Ile Glu Glu Pro Ile Lys Gly Ser

Pro Gly Leu Ser Phe Tyr Phe Lys Ile Asn Gly Phe Pro Ile Phe Leu

340 345 350

Lys Gly Ser Asn Trp Ile Pro Ala Asp Ser Phe Gln Asp Arg Val Thr

355 360 365

Ser Glu Leu Leu Arg Leu Leu Leu Gln Ser Val Val Asp Ala Asn Met

370 375 380

Asn Thr Leu Arg Val Trp Gly Gly Gly Ile Tyr Glu Gln Asp Glu Phe

Tyr Glu Leu Cys Asp Glu Leu Gly Ile Met Val Trp Gln Asp Phe Met

405 410 415

Phe Ala Cys Ala Leu Tyr Pro Thr Asp Gln Gly Phe Leu Asp Ser Val

420 425 430

Thr Ala Glu Val Ala Tyr Gln Ile Lys Arg Leu Lys Ser His Pro Ser

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

450 455 460

Asn Trp Tyr His Ile Ser Phe Thr Asp Arg Pro Ile Tyr Ile Lys Asp
 465 470 475 480
 Tyr Val Thr Leu Tyr Val Lys Asn Ile Arg Glu Leu Val Leu Ala Gly
 485 490 495
 Asp Lys Ser Arg Pro Phe Ile Thr Ser Ser Pro Thr Asn Gly Ala Glu
 500 505 510
 Thr Val Ala Glu Ala Trp Val Ser Gln Asn Pro Asn Ser Asn Tyr Phe
 515 520 525
 Gly Asp Val His Phe Tyr Asp Tyr Ile Ser Asp Cys Trp Asn Trp Lys
 530 535 540
 Val Phe Pro Lys Ala Arg Phe Ala Ser Glu Tyr Gly Tyr Gln Ser Trp
 545 550 555 560
 Pro Ser Phe Ser Thr Leu Glu Lys Val Ser Ser Thr Glu Asp Trp Ser
 565 570 575
 Phe Asn Ser Lys Phe Ser Leu His Arg Gln His His Glu Gly Gly Asn
 580 585 590
 Lys Gln Met Leu Tyr Gln Ala Gly Leu His Phe Lys Leu Pro Gln Ser
 595 600 605
 Thr Asp Pro Leu Arg Thr Phe Lys Asp Thr Ile Tyr Leu Thr Gln Val
 610 615 620
 Met Gln Ala Gln Cys Val Lys Thr Glu Thr Glu Phe Tyr Arg Arg Ser
 625 630 635 640
 Arg Ser Glu Ile Val Asp Gln Gln Gly His Thr Met Gly Ala Leu Tyr
 645 650 655
 Trp Gln Leu Asn Asp Ile Trp Gln Ala Pro Ser Trp Ala Ser Leu Glu
 660 665 670
 Tyr Gly Gly Lys Trp Lys Met Leu His Tyr Phe Ala Gln Asn Phe Phe
 675 680 685
 Ala Pro Leu Leu Pro Val Gly Phe Glu Asn Glu Asn Thr Phe Tyr Ile
 690 695 700
 Tyr Gly Val Ser Asp Leu His Ser Asp Tyr Ser Met Thr Leu Ser Val

705 710 715 720
 Arg Val His Thr Trp Ser Ser Leu Glu Pro Val Cys Ser Arg Val Thr
 725 730 735
 Glu Arg Phe Val Met Lys Gly Gly Glu Ala Val Cys Leu Tyr Glu Glu

 740 745 750
 Pro Val Ser Glu Leu Leu Arg Arg Cys Gly Asn Cys Thr Arg Glu Ser
 755 760 765
 Cys Val Val Ser Phe Tyr Leu Ser Ala Asp His Glu Leu Leu Ser Pro
 770 775 780
 Thr Asn Tyr His Phe Leu Ser Ser Pro Lys Glu Ala Val Gly Leu Cys
 785 790 795 800
 Lys Ala Gln Ile Thr Ala Ile Ile Ser Gln Gln Gly Asp Ile Phe Val

 805 810 815
 Phe Asp Leu Glu Thr Ser Ala Val Ala Pro Phe Val Trp Leu Asp Val
 820 825 830
 Gly Ser Ile Pro Gly Arg Phe Ser Asp Asn Gly Phe Leu Met Thr Glu
 835 840 845
 Lys Thr Arg Thr Ile Leu Phe Tyr Pro Trp Glu Pro Thr Ser Lys Asn
 850 855 860
 Glu Leu Glu Gln Ser Phe His Val Thr Ser Leu Thr Asp Ile Tyr

 865 870 875
 <210> 208
 <211> 466
 <212> PRT
 <213> Homo sapiens
 <400> 208
 Met Arg Ala Pro Gly Met Arg Ser Arg Pro Ala Gly Pro Ala Leu Leu
 1 5 10 15
 Leu Leu Leu Leu Phe Leu Gly Ala Ala Glu Ser Val Arg Arg Ala Gln
 20 25 30
 Pro Pro Arg Arg Tyr Thr Pro Asp Trp Pro Ser Leu Asp Ser Arg Pro

35 40 45
 Leu Pro Ala Trp Phe Asp Glu Ala Lys Phe Gly Val Phe Ile His Trp
 50 55 60
 Gly Val Phe Ser Val Pro Ala Trp Gly Ser Glu Trp Phe Trp Trp His
 65 70 75 80
 Trp Gln Gly Glu Gly Arg Pro Gln Tyr Gln Arg Phe Met Arg Asp Asn
 85 90 95
 Tyr Pro Pro Gly Phe Ser Tyr Ala Asp Phe Gly Pro Gln Phe Thr Ala
 100 105 110

 Arg Phe Phe His Pro Glu Glu Trp Ala Asp Leu Phe Gln Ala Ala Gly
 115 120 125
 Ala Lys Tyr Val Val Leu Thr Thr Lys His His Glu Gly Phe Thr Asn
 130 135 140
 Trp Pro Ser Pro Val Ser Trp Asn Trp Asn Ser Lys Asp Val Gly Pro
 145 150 155 160
 His Arg Asp Leu Val Gly Glu Leu Gly Thr Ala Leu Arg Lys Arg Asn
 165 170 175

 Ile Arg Tyr Gly Leu Tyr His Ser Leu Leu Glu Trp Phe His Pro Leu
 180 185 190
 Tyr Leu Leu Asp Lys Lys Asn Gly Phe Lys Thr Gln His Phe Val Ser
 195 200 205
 Ala Lys Thr Met Pro Glu Leu Tyr Asp Leu Val Asn Ser Tyr Lys Pro
 210 215 220
 Asp Leu Ile Trp Ser Asp Gly Glu Trp Glu Cys Pro Asp Thr Tyr Trp
 225 230 235 240

 Asn Ser Thr Asn Phe Leu Ser Trp Leu Tyr Asn Asp Ser Pro Val Lys
 245 250 255
 Asp Glu Val Val Val Asn Asp Arg Trp Gly Gln Asn Cys Ser Cys His
 260 265 270
 His Gly Gly Tyr Tyr Asn Cys Glu Asp Lys Phe Lys Pro Gln Ser Leu
 275 280 285

Pro Asp His Lys Trp Glu Met Cys Thr Ser Ile Asp Lys Phe Ser Trp
 290 295 300

Gly Tyr Arg Arg Asp Met Ala Leu Ser Asp Val Thr Glu Glu Ser Glu
 305 310 315 320

Ile Ile Ser Glu Leu Val Gln Thr Val Ser Leu Gly Gly Asn Tyr Leu
 325 330 335

Leu Asn Ile Gly Pro Thr Lys Asp Gly Leu Ile Val Pro Ile Phe Gln
 340 345 350

Glu Arg Leu Leu Ala Val Gly Lys Trp Leu Ser Ile Asn Gly Glu Ala
 355 360 365

Ile Tyr Ala Ser Lys Pro Trp Arg Val Gln Trp Glu Lys Asn Thr Thr
 370 375 380

Ser Val Trp Tyr Thr Ser Lys Gly Ser Ala Val Tyr Ala Ile Phe Leu
 385 390 395 400

His Trp Pro Glu Asn Gly Val Leu Asn Leu Glu Ser Pro Ile Thr Thr
 405 410 415

Ser Thr Thr Lys Ile Thr Met Leu Gly Ile Gln Gly Asp Leu Lys Trp
 420 425 430

Ser Thr Asp Pro Asp Lys Gly Leu Phe Ile Ser Leu Pro Gln Leu Pro
 435 440 445

Pro Ser Ala Val Pro Ala Glu Phe Ala Trp Thr Ile Lys Leu Thr Gly
 450 455 460

Val Lys
 465

<210> 209

<211> 346

<212> PRT

<213> Homo sapiens

<400> 209

Met Ala Arg Lys Ser Asn Leu Pro Val Leu Leu Val Pro Phe Leu Leu
 1 5 10 15

Cys Gln Ala Leu Val Arg Cys Ser Ser Pro Leu Pro Leu Val Val Asn

20 25 30

Thr Trp Pro Phe Lys Asn Ala Thr Glu Ala Ala Trp Arg Ala Leu Ala

35 40 45

Ser Gly Gly Ser Ala Leu Asp Ala Val Glu Ser Gly Cys Ala Met Cys

50 55 60

Glu Arg Glu Gln Cys Asp Gly Ser Val Gly Phe Gly Gly Ser Pro Asp

65 70 75 80

Glu Leu Gly Glu Thr Thr Leu Asp Ala Met Ile Met Asp Gly Thr Thr

85 90 95

Met Asp Val Gly Ala Val Gly Asp Leu Arg Arg Ile Lys Asn Ala Ile

100 105 110

Gly Val Ala Arg Lys Val Leu Glu His Thr Thr His Thr Leu Leu Val

115 120 125

Gly Glu Ser Ala Thr Thr Phe Ala Gln Ser Met Gly Phe Ile Asn Glu

130 135 140

Asp Leu Ser Thr Thr Ala Ser Gln Ala Leu His Ser Asp Trp Leu Ala

145 150 155 160

Arg Asn Cys Gln Pro Asn Tyr Trp Arg Asn Val Ile Pro Asp Pro Ser

165 170 175

Lys Tyr Cys Gly Pro Tyr Lys Pro Pro Gly Ile Leu Lys Gln Asp Ile

180 185 190

Pro Ile His Lys Glu Thr Glu Asp Asp Arg Gly His Asp Thr Ile Gly

195 200 205

Met Val Val Ile His Lys Thr Gly His Ile Ala Ala Gly Thr Ser Thr

210 215 220

Asn Gly Ile Lys Phe Lys Ile His Gly Arg Val Gly Asp Ser Pro Ile

225 230 235 240

Pro Gly Ala Gly Ala Tyr Ala Asp Asp Thr Ala Gly Ala Ala Ala Ala

245 250 255

Thr Gly Asn Gly Asp Ile Leu Met Arg Phe Leu Pro Ser Tyr Gln Ala

260 265 270
Val Glu Tyr Met Arg Arg Gly Glu Asp Pro Thr Ile Ala Cys Gln Lys

275 280 285
Val Ile Ser Arg Ile Gln Lys His Phe Pro Glu Phe Phe Gly Ala Val

290 295 300
Ile Cys Ala Asn Val Thr Gly Ser Tyr Gly Ala Ala Cys Asn Lys Leu

305 310 315 320
Ser Thr Phe Thr Gln Phe Ser Phe Met Val Tyr Asn Ser Glu Lys Asn

325 330 335
Gln Pro Thr Glu Glu Lys Val Asp Cys Ile

340 345

<210> 210

<211> 415

<212> PRT

<213> Homo sapiens

<400> 210

Met Thr Gly Glu Arg Pro Ser Thr Ala Leu Pro Asp Arg Arg Trp Gly

1 5 10 15

Pro Arg Ile Leu Gly Phe Trp Gly Gly Cys Arg Val Trp Val Phe Ala

20 25 30

Ala Ile Phe Leu Leu Leu Ser Leu Ala Ala Ser Trp Ser Lys Ala Glu

35 40 45

Asn Asp Phe Gly Leu Val Gln Pro Leu Val Thr Met Glu Gln Leu Leu

50 55 60

Trp Val Ser Gly Arg Gln Ile Gly Ser Val Asp Thr Phe Arg Ile Pro

65 70 75 80

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

85 90 95

Arg Lys Met Ser Ser Ser Asp Glu Gly Ala Lys Phe Ile Ala Leu Arg

100 105 110

Arg Ser Met Asp Gln Gly Ser Thr Trp Ser Pro Thr Ala Phe Ile Val

115 120 125
 Asn Asp Gly Asp Val Pro Asp Gly Leu Asn Leu Gly Ala Val Val Ser
 130 135 140
 Asp Val Glu Thr Gly Val Val Phe Leu Phe Tyr Ser Leu Cys Ala His
 145 150 155 160
 Lys Ala Gly Cys Gln Val Ala Ser Thr Met Leu Val Trp Ser Lys Asp
 165 170 175
 Asp Gly Val Ser Trp Ser Thr Pro Arg Asn Leu Ser Leu Asp Ile Gly

 180 185 190
 Thr Glu Val Phe Ala Pro Gly Pro Gly Ser Gly Ile Gln Lys Gln Arg
 195 200 205
 Glu Pro Arg Lys Gly Arg Leu Ile Val Cys Gly His Gly Thr Leu Glu
 210 215 220
 Arg Asp Gly Val Phe Cys Leu Leu Ser Asp Asp His Gly Ala Ser Trp
 225 230 235 240
 Arg Tyr Gly Ser Gly Val Ser Gly Ile Pro Tyr Gly Gln Pro Lys Gln

 245 250 255
 Glu Asn Asp Phe Asn Pro Asp Glu Cys Gln Pro Tyr Glu Leu Pro Asp
 260 265 270
 Gly Ser Val Val Ile Asn Ala Arg Asn Gln Asn Asn Tyr His Cys His
 275 280 285
 Cys Arg Ile Val Leu Arg Ser Tyr Asp Ala Cys Asp Thr Leu Arg Pro
 290 295 300
 Arg Asp Val Thr Phe Asp Pro Glu Leu Val Asp Pro Val Val Ala Ala

 305 310 315 320
 Gly Ala Val Val Thr Ser Ser Gly Ile Val Phe Phe Ser Asn Pro Ala
 325 330 335
 His Pro Glu Phe Arg Val Asn Leu Thr Leu Arg Trp Ser Phe Ser Asn
 340 345 350
 Gly Thr Ser Trp Arg Lys Glu Thr Val Gln Leu Trp Pro Gly Pro Ser
 355 360 365

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

370

375

380

Gln Ala Pro Gln Leu Tyr Val Leu Tyr Glu Lys Gly Arg Asn His Tyr

385

390

395

400

Thr Glu Ser Ile Ser Val Ala Lys Ile Ser Val Tyr Gly Thr Leu

405

410

415

<210> 211

<211> 480

<212> PRT

<213> Homo sapiens

<400> 211

Met Ile Arg Ala Ala Pro Pro Pro Leu Phe Leu Leu Leu Leu Leu

1

5

10

15

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

20

25

30

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

35

40

45

Gln Tyr Ser Gly Tyr Leu Lys Gly Ser Gly Ser Lys His Leu His Tyr

50

55

60

Trp Phe Val Glu Ser Gln Lys Asp Pro Glu Asn Ser Pro Val Val Leu

65

70

75

80

Trp Leu Asn Gly Gly Pro Gly Cys Ser Ser Leu Asp Gly Leu Leu Thr

85

90

95

Glu His Gly Pro Phe Leu Val Gln Pro Asp Gly Val Thr Leu Glu Tyr

100

105

110

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

115

120

125

Pro Ala Gly Val Gly Phe Ser Tyr Ser Asp Asp Lys Phe Tyr Ala Thr

130

135

140

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

145

150

155

160

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

 Ser Ser Tyr Glu Gln Asn Asp Asn Ser Leu Val Tyr Phe Ala Tyr Tyr
 210 215 220
 His Gly Leu Leu Gly Asn Arg Leu Trp Ser Ser Leu Gln Thr His Cys
 225 230 235 240
 Cys Ser Gln Asn Lys Cys Asn Phe Tyr Asp Asn Lys Asp Leu Glu Cys
 245 250 255
 Val Thr Asn Leu Gln Glu Val Ala Arg Ile Val Gly Asn Ser Gly Leu
 260 265 270

 Asn Ile Tyr Asn Leu Tyr Ala Pro Cys Ala Gly Gly Val Pro Ser His
 275 280 285
 Phe Arg Tyr Glu Lys Asp Thr Val Val Val Gln Asp Leu Gly Asn Ile
 290 295 300
 Phe Thr Arg Leu Pro Leu Lys Arg Met Trp His Gln Ala Leu Leu Arg
 305 310 315 320
 Ser Gly Asp Lys Val Arg Met Asp Pro Pro Cys Thr Asn Thr Thr Ala
 325 330 335

 Ala Ser Thr Tyr Leu Asn Asn Pro Tyr Val Arg Lys Ala Leu Asn Ile
 340 345 350
 Pro Glu Gln Leu Pro Gln Trp Asp Met Cys Asn Phe Leu Val Asn Leu
 355 360 365
 Gln Tyr Arg Arg Leu Tyr Arg Ser Met Asn Ser Gln Tyr Leu Lys Leu
 370 375 380
 Leu Ser Ser Gln Lys Tyr Gln Ile Leu Leu Tyr Asn Gly Asp Val Asp
 385 390 395 400

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

405 410 415
Gln Lys Met Glu Val Gln Arg Arg Pro Trp Leu Val Lys Tyr Gly Asp
420 425 430
Ser Gly Glu Gln Ile Ala Gly Phe Val Lys Glu Phe Ser His Ile Ala
435 440 445
Phe Leu Thr Ile Lys Gly Ala Gly His Met Val Pro Thr Asp Lys Pro
450 455 460

Leu Ala Ala Phe Thr Met Phe Ser Arg Phe Leu Asn Lys Gln Pro Tyr
465 470 475 480

<210> 212

<211> 411

<212> PRT

<213> Homo sapiens

<400> 212

Met Leu Leu Lys Thr Val Leu Leu Leu Gly His Val Ala Gln Val Leu
1 5 10 15
Met Leu Asp Asn Gly Leu Leu Gln Thr Pro Pro Met Gly Trp Leu Ala
20 25 30
Trp Glu Arg Phe Arg Cys Asn Ile Asn Cys Asp Glu Asp Pro Lys Asn

35 40 45
Cys Ile Ser Glu Gln Leu Phe Met Glu Met Ala Asp Arg Met Ala Gln
50 55 60
Asp Gly Trp Arg Asp Met Gly Tyr Thr Tyr Leu Asn Ile Asp Asp Cys
65 70 75 80
Trp Ile Gly Gly Arg Asp Ala Ser Gly Arg Leu Met Pro Asp Pro Lys
85 90 95
Arg Phe Pro His Gly Ile Pro Phe Leu Ala Asp Tyr Val His Ser Leu

100 105 110
Gly Leu Lys Leu Gly Ile Tyr Ala Asp Met Gly Asn Phe Thr Cys Met
115 120 125
Gly Tyr Pro Gly Thr Thr Leu Asp Lys Val Val Gln Asp Ala Gln Thr
130 135 140

Phe Ala Glu Trp Lys Val Asp Met Leu Lys Leu Asp Gly Cys Phe Ser
 145 150 155 160
 Thr Pro Glu Glu Arg Ala Gln Gly Tyr Pro Lys Met Ala Ala Ala Leu

 165 170 175
 Asn Ala Thr Gly Arg Pro Ile Ala Phe Ser Cys Ser Trp Pro Ala Tyr
 180 185 190
 Glu Gly Gly Leu Pro Pro Arg Val Asn Tyr Ser Leu Leu Ala Asp Ile
 195 200 205
 Cys Asn Leu Trp Arg Asn Tyr Asp Asp Ile Gln Asp Ser Trp Trp Ser
 210 215 220
 Val Leu Ser Ile Leu Asn Trp Phe Val Glu His Gln Asp Ile Leu Gln

 225 230 235 240
 Pro Val Ala Gly Pro Gly His Trp Asn Asp Pro Asp Met Leu Leu Ile
 245 250 255
 Gly Asn Phe Gly Leu Ser Leu Glu Gln Ser Arg Ala Gln Met Ala Leu
 260 265 270
 Trp Thr Val Leu Ala Ala Pro Leu Leu Met Ser Thr Asp Leu Arg Thr
 275 280 285
 Ile Ser Ala Gln Asn Met Asp Ile Leu Gln Asn Pro Leu Met Ile Lys

 290 295 300
 Ile Asn Gln Asp Pro Leu Gly Ile Gln Gly Arg Arg Ile His Lys Glu
 305 310 315 320
 Lys Ser Leu Ile Glu Val Tyr Met Arg Pro Leu Ser Asn Lys Ala Ser
 325 330 335
 Ala Leu Val Phe Phe Ser Cys Arg Thr Asp Met Pro Tyr Arg Tyr His
 340 345 350
 Ser Ser Leu Gly Gln Leu Asn Phe Thr Gly Ser Val Ile Tyr Glu Ala

 355 360 365
 Gln Asp Val Tyr Ser Gly Asp Ile Ile Ser Gly Leu Arg Asp Glu Thr
 370 375 380
 Asn Phe Thr Val Ile Ile Asn Pro Ser Gly Val Val Met Trp Tyr Leu

385 390 395 400
 Tyr Pro Ile Lys Asn Leu Glu Met Ser Gln Gln
 405 410
 <210> 213
 <211> 1256
 <212> PRT
 <213> Homo sapiens
 <400> 213
 Met Leu Phe Lys Leu Leu Gln Arg Gln Thr Tyr Thr Cys Leu Ser His

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

 65 70 75 80
 Trp Val Asn Gly Thr Asp Leu Glu Leu Leu Lys Glu Leu Gln Gln Val
 85 90 95
 Arg Glu Gln Met Glu Glu Glu Gln Lys Ala Met Arg Glu Ile Leu Gly
 100 105 110
 Lys Asn Thr Thr Glu Pro Thr Lys Lys Ser Glu Lys Gln Leu Glu Cys
 115 120 125
 Leu Leu Thr His Cys Ile Lys Val Pro Met Leu Val Leu Asp Pro Ala

 130 135 140
 Leu Pro Ala Asn Ile Thr Leu Lys Asp Leu Pro Ser Leu Tyr Pro Ser
 145 150 155 160
 Phe His Ser Ala Ser Asp Ile Phe Asn Val Ala Lys Pro Lys Asn Pro
 165 170 175
 Ser Thr Asn Val Ser Val Val Val Phe Asp Ser Thr Lys Asp Val Glu
 180 185 190

Asp Ala His Ser Gly Leu Leu Lys Gly Asn Ser Arg Gln Thr Val Trp

195 200 205

Arg Gly Tyr Leu Thr Thr Asp Lys Glu Val Pro Gly Leu Val Leu Met

210 215 220

Gln Asp Leu Ala Phe Leu Ser Gly Phe Pro Pro Thr Phe Lys Glu Thr

225 230 235 240

Asn Gln Leu Lys Thr Lys Leu Pro Glu Asn Leu Ser Ser Lys Val Lys

245 250 255

Leu Leu Gln Leu Tyr Ser Glu Ala Ser Val Ala Leu Leu Lys Leu Asn

260 265 270

Asn Pro Lys Asp Phe Gln Glu Leu Asn Lys Gln Thr Lys Lys Asn Met

275 280 285

Thr Ile Asp Gly Lys Glu Leu Thr Ile Ser Pro Ala Tyr Leu Leu Trp

290 295 300

Asp Leu Ser Ala Ile Ser Gln Ser Lys Gln Asp Glu Asp Ile Ser Ala

305 310 315 320

Ser Arg Phe Glu Asp Asn Glu Glu Leu Arg Tyr Ser Leu Arg Ser Ile

325 330 335

Glu Arg His Ala Pro Trp Val Arg Asn Ile Phe Ile Val Thr Asn Gly

340 345 350

Gln Ile Pro Ser Trp Leu Asn Leu Asp Asn Pro Arg Val Thr Ile Val

355 360 365

Thr His Gln Asp Val Phe Arg Asn Leu Ser His Leu Pro Thr Phe Ser

370 375 380

Ser Pro Ala Ile Glu Ser His Ile His Arg Ile Glu Gly Leu Ser Gln

385 390 395 400

Lys Phe Ile Tyr Leu Asn Asp Asp Val Met Phe Gly Lys Asp Val Trp

405 410 415

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

420 425 430

Trp Pro Val Pro Asn Cys Ala Glu Gly Cys Pro Gly Ser Trp Ile Lys

435 440 445
Asp Gly Tyr Cys Asp Lys Ala Cys Asn Asn Ser Ala Cys Asp Trp Asp

450 455 460
Gly Gly Asp Cys Ser Gly Asn Ser Gly Gly Ser Arg Tyr Ile Ala Gly
465 470 475 480

Gly Gly Gly Thr Gly Ser Ile Gly Val Gly Gln Pro Trp Gln Phe Gly
485 490 495

Gly Gly Ile Asn Ser Val Ser Tyr Cys Asn Gln Gly Cys Ala Asn Ser
500 505 510

Trp Leu Ala Asp Lys Phe Cys Asp Gln Ala Cys Asn Val Leu Ser Cys

515 520 525
Gly Phe Asp Ala Gly Asp Cys Gly Gln Asp His Phe His Glu Leu Tyr

530 535 540
Lys Val Ile Leu Leu Pro Asn Gln Thr His Tyr Ile Ile Pro Lys Gly

545 550 555 560
Glu Cys Leu Pro Tyr Phe Ser Phe Ala Glu Val Ala Lys Arg Gly Val

565 570 575
Glu Gly Ala Tyr Ser Asp Asn Pro Ile Ile Arg His Ala Ser Ile Ala

580 585 590
Asn Lys Trp Lys Thr Ile His Leu Ile Met His Ser Gly Met Asn Ala

595 600 605
Thr Thr Ile His Phe Asn Leu Thr Phe Gln Asn Thr Asn Asp Glu Glu

610 615 620
Phe Lys Met Gln Ile Thr Val Glu Val Asp Thr Arg Glu Gly Pro Lys

625 630 635 640
Leu Asn Ser Thr Ala Gln Lys Gly Tyr Glu Asn Leu Val Ser Pro Ile

645 650 655
Thr Leu Leu Pro Glu Ala Glu Ile Leu Phe Glu Asp Ile Pro Lys Glu

660 665 670
Lys Arg Phe Pro Lys Phe Lys Arg His Asp Val Asn Ser Thr Arg Arg

675 680 685

Ala Gln Glu Glu Val Lys Ile Pro Leu Val Asn Ile Ser Leu Leu Pro
690 695 700

Lys Asp Ala Gln Leu Ser Leu Asn Thr Leu Asp Leu Gln Leu Glu His

705 710 715 720

Gly Asp Ile Thr Leu Lys Gly Tyr Asn Leu Ser Lys Ser Ala Leu Leu
725 730 735

Arg Ser Phe Leu Met Asn Ser Gln His Ala Lys Ile Lys Asn Gln Ala
740 745 750

Ile Ile Thr Asp Glu Thr Asn Asp Ser Leu Val Ala Pro Gln Glu Lys
755 760 765

Gln Val His Lys Ser Ile Leu Pro Asn Ser Leu Gly Val Ser Glu Arg

770 775 780

Leu Gln Arg Leu Thr Phe Pro Ala Val Ser Val Lys Val Asn Gly His
785 790 795 800

Asp Gln Gly Gln Asn Pro Pro Leu Asp Leu Glu Thr Thr Ala Arg Phe
805 810 815

Arg Val Glu Thr His Thr Gln Lys Thr Ile Gly Gly Asn Val Thr Lys
820 825 830

Glu Lys Pro Pro Ser Leu Ile Val Pro Leu Glu Ser Gln Met Thr Lys

835 840 845

Glu Lys Lys Ile Thr Gly Lys Glu Lys Glu Asn Ser Arg Met Glu Glu
850 855 860

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

Leu Gln His Tyr Thr Asp Ser Tyr Leu Gly Phe Leu Pro Trp Glu Lys
885 890 895

Lys Lys Tyr Phe Gln Asp Leu Leu Asp Glu Glu Glu Ser Leu Lys Thr

900 905 910

Gln Leu Ala Tyr Phe Thr Asp Ser Lys Asn Thr Gly Arg Gln Leu Lys
915 920 925

Asp Thr Phe Ala Asp Ser Leu Arg Tyr Val Asn Lys Ile Leu Asn Ser

930 935 940
 Lys Phe Gly Phe Thr Ser Arg Lys Val Pro Ala His Met Pro His Met
 945 950 955 960
 Ile Asp Arg Ile Val Met Gln Glu Leu Gln Asp Met Phe Pro Glu Glu

 965 970 975
 Phe Asp Lys Thr Ser Phe His Lys Val Arg His Ser Glu Asp Met Gln
 980 985 990
 Phe Ala Phe Ser Tyr Phe Tyr Tyr Leu Met Ser Ala Val Gln Pro Leu
 995 1000 1005
 Asn Ile Ser Gln Val Phe Asp Glu Val Asp Thr Asp Gln Ser Gly
 1010 1015 1020
 Val Leu Ser Asp Arg Glu Ile Arg Thr Leu Ala Thr Arg Ile His

 1025 1030 1035
 Glu Leu Pro Leu Ser Leu Gln Asp Leu Thr Gly Leu Glu His Met
 1040 1045 1050
 Leu Ile Asn Cys Ser Lys Met Leu Pro Ala Asp Ile Thr Gln Leu
 1055 1060 1065
 Asn Asn Ile Pro Pro Thr Gln Glu Ser Tyr Tyr Asp Pro Asn Leu
 1070 1075 1080
 Pro Pro Val Thr Lys Ser Leu Val Thr Asn Cys Lys Pro Val Thr

 1085 1090 1095
 Asp Lys Ile His Lys Ala Tyr Lys Asp Lys Asn Lys Tyr Arg Phe
 1100 1105 1110
 Glu Ile Met Gly Glu Glu Glu Ile Ala Phe Lys Met Ile Arg Thr
 1115 1120 1125
 Asn Val Ser His Val Val Gly Gln Leu Asp Asp Ile Arg Lys Asn
 1130 1135 1140
 Pro Arg Lys Phe Val Cys Leu Asn Asp Asn Ile Asp His Asn His

 1145 1150 1155
 Lys Asp Ala Gln Thr Val Lys Ala Val Leu Arg Asp Phe Tyr Glu
 1160 1165 1170

Ser Met Phe Pro Ile Pro Ser Gln Phe Glu Leu Pro Arg Glu Tyr
1175 1180 1185

Arg Asn Arg Phe Leu His Met His Glu Leu Gln Glu Trp Arg Ala
1190 1195 1200

Tyr Arg Asp Lys Leu Lys Phe Trp Thr His Cys Val Leu Ala Thr
1205 1210 1215

Leu Ile Met Phe Thr Ile Phe Ser Phe Phe Ala Glu Gln Leu Ile
1220 1225 1230

Ala Leu Lys Arg Lys Ile Phe Pro Arg Arg Arg Ile His Lys Glu
1235 1240 1245

Ala Ser Pro Asn Arg Ile Arg Val
1250 1255

<210> 214

<211> 367

<212> PRT

<213> Homo sapiens

<400> 214

Met Ile Arg Asn Trp Leu Thr Ile Phe Ile Leu Phe Pro Leu Lys Leu

1 5 10 15
Val Glu Lys Cys Glu Ser Ser Val Ser Leu Thr Val Pro Pro Val Val
20 25 30

Lys Leu Glu Asn Gly Ser Ser Thr Asn Val Ser Leu Thr Leu Arg Pro
35 40 45

Pro Leu Asn Ala Thr Leu Val Ile Thr Phe Glu Ile Thr Phe Arg Ser
50 55 60

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

65 70 75 80
Gly Val Thr Asn Ser Ser Phe Gln Val Thr Ser Gln Asn Val Gly Gln
85 90 95

Leu Thr Val Tyr Leu His Gly Asn His Ser Asn Gln Thr Gly Pro Arg
100 105 110

Ile Arg Phe Leu Val Ile Arg Ser Ser Ala Ile Ser Ile Ile Asn Gln

115 120 125
Val Ile Gly Trp Ile Tyr Phe Val Ala Trp Ser Ile Ser Phe Tyr Pro

130 135 140
Gln Val Ile Met Asn Trp Arg Arg Lys Ser Val Ile Gly Leu Ser Phe
145 150 155 160

Asp Phe Val Ala Leu Asn Leu Thr Gly Phe Val Ala Tyr Ser Val Phe
165 170 175

Asn Ile Gly Leu Leu Trp Val Pro Tyr Ile Lys Glu Gln Phe Leu Leu
180 185 190

Lys Tyr Pro Asn Gly Val Asn Pro Val Asn Ser Asn Asp Val Phe Phe

195 200 205
Ser Leu His Ala Val Val Leu Thr Leu Ile Ile Ile Val Gln Cys Cys

210 215 220
Leu Tyr Glu Arg Gly Gly Gln Arg Val Ser Trp Pro Ala Ile Gly Phe

225 230 235 240
Leu Val Leu Ala Trp Leu Phe Ala Phe Val Thr Met Ile Val Ala Ala

245 250 255
Val Gly Val Thr Thr Trp Leu Gln Phe Leu Phe Cys Phe Ser Tyr Ile

260 265 270
Lys Leu Ala Val Thr Leu Val Lys Tyr Phe Pro Gln Ala Tyr Met Asn

275 280 285
Phe Tyr Tyr Lys Ser Thr Glu Gly Trp Ser Ile Gly Asn Val Leu Leu

290 295 300
Asp Phe Thr Gly Gly Ser Phe Ser Leu Leu Gln Met Phe Leu Gln Ser

305 310 315 320
Tyr Asn Asn Asp Gln Trp Thr Leu Ile Phe Gly Asp Pro Thr Lys Phe

325 330 335
Gly Leu Gly Val Phe Ser Ile Val Phe Asp Val Val Phe Phe Ile Gln

340 345 350
His Phe Cys Leu Tyr Arg Lys Arg Pro Gly Tyr Asp Gln Leu Asn

355 360 365

<210> 215

<211> 495

<212> PRT

<213> Homo sapiens

<400> 215

Met Arg Ser Pro Val Arg Asp Leu Ala Arg Asn Asp Gly Glu Glu Ser

1 5 10 15

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

20 25 30

Pro Val Cys Cys Ser Ala Arg Tyr Asn Leu Ala Ile Leu Ala Phe Phe

35 40 45

Gly Phe Phe Ile Val Tyr Ala Leu Arg Val Asn Leu Ser Val Ala Leu

50 55 60

Val Asp Met Val Asp Ser Asn Thr Thr Leu Glu Asp Asn Arg Thr Ser

65 70 75 80

Lys Ala Cys Pro Glu His Ser Ala Pro Ile Lys Val His His Asn Gln

85 90 95

Thr Gly Lys Lys Tyr Gln Trp Asp Ala Glu Thr Gln Gly Trp Ile Leu

100 105 110

Gly Ser Phe Phe Tyr Gly Tyr Ile Ile Thr Gln Ile Pro Gly Gly Tyr

115 120 125

Val Ala Ser Lys Ile Gly Gly Lys Met Leu Leu Gly Phe Gly Ile Leu

130 135 140

Gly Thr Ala Val Leu Thr Leu Phe Thr Pro Ile Ala Ala Asp Leu Gly

145 150 155 160

Val Gly Pro Leu Ile Val Leu Arg Ala Leu Glu Gly Leu Gly Glu Gly

165 170 175

Val Thr Phe Pro Ala Met His Ala Met Trp Ser Ser Trp Ala Pro Pro

180 185 190

Leu Glu Arg Ser Lys Leu Leu Ser Ile Ser Tyr Ala Gly Ala Gln Leu

195 200 205

Gly Thr Val Ile Ser Leu Pro Leu Ser Gly Ile Ile Cys Tyr Tyr Met
 210 215 220
 Asn Trp Thr Tyr Val Phe Tyr Phe Phe Gly Thr Ile Gly Ile Phe Trp
 225 230 235 240
 Phe Leu Leu Trp Ile Trp Leu Val Ser Asp Thr Pro Gln Lys His Lys
 245 250 255
 Arg Ile Ser His Tyr Glu Lys Glu Tyr Ile Leu Ser Ser Leu Arg Asn
 260 265 270

 Gln Leu Ser Ser Gln Lys Ser Val Pro Trp Val Pro Ile Leu Lys Ser
 275 280 285
 Leu Pro Leu Trp Ala Ile Val Val Ala His Phe Ser Tyr Asn Trp Thr
 290 295 300
 Phe Tyr Thr Leu Leu Thr Leu Leu Pro Thr Tyr Met Lys Glu Ile Leu
 305 310 315 320
 Arg Phe Asn Val Gln Glu Asn Gly Phe Leu Ser Ser Leu Pro Tyr Leu
 325 330 335

 Gly Ser Trp Leu Cys Met Ile Leu Ser Gly Gln Ala Ala Asp Asn Leu
 340 345 350
 Arg Ala Lys Trp Asn Phe Ser Thr Leu Cys Val Arg Arg Ile Phe Ser
 355 360 365
 Leu Ile Gly Met Ile Gly Pro Ala Val Phe Leu Val Ala Ala Gly Phe
 370 375 380
 Ile Gly Cys Asp Tyr Ser Leu Ala Val Ala Phe Leu Thr Ile Ser Thr
 385 390 395 400

 Thr Leu Gly Gly Phe Cys Ser Ser Gly Phe Ser Ile Asn His Leu Asp
 405 410 415
 Ile Ala Pro Ser Tyr Ala Gly Ile Leu Leu Gly Ile Thr Asn Thr Phe
 420 425 430
 Ala Thr Ile Pro Gly Met Val Gly Pro Val Ile Ala Lys Ser Leu Thr
 435 440 445
 Pro Asp Asn Thr Val Gly Glu Trp Gln Thr Val Phe Tyr Ile Ala Ala

450						455						460					
Ala	Ile	Asn	Val	Phe	Gly	Ala	Ile	Phe	Phe	Thr	Leu	Phe	Ala	Lys	Gly		
465						470						475					
Glu	Val	Gln	Asn	Trp	Ala	Leu	Asn	Asp	His	His	Gly	His	Arg	His			
					485						490						
<210> 216																	
<211> 438																	
<212> PRT																	
<213> Homo sapiens																	
<400> 216																	
Met	Gly	Gly	Cys	Ala	Gly	Ser	Arg	Arg	Arg	Phe	Ser	Asp	Ser	Glu	Gly		
1						5						10					
Glu	Glu	Thr	Val	Pro	Glu	Pro	Arg	Leu	Pro	Leu	Leu	Asp	His	Gln	Gly		
					20						25						
Ala	His	Trp	Lys	Asn	Ala	Val	Gly	Phe	Trp	Leu	Leu	Gly	Leu	Cys	Asn		
					35						40						
Asn	Phe	Ser	Tyr	Val	Val	Met	Leu	Ser	Ala	Ala	His	Asp	Ile	Leu	Ser		
					50						55						
His	Lys	Arg	Thr	Ser	Gly	Asn	Gln	Ser	His	Val	Asp	Pro	Gly	Pro	Thr		
65						70						75					
Pro	Ile	Pro	His	Asn	Ser	Ser	Ser	Arg	Phe	Asp	Cys	Asn	Ser	Val	Ser		
					85						90						
Thr	Ala	Ala	Val	Leu	Leu	Ala	Asp	Ile	Leu	Pro	Thr	Leu	Val	Ile	Lys		
					100						105						
Leu	Leu	Ala	Pro	Leu	Gly	Leu	His	Leu	Leu	Pro	Tyr	Ser	Pro	Arg	Val		
					115						120						
Leu	Val	Ser	Gly	Ile	Cys	Ala	Ala	Gly	Ser	Phe	Val	Leu	Val	Ala	Phe		
					130						135						
Ser	His	Ser	Val	Gly	Thr	Ser	Leu	Cys	Gly	Val	Val	Phe	Ala	Ser	Ile		
					145						150						
Ser	Ser	Gly	Leu	Gly	Glu	Val	Thr	Phe	Leu	Ser	Leu	Thr	Ala	Phe	Tyr		

165 170 175
 Pro Arg Ala Val Ile Ser Trp Trp Ser Ser Gly Thr Gly Gly Ala Gly
 180 185 190
 Leu Leu Gly Ala Leu Ser Tyr Leu Gly Leu Thr Gln Ala Gly Leu Ser
 195 200 205
 Pro Gln Gln Thr Leu Leu Ser Met Leu Gly Ile Pro Ala Leu Leu Leu

 210 215 220
 Ala Ser Tyr Phe Leu Leu Leu Thr Ser Pro Glu Ala Gln Asp Pro Gly
 225 230 235 240
 Gly Glu Glu Glu Ala Glu Ser Ala Ala Arg Gln Pro Leu Ile Arg Thr
 245 250 255
 Glu Ala Pro Glu Ser Lys Pro Gly Ser Ser Ser Ser Leu Ser Leu Arg
 260 265 270
 Glu Arg Trp Thr Val Phe Lys Gly Leu Leu Trp Tyr Ile Val Pro Leu

 275 280 285
 Val Val Val Tyr Phe Ala Glu Tyr Phe Ile Asn Gln Gly Leu Phe Glu
 290 295 300
 Leu Leu Phe Phe Trp Asn Thr Ser Leu Ser His Ala Gln Gln Tyr Arg
 305 310 315 320
 Trp Tyr Gln Met Leu Tyr Gln Ala Gly Val Phe Ala Ser Arg Ser Ser
 325 330 335
 Leu Arg Cys Cys Arg Ile Arg Phe Thr Trp Ala Leu Ala Leu Leu Gln

 340 345 350
 Cys Leu Asn Leu Val Phe Leu Leu Ala Asp Val Trp Phe Gly Phe Leu
 355 360 365
 Pro Ser Ile Tyr Leu Val Phe Leu Ile Ile Leu Tyr Glu Gly Leu Leu
 370 375 380
 Gly Gly Ala Ala Tyr Val Asn Thr Phe His Asn Ile Ala Leu Glu Thr
 385 390 395 400
 Ser Asp Glu His Arg Glu Phe Ala Met Ala Ala Thr Cys Ile Ser Asp

 405 410 415

Thr Leu Gly Ile Ser Leu Ser Gly Leu Leu Ala Leu Pro Leu His Asp
420 425 430

Phe Leu Cys Gln Leu Ser
435

<210> 217

<211> 306

<212> PRT

<213> Homo sapiens

<400> 217

Met Ala Ser Pro Gly Cys Leu Trp Leu Leu Ala Val Ala Leu Leu Pro
1 5 10 15

Trp Thr Cys Ala Ser Arg Ala Leu Gln His Leu Asp Pro Pro Ala Pro
20 25 30

Leu Pro Leu Val Ile Trp His Gly Met Gly Asp Ser Cys Cys Asn Pro
35 40 45

Leu Ser Met Gly Ala Ile Lys Lys Met Val Glu Lys Lys Ile Pro Gly
50 55 60

Ile Tyr Val Leu Ser Leu Glu Ile Gly Lys Thr Leu Met Glu Asp Val
65 70 75 80

Glu Asn Ser Phe Phe Leu Asn Val Asn Ser Gln Val Thr Thr Val Cys
85 90 95

Gln Ala Leu Ala Lys Asp Pro Lys Leu Gln Gln Gly Tyr Asn Ala Met
100 105 110

Gly Phe Ser Gln Gly Gly Gln Phe Leu Arg Ala Val Ala Gln Arg Cys
115 120 125

Pro Ser Pro Pro Met Ile Asn Leu Ile Ser Val Gly Gly Gln His Gln
130 135 140

Gly Val Phe Gly Leu Pro Arg Cys Pro Gly Glu Ser Ser His Ile Cys
145 150 155 160

Asp Phe Ile Arg Lys Thr Leu Asn Ala Gly Ala Tyr Ser Lys Val Val
165 170 175

Gln Glu Arg Leu Val Gln Ala Glu Tyr Trp His Asp Pro Ile Lys Glu

180 185 190
 Asp Val Tyr Arg Asn His Ser Ile Phe Leu Ala Asp Ile Asn Gln Glu
 195 200 205
 Arg Gly Ile Asn Glu Ser Tyr Lys Lys Asn Leu Met Ala Leu Lys Lys

 210 215 220
 Phe Val Met Val Lys Phe Leu Asn Asp Ser Ile Val Asp Pro Val Asp
 225 230 235 240
 Ser Glu Trp Phe Gly Phe Tyr Arg Ser Gly Gln Ala Lys Glu Thr Ile
 245 250 255
 Pro Leu Gln Glu Thr Ser Leu Tyr Thr Gln Asp Arg Leu Gly Leu Lys
 260 265 270
 Glu Met Asp Asn Ala Gly Gln Leu Val Phe Leu Ala Thr Glu Gly Asp

 275 280 285
 His Leu Gln Leu Ser Glu Glu Trp Phe Tyr Ala His Ile Ile Pro Phe
 290 295 300
 Leu Gly
 305
 <210> 218
 <211> 580
 <212> PRT
 <213> Homo sapiens
 <400> 218
 Met Thr Ala Pro Ala Gly Pro Arg Gly Ser Glu Thr Glu Arg Leu Leu
 1 5 10 15
 Thr Pro Asn Pro Gly Tyr Gly Thr Gln Ala Gly Pro Ser Pro Ala Pro
 20 25 30

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

Ile Leu Phe Gly Leu Ser Asn Gln Leu Ala Val Thr Phe Arg Glu Glu
85 90 95

Asn Thr Ile Ala Phe Arg His Leu Phe Leu Leu Gly Tyr Ser Asp Gly
100 105 110

Ala Asp Asp Thr Phe Ala Ala Tyr Thr Arg Glu Gln Leu Tyr Gln Ala
115 120 125

Ile Phe His Ala Val Asp Gln Tyr Leu Ala Leu Pro Asp Val Ser Leu
130 135 140

Gly Arg Tyr Ala Tyr Val Arg Gly Gly Gly Asp Pro Trp Thr Asn Gly
145 150 155 160

Ser Gly Leu Ala Leu Cys Gln Arg Tyr Tyr His Arg Gly His Val Asp
165 170 175

Pro Ala Asn Asp Thr Phe Asp Ile Asp Pro Met Val Val Thr Asp Cys
180 185 190

Ile Gln Val Asp Pro Pro Glu Arg Pro Pro Pro Pro Pro Ser Asp Asp
195 200 205

Leu Thr Leu Leu Glu Ser Ser Ser Ser Tyr Lys Asn Leu Thr Leu Lys
210 215 220

Phe His Lys Leu Val Asn Val Thr Ile His Phe Arg Leu Lys Thr Ile
225 230 235 240

Asn Leu Gln Ser Leu Ile Asn Asn Glu Ile Pro Asp Cys Tyr Thr Phe
245 250 255

Ser Val Leu Ile Thr Phe Asp Asn Lys Ala His Ser Gly Arg Ile Pro
260 265 270

Ile Ser Leu Glu Thr Gln Ala His Ile Gln Glu Cys Lys His Pro Ser
275 280 285

Val Phe Gln His Gly Asp Asn Ser Phe Arg Leu Leu Phe Asp Val Val
290 295 300

Val Ile Leu Thr Cys Ser Leu Ser Phe Leu Leu Cys Ala Arg Ser Leu
305 310 315 320

Leu Arg Gly Phe Leu Leu Gln Asn Glu Phe Val Gly Phe Met Trp Arg

325 330 335
 Gln Arg Gly Arg Val Ile Ser Leu Trp Glu Arg Leu Glu Phe Val Asn
 340 345 350

 Gly Trp Tyr Ile Leu Leu Val Thr Ser Asp Val Leu Thr Ile Ser Gly
 355 360 365
 Thr Ile Met Lys Ile Gly Ile Glu Ala Lys Asn Leu Ala Ser Tyr Asp
 370 375 380
 Val Cys Ser Ile Leu Leu Gly Thr Ser Thr Leu Leu Val Trp Val Gly
 385 390 395 400
 Val Ile Arg Tyr Leu Thr Phe Phe His Asn Tyr Asn Ile Leu Ile Ala
 405 410 415

 Thr Leu Arg Val Ala Leu Pro Ser Val Met Arg Phe Cys Cys Cys Val
 420 425 430
 Ala Val Ile Tyr Leu Gly Tyr Cys Phe Cys Gly Trp Ile Val Leu Gly
 435 440 445
 Pro Tyr His Val Lys Phe Arg Ser Leu Ser Met Val Ser Glu Cys Leu
 450 455 460
 Phe Ser Leu Ile Asn Gly Asp Asp Met Phe Val Thr Phe Ala Ala Met
 465 470 475 480

 Gln Ala Gln Gln Gly Arg Ser Ser Leu Val Trp Leu Phe Ser Gln Leu
 485 490 495
 Tyr Leu Tyr Ser Phe Ile Ser Leu Phe Ile Tyr Met Val Leu Ser Leu
 500 505 510
 Phe Ile Ala Leu Ile Thr Gly Ala Tyr Asp Thr Ile Lys His Pro Gly
 515 520 525
 Gly Ala Gly Ala Glu Glu Ser Glu Leu Gln Ala Tyr Ile Ala Gln Cys
 530 535 540

 Gln Asp Ser Pro Thr Ser Gly Lys Phe Arg Arg Gly Ser Gly Ser Ala
 545 550 555 560
 Cys Ser Leu Leu Cys Cys Cys Gly Arg Asp Pro Ser Glu Glu His Ser
 565 570 575

Leu Leu Val Asn

580

<210> 219

<211> 524

<212> PRT

<213> Homo sapiens

<400> 219

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> 220
 <211> 726
 <212> PRT
 <213> Homo sapiens
 <400> 220
 Met Arg Pro Leu Leu Leu Leu Ala Leu Leu Gly Trp Leu Leu Leu Ala

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

 65 70 75 80
 Leu Lys Lys Ala Leu Glu Lys His Ala Asp Lys Glu Asp Leu Val Ile
 85 90 95
 Leu Phe Ala Asp Ser Tyr Asp Val Leu Phe Ala Ser Gly Pro Arg Glu
 100 105 110
 Leu Leu Lys Lys Phe Arg Gln Ala Arg Ser Gln Val Val Phe Ser Ala
 115 120 125

Glu Glu Leu Ile Tyr Pro Asp Arg Arg Leu Glu Thr Lys Tyr Pro Val

130 135 140

Val Ser Asp Gly Lys Arg Phe Leu Gly Ser Gly Gly Phe Ile Gly Tyr

145 150 155 160

Ala Pro Asn Leu Ser Lys Leu Val Ala Glu Trp Glu Gly Gln Asp Ser

165 170 175

Asp Ser Asp Gln Leu Phe Tyr Thr Lys Ile Phe Leu Asp Pro Glu Lys

180 185 190

Arg Glu Gln Ile Asn Ile Thr Leu Asp His Arg Cys Arg Ile Phe Gln

195 200 205

Asn Leu Asp Gly Ala Leu Asp Glu Val Val Leu Lys Phe Glu Met Gly

210 215 220

His Val Arg Ala Arg Asn Leu Ala Tyr Asp Thr Leu Pro Val Leu Ile

225 230 235 240

His Gly Asn Gly Pro Thr Lys Leu Gln Leu Asn Tyr Leu Gly Asn Tyr

245 250 255

Ile Pro Arg Phe Trp Thr Phe Glu Thr Gly Cys Thr Val Cys Asp Glu

260 265 270

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

275 280 285

Leu Val Gly Val Phe Ile Glu Gln Pro Thr Pro Phe Val Ser Leu Phe

290 295 300

Phe Gln Arg Leu Leu Arg Leu His Tyr Pro Gln Lys His Met Arg Leu

305 310 315 320

Phe Ile His Asn His Glu Gln His His Lys Ala Gln Val Glu Glu Phe

325 330 335

Leu Ala Gln His Gly Ser Glu Tyr Gln Ser Val Lys Leu Val Gly Pro

340 345 350

Glu Val Arg Met Ala Asn Ala Asp Ala Arg Asn Met Gly Ala Asp Leu

355 360 365

Cys Arg Gln Asp Arg Ser Cys Thr Tyr Tyr Phe Ser Val Asp Ala Asp

370 375 380
 Val Ala Leu Thr Glu Pro Asn Ser Leu Arg Leu Leu Ile Gln Gln Asn

 385 390 395 400
 Lys Asn Val Ile Ala Pro Leu Met Thr Arg His Gly Arg Leu Trp Ser
 405 410 415
 Asn Phe Trp Gly Ala Leu Ser Ala Asp Gly Tyr Tyr Ala Arg Ser Glu
 420 425 430
 Asp Tyr Val Asp Ile Val Gln Gly Arg Arg Val Gly Val Trp Asn Val
 435 440 445
 Pro Tyr Ile Ser Asn Ile Tyr Leu Ile Lys Gly Ser Ala Leu Arg Gly

 450 455 460
 Glu Leu Gln Ser Ser Asp Leu Phe His His Ser Lys Leu Asp Pro Asp
 465 470 475 480
 Met Ala Phe Cys Ala Asn Ile Arg Gln Gln Asp Val Phe Met Phe Leu
 485 490 495
 Thr Asn Arg His Thr Leu Gly His Leu Leu Ser Leu Asp Ser Tyr Arg
 500 505 510
 Thr Thr His Leu His Asn Asp Leu Trp Glu Val Phe Ser Asn Pro Glu

 515 520 525
 Asp Trp Lys Glu Lys Tyr Ile His Gln Asn Tyr Thr Lys Ala Leu Ala
 530 535 540
 Gly Lys Leu Val Glu Thr Pro Cys Pro Asp Val Tyr Trp Phe Pro Ile
 545 550 555 560
 Phe Thr Glu Val Ala Cys Glu Leu Val Glu Glu Met Glu His Phe Gly
 565 570 575
 Gln Trp Ser Leu Gly Asn Asn Lys Asp Asn Arg Ile Gln Gly Gly Tyr

 580 585 590
 Glu Asn Val Pro Thr Ile Asp Ile His Met Asn Gln Ile Gly Phe Glu
 595 600 605
 Arg Glu Trp His Lys Phe Leu Leu Glu Tyr Ile Ala Pro Met Thr Glu
 610 615 620

Lys Leu Tyr Pro Gly Tyr Tyr Thr Arg Ala Gln Phe Asp Leu Ala Phe
 625 630 635 640
 Val Val Arg Tyr Lys Pro Asp Glu Gln Pro Ser Leu Met Pro His His

 645 650 655
 Asp Ala Ser Thr Phe Thr Ile Asn Ile Ala Leu Asn Arg Val Gly Val
 660 665 670
 Asp Tyr Glu Gly Gly Gly Cys Arg Phe Leu Arg Tyr Asn Cys Ser Ile
 675 680 685
 Arg Ala Pro Arg Lys Gly Trp Thr Leu Met His Pro Gly Arg Leu Thr
 690 695 700
 His Tyr His Glu Gly Leu Pro Thr Thr Arg Gly Thr Arg Tyr Ile Ala

 705 710 715 720
 Val Ser Phe Val Asp Pro
 725
 <210> 221
 <211> 357
 <212> PRT
 <213> Homo sapiens
 <400> 221
 Met Glu Glu Gly Met Asn Val Leu His Asp Phe Gly Ile Gln Ser Thr
 1 5 10 15
 His Tyr Leu Gln Val Asn Tyr Gln Asp Ser Gln Asp Trp Phe Ile Leu
 20 25 30
 Val Ser Val Ile Ala Asp Leu Arg Asn Ala Phe Tyr Val Leu Phe Pro

 35 40 45
 Ile Trp Phe His Leu Gln Glu Ala Val Gly Ile Lys Leu Leu Trp Val
 50 55 60
 Ala Val Ile Gly Asp Trp Leu Asn Leu Val Phe Lys Trp Ile Leu Phe
 65 70 75 80
 Gly Gln Arg Pro Tyr Trp Trp Val Leu Asp Thr Asp Tyr Tyr Ser Asn
 85 90 95

Thr Ser Val Pro Leu Ile Lys Gln Phe Pro Val Thr Cys Glu Thr Gly

100 105 110

Pro Gly Ser Pro Ser Gly His Ala Met Gly Thr Ala Gly Val Tyr Tyr

115 120 125

Val Met Val Thr Ser Thr Leu Ser Ile Phe Gln Gly Lys Ile Lys Pro

130 135 140

Thr Tyr Arg Phe Arg Cys Leu Asn Val Ile Leu Trp Leu Gly Phe Trp

145 150 155 160

Ala Val Gln Leu Asn Val Cys Leu Ser Arg Ile Tyr Leu Ala Ala His

165 170 175

Phe Pro His Gln Val Val Ala Gly Val Leu Ser Gly Ile Ala Val Ala

180 185 190

Glu Thr Phe Ser His Ile His Ser Ile Tyr Asn Ala Ser Leu Lys Lys

195 200 205

Tyr Phe Leu Ile Thr Phe Phe Leu Phe Ser Phe Ala Ile Gly Phe Tyr

210 215 220

Leu Leu Leu Lys Gly Leu Gly Val Asp Leu Leu Trp Thr Leu Glu Lys

225 230 235 240

Ala Gln Arg Trp Cys Glu Gln Pro Glu Trp Val His Ile Asp Thr Thr

245 250 255

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

260 265 270

Leu Ala Leu Asn Ser Ser Met Tyr Arg Glu Ser Cys Lys Gly Lys Leu

275 280 285

Ser Lys Trp Leu Pro Phe Arg Leu Ser Ser Ile Val Ala Ser Leu Val

290 295 300

Leu Leu His Val Phe Asp Ser Leu Lys Pro Pro Ser Gln Val Glu Leu

305 310 315 320

Val Phe Tyr Val Leu Ser Phe Cys Lys Ser Ala Val Val Pro Leu Ala

325 330 335

Ser Val Ser Val Ile Pro Tyr Cys Leu Ala Gln Val Leu Gly Gln Pro

340 345 350
 His Lys Lys Ser Leu
 355
 <210
 > 222
 <211> 429
 <212> PRT
 <213> Homo sapiens
 <400> 222
 Met Ala Ala Gln Gly Tyr Gly Tyr Tyr Arg Thr Val Ile Phe Ser Ala
 1 5 10 15
 Met Phe Gly Gly Tyr Ser Leu Tyr Tyr Phe Asn Arg Lys Thr Phe Ser
 20 25 30
 Phe Val Met Pro Ser Leu Val Glu Glu Ile Pro Leu Asp Lys Asp Asp
 35 40 45
 Leu Gly Phe Ile Thr Ser Ser Gln Ser Ala Ala Tyr Ala Ile Ser Lys
 50 55 60
 Phe Val Ser Gly Val Leu Ser Asp Gln Met Ser Ala Arg Trp Leu Phe
 65 70 75 80
 Ser Ser Gly Leu Leu Leu Val Gly Leu Val Asn Ile Phe Phe Ala Trp
 85 90 95
 Ser Ser Thr Val Pro Val Phe Ala Ala Leu Trp Phe Leu Asn Gly Leu
 100 105 110
 Ala Gln Gly Leu Gly Trp Pro Pro Cys Gly Lys Val Leu Arg Lys Trp
 115 120 125
 Phe Glu Pro Ser Gln Phe Gly Thr Trp Trp Ala Ile Leu Ser Thr Ser
 130 135 140
 Met Asn Leu Ala Gly Gly Leu Gly Pro Ile Leu Ala Thr Ile Leu Ala
 145 150 155 160
 Gln Ser Tyr Ser Trp Arg Ser Thr Leu Ala Leu Ser Gly Ala Leu Cys
 165 170 175
 Val Val Val Ser Phe Leu Cys Leu Leu Leu Ile His Asn Glu Pro Ala

180 185 190
 Asp Val Gly Leu Arg Asn Leu Asp Pro Met Pro Ser Glu Gly Lys Lys
 195 200 205
 Gly Ser Leu Lys Glu Glu Ser Thr Leu Gln Glu Leu Leu Ser Pro
 210 215 220
 Tyr Leu Trp Val Leu Ser Thr Gly Tyr Leu Val Val Phe Gly Val Lys
 225 230 235 240
 Thr Cys Cys Thr Asp Trp Gly Gln Phe Phe Leu Ile Gln Glu Lys Gly

 245 250 255
 Gln Ser Ala Leu Val Gly Ser Ser Tyr Met Ser Ala Leu Glu Val Gly
 260 265 270
 Gly Leu Val Gly Ser Ile Ala Ala Gly Tyr Leu Ser Asp Arg Ala Met
 275 280 285
 Ala Lys Ala Gly Leu Ser Asn Tyr Gly Asn Pro Arg His Gly Leu Leu
 290 295 300
 Leu Phe Met Met Ala Gly Met Thr Val Ser Met Tyr Leu Phe Arg Val

 305 310 315 320
 Thr Val Thr Ser Asp Ser Pro Lys Leu Trp Ile Leu Val Leu Gly Ala
 325 330 335
 Val Phe Gly Phe Ser Ser Tyr Gly Pro Ile Ala Leu Phe Gly Val Ile
 340 345 350
 Ala Asn Glu Ser Ala Pro Pro Asn Leu Cys Gly Thr Ser His Ala Ile
 355 360 365
 Val Gly Leu Met Ala Asn Val Gly Gly Phe Leu Ala Gly Leu Pro Phe

 370 375 380
 Ser Thr Ile Ala Lys His Tyr Ser Trp Ser Thr Ala Phe Trp Val Ala
 385 390 395 400
 Glu Val Ile Cys Ala Ala Ser Thr Ala Ala Phe Phe Leu Leu Arg Asn
 405 410 415
 Ile Arg Thr Lys Met Gly Arg Val Ser Lys Lys Ala Glu
 420 425
 <210> 223

<211> 507

<212> PRT

<213> Homo sapiens

<400> 223

Met Glu Lys Trp Tyr Leu Met Thr Val Val Val Leu Ile Gly Leu Thr

1 5 10 15

Val Arg Trp Thr Val Ser Leu Asn Ser Tyr Ser Gly Ala Gly Lys Pro

20 25 30

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

35 40 45

Phe Asn Leu Pro Val Lys Gln Trp Tyr Phe Asn Ser Ser Asp Asn Asn

50 55 60

Leu Gln Tyr Trp Gly Leu Asp Tyr Pro Pro Leu Thr Ala Tyr His Ser

65 70 75 80

Leu Leu Cys Ala Tyr Val Ala Lys Phe Ile Asn Pro Asp Trp Ile Ala

85 90 95

Leu His Thr Ser Arg Gly Tyr Glu Ser Gln Ala His Lys Leu Phe Met

100 105 110

Arg Thr Thr Val Leu Ile Ala Asp Leu Leu Ile Tyr Ile Pro Ala Val

115 120 125

Val Leu Tyr Cys Cys Cys Leu Lys Glu Ile Ser Thr Lys Lys Lys Ile

130 135 140

Ala Asn Ala Leu Cys Ile Leu Leu Tyr Pro Gly Leu Ile Leu Ile Asp

145 150 155 160

Tyr Gly His Phe Gln Tyr Asn Ser Val Ser Leu Gly Phe Ala Leu Trp

165 170 175

Gly Val Leu Gly Ile Ser Cys Asp Cys Asp Leu Leu Gly Ser Leu Ala

180 185 190

Phe Cys Leu Ala Ile Asn Tyr Lys Gln Met Glu Leu Tyr His Ala Leu

195 200 205

Pro Phe Phe Cys Phe Leu Leu Gly Lys Cys Phe Lys Lys Gly Leu Lys

210 215 220
 Gly Lys Gly Phe Val Leu Leu Val Lys Leu Ala Cys Ile Val Val Ala
 225 230 235 240
 Ser Phe Val Leu Cys Trp Leu Pro Phe Phe Thr Glu Arg Glu Gln Thr
 245 250 255
 Leu Gln Val Leu Arg Arg Leu Phe Pro Val Asp Arg Gly Leu Phe Glu

 260 265 270
 Asp Lys Val Ala Asn Ile Trp Cys Ser Phe Asn Val Phe Leu Lys Ile
 275 280 285
 Lys Asp Ile Leu Pro Arg His Ile Gln Leu Ile Met Ser Phe Cys Ser
 290 295 300
 Thr Phe Leu Ser Leu Leu Pro Ala Cys Ile Lys Leu Ile Leu Gln Pro
 305 310 315 320
 Ser Ser Lys Gly Phe Lys Phe Thr Leu Val Ser Cys Ala Leu Ser Phe

 325 330 335
 Phe Leu Phe Ser Phe Gln Val His Glu Lys Ser Ile Leu Leu Val Ser
 340 345 350
 Leu Pro Val Cys Leu Val Leu Ser Glu Ile Pro Phe Met Ser Thr Trp
 355 360 365
 Phe Leu Leu Val Ser Thr Phe Ser Met Leu Pro Leu Leu Leu Lys Asp
 370 375 380
 Glu Leu Leu Met Pro Ser Val Val Thr Thr Met Ala Phe Phe Ile Ala

 385 390 395 400
 Cys Val Thr Ser Phe Ser Ile Phe Glu Lys Thr Ser Glu Glu Glu Leu
 405 410 415
 Gln Leu Lys Ser Phe Ser Ile Ser Val Arg Lys Tyr Leu Pro Cys Phe
 420 425 430
 Thr Phe Leu Ser Arg Ile Ile Gln Tyr Leu Phe Leu Ile Ser Val Ile
 435 440 445
 Thr Met Val Leu Leu Thr Leu Met Thr Val Thr Leu Asp Pro Pro Gln

 450 455 460

Lys Leu Pro Asp Leu Phe Ser Val Leu Val Cys Phe Val Ser Cys Leu
 465 470 475 480
 Asn Phe Leu Phe Phe Leu Val Tyr Phe Asn Ile Ile Ile Met Trp Asp
 485 490 495
 Ser Lys Ser Gly Arg Asn Gln Lys Lys Ile Ser
 500 505
 <210> 224
 <211> 438
 <212> PRT
 <213> Homo sapiens
 <400> 224
 Met Ala Ala Gly Leu Arg Lys Arg Gly Arg Ser Gly Ser Ala Ala Gln

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

 65 70 75 80
 Val Ile Asn Gly Thr Tyr Asp Tyr Thr Gln Leu Gln Gly Asp Thr Gly
 85 90 95
 Pro Leu Val Tyr Pro Ala Gly Phe Val Tyr Ile Phe Met Gly Leu Tyr
 100 105 110
 Tyr Ala Thr Ser Arg Gly Thr Asp Ile Arg Met Ala Gln Asn Ile Phe
 115 120 125
 Ala Val Leu Tyr Leu Ala Thr Leu Leu Leu Val Phe Leu Ile Tyr His

 130 135 140
 Gln Thr Cys Lys Val Pro Pro Phe Val Phe Phe Phe Met Cys Cys Ala
 145 150 155 160
 Ser Tyr Arg Val His Ser Ile Phe Val Leu Arg Leu Phe Asn Asp Pro

165 170 175
 Val Ala Met Val Leu Leu Phe Leu Ser Ile Asn Leu Leu Leu Ala Gln
 180 185 190
 Arg Trp Gly Trp Gly Cys Cys Phe Phe Ser Leu Ala Val Ser Val Lys

 195 200 205
 Met Asn Val Leu Leu Phe Ala Pro Gly Leu Leu Phe Leu Leu Leu Thr
 210 215 220
 Gln Phe Gly Phe Arg Gly Ala Leu Pro Lys Leu Gly Ile Cys Ala Gly
 225 230 235 240
 Leu Gln Val Val Leu Gly Leu Pro Phe Leu Leu Glu Asn Pro Ser Gly
 245 250 255
 Tyr Leu Ser Arg Ser Phe Asp Leu Gly Arg Gln Phe Leu Phe His Trp

 260 265 270
 Thr Val Asn Trp Arg Phe Leu Pro Glu Ala Leu Phe Leu His Arg Ala
 275 280 285
 Phe His Leu Ala Leu Leu Thr Ala His Leu Thr Leu Leu Leu Leu Phe
 290 295 300
 Ala Leu Cys Arg Trp His Arg Thr Gly Glu Ser Ile Leu Ser Leu Leu
 305 310 315 320
 Arg Asp Pro Ser Lys Arg Lys Val Pro Pro Gln Pro Leu Thr Pro Asn

 325 330 335
 Gln Ile Val Ser Thr Leu Phe Thr Ser Asn Phe Ile Gly Ile Cys Phe
 340 345 350
 Ser Arg Ser Leu His Tyr Gln Phe Tyr Val Trp Tyr Phe His Thr Leu
 355 360 365
 Pro Tyr Leu Leu Trp Ala Met Pro Ala Arg Trp Leu Thr His Leu Leu
 370 375 380
 Arg Leu Leu Val Leu Gly Leu Ile Glu Leu Ser Trp Asn Thr Tyr Pro

 385 390 395 400
 Ser Thr Ser Cys Ser Ser Ala Ala Leu His Ile Cys His Ala Val Ile
 405 410 415

Leu Leu Gln Leu Trp Leu Gly Pro Gln Pro Phe Pro Lys Ser Thr Gln

420

425

430

His Ser Lys Lys Ala His

435

<210> 225

<211> 447

<212> PRT

<213> Homo sapiens

<400> 225

Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val

1

5

10

15

Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg

20

25

30

Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg

35

40

45

Gly Ala Gly Gly Arg Gly Gly Asp His Pro Ser Val Ala Val Gly Ile

50

55

60

Arg Arg Val Ser Asn Val Ser Ala Ala Ser Leu Val Pro Ala Val Pro

65

70

75

80

Gln Pro Glu Ala Asp Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr

85

90

95

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

100

105

110

Trp Ala Pro Arg Glu Leu Val Leu Val Val Gln Val His Asn Arg Pro

115

120

125

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

130

135

140

Asp Asn Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Thr Glu Ile

145

150

155

160

Asn Gln Leu Ile Ala Gly Val Asn Phe Cys Pro Val Leu Gln Val Phe

165

170

175

Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp

180 185 190
Pro Arg Asp Cys Pro Arg Asp Leu Pro Lys Asn Ala Ala Leu Lys Leu

195 200 205
Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu

210 215 220
Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe

225 230 235 240
Val Trp Glu Arg Val Lys Ile Leu Arg Asp Tyr Ala Gly Leu Ile Leu

245 250 255
Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe

260 265 270
Lys Lys Met Trp Lys Leu Lys Gln Gln Glu Cys Pro Glu Cys Asp Val

275 280 285
Leu Ser Leu Gly Thr Tyr Ser Ala Ser Arg Ser Phe Tyr Gly Met Ala

290 295 300
Asp Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly

305 310 315 320
Leu Ala Leu Thr Arg Asn Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp

325 330 335
Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr

340 345 350
Leu Thr Val Ser Cys Leu Pro Lys Phe Trp Lys Val Leu Val Pro Gln

355 360 365
Ile Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His His Lys Lys

370 375 380
Thr Cys Arg Pro Ser Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn

385 390 395 400
Asn Asn Lys Gln Tyr Met Phe Pro Glu Thr Leu Thr Ile Ser Glu Lys

405 410 415
Phe Thr Val Val Ala Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly

420 425 430

Asp Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln

435 440 445

<210> 226

<211> 837

<212> PRT

<213> Homo sapiens

<400> 226

Met Ala Arg Gly Glu Arg Arg Arg Arg Ala Val Pro Ala Glu Gly Val

1 5 10 15

Arg Thr Ala Glu Arg Ala Ala Arg Gly Gly Pro Gly Arg Arg Asp Gly

20 25 30

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

35 40 45

Val Val Leu Ser Leu Ala Leu Gly Met Ser Gly Arg Trp Val Leu Ala

50 55 60

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

65 70 75 80

Leu Pro Ala Asp Ser Ser Ser Pro Ala Val Ala Pro Asp Leu Phe Trp

85 90 95

Gly Thr Tyr Arg Pro His Val Tyr Phe Gly Met Lys Thr Arg Ser Pro

100 105 110

Lys Pro Leu Leu Thr Gly Leu Met Trp Ala Gln Gln Gly Thr Thr Pro

115 120 125

Gly Thr Pro Lys Leu Arg His Thr Cys Glu Gln Gly Asp Gly Val Gly

130 135 140

Pro Tyr Gly Trp Glu Phe His Asp Gly Leu Ser Phe Gly Arg Gln His

145 150 155 160

Ile Gln Asp Gly Ala Leu Arg Leu Thr Thr Glu Phe Val Lys Arg Pro

165 170 175

Gly Gly Gln His Gly Gly Asp Trp Ser Trp Arg Val Thr Val Glu Pro

180 185 190

Gln Asp Ser Gly Thr Ser Ala Leu Pro Leu Val Ser Leu Phe Phe Tyr
195 200 205

Val Val Thr Asp Gly Lys Glu Val Leu Leu Pro Glu Val Gly Ala Lys
210 215 220

Gly Gln Leu Lys Phe Ile Ser Gly His Thr Ser Glu Leu Gly Asp Phe
225 230 235 240

Arg Phe Thr Leu Leu Pro Pro Thr Ser Pro Gly Asp Thr Ala Pro Lys
245 250 255

Tyr Gly Ser Tyr Asn Val Phe Trp Thr Ser Asn Pro Gly Leu Pro Leu
260 265 270

Leu Thr Glu Met Val Lys Ser Arg Leu Asn Ser Trp Phe Gln His Arg
275 280 285

Pro Pro Gly Ala Pro Pro Glu Arg Tyr Leu Gly Leu Pro Gly Ser Leu
290 295 300

Lys Trp Glu Asp Arg Gly Pro Ser Gly Gln Gly Gln Gly Gln Phe Leu
305 310 315 320

Ile Gln Gln Val Thr Leu Lys Ile Pro Ile Ser Ile Glu Phe Val Phe
325 330 335

Glu Ser Gly Ser Ala Gln Ala Gly Gly Asn Gln Ala Leu Pro Arg Leu
340 345 350

Ala Gly Ser Leu Leu Thr Gln Ala Leu Glu Ser His Ala Glu Gly Phe
355 360 365

Arg Glu Arg Phe Glu Lys Thr Phe Gln Leu Lys Glu Lys Gly Leu Ser
370 375 380

Ser Gly Glu Gln Val Leu Gly Gln Ala Ala Leu Ser Gly Leu Leu Gly
385 390 395 400

Gly Ile Gly Tyr Phe Tyr Gly Gln Gly Leu Val Leu Pro Asp Ile Gly
405 410 415

Val Glu Gly Ser Glu Gln Lys Val Asp Pro Ala Leu Phe Pro Pro Val
420 425 430

Pro Leu Phe Thr Ala Val Pro Ser Arg Ser Phe Phe Pro Arg Gly Phe

435 440 445
 Leu Trp Asp Glu Gly Phe His Gln Leu Val Val Gln Arg Trp Asp Pro
 450 455 460
 Ser Leu Thr Arg Glu Ala Leu Gly His Trp Leu Gly Leu Leu Asn Ala
 465 470 475 480
 Asp Gly Trp Ile Gly Arg Glu Gln Ile Leu Gly Asp Glu Ala Arg Ala
 485 490 495
 Arg Val Pro Pro Glu Phe Leu Val Gln Arg Ala Val His Ala Asn Pro
 500 505 510

 Pro Thr Leu Leu Leu Pro Val Ala His Met Leu Glu Val Gly Asp Pro
 515 520 525
 Asp Asp Leu Ala Phe Leu Arg Lys Ala Leu Pro Arg Leu His Ala Trp
 530 535 540
 Phe Ser Trp Leu His Gln Ser Gln Ala Gly Pro Leu Pro Leu Ser Tyr
 545 550 555 560
 Arg Trp Arg Gly Arg Asp Pro Ala Leu Pro Thr Leu Leu Asn Pro Lys
 565 570 575

 Thr Leu Pro Ser Gly Leu Asp Asp Tyr Pro Arg Ala Ser His Pro Ser
 580 585 590
 Val Thr Glu Arg His Leu Asp Leu Arg Cys Trp Val Ala Leu Gly Ala
 595 600 605
 Arg Val Leu Thr Arg Leu Ala Glu His Leu Gly Glu Ala Glu Val Ala
 610 615 620
 Ala Glu Leu Gly Pro Leu Ala Ala Ser Leu Glu Ala Ala Glu Ser Leu
 625 630 635 640

 Asp Glu Leu His Trp Ala Pro Glu Leu Gly Val Phe Ala Asp Phe Gly
 645 650 655
 Asn His Thr Lys Ala Val Gln Leu Lys Pro Arg Pro Pro Gln Gly Leu
 660 665 670
 Val Arg Val Val Gly Arg Pro Gln Pro Gln Leu Gln Tyr Val Asp Ala
 675 680 685

Leu Gly Tyr Val Ser Leu Phe Pro Leu Leu Leu Arg Leu Leu Asp Pro
690 695 700

Thr Ser Ser Arg Leu Gly Pro Leu Leu Asp Ile Leu Ala Asp Ser Arg
705 710 715 720

His Leu Trp Ser Pro Phe Gly Leu Arg Ser Leu Ala Ala Ser Ser Ser
725 730 735

Phe Tyr Gly Gln Arg Asn Ser Glu His Asp Pro Pro Tyr Trp Arg Gly
740 745 750

Ala Val Trp Leu Asn Val Asn Tyr Leu Ala Leu Gly Ala Leu His His
755 760 765

Tyr Gly His Leu Glu Gly Pro His Gln Ala Arg Ala Ala Lys Leu His
770 775 780

Gly Glu Leu Arg Ala Asn Val Val Gly Asn Val Trp Arg Gln Tyr Gln
785 790 795 800

Ala Thr Gly Phe Leu Trp Glu Gln Tyr Ser Asp Arg Asp Gly Arg Gly
805 810 815

Met Gly Cys Arg Pro Phe His Gly Trp Thr Ser Leu Val Leu Leu Ala
820 825 830

Met Ala Glu Asp Tyr
835

<210> 227

<211> 2523

<212> DNA

<213> Homo sapiens

<400> 227

agtcaagtga cggaagcgg ccggcctggg cgccgactgc agagccggga ggctggtggt	60
catgccgggg ttcttggttc gcatcctccc tctgttgctg gttctgctgc ttctgggccc	120
tacgcgcggc ttgcgaatg ccaccagag gatgtttgaa attgactata gccgggactc	180
cttctcaag gatggccagc catttcgcta catctcagga agcattcact actcccgtgt	240
gccccgttc tactggaagg accggtgct gaagatgaag atggctgggc tgaacccat	300
ccagacgtat gtgccctgga actttcatga gccctggcca ggacagtacc agttttctga	360

ggacatgat gtggaatatt ttcttcggct ggctcatgag ctgggactgc tggttatcct 420
 gaggcccggt ccctacatct gtgcagagtg ggaaatggga ggattacctg cttggctgct 480
 agagaaagag tctattcttc tccgtctctc cgaccagat tacctggcag ctgtggacaa 540
 gtggttggga gtccittctgc ccaagatgaa gcctctctc tatcagaatg gagggccagt 600
 tataacagt caggttgaat atgaatatgg cagctacttt gcctgtgatt ttgactacct 660
 gcgttctctg cagaagcgtt ttccccacca tctgggggat gatgtggttc tgtttaccac 720
 tgatggagca cataaaacat tctgaaatg tggggccctg cagggcctct acaccacggt 780

 ggactttgga acaggcagca acatcacaga tgctttccta agccagagga agtgtgagcc 840
 caaaggaccc ttgatcaatt ctgaattctt tactggctgg ctagatcact ggggccaacc 900
 tctctccaca atcaagaccg aagcagtggc ttctctctc tatgatatac ttgccctggg 960
 ggcgagtgtg aacttgtaca tgtttatagg tgggaccaat ttgcctatt ggaatggggc 1020
 caactaccc taigcagcac agcccaccag ctacgactat gatgccccac tgagtgaggc 1080
 tggggacctc actgagaagt attttctct gcgaaacatc atccagaagt ttgaaaaagt 1140
 accagaaggt cctatccctc catctacacc aaagtttgca tatggaaagg tcactttgga 1200

 aaagttaaag acagtgggag cagctctgga cattctgtgt ccctctgggc ccatcaaaag 1260
 cctttatccc ttgacattta tccaggtgaa acagcattat gggtttgtgc tgtaccggac 1320
 aacatttct caagattgca gcaaccagc acctctctct tcacctca atggagtcca 1380
 cgatcgagca tatgttgctg tggatgggat cccccaggga gtccttgagc gaaacaatgt 1440
 gatcactctg aacataacag ggaaagctgg agccactctg gaccttctgg tagagaacat 1500
 gggacgtgtg aactatggtg catatatcaa cgattttaag ggtttggttt ctaacctgac 1560
 tctcagttcc aatatcctca cggactggac gatctttcca ctggacactg aggatgcagt 1620

 gtgcagccac ctggggggct ggggacaccg tgacagtggc caccatgatg aagcctgggc 1680
 ccacaactca tccaactaca cgctcccggc cttttatatg gggaacttct ccattcccag 1740
 tgggatccca gacttgcctc aggacacctt tatccagttt cctggatgga ccaagggcc 1800
 ggtctggatt aatggcttta accttgccg ctattggcca gcccggggcc ctcagttgac 1860
 cttgtttgtg cccagcaca tctgatgac ctgcgcccc aacaccatca ccgtgctgga 1920
 actggagtgg gcacctgca gcagtatga tccagaacta tgtgtgtgta cgttcgtgga 1980
 caggccagtt attggtcat ctgtgacctt cgatcatccc tccaaacctg ttgaaaaaag 2040

 actcatgccc ccacccccgc aaaaaaaca agattcatgg ctggacctg tatgatgatg 2100
 aaagcctgtg tctttgaggg attctacct gaacatacct cacagatcct ccctgtcatg 2160
 ccacatttca ctgattggaa tgttgaaatg gaaaaggaat ttaggatgtg cattttcacc 2220

tgaggtttcc ctgcatccct gcagtgccaa agccccacct tcagggacca cctggaatgt	2280
gtgaggggct gacagcacag taacgtgcat acatatctgc agggctggaa tggaagcttt	2340
aaagtggtga gtgattttta ttttggaga atcatgttac ctttttgta aataaaattt	2400
gtactcaaat gatgatgta ctgtttttaa tgtgcaggta ttgaattata tggctgact	2460
taaatcataa ctagacttga gtgggctgaa taaaccactt cactaacttg aagttcaaaa	2520
gga	2523
<210> 228	
<211> 4785	
<212> DNA	
<213> Homo sapiens	
<400> 228	
ctcacgtggc cagccccctc cgagagggga gaccagcggg ccatgacaag ctccaggctt	60
tggttttcgc tgctgctggc ggcagcggtc gcaggacggg cgacggccct ctggccctgg	120
cctcagaact tccaaacctc cgaccagcgc tacgtccttt acccgaacaa ctttcaattc	180
cagtacgatg tcagctcggc cgcgagccc ggctgctcag tcctcgacga ggccttccag	240
cgctatcgtg acctgctttt cggttccggg tcttggcccc gtccttacct cacagggaaa	300
cggcatacac tggagaagaa tgtgttggtt gtctctgtag tcacacctgg atgtaaccag	360
cttctactt tggagtcagt ggagaattat accctgacca taaatgatga ccagtgttta	420
ctctctctg agactgtctg gggagctctc cgaggctctg agacttttag ccagcttggt	480
tggaaatctg ctgagggcac attctttatc aacaagactg agattgagga ctttccccgc	540
tttctcacc ggggcttgct gttggataca tctcgccatt acctgccact ctctagcatc	600
ctggacactc tggatgtcat ggctacaat aaattgaacg tgttccactg gcatctggtg	660
gatgatcctt cttcccata tgagagcttc acttttccag agctcatgag aaaggggtcc	720
tacaacctg taccacat ctacacagca caggatgtga aggaggtcat tgaatacgca	780
cggctccggg gtatccgtgt gcttcagag tttgacactc ctggccacac tttgtcctgg	840
ggaccaggta tccctggatt actgactcct tgctactctg ggtctgagcc ctctggcacc	900
tttggaccag tgaatccag tctcaataat acctatgagt tcatgagcac attcttctta	960
gaagtcagct ctgtcttccc agatttttat cttcatcttg gaggagatga ggttgatttc	1020
acctgctgga agtccaacct agagatccag gactttatga ggaagaaagg cttcggtgag	1080
gacttcaagc agctggagtc cttctacatc cagacgtgc tggacatcgt ctcttcttat	1140
ggcaagggct atgtggtgtg gcaggagggtg tttgataata aagtaaagat tcagccagac	1200

acaatcatac aggtgtggcg agaggatatt ccagtgaact atatgaagga gctggaactg	1260
gtcaccaagg ccggcttccg ggcccttctc tctgccccct ggtacctgaa ccgtatatcc	1320
tatggccctg actggaagga tttctacata gtggaacccc tggcatttga aggtaccctt	1380
gagcagaagg ctctggtgat tgggtggagag gcttgtatgt ggggagaata tgtggacaac	1440
acaaacctgg tccccaggct ctggcccaga gcaggggctg ttgccgaaag gctgtggagc	1500
aacaagttag catctgacct gacatttggc tatgaacgtt tgtcacactt ccgctgtgaa	1560
ttgctgaggc gaggtgtcca ggcccaaccc ctcaatgtag gcttctgtga gcaggagttt	1620
gaacagacct gagccccagg caccgaggag ggtgctggct gtaggtgaat gtagtggag	1680
ccaggcttcc actgcatcct ggccagggga cggagccctt tgccttcgtg ccccttgcct	1740
gcgtgcccc tgccttggag agaaaggggc cgggtgctggc gctcgcatc aataaagagt	1800
aatgtggcat ttttctataa taaacatgga ttacctgtgt ttaaaaaaa aagtgtgaat	1860
ggcgttaggg taagggcaca gccaggctgg agtcagtgtc tgcccctgag gtcctttaag	1920
ttgagggtg ggaatgaaac ctatagcctt tgtgctgttc tgccttgcct gtgagctatg	1980
tcactccctt cccactctg accatattcc agacacctgc cctaatectc agcctgtca	2040
cttcaattct gcattatata tccaaggcgt tggatatatg aaaaagatgt aggggcttgg	2100
aggtgttctg gacagtgggg agggctccag acccaacctg gtcacagaag agcctctccc	2160
ccatgcatac tcatccacct ccctccccta gagctattct cctttgggtt tcttgcctgt	2220
tcaattttat acaaccatta tttaaatatt attaaacaca tattgttctc taggcactgt	2280
ggtagtgggt tttttgttg tttttgttt tgagactgtc tcaaaaactc tgtcgcccag	2340
gctgacagtg cagtggcaca atcttggctc actgcagcct ctgcctcctg ggttcaagcg	2400
attctcgtgc atcagcctcc tgagtaactg gaattaatag gcacgtgcca ccatgtccat	2460
ctaattcata tatatatatt tttttttct gagacggagt ctcaactgtca cacaggctgg	2520
agtgcagtgg cagatctcg actcactgca agctccacct cctgggttca cgccattctc	2580
ctgcctcagc ctccccagta gctgggacta caggcgcccg ccaccagcc cggttaattt	2640
tttgatattt tagtagagat ggggtttctc cgtgttagcc aggatggtct cgatctcctg	2700
acctcgtgat ccgccccctt tggcctccca aagtgtggg attacaggcg tgagccaccg	2760
cgcccgccg aattcatcta ttttagtag agatggggtt tcaatatatt ggccaggctg	2820
gtcttgaact catgacctca gatgttcaact tgccttggcc tcccaaagt ctgggattag	2880
aggcgtgagc caccgcacgc gggcctgtgg taaattgttg aatttgaagg actcagaggc	2940
cctggtcaat tccaaaataa cgtaggcgac ttccatcccc ctctcccaa ccattttcag	3000
cccaaagcat ctctgcaggg aatggatggc tgcgcggagg tggcggtgg ctctggagag	3060

ggtcttttgcagggtgatatttctctagaaggaaatgtctcgtctggaccagactgccc3120
 cctcctgggttcagatgcagaagtgatactgtaagccagaggcgggggcaataatgcatc3180

 gcagccatttaggtgaggaattccttggcggttatttgttaagtctttggctgggccc3240
 tgggctggggtaacaatggaagggtccagcattttttcagaaagctccagtgtagt3300
 ggatacagaaacttcaggaaaggcagggtcagaaggatctgagtaaaactcggtccttca3360
 acaccatccttcagccctgggtcatgttccttcagggtcctgggtgggaggtagacaagc3420
 ctacgcttgtgctgttctgtaaggacagggtgggcattttctaccaacaagaattcttgg3480
 aattttcacaagccccagcagccaagtcagggtctatagccagatacaagaattaaag3540
 gtcccagcacgggcacccacacaggagccacctacctctattaccagaaagctttag3600

 gaggggtggtccgcagacaaaggacctgcaagggtgcgacctgtctcccctctggtcat3660
 aactttcatgttactattgctggggataatgttaagtaaaaatagcagacactgagtttt3720
 aagtctcaagtggatgaaggcagagatcgtgatacacttgagttaaagcagtaggggttct3780
 gtcattttctattcctgtgttaaacattttctttaatgttattatttttccactaaact3840
 aacgtggcctggtcacgactttcattggtaagtgtgctgttctcaccctccaccgttg3900
 ctcttttggctcactgatcataagagcatttacctgaaggctgtcagaccctgaatgccca3960
 acaggtcaactgcagtggtcgcagttaccacctagctgttccaatgaacagaatcgt4020

 gtgccccaaatcatctcccctcacctaggctgtaaatgaaagtcaccacctgagcgg4080
 aacacaggccatctgtgtgtctgtgcaccaagggggggtgggaagtctcagactgact4140
 tcttggtctcagtcacctaaggaaaagagtctccagtcgtccccaccctcacccttc4200
 ccattccaggagictattaaaggaggcaaaggaggcctaagggtatcaaaagaaaggagt4260
 gaatggagactgggagagcttcaacctctctctccttgtaggagctgaggctgcatg4320
 ccaggctaccttcccttcgaggaatctaataagctagggtcactggtgttttcagggtgctt4380
 ctcaaaggatgtccgtagggtaggatatacaggatgtgggagcacagggtccaccacagc4440

 actagtgatggagatcattgccctagacttctgggacagtgagactgtgaggaaagct4500
 gaaatgatactgggaaagggtgaaagaaaggatgtaggtggaattttattagttatattg4560
 taggtacacataccttatggcaacattcctagcactctaaattctagatttgtatagttt4620
 ctgtcaatatcttttgtaagcttaataacacaggcatgacaagtatgtgtcacatact4680
 tttttttccaagagaaaaaataaagtaggaattgggtgctttgtttatcaaaatttgt4740
 atttcctttaataataaactttgaaataaagggttgaaaatagta4785

<210> 229

<211> 4785

<212> DNA

<213> Homo sapiens

<400> 229

ctcacgtggc cagccccctc cgagagggga gaccagcggg ccatgacaag ctccaggctt 60
 tggttttcgc tgcctgctggc ggcagcgttc gcaggacggg cgacggccct ctggccctgg 120
 cctcagaact tccaaacctc cgaccagcgc tacgtccttt acccgaacaa ctttcaattc 180
 cagtacgatg tcagctcggc cgcgacggc ggctgctcag tcctcgacga ggccttcag 240
 cgctatcgtg acctgctttt cggttccggg tcttggtccc gtccttacct cacagggaaa 300
 cggcatacac tggagaagaa tgtgttggtt gtctctgtag tcacacctgg atgtaaccag 360
 cttcctactt tggagtcagt ggagaattat accctgacca taaatgatga ccagtgttta 420

ctcctctctg agactgtctg gggagctctc cgaggtctgg agacttttag ccagcttgtt 480
 tggaaatctg ctgagggcac attctttatc aacaagactg agattgagga ctttccccgc 540
 tttcctcacc ggggcttctt gttggataca tctgccatt acctgccact ctctagcacc 600
 ctggacactc tggatgtcat ggcgtacaat aaattgaacg tgttccactg gcatctggta 660
 gatgatcctt ctttccata tgagagcttc acttttccag agctcatgag aaaggggtcc 720
 tacaacctg tcacccacat ctacacagca caggatgtga aggaggtcat tgaatacgca 780
 cggctccggg gtatccgtgt gcttcagag tttgacactc ctggccacac tttgtcctgg 840

ggaccaggta tccctggatt actgactcct tgctactctg ggtctgagcc ctctggcacc 900
 tttggaccag tgaatccag tctcaataat acctatgagt tcatgagcac attcttctta 960
 gaagtacgt ctgtcttccc agatttttat cttcatcttg gaggagatga ggttgatttc 1020
 acctgctgga agtccaaccc agagatccag gactttatga ggaagaaagg cttcgtgag 1080
 gacttcaagc agctggagtc cttctacatc cagacgtgc tggacatcgt ctcttcttat 1140
 ggcaagggt atgtggtgtg gcaggaggtg tttgataata aagtaaagat tcagccagac 1200
 acaatcatc aggtgtggcg agaggatatt ccagtgaact atatgaagga gctggaactg 1260

gtcaccaagg cgggcttccg ggcccttctc tctgccccct ggtacctgaa ccgtatatcc 1320
 tatggccctg actggaagga tttctacata gtggaacccc tggcatttga aggtaccct 1380
 gagcagaagg ctctggtgat tgggtggagag gcttgtatgt ggggagaata tgtggacaac 1440
 acaaacctgg tccccaggct ctggcccaga gcaggggctg ttgccgaaag gctgtggagc 1500
 aacaagtga catctgacct gacatttgcc tatgaacgtt tgtcacactt ccgctgtgaa 1560
 ttgctgagge gaggtgtcca ggcccaaccc ctcaatgtag gcttctgtga gcaggagttt 1620

gaacagacct gagccccagg caccgaggag ggtgctggct gtaggtgaat ggtagtggag	1680
ccaggcttcc actgcatcct ggccagggga cggagccctc tgccttcgtg ccccttgcct	1740
gcgtgccccct gtgcttggag agaaaggggc cggtgctggc gctcgcatc aataaagagt	1800
aatgtggcat ttttctataa taaacatgga ttacctgtgt ttaaaaaaa aagtgtgaat	1860
ggcgttaggg taagggcaca gccaggctgg agtcagtgtc tgcccctgag gtcttttaag	1920
ttgagggctg ggaatgaaac ctatagcctt tgtgctgttc tgccttgcct gtgagctatg	1980
tcactccctc cccactcctg accatattcc agacacctgc cctaactctc agcctgctca	2040
cttcacttct gcattatata tccaaggcgt tggatatatg aaaaagatgt aggggcttgg	2100
agggtgtctg gacagtgggg agggctccag acccaacctg gtcacagaag agcctctccc	2160
ccatgcatac tcatccacct cctccccccta gagctattct cctttgggtt tcttgcctgt	2220
tcaattttat acaaccatta tttaaatatt attaaacaca tattgttctc taggcactgt	2280
ggtagtgggt tttttgttg tttttgttt tgagactgtc tcaaaaactc tgtcgcccag	2340
gctgacagtg cagtggcaca atcttggtc actgcagcct ctgcctcctg ggttcaagcg	2400
attctcgtgc atcagcctcc tgagtaactg gaattaatag gcacgtgcca ccatgtccat	2460
ctaattcata tatatatatt tttttttct gagacggagt ctactgtca cacaggctgg	2520
agtgcagtgg cagatctcg actcactgca agctccacct cctgggttca cgccattctc	2580
ctgcctcagc ctcccagta gctgggacta caggcgcccg ccaccacgcc cggctaattt	2640
tttgtatttt tagtagagat ggggtttctc cgtgttagcc aggatggtct cgatctcctg	2700
acctcgtgat ccgcccgcct tggcctccca aagtgtctggg attacaggcg tgagccaccg	2760
cgcccgcccg aattcatcta tttttagtag agatgggggt tcaactatatt ggccaggctg	2820
gtcttgaact catgacctca gatgttact tgtcttggcc tcccaaagt ctgggattag	2880
aggcgtgagc caccgcacgc gggcctgtgg taaattgttg aatttgaagg actcagaggc	2940
cctggtcaat tcaaaaataa cgtaggcgac ttccatcccc ctctcccaa ccattttcag	3000
cccaaagcat cttgcaggg aatggatggc tgcgcggagg tggcggttgg ctctggagag	3060
ggtcttttga ggigtgattt tctctagaag gaaatgtctc gtcgtggacc cagactgccc	3120
cctcctggtt tcagatgcag aagtgatact gtaagccaga ggcgggggca gtaatgcatc	3180
gcagccattt taggtgagga tttccttggc ggttatttgt taagtctttt ggctggggcc	3240
tgggctgggg taacaatgga caggttccag gcattttttt cagaaagctt ccagtgtagt	3300
ggatacagaa acttcaggaa ggcagggtct agaaggatct gagtaaaact cggtccttca	3360

acaccatcct tcagccctg ggtcatgttc cttcgaggtc ctggtgggag gtagacaagc 3420
ctagccttgt gctgttcctg taaggacagg gtgggcattt tctaccaaca gaattcttgg 3480
aattttcaca cagcccagcc tagccaagtc cagggctata gccagatac acaagttaag 3540
gtcccagcac tggcaccac cacaggagcc cccttacctc tattaccag aagctttag 3600
gaggggtggt ccgcagacaa ggacctgca caggtgagc cctgcttccc tctgtgcat 3660
aactttcatg ttactattgc ttgggataat gtttaagtaa aatagcagac actgagtttt 3720
aagtctcaag tggatgaagg cagagatcgt gatacacttg agttaaagca gtagggttct 3780

gtcattttct attcctgttg taaacatttt ctttaagtgt attattttta ccactaaact 3840
aacgtggcct ggtcacgact ttcatgtgta aagtgtgctg ttcctcacc tccaccgttg 3900
ctcctttggt ccactgatca taagagcatt tacctgaagg tcgtcagacc tcgaatgcca 3960
acaggtcaac tgcagtggcc tgcagttacc acccagctg ttccaatgaa cagaatcgt 4020
gttgcccaaa ctcatctccc ttacactagg ctgtaaattg aaagtccac ccctgagcgg 4080
aacacaggcc atcttgtgtg ctgtgcacca ccagggggtg gggaagtttc cagactgact 4140
tcttggtcc agtcacctta ggaaaagagt tctccagtcg ctccccaccc ccacccttc 4200

ccattccagg agtctattaa ggaggcaaag caggcctaac gggtatcaaa gcaaaggagt 4260
gaatggagac tgggagagtc ttcaacctct cctctccttg gtaggagctg aggctgcatg 4320
ccaggtaacct tccttcgag gaatctaata aagctaggtc actggtgttt tcaggtgctt 4380
ctcaaaggat tgccgtaggg gtaggatatc aggatgtggg agcacagggtg ccaccacagc 4440
actagtgtg gagagtcatt gccctagac ttctgggaca gtgagactgt gaggaaagct 4500
gaaatgatac tgggaaaggg tgaaagaaag gatgtaggtg gaatttattt agtattaatg 4560
taggtacaca taccttatgg caacattcct agcactctaa attctagatt tgtatagttt 4620

ctgtcaatat cttttgtaag cttaatcaat acagggcatg acaagtatgt gtcacatact 4680
ttttttcca cgaagaaaaa aaataagtag gaattgggtg ctttgtttat caaaatttgt 4740
atttccttta taaataaact ttgaaataaa ggttgaaaat tagta 4785

<210> 230
<211> 1812
<212> DNA
<213> Homo sapiens
<400> 230

agaccgggag gaaagcagcc gagcggccat ggagctgtgc gggctgggag tgccccggcc 60
gcccagctgt ctggcgctgc tgttggcgac actgctggcg gcgatgttgg cgctgctgac 120

tcaggtggcg ctggtggtgc aggtggcgga ggcggtcgg gccccgagcg tctcgccaa	180
gccggggccg gcgtgtggc cctgccgct ctggtgaag atgacccga acctgtgca	240
tctgccccg gagaacttct acatcagcca cagcccaat tccacggcg gcccctcctg	300
cacctgtctg gaggaagcgt ttcgacgata tcatggctat attttgggt tctacaagt	360
gcatcatgaa cctgctgaat tccaggctaa aaccaggtt cagcaactt ttgtctaat	420
cacccttcag tcagagtgtg atgcttccc caacatatct tcagatgagt cttatacttt	480
acttgtgaaa gaaccagtgg ctgtccttaa ggccaacaga gtttggggag cattacgagg	540
tttagagacc ttagccagt tagtttatca agattcttat ggaactttca ccatcaatga	600
atccaccatt atigattctc caaggttttc tcacagagga attttgattg atacatccag	660
acattatctg ccagttaaga ttattcttaa aactctggat gccatggctt ttaataagtt	720
taatgttctt cactggcaca tagttgatga ccagtcttc ccatatcaga gcatcacttt	780
tcctgagtta agcaataaag gaagctattc ttgtctcat gtttatacac caaatgatgt	840
ccgtatggtg attgaatatg ccagattacg aggaattcga gtcctgccag aatttgatac	900
ccctgggcat acactatctt ggggaaaagg tcagaaagac ctctgactc catgttacag	960
tagacaaaac aagtggact cttttggacc tataaacctt actctgaata caacatacag	1020
cttcttact acatttttca aagaaattag tgaggtgtt ccagatcaat tcattcattt	1080
gggaggagat gaagtggat ttaaatgttg ggaatcaaat ccaaaaattc aagatttcat	1140
gaggcaaaaa ggctttggca cagattttta gaaactagaa tctttctaca ttcaaaaggt	1200
tttgatatt attgcaacca taaacaagg atccattgtc tggcaggagg tttttgatga	1260
taaagcaaag ctigcgccgg gcacaatagt tgaagtatgg aaagacagcg catatcctga	1320
ggaactcagt agagtcacag catctggctt cctgtaatc ctttctgctc cttggtactt	1380
agatttgatt agctatggac aagattggag gaaatactat aaagtgaac ctcttgattt	1440
tggcgttact cagaaacaga aacaactttt cattggtgga gaagcttgtc tatggggaga	1500
atatgtggat gcaactaacc tcaatccaag attatggcct cgggcaagtg ctgttggtga	1560
gagactctgg agttccaaag atgtcagaga tatggatgac gcctatgaca gactgacaag	1620
gcaccgtgc aggatggtcg aacgtggaat agctgcacaa cctctttatg ctggatattg	1680
taacatgag aacatgtaaa aaatggaggg gaaaaaggcc acagcaatct gtactacaat	1740
caactttatt ttgaaatcat gtaaaataag atattagact gttttttgaa taaaatattt	1800
ttattgattg aa	1812

<210> 231

<211> 1318

<212> DNA

<213> Homo sapiens

<400> 231

atgctgtccg gtcaccgtga caatgcagct gaggaacca gaactacatc tgggctgcgc	60
gcttgcgctt cgcttcctgg ccctcgtttc ctgggacatc cctggggcta gagcactgga	120
caatggattg gcaaggacgc ctaccatggg ctggctgcac tgggagcgct tcatgtgcaa	180
ccttgactgc caggaagagc cagattcctg catcagttag aagctcttca tggagatggc	240
agagctcatg gtctcagaag gctggaagga tgcaggttat gactacctct gcattgatga	300
ctgttggatg gctcccaaaa gagattcaga aggcagactt caggcagacc ctacgcgctt	360
tcctcatggg attcgccagc tagctaatta tgttcacagc aaaggactga agctagggat	420
ttatgcagat gttggaaata aaacctgcgc aggccttcctt gggagttttg gatactacga	480
cattgatgcc cagacctttg ctgactgggg agtagatctg ctaaaatttg atggttgta	540
ctgtgacagt ttggaaaatt tggcagatgg ttataagcac atgtccttgg ccctgaatag	600
gactggcaga agcatttgtt actcctgtga gtggcctctt tataatgtggc cctttcaaaa	660
gccccattat acagaaatcc gacagtactg caatcactgg cgaaattttg ctgacattga	720
tgattcctgg aaaagtataa agagtatctt ggactggaca tcttttaacc aggagagaat	780
tgttgatgtt gctggaccag ggggttggaa tgaccagat atgttagtga ttggcaactt	840
tggcctcagc tggaatcagc aagtaactca gatggccctc tgggctatca tggctgctcc	900
tttattcatg tctaagtacc tccgacacat cagccctcaa gccaaagctc tccttcagga	960
taaggacgta attgccatca atcaggaccc ctggggcaag caagggtacc agcttagaca	1020
gggagacaac ttigaagtgt gggaacgacc tctctcaggc ttagcctggg ctgtagctat	1080
gataaaccgg caggagattg gtggacctcg ctcttatacc atcgagtgtt cttccctggg	1140
taaaggagtg gctgtaatc ctgcttctt catcacacag ctctccctg tgaaggaa	1200
gctagggttc tatgaatgga cttcaaggtt aagaagtcac ataaatcca caggcactgt	1260
tttgcttcag ctagaaaata caatgcagat gtcattaaaa gacttacttt aaaatgtt	1318

<210> 232

<211> 2291

<212> DNA

<213> Homo sapiens

<400> 232

ggaattactt gcagggctaa cctagtcct atagctaagg caggtacctg catccttgtt 60
 tttgtttagt ggatcctcta tccttcagag actctggaac cctgttggtc ttctcttcat 120
 ctaatgacct tgaggggatg gagttttcaa gtccttccag agaggaatgt cccaagcctt 180
 tgagtagggt aagcatcatg gctggcagcc tcacaggatt gcttctactt caggcagtgt 240

 cgtgggcac cagggtccc cctgcatcc ctaaaagctt cggctacagc tcggtggtgt 300
 gtgtctgcaa tgccacatac tgtgactcct ttgaccccc gacctttcct gcccttggtg 360
 ccttcagccg ctatgagagt acacgcagtg ggcgcaggat ggagctgagt atggggccca 420
 tccagcttaa tcacacgggc acaggcctgc tactgacct gcagccagaa cagaagtcc 480
 agaaagtga gggatttga ggggccatga cagatgctgc tgctctcaac atccttgccc 540
 tgtaccccc tgcccaaat ttgctactta aatcgtactt ctctgaagaa ggaatcggat 600
 ataacatcat ccgggtacct atggccagct gtgacttctc catccgcacc tacacctatg 660

 cagacacccc tgatgatttc cagttgcaca acttcagcct cccagaggaa gataccaagc 720
 tcaagatacc cctgattcac cgagccctgc agttggcca gcgtcccgtt tcactccttg 780
 ccagcccctg gacatcaccc acttggctca agaccaatgg agcggatgaat ggggaagggt 840
 cactcaaggg acagcccga gacatctacc accagacctg ggccagatac tttgtgaagt 900
 tcctggatgc ctatgctgag cacaagttac agttctgggc agtgacagct gaaaatgagc 960
 ctctgctgg gctgttagt ggataccct tccagtcct gggttcacc cctgaacatc 1020
 agcgagactt cattgccctg gacctaggtc ctaccctgc caacagtact caccacaatg 1080

 tccgcctact catgctggat gaccaacgt tctgctgcc ccaactggga aaggtggtac 1140
 tgacagacc agaagcagct aaatatgttc atggcattgc tgtacattgg tacctggact 1200
 ttctggtcc agccaaagcc accctagggg agacacaccg cctgttccc aacaccatgc 1260
 tctttgcctc agaggcctgt gtgggctcca agttctggga gcagagtgtg cggctaggt 1320
 cctgggatcg agggatgcag tacagccaca gcatcatcac gaacctcctg taccatgtgg 1380
 tcggctggac cgactggaac cttgccctga accccgaagg aggaccaat tgggtgcgta 1440
 actttgtcga cagtcccatc attgtagaca tcaccaagga cacgttttac aaacagccca 1500

 tgttctacca ccttggccac ttcagcaagt tcattcctga gggctcccag agagtggggc 1560
 tggttgccag tcagaagaac gacctggacg cagtggcact gatgcatccc gatggctctg 1620
 ctgttgtgtt cgtgctaaac cgctcctcta aggatgtgcc tcttaccatc aaggatcctg 1680
 ctgtgggctt cctggagaca atctcacctg gctactccat tcacacctac ctgtggcgtc 1740
 gccagtgatg gagcagatac tcaaggaggc actgggctca gcctgggcat taaagggaca 1800
 gagtgcgtc acacgtgtc tgtactaaa gagggcacag cagggccagt gtgagcttac 1860

agcgacgtaa gccacggggc aatggtttgg gtgactcact ttccctcta ggtggtgcca 1920

ggggctggag gccctagaa aaagatcagt aagccccagt gtccccccag ccccatgct 1980

tatgtgaaca tgcgctgtgt gctgcttgct ttggaaactg ggcttggtc caggcctagg 2040

gtgagctcac tgtccgtaca aacacaagat cagggtgag ggtaaggaaa agaagagact 2100

aggaaagctg ggcccaaac tggagactgt ttgtctttcc tggagatgca gaactgggcc 2160

cgtggagcag cagtgtcagc atcaggcgagg aagccttaaa gcagcagcgg gtgtgcccag 2220

gcaccagat gattcctatg gcaccagcca ggaaaaatgg cagctcttaa aggagaaaat 2280

gtttgagccc a 2291

<210> 233

<211> 4294

<212> DNA

<213> Homo sapiens

<400> 233

gaggcggaag cgcccgagc cgggtaccgg ctctctctgg gctccctcta gcgccttccc 60

cccggcccga ctccgctggt cagcgccaag tgacttacgc ccccgaccct gagcccgagc 120

cgctaggcga ggaggatcag atctccgctc gagaatctga aggtgccctg gtcctggagg 180

agttccgtcc cagcccgagg tctcccggtc ctgtcgggccc cgggccctct ggagcttcag 240

gaggcgcccg tcagggtcgg ggagtatttg ggtccggggt ctgagggaag ggcggcgcct 300

gggtctgcgg tatcggaag agcctgctgg agccaagtag cctccctct ctgggacag 360

acccctcggc cccatgtcca tgggggcacc gcggtccctc ctcttgcccc tggctgctgg 420

cctggccgtt gccgctccg ccaacatcgt gctgatcttt gccgacgacc tcggtatagg 480

ggacctgggc tgetatgggc acccagctc taccactccc aacctggacc agctggcggc 540

gggagggtg cggttcacag acttctacgt gctgtgtct ctgtgcacac cctctagggc 600

cgccctctg accggccggc tcccgttcg gatgggcatg taccctggcg tctgtgtgcc 660

cagctcccgg gggggcctgc ccctggagga ggtgaccgtg gccgaagtcc tggctgcccc 720

aggctacctc acaggaatgg ccggcaagtg gcaccttggg gtggggcctg agggggcctt 780

cctgcccccc catcagggtc tccatcgatt tctaggcacc ccgtactccc acgaccaggg 840

cccctgccag aacctgacct gcttcccgc gccactcct tgcgacggtg gctgtgacca 900

gggcctggtc cccatccac tgttgccaa cctgtccgtg gaggcgcagc cccctggct 960

gcccggacta gagggccgtc acatggcttt cgcccatgac ctcatggcgg acgcccagcg 1020

ccaggatcgc ccttcttcc tgtactatgc ctctcaccac accactacc ctgattcag 1080

tgggcagagc ttgcagagc gttcaggccg cgggccatit ggggactccc tgatggagct	1140
ggatgcagct gtggggaccc tgatgacagc cataggggac ctggggctgc ttgaagagac	1200
gctggtcac ttcactgcag acaatggacc tgagaccatg cgtatgtccc gaggcggctg	1260
ctccggtctc ttgcggtgtg gaaaggggaa gacctacgag ggcggtgtcc gagagcctgc	1320
cttggccttc tggccaggtc atatcgctcc cggcgtgacc cacgagctgg ccagctccct	1380
ggacctgctg cctaccctgg cagccctggc tggggcccca ctgccaatg tcaccttgga	1440
tggctttgac ctgagccccc tgctgctggg cacaggcaag agccctcggc agtctctctt	1500
cttctacccg tcttaccag acgaggtccg tggggttttt gctgtgcgga ctggaaagta	1560
caaggctcac ttcttacc cagggtctgc ccacagtga accactgcag accctgcctg	1620
ccacgcctcc agctctctga ctgctcatga gcccccgctg ctctatgacc tgtccaagga	1680
ccctggtgag aactacaacc tgctggggggg tgtggccggg gccacccag aggtgctgca	1740
agccctgaaa cagcttcagc tgctcaaggc ccagttagac gcagctgtga ccttcggccc	1800
cagccaggtg gcccggggcg aggacccgc cctgcagatc tgctgtcac ctggctgcac	1860
ccccgcccc gcttctgccc attgcccaga tccccatgcc tgagggcccc tgggtggcc	1920
tgggcatgtg atggctctc actgggagcc tgtgggggag gctcaggtgt ctggaggggg	1980
tttgtcctg ataacgtaat aacaccagtg gagacttgca gatgtgacaa ttcgtccaat	2040
cctggggtaa tgctgtgtgc tgggtgccgt ccctgtggt acgaatgagg aaactgaggt	2100
gcagagaggt tcaggacttg tacaagatca ccagccaga aagaggttgg gctgggattt	2160
gaacctggt gtctgtgctc tggaaagtgc cctggcgctt tggatgctg cgtgggtcag	2220
tgcacacagg cacacgtcag cctcaaggac atgggcacat ctgttcacag gagcagcgcc	2280
acgtgccttt gagtgccagg aacggggtgg gaggtggga ggggtgtagg gccagaagac	2340
tcagaagatg caaagtgcct gagagagacg ggatattccc ccagaagaag cattcttaga	2400
gacacaggca ctggacctcc ttggttctta taagaaacct gtctgaagct gggatgatgag	2460
ttgcacactc caggtggggc taaggggcct ggagccccctg ctggctccta ggaaggcaca	2520
gcagcaggcc ctgagacggc tctctggggg cccctccacc ctcccaggcc tctgcatttc	2580
acctgtgccc acacttctgt ctctgcctt caccttttga cccactacta acgattctcc	2640
accagcaga caaagtgate tcttaaaaat atctgttggc tgggcacggt ggctcacgcc	2700
tgtaatccca gcactttagg aagccgaggc ggggtgatca cctgaggtcg ggagttcgag	2760
accagcctga ccaacatgga gaaaccccat ctctactaaa aatacaaaat tagccaggtg	2820

tagtggtgca tcctgtaat ccagctact tgggagtctg aggctggaga atcacttgaa 2880

cctgggaggc ggtggttgca gtgagccgag atgcacccat tgcactccag cctgggcaac 2940

aagagaaaaa ctctgtctca aaaaacaaaa aatctgttag gctgcacacg gcgattcact 3000

cctgtattcc cagtgcittg ggaggctgag gtgagaggat gccagaggcc aggaattcag 3060

accagcctgg gcaacatagt gagaccccag ctctaaagat ttgtttttgt tttttttttt 3120

tttttttttt tttttttttt tttttttttg agacggagtc tcgctctgtc gccaggcta 3180

gagtgcagtg gtaccatctc cgctcactgc aacctccgcc tcccgggttc cagggtattct 3240

cctgcctcag cctccctagt agctggaact acaggtgtgt gctgccatgc ccagctaatt 3300

tttttttatt taatagagac aagatttcac catgttggcc aggctggtct caaacctctg 3360

acctcaggtg atccaccgc ctcagcctcc caaagtctg ggattacagg tgtgaaccac 3420

cacacctggc caacaatatt tgttttaatt agccaggcgt ggtagcattt gtcctagcaa 3480

tttgggaggt tgaggtggga gaatcacttc agcccactag gtcgaggctg tagtgagcta 3540

taattgtacc actgcactcc agcctcgggg acagagttag accctgtctg caaataaaca 3600

aataaaaacat caggctgggc ttgagcatct attcctgtctc aaaatttcgc aggccttctca 3660

gaagaaaatc caaacccctt acagtgacct agtttgccct tgaggcctcc acccacacc 3720

ccttcccccc agtcttaggg ggtggcctgg ctgttccctt caacggcaac gctctgcctc 3780

cattgttggc ctctctgca gggagggact gtctgagcac ctgcccgtgt ctgtgcagca 3840

tggcacactg agtcaggcc cactgcatg cccagggtgc cagtacacg ccagggtctc 3900

cctcagtgtt ggccaagtga gaggagcaca ccttccgggc gttcagacac ctccccgtgg 3960

cagacaccgt tcgttgctac caaacagcca cctccttctt aatgggtctc cattttcag 4020

tgctgggcaa aggtcccttg atcttggagt tgcagcctct ttctctcaa ggaggcggt 4080

gaccagcctg agccagtcaa tccagtgatt ggttcaggag tagcctgtga ccaggagtec 4140

tggtagttaa cgactggggc agccctgggg gtgaggacct tgcgcagccg tcacaggccc 4200

tgattggaca ctgggcagct gctaaccag tgtctccagc tgcctacctg gagagctcca 4260

agcgtaagaa aataaacct gcctgttgaa gcca 4294

<210> 234

<211> 3794

<212> DNA

<213> Homo sapiens

<400> 234

agtcatgtga cccacacaat ggctgagtggt ctactctcgg cttcctggca acgccgagcg 60

aaagctatga ctgcggccgc gggttcggcg ggccgcgcg cggtgccctt gctgctgtgt 120
 gcgctgctgg cggccggcgg cgcgtacgtg ctgcacgact ccgacgggct gggccgggag 180

 ttcgacggca tcggcgcggt cagcggcggc ggggcaacct cccgacttct agtaaattac 240
 ccagagccct atcgttctca gatatggat tatctcttta agccgaattt tggcgcctct 300
 ttgcatattt taaaagtga aataggtggt gatgggcaga caacagacgg cactgagccc 360
 tcccacatgc attatgact agatgagaat tatttccgag gatacgagtg gtggttgatg 420
 aaagaagcta agaagaggaa tccaatatt acactcattg ggttgccatg gtcattccct 480
 ggatggctgg gaaaaggttt cgactggcct tatgtcaatc ttcagctgac tgcctattat 540
 gtcgtgacct ggattgtggg cgccaagcgt taccatgatt tggacattga ttatatggga 600

 atttggatg agaggtcata taatgccaat tatattaaga tattaagaaa aatgctgaat 660
 tatcaaggtc tccagcgagt gaaaatcata gcaagtata atctctggga gtccatctct 720
 gcatccatgc tccttgatgc cgaactcttc aaggtggttg atgttatagg ggctcattat 780
 cctggaaccc attcagcaaa agatgcaaag ttgactggga agaagctttg gtcttctgaa 840
 gactttagca ctttaaatag tgacatgggt gcaggctgct ggggtcgcat tttaaatcag 900
 aattatatca atggctatat gacttcaca atcgcatgga atttagtggc tagttactat 960
 gaacagttgc cttatgggag atgcgggttg atgacggccc aggagccatg gagtgggcac 1020

 tacgtggtag aatctctgt ctgggtatca gtcatacca ctgagtttac tcaacctggc 1080
 tggattacc tgaagacagt tggccattta gagaaaggag gaagctacgt agctctgact 1140
 gatggcttag ggaacctcac catcatcatt gaaacatga gtcataaaca ttctaagtgc 1200
 atacggccat ttcttctta ttcaatgtg tcacaacaat ttgccacctt tgttcttaag 1260
 ggatctttta gtgaaatacc agagctacag gtatggtata ccaaacttgg aaaaacatcc 1320
 gaaagatttc ttttaagca gctggattct ctatggtcc ttgacagcga tggcagtttc 1380
 aactgagcc tgcatgaaga tgagctgttc aactcacca ctctcaccac tggtcgcaaa 1440

 ggcagctacc cgcttctcc aaaatcccag cccttcccaa gtacctataa ggatgatttc 1500
 aatgttgatt acccatTTTT tagtgaagct ccaaactttg ctgatcaaac tgggtgattt 1560
 gaatatTTTA caaatattga agacctggc gagcatcact tcacgtacg ccaagtctc 1620
 aaccagagac ccattacatg ggctgccgat gcatccaaca caatcagtat tataggagac 1680
 tacaactgga ccaatctgac tataaagtgt gatgtataca tagagacccc tgacacagga 1740
 ggtgtgttca ttgcaggaag agtaataaaa ggtgggtattt tgattagaag tgccagagga 1800

atctctctct ggatttttgc aaatggatct tacagggtta caggtgattt agctggatgg	1860
attatataatg ctttaggacg tgttgaagt acagcaaaaa aatggtatac actcacgtta	1920
actattaagg gtcatttcac ctctggcatg ctgaatgaca agtctctgtg gacagacatc	1980
cctgtgaatt ttccaaagaa tggctgggct gcaattggaa ctactcctt tgaatttgca	2040
cagtttgaca actttcttgt ggaagccaca cgctaatact taacagggca tcatagaata	2100
ctctggattt tcttcccttc tttttggtt tggttcagag ccaattcttg tttcattgga	2160
acagtatatg aggcttttga gactaaaaat aatgaagagt aaaaggggag agaaatttat	2220
ttttaattta cctgtggaa gattttatta gaattaattc caaggggaaa actggtgaat	2280
ctttaacatt acctgggtgtg ttccctaaca ttcaaactgt gcattggcca tacccttagg	2340
agtggtttga gtagtacaga cctcgaagcc ttgctgctaa cactgaggta gctctcttca	2400
tcttatttgc aagcggctct gtagatggca gtaacttgat catcactgag atgtatttat	2460
gcatgctgac cgtgtgtcca agtgagccag tgccttcac acaagatgat gctgccataa	2520
tagaaagctg aagaacacta gaagtagctt ttgaaaacc acttcaacct gttatgcttt	2580
atgctctaaa aagtattttt ttattttcct ttttaagatg atacttttga aatgcaggat	2640
atgatgagt ggtatgattt aaaaatgcct ctttaataaa ctacctctaa cactatttct	2700
gcagtaatag atattagcag attaatggg ttatttgcac tatttaattt ttttgattcc	2760
aagtttttgt ctgttaacca ctataactct ctgtgaacgt tttccaggt ggctggaaga	2820
aggaagaaaa cctgatatag ccaatgctgt tgtagtcgtt tctcagcct catctcactg	2880
tgtgtggtc tgtcctcaca tgtgactgg taacagactc acacagctga tgaatgcttt	2940
tctctcctta tgtgtggaag gaggggagca cttagacatt tgctaactcc cagaattgga	3000
tcctctccta agatgtactt actttttaaa gtccaaatat gtttatattt aaatatacgt	3060
gagcgtgttc atcatgttgt atgatttata ctaagcatta atgtggctct atgtagcaaa	3120
tcagttattc atgtaggtaa agtaaatcta gaattattta taagaattac tcattgaact	3180
aattctacta tttaggaatt tgtaagagtc taacataggc ttagctacag tgaagttttg	3240
cattgctttt gaagacaaga agataagtgc tagaataaat aagattacag agaaaatttt	3300
ttgttaaaac caagtattt ccagctgatg tatctaatat tttttaaaac gaacattata	3360
gagggtgaat ttatttacia taaaatgttc ctactttaaa tatacaattc agtgagtttt	3420
gataaattga tataccatg taaccaacac tccagtcaag cttcagaata tttccatcac	3480
cccagaaggt tctcttgtat acctgctcag tcagttcctt tcactcccaa ttgttggcag	3540

ccattgatag gaattctatc actatagggt agttttcttt gttccagaac atcatgaaag 3600
 cggcgctatg tactgtgtat tcttatgaat ggtttctttc catcagcata atgatttgag 3660
 atttgtccat gttgtgtgat tcagtggttt gttccttctt atttctgaag agttttccat 3720
 tgtatgaata taccacaatt tgtttcctcc ccaccagttt ctgatactac aattaaaact 3780
 gtctacattt acaa 3794
 <210> 235
 <211> 2410
 <212> DNA
 <213> Homo sapiens
 <400> 235
 agtcagccga ctacagagaa gggtaatcgg gtgtccccgg cgccgcccg ggccctgagg 60

 gctggctagg gtccaggccg ggggggacgg gacagacgaa ccagccccgt gtaggaagcg 120
 cgacaatgcc ccgtacgga gcgtcactcc gccagagctg cccaggtcc ggccgggagc 180
 agggacaaga cgggaccgcc ggagcccccg gactcctttg gatgggctg gtgctggcgc 240
 tggcgctggc gctggcgctg gcgctggctc tgtctgactc tcgggttctc tgggctccgg 300
 cagaggtca cctcttttct cccaaggcc atcctgccag gttacatcgc atagtcccc 360
 ggctccgaga tgtcttttgg tgggggaacc tcacctgccc aatctgcaa ggtctattca 420
 ccgccatcaa cctcgggctg aagaaggaac ccaatgtggc tcgctgggc tccgtggcca 480

 tcaagctgtg caatctgtg aagatagcac cacctgccgt gtgccaatcc attgtccacc 540
 tctttgagga tgacatggtg gaggtgtgga gacgtcagt gctgagcca tctgaggcct 600
 gtggcctgct cctgggctcc acctgtgggc actgggacat tttctcatct tggaacatct 660
 ctttgectac tgtgccgaag ccgccccca aacccctag ccccccagcc ccaggtgccc 720
 ctgtcagccg catcctcttc ctactgacc tgcactggga tcatgactac ctggagggca 780
 cggacctga ctgtgcagac ccaatgtgct gccgccgggg ttctggcctg ccgccgcat 840
 cccggccagg tgccggatc tggggcgaat acagcaagtg tgacctgccc ctgaggaccc 900

 tggagagcct gttgagtggg ctgggcccag ccggcccttt tgatatggtg tactggacag 960
 gagacatccc cgcacatgat gtctggcacc agactcgtca ggaccaactg cgggccctga 1020
 ccaccgtcac agcacttgtg aggaagtcc tggggccagt gccagtgtac cctgctgtgg 1080
 glaaccatga aagcacacct gtcaatagct tcctcccc cttcattgag ggcaaccact 1140
 cctcccgtg gctctatgaa gcgatggcca agccttggga gccctggctg cctgccgaag 1200
 ccctgcgcac cctcagaatt ggggggttct atgctcttcc cccatacccc ggtctccgcc 1260

tcattctctt caatatgaat ttttgttccc gtgagaactt ctggctcttg atcaactcca 1320

cggatccgc aggcagctc cagtggctgg tgggggagct tcaggctgct gaggatcgag 1380

gagacaaagt gcatataatt ggccacattc ccccagggca ctgtctgaag agctggagct 1440

ggaattatta ccgaattgta gccaggtaig agaacaccct ggctgctcag ttctttggcc 1500

acactcatgt ggatgaattt gaggtcttct atgatgaaga gactctgagc cgcccgtgg 1560

ctgtagcctt cctggcaccc agtgcaacta cctacatcgg ccttaatcct ggttaccgtg 1620

tgtacaaat agatggaac tactccggga gctctcacgt ggtcctggac catgagacct 1680

acatcctgaa tctgaccag gcaaacatac cgggagccat accgcactgg cagcttctct 1740

acagggtcgc agaaacctat gggctgcccc acacactgcc taccgcctgg cacaacctgg 1800

tatatcgcat gcggggcgac atgcaacttt tccagacctt ctggtttctc taccataagg 1860

gccaccacc ctcggagccc tgtggcacgc cctgccgtct ggctactctt tgtgccagc 1920

tctctgcccg tgctgacagc cctgctctgt gccgccacct gatgccagat gggagcctcc 1980

cagaggccca gagcctgtgg ccaaggccac tgttttgcta gggccccagg gccacattt 2040

gggaaagtgc ttgatgtagg aaagggtgaa aaagcccaaa tgctgctgtg gttcaaccag 2100

gcaagatcat ccgtgaaa aaccagtccc tgggccccaa ggatgccggg gaaacaggac 2160

cttctccttt cctggagctg gtttagctgg atatgggagg gggtttgct gcctgtgccc 2220

aggagctaga ctgccttgag gctgctgtcc ttccacagcc atggagtaga ggcctaagtt 2280

gacactgccc tgggcagaca agacaggagc tctcgcccca ggctgtgct gccagccag 2340

gaacctgta ctgctgtgc gacctgatgc tgccagtctg ttaaaataaa gataagagac 2400

ttggactcca 2410

<210> 236

<211> 4760

<212> DNA

<213> Homo sapiens

<400> 236

cttctgacc ggcgcgcgca gcctgctgcc gcggtcagcg cctgctcctg ctctccgct 60

cctctgccc ggggtgctga aacagcccgg ggaagtagag ccgcctccgg ggagcccaac 120

cagccgaacg ccgccggcgt cagcagcctt gcgcggccac agcatgaccg ctgcggcct 180

ggccttggc ctctctgc tgctactgtg tccagcgcag gtgttttcac agtcctgtgt 240

ttggtatgga gagtgtgga ttgcatatgg ggacaagagg tacaattgcg aatattctgg 300

cccacaaaa ccattgccaa aggatggata tgacttagtg caggaaactct gtccaggatt 360

cttcttttggc aatgtcagtc tctgtttgtga tgttcggcag cttcagacac taaaagacaa	420
cctgcagctg cctctacagt ttctgtccag atgtccatcc tgtttttata acctactgaa	480
cctgttttgt gagctgacat gtagccctcg acagagtcag tttttgaatg ttacagctac	540
tgaagattat gtigatcctg ttacaaacca gacgaaaaca aatgtgaaag agttacaata	600
ctacgtcgga cagagttttg ccaatgcaat gtacaatgcc tgccgggatg tggaggcccc	660
ctcaagtaat gacaaggccc tgggactcct gtgtgggaag gacgtgacg cctgtaatgc	720
caccaactgg attgaatata tgttcaataa ggacaatgga caggcacctt ttaccatcac	780
tcctgtgttt tcagattttc cagtccatgg gatggagccc atgaacaatg ccaccaaagg	840
ctgtgacgag tcgtgggatg aggtcacagc accatgtagc tgccaagact gctctattgt	900
ctgtggcccc aagccccagc cccacactcc tctgtctccc tggacgatcc ttggcttgga	960
cgccatgtat gtcatcatgt ggatcaccta catggcgttt ttgcttgtgt tttttggagc	1020
atTTTTtGca gtTgTgTgct acagaaaacg gtattttgtc tccgagtaca ctcccatcga	1080
tagcaatata gctttttctg ttaatgcaag tgacaaagga gaggcgtcct gctgtgaccc	1140
tgtcagcgca gcatTTgagg gctgcttgag gcggctgttc acacgctggg ggtctttctg	1200
cgtccgaaac cctggctgtg tcatTTtctt ctcgctggtc ttattactg cgtgttcgtc	1260
aggcctgggtg tttgtccggg tcacaaccaa tccagttgac ctctggtcag cccccagcag	1320
ccaggctcgc ctggaaaaag agtactttga ccagcacttt gggcctttct tccggacgga	1380
gcagctcacc atccggggccc ctctcactga caaacacatt taccagccat acccttcggg	1440
agctgatgta ccttttggac ctccgcttga catacagata ctgcaccagg ttcttgactt	1500
acaaatagcc atcgaaaaca ttactgcctc ttatgacaat gagactgtga cacttcaaga	1560
catctgcttg gccctcttt caccgtataa caggaactgc accattttga gtgtgttaaa	1620
ttacttccag aacagccatt ccgtgctgga ccacaagaaa ggggacgact tctttgtgta	1680
tgccgattac cacacgcact ttctgtactg cgtacgggct cctgcctctc tgaatgatac	1740
aagtttgctc catgaccctt gtctgggtac gtttgggtgga ccagtgttcc cgtggcttgt	1800
gttgggaggc tatgatgac aaaactacaa taacgccact gcccttgtga ttacctccc	1860
tgtcaataat tactataatg atacagagaa gctccagagg gcccaggcct gggaaaaaga	1920
gtttattaat ttttgtaaaa actacaagaa tcccaatctg accatttcct tactgtctga	1980
acgaagtatt gaagatgaac taaatcgtga aagtgcagct gatgtcttca ccgttgtaat	2040
tagctatgcc atcatgtttc tataatattc cctagccttg gggcacatga aaagctgtcg	2100

caggcttctg gtggattcga aggtctcact aggcacgcg ggcaccttga tcgtgctgag 2160

ctcgggtggct tgctccttgg gtgtcttcag ctacattggg ttgcccttga ccctcattgt 2220

gattgaagtc atcccgttcc tgggtctggc tgttggagtg gacaacatct tcattctggt 2280

gcaggcctac cagagagatg aacgtcttca aggggaaacc ctggatcagc agctgggcag 2340

ggtcctagga gaagtggctc ccagtatgtt cctgtcatcc ttttctgaga ctgtagcatt 2400

tttcttagga gcattgtccg tgatgccagc cgtgcacacc ttctctctct ttgcgggatt 2460

ggcagtcctt attgacttcc ttctgcagat tacctgttcc gtgagtctct tggggttaga 2520

cattaaacgt caagagaaaa atcggctaga catcttttgc tgtgtcagag gtgctgaaga 2580

tggaacaagc gtccaggcct cagagagctg tttgtttcgc ttcttcaaaa actcciatte 2640

tccacttctg ctaaaggact ggatgagacc aattgtgata gcaatatttg tgggtgttct 2700

gtcattcagc atcgcagtc tgaacaaagt agatattgga ttggatcagt ctctttcgat 2760

gccagatgac tctacatgg tggattatit caaatccatc agtcagtacc tgcacgcggg 2820

tccgcctgtg tactttgtcc tggaggaagg gcacgactac acttcttcca aggggcagaa 2880

catggtgtgc ggcggcatgg gctgcaacaa tgattccctg gtgcagcaga tatitaaacgc 2940

ggcgacagctg gacaactata cccgaatagg ctccgcccc tcgtcctgga tcgacgatta 3000

tttcgactgg gtgaagccac agtcgtcttg ctgtcagtg gacaatatca ctgaccagtt 3060

ctgcaatgct tcagtgggtg accctgcctg cgttcgtgc aggcctctga ctccggaagg 3120

caaacagagg cctcaggggg gagacttcat gagattcctg cccatgttcc tttcggataa 3180

ccctaacccc aagtgtggca aagggggaca tctgcctat agttctgcag ttaacatcct 3240

ccttggccat ggcaccaggg tcggagccac gtacttcatg acctaccaca ccgtgctgca 3300

gacctctgct gactttattg acgctctgaa gaaagcccga cttatagcca gtaatgtcac 3360

cgaacacatg ggcattaacg gcagtgccta ccgagtatit cttacagtg tgttttatgt 3420

cttctacgaa caglacctga ccatcattga cgacactatc ttcaacctcg gtgtgtccct 3480

gggcgcgata tttctgggtg ccatggtcct cctgggctgt gagctctggt ctgcagtcat 3540

catgtgtgcc accatcgcca tggctttggt caacatgttt ggagtatatgt ggctctgggg 3600

catcagtctg aacgtgtgat ccttgggtcaa cctgggtgat agctgtggca tctccgtgga 3660

gttctgcagc cacataacca gagcgttcac ggtgagcatg aaaggcagcc gcgtggagcg 3720

cgcggaagag gcacttggcc acatgggcag ctccgtgttc agtggaaatca cacttacaaa 3780

atttggaggg attgtgggtg tggcttttgc caaatctcaa attttccaga tattctactt 3840

caggatgtat ttggccatgg tcttactggg agccactcac ggattaatat ttctccctgt 3900
 cttactcagt tacatagggc catcagtaaa taaagccaaa agttgtgcca ctgaagagcg 3960
 atacaaagga acagagcgcg aacggcttct aaatttctag ccctctcgca gggcatcctg 4020
 actgaactgt gtctaagggt cggtcgggtt accactggac ggggtgctgca tcggcaaggc 4080
 caagttgaac accggatggt gccaaaccatc ggttggttgg cagcagcttt gaacgtagcg 4140
 cctgtgaact caggaatgca cagttgactt gggaagcagt attactagat ctggaggcaa 4200
 ccacaggaca ctaaacttct ccagcctct tcaggaaaga aacctcattc ttggcaagc 4260

aggaggtgac actagatggc tgtgaatgtg atccgctcac tgacactctg taaaggccaa 4320
 tcaatgcact gtctgtctct ccttttagga gtaagccatc ccacaagttc tataccatat 4380
 ttttagtgac agttgaggtt gtagatacac ttataacat ttatagttt aaagagcttt 4440
 attaatgcaa taaattaact ttgtacacat tttatataa aaaaacagca agtgatttca 4500
 gaatgttgta ggccctatta gagcttggtc tccaaaaatc tgtttgaaaa aagcaacatg 4560
 ttcttcacag tgttcccta gaaaggaaga gatttaattg ccagttagat gtggcatgaa 4620
 atgagggaca aagaaagcat ctctgaggtg tgtctactgg gttttaactt atttttcttt 4680

aataaaatac attgttttcc taagttttgg ggttaccta tctgctttga gagacaaata 4740
 caaaagctaa atggaagaga 4760

<210> 237

<211> 821

<212> DNA

<213> Homo sapiens

<400> 237

ctttcccgag cttggaactt cgttatccgc gatgcgttcc ctggcagcta cattcctgct 60
 cctggcgctc agcaccgctg ccagggccga accggtgcag ttcaaggact gcggttctgt 120
 ggatggagtt ataaaggaag tgaatgtgag cccatgcccc acccaaccct gccagctgag 180
 caaaggacag tcttacagcg tcaatgtcac cttcaccagc aatattcagt ctaaaagcag 240

caaggccgtg gtgcatggca tcctgatggg cgtcccagtt cccittccca ttctgagcc 300
 tgatggttgt aagagtggaa ttaactgccc tatccaaaaa gacaagacct atagctacct 360
 gaataaacta ccagtgaata gcgaatatcc ctctataaaa ctggtggtgg agtggcaact 420
 tcaggatgac aaaaaccaa gtctcttctg ctgggaaatc ccagtacaga tcgtttctca 480
 tctctaagt cctcattgag ttccgtgcat ctggccaatg agtctgctga gactcttgac 540
 agcacctcca gctctgctgc ttcaacaaca gtgacttgct ctccaatggt atccagtgat 600

tcgttgaaga ggaggtgctc tgtagcagaa actgagctcc gggtagctgg ttctcagtgg 660

ttgtctcatg tctctttttc tgtcttaggt ggtttcatta aatgcagcac ttggtttagca 720

gatgtttaat ttttttttta acaacattaa cttgtggcct ctttctacac ctggaaattt 780

actcttgaat aaataaaaac tcgtttgtct tgtcttctgc a 821

<210> 238

<211> 2938

<212> DNA

<213> Homo sapiens

<400> 238

gaaatccaac ccggtcacct acccgcgcca ctgtgtccac ggatggcacg aaagccaagc 60

gagtcacctt gccgagctac tcgctccgc ctctcccaa gctgagctct gctccgcca 120

cctgagtcct tcgccagtta ggaggaaaca cagccgctta atgaactgct gcacgggct 180

gggagagaaa gctcgcgggt cccaccgggc ctctaccca agtctcagcg cgcttttcac 240

cgaggcctca attctgggat ttggcagctt tgtgtgaaa gcccaatgga cagaggactg 300

cagaaaaatca acctatctc cttcaggacc aacgtacaga ggtgcagttc catggtacac 360

cataaatctt gacttaccac cctacaaaag atggcatgaa ttgatgcttg acaaggcacc 420

agtgtctaaag gttatagtga attctctgaa gaatatgata aatacattcg tgccaagtgg 480

aaaaattatg caggtgggtg atgaaaaatt gcctggccta cttggcaact ttcctggccc 540

ttttgaagag gaaatgaagg gtattgccgc tgttactgat atacctttag gagagattat 600

ttcattcaat attttttatg aattatttac catttgtact tcaatagtag cagaagacaa 660

aaaaggtcat ctaatacatg ggagaaacat ggattttgga gtatttcttg ggtggaacat 720

aaataatgat acctgggtca taactgagca actaaaacct ttaacagtga atttggattt 780

ccaaagaaac acaaaaactg tcttcaaggc ttcaagcttt gctggctatg tgggcatgtt 840

aacaggattc aaaccaggac tgttcagtct tacactgaat gaacgtttca gtataaatgg 900

tggttatctg ggtattctag aatggattct gggaaagaaa gatgtcatgt ggatagggtt 960

cctcactaga acagttctgg aaaatagcac aagttatgaa gaagccaaga atttattgac 1020

caagaccaag atattggccc cagcctactt tatcctggga ggcaaccagt ctggggaagg 1080

ttgtgtgatt acacgagaca gaaaggaatc attggatgta tatgaactcg atgctaagca 1140

gggtagatgg tatgtgttac aaacaaatta tgaccgttgg aaacatccct tcttccttga 1200

tgatcgaga acgctgcaa agatgtgtct gaaccgcacc agccaagaga atatctcatt 1260

tgaaccatg tatgatgtcc tgtcaacaaa acctgtctc aacaagctga ccgtatacac 1320

aaccttgata gatgttacca aagggtcaatt cgaaccttac ctgcgggact gccctgaccc 1380
 ttgtataggt tgggtgagcac acgtctggcc tacagaatgc gccctctgag acatgaagac 1440

 accatctcca tgtgaccgaa cactgcagct gtctgacctt ccaaagacta agactcgagg 1500
 caggttctct ttgagtcaat agcttgtctt cgtccatctg ttgacaaatg acagatcttt 1560
 tttttttccc cctatcagtt gatttttctt atttacagat aacttcttta ggggaagtaa 1620
 aacagtcate tagaattcac tgagttttgt ttcactttga catttgggga tctggtgggc 1680
 agtcgaacca tggatgaactc caccctccgtg gaataaatgg agattcagcg tgggtgttga 1740
 atccagcacg tctgtgtgag taacgggaca gtaaacactc cacattcttc agtttttcac 1800
 ttctacctac atatgtgat gtttttctgt ataacagcct tttcttctg gtcttaactg 1860

 ctgttaaaat taatatafca ttatctttgc tgttattgac agcgatataa ttttattaca 1920
 tatgattaga gggatgagac agacattcac ctgtatatit cttttaatgg gcacaaaatg 1980
 ggcccttgcc tctaaatagc actttttggg gtccaagaag taatcagtat gcaaagcaat 2040
 cttttataca ataattgaag tgttcccttt ttcataatta ctctacttcc cagtaaccct 2100
 aaggaagttg ctaacttaaa aaactgcate ccacgttctg ttaatttagt aaataaacia 2160
 gtcaaagact tgtggaaaat aggaagtga cccatatit aaattctcat aagtagcatt 2220
 catgtaataa acaggttttt agtttgttct tcagattgat agggagtttt aaagaaattt 2280

 tagtagttac taaaattatg ttactgtatt tttcagaaat caaactgctt atgaaaagta 2340
 ctaatagaac ttgttaacct ttctaacctt cagattaac tgtgaaatgt acgtcatttg 2400
 tgcaagaccg tttgtccact tcatittgta taatcacagt tgtgttccg acaactcaata 2460
 aacagtcact ggaaagagtg ccagtcagca gtcatgcacg ctgattgggt gtgtacatct 2520
 gtgttggttt ggaggagggg gggtattatc aagaaatata agctgaacat cgtatctgaa 2580
 ttaaataaca agggttatgt aagtgggtta ttaaatatga ctaaaccgtc acaatcttac 2640
 tcttacatga ctgataccac ctcttaacca gatactatca acccaatttg caggtgttta 2700

 aaggcagttt ttccccaga taggtggtgc ttccacaaaa ctttagtgat tccaagatta 2760
 ctttagtgaa acgaggaatt ataagaggat ttacattaaa ataggattta attaatgaaa 2820
 aaaaaacttt actcagattt tctacctctt acataaatga agtcaggtca acttgatgta 2880
 ttttttgggt aatgtaataa ctttgatat gcctgatagt aaagaggtga gatgctaa 2938

 <210> 239
 <211> 2496
 <212> DNA
 <213> Homo sapiens

<400> 239

actgcgactc gagacagcgg cccggcagga cagctccaga atgaaaatgc ggttcttggg	60
gttgggtggtc tgtttggttc tctggaccct gcattctgag gggctctggag ggaaactgac	120
agctgtggat ccigaaacaa acatgaatgt gagtgaaatt atctcttact ggggattccc	180
tagtgaggaa tacctagtig agacagaaga tggatatatt ctgtgcctta accgaattcc	240
tcatgggagg aagaaccatt ctgacaaagg tcccaaacca gttgtcttcc tgcaacatgg	300
cttgcctggca gattctagta actgggtcac aaaccttgcc aacagcagcc tgggcttcat	360
tcttgcctgat gctggttttg acgtgtggat gggcaacagc agaggaaata cctggtctcg	420
gaaacataag acactctcag tttctcagga tgaattctgg gctttcagtt atgatgagat	480
ggcaaaatat gacctaccag cttccattaa cttcattctg aataaaactg gccaagaaca	540
agtgatttat gtgggtcatt ctcaaggcac cactataggt tttatagcat tttcacagat	600
ccctgagctg gctaaaagga ttaaaatgtt ttttgccttg ggtcctgtgg cttccgtcgc	660
cttctgtact agccctatgg ccaaattagg acgattacca gatcatctca ttaaggactt	720
atttggagac aaagaatttc tccccagag tgcgtttttg aagtggctgg gtaccacgt	780
ttgcactcat gtcatactga aggagctctg tggaaatctc tgttttcttc tgtgtggatt	840
taatgagaga aatttaaata tgtctagagt ggatgtatat acaacacatt ctctgctgg	900
aacttctgtg caaaacatgt tacactggag ccaggctgtt aaattccaaa agtttcaagc	960
ctttgactgg ggaagcagtg ccaagaatta ttttcattac aaccagagtt atcctccac	1020
atacaatgtg aaggacatgc ttgtcccgac tgcagtctgg agcgggggtc acgactggct	1080
tgcagatgtc tacgacgtca atatcttact gactcagatc accaacttgg tgttccatga	1140
gagcattccg gaatgggagc atcttgactt catttggggc ctggatgccc cttggaggct	1200
ttataataaa attattaatc taatgaggaa atatcagtga aagctggact tgagctgtgt	1260
accaccaagt caatgattat gtcatgtgaa aatgtgtttg cttcatttct gtaaaacact	1320
tgtttttctt tcccaggctt tttgtttttt tatatccaag aaaatgataa ctttgaagat	1380
gcccagttca ctctagtttc aattagaaac atactagcta ttttttcttt aattagggt	1440
ggaataggaa gccagtgtct caaccatagt attgtctctt taagtctttt aaatatcact	1500
gatgtgtaaa aaggtcatta tatccattct gtttttaaaa tttaaatat attgactttt	1560
tgcccttcat aggacaaagt aatatatgtg ttggaatttt aaaattgtgt tgtcattgg	1620
aaatctgtca ctgacttaag cgaggtataa aagtacgcag ttttcatgtc cttgccttaa	1680
agagctctct agtctaacgg tctttagtct agagatctaa atgacatttt atcatgtttt	1740

cctgcagcag gtgcatagtc aaatccagaa atatcacagc tgtgccagta ataaggatgc 1800

taacaattaa ttttatcaaa cctaactgtg acagctgtga tttgacacgt ttttaattgct 1860

caggttaaat gaaatagttt tccggcgtct tcaaaaacaa attgcactga taaaacaaaa 1920

acaaaagtat gttttaaatg ctttgaagac tgatacactc aaccatctat attcatgagc 1980

tctcaatttc atggcaggcc atagtcttac ttatctgaga agcaaattccc tgtggagact 2040

ataccactat tttttctgag attaatgtac tcttggagcc cgctactgtc gttattgac 2100

acatctgtgt gaagccaaag ccccggtggt gcccatgaga agtgccttg ttcattttca 2160

cccaaatgaa gtgtgaacgt gatgttttcg gatgcaaact cagctcaggg attcattttg 2220

tgtcttagtt ttatatgcat ctttattttt aatacacctg cttcacgtcc ctatgttggg 2280

aagtccatat ttgtctgctt ttcttgcagc atcatttctt tacaatactg tccggtggac 2340

aaaatgacaa ttgatatgtt tttctgatat aattacttta gctgcactaa cagtacaatg 2400

cttgtaaatg gttaatatag gcaggcgaa tactactttg taacttttaa agtcttaaac 2460

ttttcaataa aattgagtga gacttatagg cccaaa 2496

<210> 240

<211> 2174

<212> DNA

<213> Homo sapiens

<400> 240

agtgcagccc gaagccccgc agtccccgag cacgcgtggc catgcgtccc ctgcgcccc 60

gcgcgcgcgt gctggcgctc ctggcctcgc tcttggccgc gcccccggtg gccccggccg 120

aggccccgca cctgggtgat gtggacgcgg cccgcgcgt gtggccccgt cggcgcttct 180

ggaggagcac aggtttctgc cccccgtgc cacacagcca ggctgaccag tacgtcctca 240

gctgggacca gcagctcaac ctgcctatg tgggcgccgt ccctcaccgc ggcataaagc 300

aggtccggac cactggctg ctggagcttg tcaccaccag ggggtccact ggacggggcc 360

tgagctacaa cttcaccac ctggacgggt acctggacct tctcaggag aaccagctcc 420

tcccagggtt tgagctgatg ggcagcgct cgggccactt cactgacttt gaggacaagc 480

agcaggtgtt tgagtgaag gacttggctt ccagcctggc caggagatac atcggtaggt 540

acggactggc gcatgtttcc aagtggaact tcgagacgtg gaatgagca gaccaccacg 600

actttgacaa cgtctccatg accatgcaag gcttctgaa ctactacgat gcctgctcgg 660

agggtctcgc cgccgccag cccgccctgc ggttgggagg ccccggcgac tccttcaca 720

ccccaccgcg atccccgctg agctggggcc tctgcgcca ctgccacgac ggtaccaact 780

tcttctactgg ggaggcgggc gtgcggctgg actacatctc cctccacagg aagggtgcgc	840
gcagctccat ctccatcctg gagcaggaga aggtcgtcgc gcagcagatc cggcagctct	900
tccccaagtt cgcgacaccc cccatttaca acgacgaggc ggacccgctg gtgggctggt	960
ccctgccaca gccgtggagg gcggacgtga cctacgcggc catggtggtg aaggtcatcg	1020
cgcagcatca gaacctgcta ctggccaaca ccacctccgc cttccctac gcgtcctga	1080
gcaacgacaa tgccttcctg agctaccacc cgcacccctt cgcgacgagc acgtcacccg	1140
cgcgcttcca ggtaacaac acccgcccg gcacacgtga gctgttgcgc aagccggtgc	1200
tcacggccat ggggctgctg gcgctgctgg atgaggagca gctctgggcc gaagtgtcgc	1260
aggccgggac cgtcctggac agcaaccaca cgggtggcggt cctggccagc gcccaccgcc	1320
cccagggccc ggccgacgcc tggcgcgccc cgggtgctgat ctacgcgagc gacgacaccc	1380
gcgcccaccc caaccgcagc gtgcgggtga cctgcgggt gcgcgggggtg cccccggcc	1440
cgggcctggt ctacgtcacg cgctacctgg acaacgggct ctgcagcccc gacggcgagt	1500
ggcggcgccct gggccggccc gtcttcccca cggcagagca gttccggcgc atgcgcgagg	1560
ctgaggaccc ggtggccgag gcgccccgcc cttaccgcgc cggcggccgc ctgacctgc	1620
gccccgcgct gcggctgccc tcgcttttgc tgggtgcacgt gtgtgcgcgc cccgagaagc	1680
cggccgggca ggtcacgcgg ctccgcgcc tgccttgac ccaagggcag ctggttctgg	1740
tctggctcga tgaacacgtg ggctccaagt gcctgtggac atacgagatc cagttctctc	1800
aggacggtaa ggcttacacc ccggtcagca ggaagccatc gaccttaac ctctttgtgt	1860
tcagcccaga cacaggtgct gtctctggct cctaccgagt tcgagccctg gactactggg	1920
cccgaccagg ccccttctcg gacctgtgc cgtacctgga ggtccctgtg ccaagagggc	1980
ccccatccc gggaatcca tgagcctgtg ctgagcccca gtgggttgca cctccaccgg	2040
cagtcagcga gctggggtg cactgtgcc atgctgccct cccatcacc cttttgcaat	2100
atatttttat atttattat tttctttat atcttggtac caacgcccc tttaaagcgg	2160
ctttgcacag gtca	2174
<210> 241	
<211> 7580	
<212> DNA	
<213> Homo sapiens	
<400> 241	
actcgcgctg cggccagcgc ccgggcctgc gggcccgggc ggcggtgtg ttgcgcagtc	60
ttcatgggtt cccgacgagg aggtctctgt ggctgcggcg gcggtgcta actgcgccac	120

ctgctgcagc ctgtccccgc cgctctgaag cggccgcgtc gaagccgaaa tgccgccacc	180
cggaccggc cgagggcttc tctggctggg tctggttctg agctccgtct gcgtcgccct	240
cggatccgaa acgcaggcca actcgaccac agatgctctg aacgttcttc tcatcatcgt	300
ggatgacctg cgccccctcc tgggctgtta tggggataag ctggtgaggt ccccaaatat	360
tgaccaactg gcatccaca gcctcctctt ccagaatgcc tttgcgcagc aagcagtgtg	420
cgccccgagc cgcgtttctt tcctcactgg caggagacct gacaccacc gcctgtacga	480
cttcaactcc tactggaggg tgcacgctgg aaacttctcc accatcccc agtacttcaa	540
ggagaatggc tatgtgacca tgtcggtggg aaaagtcttt caccctggga tatcttctaa	600
ccataccgat gattctccgt atagctggc tttccacct tatcaccct cctctgagaa	660
glatgaaaac actaagacat gtcgagggcc agatggagaa ctccatgcca acctgctttg	720
ccctgtggat gtgctggatg ttcccagggg caccctgcct gacaaacaga gcactgagca	780
agccatacag ttgttgaaa agatgaaaac gtcagccagt ctttcttcc tggccgttgg	840
glatcataag ccacacatcc cttcagata cccaaggaa ttccagaagt tgtatccctt	900
ggagaacatc acctggccc cggatcccga ggtccctgat ggccctaccc ctgtggccta	960
caaccctgg atggacatca ggcaacggga agacgtccaa gccttaaaca tcagtgtgcc	1020
glatggtcca attcctgtgg actttcagcg gaaaatccgc cagagctact ttgcctctgt	1080
glatatattg gatacacagg tcggccgcct cttagagtgt ttggacgac ttcagctggc	1140
caacagcacc atcattgcat ttacctgga tcatgggtgg gctctaggtg aacatggaga	1200
atgggcaaaa tacagcaatt ttgatgttc taccatgtt cccctgatat tctatgttcc	1260
tggaggagc gcttcacttc cggaggcagg cgagaagctt ttcccttacc tcgacccttt	1320
tgattccgc tcacagtga tggagccagg caggcaatcc atggaccttg tggaaactgt	1380
gtctcttttt cccacgtgg ctggacttgc aggactgcag gttccacctc gtgccccgt	1440
tccttcattt cacgttgagc tgtgcagaga aggcaagaac cttctgaagc attttcgatt	1500
ccgtgacttg gaagaggatc cgtacctccc tggtaatccc cgtgaactga ttgcctatag	1560
ccagtatccc cggccttcag acatccctca gtggaattct gacaagccga gtttaaaaga	1620
tataaagatc atgggctatt ccatacgac catagactat aggtatactg tgtgggttgg	1680
cttcaatcct gatgaatttc tagctaactt ttctgacatc catgcagggg aactgtattt	1740
tgtggattct gaccattgc aggatcacia tatgtataat gattcccaag gtggagatct	1800
tttccagtgt ttgatgcctt gagttttgcc aacctggat ggcaaatgtg atgtgtccc	1860
ttccagctgg tgagaggagg agtttagact ggtcgttttg tgattacca taatattgga	1920

agcagcctga gggctagtta atccaaacat gcatcaacaa tttggcctga gaatatgtaa 1980

cagccaaacc ttttcgttta gtctttatta aaattttataa ttggtaattg gaccagtttt 2040

ttttttaatt tccctctttt taaaacagtt acggcttatt tactgaataa atacaaagca 2100

aacaaactca agttatgtca tacctttgga tacgaagacc atacataata accaaacata 2160

acattataca caaagaatac tttcattatt tgtggaattt agtgcatttc aaaaagtaat 2220

catatatcaa actaggcacc acactaagtt cctgattatt ttgtttataa ttttaataata 2280

tatcttatga gccctatata ttcaaaatat tatgttaaca tgtaatccat gtttcttttt 2340

caaatctaaa gttaaaaaaa aatagcagaa gccagtgtct taaagtctat cttttgtttc 2400

taagaccatg ggatttcata atctcaagat aaaatatgta tgaagtaatt aatgtagaat 2460

ttttacacca aataataaat aatgcttaat aaactagaga tatgagatgt gtaggaaatt 2520

tggttaaact ttttcagat actttctggc ccaataataa atttgtagc aaataatatg 2580

accttgaac tcaatggcca tctattaaaa gactgttgtt cacttgga aacatttaaa 2640

gatgtgacta tatccatggg tggattgaat cactcaaaat atattagtat ctttctttag 2700

ggatggttgg ttacagacat gtattttattc aggaggcaga aaatattcca ttttaattgc 2760

ttattaaaga aaacattaaa ttctaaatta ttttgaggac tgtgaagact tttcattagt 2820

glaatattag gtcattgtca atctcccaga atgtagtctt atattctcta aatatgaaag 2880

tatccagaaa ggccagtggg agtaaaaagc ttagtgtata taatctcaa agggatggaa 2940

tatttacaac tcatatttat aacatgttga atcttctcag ttatcagtag tcatcagaag 3000

tgtcaatagc tttctaaata aatattaaat atctactgtc ctgtagtgaa ggagtaattt 3060

ttagtaattt tctctttaca aagtctccag tgtttccagg taaatatttg tgaacaaaa 3120

tacagcaaac tacattgtta cttcagtgtt ttgttgccaa aaatgacaag atattatatt 3180

aaaatcagta aatttttagac agattttaaa aattaattag cctacaatag aggttatatg 3240

glaacacggg gatcttctaa gcagttaagt gactgactgt tctggcaaca acgacttctc 3300

cgtgactgaa gggccctgtt catttctga tctgaagct cgtctctctt ttgagcctcc 3360

gcttgctttg gtcgatgggt tccctcagct ttttctttgc tgttcttcat cctcgttgtt 3420

gctgtcatca tgttcactgt ggcttttaca atacagcctg taaattcctt atgacatagt 3480

tcagtgcatt tggttttatt gctgctcca cagtcttita cctttacttg gcttagagaa 3540

actgtatctt tgttgcttca tataaccttt cccaacccc actaagctgg acataactta 3600

ttagtggtcc tcccgctact ttattttagt aaatctctct ttcacatgag caggggttct 3660

ttcatgtggt ttagctgaca gcagaactag tgattctaga cattttgcat ggccctcatt	3720
cagtggctca caaacatgag ggagcatcag aactacttga ggggcttggt aaaaccagtt	3780
gcgttagaag tcggatgcgg tggctcacac ctgtaatccc agcactttgg gaggcccagg	3840
caggcggatc acttgagggt aggagttaa gaccagcctg gccaacatgg tgaaaccccg	3900
tctctactaa aaatacaaaa gttagccggg tgtggtggtg catgcctgta atcccagctt	3960
cttgggaggc caaggcacia gaatcgcttg aaccaggaga cggaggtttc agtgaatgaa	4020
gatcgtgcca ttgtattcca gcctcggcaa cacagcagga ctgtgatttt ctttgagac	4080
tcctagattt tctgtggttt tgaactgaat ttgttgatg ttggcaagtg cctcttatga	4140
gctgtttctt tatcctgcat ttgccccaca aagacttatt tggagggtgag caaagtatgt	4200
ttggtagtga ggtcacaaag gcaatcagcc ccttctctcc cactccatt gccatcttct	4260
cagtcttctt ccttttctt ccaagtagtt taccacccc tctctttcc tcccgtgctc	4320
ctaaaataat ccacgtgtct tcctaaaatc tctctttgat cctgtccttt gataacaccg	4380
tcagtgccta ctactgggtc tagacagacc tctgttgagc agtcagagtc ttccctgact	4440
ccacaatgcc cttttcttg gctgaccagt atgactactg gtcccccact ttcccttgcc	4500
tatccctacc tccctcctac taggttgctc catccctctc ttacccatt cattcatgac	4560
catttttcac taccaagctc cccccctccc gaaggaggct gaggtttttg tgactctcta	4620
gactctattg tgggatggaa tgaacattgc taaagaatct tgtgttcgct ttactttaaa	4680
aaggtatttt tttcctaatt ataaaactga tgtgtcagtt acggaaaaat tagaaatgca	4740
gcacaaatac atgaatatat taccacaaaa ttgcatata atatcttgct ttttttgggg	4800
gtgtgaattt ttgcatgtg tctggtcata ttttttatca tgtaatttat gtctttttt	4860
actaagtatt atgtgtggtt attatagatt ttcacaaaga tatattgctg gtaatatatt	4920
ttattgtgta gtcttataat ttacttaacc ttttttcaat tgttagaaat ttaggtatt	4980
tccagatttt cagtattgta aataatgctg tgatgaccaa ttttgtgaat aaaatgtttt	5040
tatgtatttc agattattcc cttaggatag tctctcagtg ccaagttgct aaaaacatct	5100
ctattttgct tatcttctg ctctcttgct gccttagggg gtagtaaaact gaaacataaa	5160
glaaacatgc atacaaataa aaaacataaa acaaaaataa gcaacctgat ggtaataggt	5220
gaaagtggta acctgtttta actttgaatt ctgcccgggc gcggtggctc acgcctgtaa	5280
tcccagcact ttgggaggct gaggcgggtg gatcacgagg tcaggagtgc aaaaccagcc	5340
tggccaagat ggtgaaatcc cgtctctact aaaaatacaa aaattagccg ggcgtggtgg	5400
cgggcgcctg taatcccage tacttgggag gctgaggcag agaattgctt gaaccagga	5460
ggcggagggt gcagtgagcc aagatcgctc cactgcactc cagcctgggt gacagagcga	5520

gactccgtct caaataaaaa acaacaaaaa acaaaaaaaa cttaaaattc ttgcttgtt 5580
 agtgaccttg atcatggttc tctttgtacg atagtgggc atctgtattt ccacttgtgt 5640
 gaatttgcct ttaaattttg gttatgggtt tcacctttta aaataatcaa acatatttat 5700
 cttttcctgt gtgatagggt tttttctgta tcttttcctg ttaaacacac agaccctcc 5760

ccaatctgga cattgaataa atattcattt tcctttgcat tgtttgggtt ttggttttat 5820
 gcttactgtg tctctccacc ccattatcag agtaatatat gctcattgca taaaatctgg 5880
 aacatgtgaa aactggagat gaaaagaagc ccagcatttg agacataccc actgctgaca 5940
 ttgtatatt tcctttcatt cttttacctg cctaaagaca tatatgcaga tatggacata 6000
 tttagcttta catacatgca gactaccaca tataatctat atcttgtttt ttgatctaa 6060
 cacaatgta gcttggtttt ttatgtcatt gacagtctt gattaacagc tgtaatggct 6120
 gcacagccct gcttcattaa tttttttttt tggactgaaa tgtatgtacc acgttccgt 6180

tatccattca tccaccagtg accacttggg tcatttctac cttttggcca ttgcaataat 6240
 gttgctgtca acaaggatgt acaaatatct gtctgagtc ctgctcttaa ttgttttggg 6300
 tatacaccca acctacatta atctttaatc acttaagatg ttgggggttg tgctgggagg 6360
 taaccaagtg actctggagc agcctcacc tagctacaga gggaaactga ctcttcctc 6420
 tctgcccatt cttttgccgc ctgtcagggt cccaagctgt ccataacctta ctgtgtgaag 6480
 cgctttgggt ataaactccc aaatagaggc tcagccctt ttcctttccc aagtcatecc 6540
 tgcaggttat agctcttcag aggaatctct tgaattctgg cctgggatct gggattaaag 6600

tagaaccag ggtgtccctt ggggccttag gcatagcccg ttgtggtgga acctgccatc 6660
 agccccactg cactcagtgg tcccctggta cttcaggaca gtttctctt tctgcacatg 6720
 agaacttcaa gcagagagga ctgtcagttg ttgccttcac atcaccaatc gtttggccat 6780
 aatggccaag agccctcact gctcttgggt cttgccttc cacatgttac tgaaagctec 6840
 cgttgggtgc cctgtccact aactcacctg atgggacaga tgccacttc ctctgagtgt 6900
 ctgataaaac acagtctgtt gtggggagcc cacgtcaag ggttttctt gggatcagag 6960
 ttcccggtt attactccat gccctgcctg acatcactgt catgtatggt gctaatcacc 7020

aaaggatgcc ctgttttgat tctggattct aggtgatagt ccttatectc tgatggggga 7080
 agagtttggg tgggagagaa caaacgaca agttggttcc agatgaggtg gctgggagcc 7140
 aacctcagg tacttggctc tgagtaatgt gacggcccag gcaagagggc tccactgcc 7200
 aatgaggaac aaaagctcaa ggtctagtag tgtattcagc acatgctagg aaacaagagt 7260
 tgggacgtgg acaaggagag gagcaaaggc caggggttaa ctgcttttgg tttttcctct 7320
 ataaaagtca gggactatga tttatatggg tgagaagggt tggagtagga ctgaagctgc 7380

catttacaaa ccccttgccc taatttattg tagttttggg tgtgtttctg aaagaacctt	7440
ttgcagggtga tttttattca gtactcatta acttacagct tttatgaatt cacatcttct	7500
acactattga aaaacataag tttgtattgt gtttaagtga gcatgagtta ttatatgaat	7560
aaatcacatt aaatttgcca	7580
<210> 242	
<211> 2705	
<212> DNA	
<213> Homo sapiens	
<400> 242	
gagtccgagc cgccgccgcc atgagctgcc ccgtgccgcg ctgctgcgcg ctgctgctag	60
tcctggggct ctgccggcg cgccccgga acgcaactgct gctcctcgcg gatgacggag	120
gctttgagag tggcgcgtac aacaacagcg ccatcgccac cccgcacctg gacgccttgg	180
cccgccgag cctcctcttt cgcaatgcct tcacctcggg cagcagctgc tctcccagcc	240
gcgccagcct cctcaactgg ctgccccagc atcagaatgg gatgtacggg ctgcaccagg	300
acgtgcacca ctcaactcc ttcgacaagg tgcggagcct gccgctgctg ctgagccaag	360
ctggtgtgcg cacaggcatc atcgggaaga agcacgtggg gccggagacc gtgtaccgt	420
ttgactttgc gtacacggag gagaatggct ccgtcctcca ggtggggcgg aacatcacta	480
gaattaagct gctcgtccgg aaattcctgc agactcagga tgaccggcct ttcttctct	540
acgtcgctt ccacgacccc caccgctgtg ggcaactcca gcccagtag ggaaccttct	600
gtgagaagtt tggcaacgga gagagcggca tgggtcgtat ccagactgg acccccagg	660
cctacgaccc actggacgtg ctggtgcctt acttcgtccc caacacccg gcagcccag	720
ccgacctggc cgtcagtag accaccgtcg gcgcgatga ccaaggagtt ggactggtgc	780
tccaggagct gcgtgacgcc ggtgtcctga acgacacact ggtgatcttc acgtccgaca	840
acgggatccc ctccccagc ggcaggacca acctgtactg gccgggact gctgaacct	900
tactggtgct atccccggag caccacaaac gctggggcca agtcagcgag gcctacgtga	960
gcctcctaga cctacgccc accatcttgg attggttctc gatcccgtag cccagctacg	1020
ccatctttgg ctgcaagacc atccacctca ctggccggtc cctcctgccg gcgtggagg	1080
ccgagccct ctgggccacc gtctttggca gccagagcca ccacagggtc accatgtcct	1140
accccatgcg ctccgtgcag caccggcact tccgcctcgt gcacaacctc aacttcaaga	1200
tgccctttcc catcgaccag gacttctacg tctacccac cttccaggac ctctgaacc	1260
gcaccacagc tggtcagccc acgggctggt acaaggacct ccgtcattac tactaccggg	1320

cgcgctggga gctctacgac cggagccggg acccccacga gaccagaaac ctggccaccg	1380
acccgcgctt tgctcagctt ctggagatgc ttcgggacca gctggccaag tggcagtggg	1440
agaccacga cccctgggtg tgcgccccg acggcgtcct ggaggagaag ctctctcccc	1500
agtgccagcc cctccacaat gagctgtgac catcccagga ggctgtgca cacatcccag	1560
gcatgtccca gacacatccc acacgtgtcc gtgtggccgg ccagcctggg gagtagtggc	1620
aacagccctt cgtccacac tcccatccaa ggagggttct tcttctctgt ggggtcactc	1680
ttgccattgc ctggaggggg accagagcat gtgaccagag catgtgccca gcccctccac	1740
caccaggggc actgccgtca tggcagggga cacagtgtc cttgtgtctg aacctgtcc	1800
cagcacggga attctagaca tacgtggtct gcggacaggg cagcgcccc agcccatgac	1860
aaggaggtct tgttttctgg cttggtttgg ggacctgcaa atgggaggcc tgaggccctc	1920
ttcagccttt ggcagccaca gatacttctg aaccttcac agagagcagg caggggcttc	1980
ggtgccgctt gggcagtacg caggctccac cgacactcac ctgggagcac ggcgcctggc	2040
tcttaccagc gtctggccta gaggaagcct ttgagcgacc tttgggcagg tttctgcttc	2100
ttctgttttg ccccatggtc aagtcctgt tcccaggca ggtttcagct gattggcagc	2160
aggctccctg agtcatgagc ttgaacctgt ggtgtttctg ggcagaagct tatcttttt	2220
gagagtgtcc gaagatgaag gcatggcgat gcccgctctc tggcttgggt taattcttcg	2280
gtgacactgg cattgctggg tggatgatgc cgtcctctgg cttgggttaa tcttcgggtg	2340
acactggcgt tgctgggtgg caatgcccc cctctgcctt gggttaattc ttcggtgaca	2400
ctggcgttgc tgggtggcga tgcccgtcct ctggcttggg ttaattcttg gatgacgtcg	2460
gcgttctggt gagaatgtgc cgttctctgc ctgcctccac ccacctcggg agcagaagcc	2520
cggcctggac acccctcggc ctggacaccc ctggaaggag agggcgcttc cttgagtagg	2580
tgggtcctcc ttgcccttcc ctccctatca ctccatactg ggggtgggtg gaggaggcca	2640
caggccagct attgtaaaag ctttttattt tagtaaaata tacagaagtt ctttttctga	2700
actca	2705
<210> 243	
<211> 2475	
<212> DNA	
<213> Homo sapiens	
<400> 243	
ctggccgtcg cgggaccgcg aggaactgaga ccatggaggc ggtggcgggtg gccgcggcgg	60
tgggggtcct tctcctggcc ggggcccggg gcgcggcagg cgacgaggcc cgggaggcgg	120

cgcccgctgcg ggcgctcgtg gcccggctgc tggggccagg ccccgcgcc gactttctccg	180
tgtcgggtgga gcgcgctctg gctgccaagc cgggcttggga cacctacagc ctgggcggcg	240
gcggcgccggc gcgcgtgcgg gtgcgcggct ccacgggcgt ggccggccgcc gcggggctgc	300
accgctacct gcgcgacttc tgttgctgcc acgtggcctg gtccggctct cagctgcgcc	360
tgcccgggcc actgccagcc gtgccggggg agctgaccga ggccacgccc aacaggtacc	420
gctattacca gaatgtgtgc acgcaaagct actctttcgt gtggtgggac tgggcccgct	480
gggagcgaga gatagactgg atggcgctga atggcatcaa cctggcactg gcctggagcg	540
gccaggaggc catctggcag cgggtgtacc tggccttggg cctgaccag gcagagatca	600
atgagttctt tactggctct gccttcctgg cctggggcg aatgggcaac ctgcacacct	660
gggatggccc cctgcccccc tcttggcaca tcaagcagct ttacctgcag caccgggtcc	720
tggaccagat gcgctccttc ggcatgacct cagtgtctgc tgcattcgcg gggcatgttc	780
ccgaggctgt caccagggtg ttccctcagg tcaatgtcac gaagatgggc agttggggcc	840
actttaactg ttctactcc tgcctcttc ttctggctcc ggaagacccc atattccca	900
tcatcgggag cctcttctg cgagagctga tcaaagagtt tggcacagac cacatctatg	960
gggccgacac tttaaatgag atgcagccac ctctctcaga gccctctac cttgcccgag	1020
ccaccactgc cgtctatgag gccatgactg cagtggatac tgaggctgtg tggtctctcc	1080
aaggctggct ctccagcac cagccgcagt tctgggggcc cgccagatc agggctgtgc	1140
tgggagctgt gcccgtggc cgctctctgg ttctggacct gtttctgag agccagcctg	1200
tgtatacccg cactgcctcc ttccagggcc agcccttcat ctggtgcatg ctgcacaact	1260
ttgggggaaa ccatggtctt ttggagccc tagaggctgt gaacggaggc ccagaagctg	1320
ccgcctctt ccccaactcc accatggtag gcacgggcgt ggcccccgag ggcacagcc	1380
agaacgaagt ggtctattcc ctcatggctg agctgggctg gcgaaaggac ccagtcccag	1440
atttggcagc ctgggtgacc agctttgccg ccggcggta tggggtctcc caccggacg	1500
caggggcagc gtggaggcta ctgctccgga gtgtgtacaa ctgctccggg gaggcctgca	1560
ggggccacaa tcgtagcccc ctggtcaggc ggccgtccct acagatgaat accagcatct	1620
ggtacaaccg atctgatgtg tttaggcct ggccgctgct gctcacatct gctccctccc	1680
tggccaccag ccccgcttc cgctacgacc tctggacct cactcggcag gcagtgcagg	1740
agctggtcag ctgtactat gaggaggcaa gaagcgcta cctgagcaag gagctggcct	1800
ccctgttagg ggtcggaggc gtcctggcct atgagctgct gccggcactg gacgaggtgc	1860
tggctagtga cagccgttc ttgctgggca gctggctaga gcaggcccga gcagcggcag	1920

tcagtgaggc cgaggccgat ttctacgagc agaacagccg ctaccagctg accttggtgg	1980
ggccagaagg caacatcctg gactatgcc acaagcagct ggcggggttg gtggccaact	2040
actacacccc tcgctggcgg cttttcctgg aggcgctggt tgacagtgtg gccagggca	2100
tccctttcca acagcaccag ttgacaaaa atgtcttcca actggagcag gccttcgttc	2160
tcagcaagca gaggtacccc agccagccgc gaggagacac tgtggacctg gccaagaaga	2220
tcttctcaa atattacccc cgctgggtgg ccggctcttg gtgatagatt cgccaccact	2280
gggccttggt ttccgcta atccagggcag attccagggc ccagagctgg acagacatca	2340
caggataacc caggcctggg aggaggcccc acggcctgct ggtgggtctt gacctggggg	2400
gattggaggg aaatgacctg ccctccacca ccacccaaag tgtgggatta aagtactgtt	2460
ttctttccac ttaaa	2475
<210> 244	
<211> 5227	
<212> DNA	
<213> Homo sapiens	
<400> 244	
gggcaggcaa ggccggccga gcggcgccgc ggcatgagcg ggccgggcag ggcgtggcc	60
gcgtgctgc tggccgcgc cgtgctgagc gccgcgtgc tggccccgg cggctcttcg	120
ggcgcgatg ccaggccgc gccccacga gacttagaca aaaaaagaca tgcagagctg	180
aagatggatc aggccttgct actcatccat aatgaacttc tctggaccaa ctgaccgtc	240
tactggaaat ctgaatgctg ttatcactgc ttgtttcagg ttctggtaaa cgttcctcag	300
agtccaaaag cagggaagcc tagtgctgca gctgcctctg tcagcacca gcacggatct	360
atcctgcagc tgaacgacac ctggaagag aaagaagttt gtaggttga atacagattt	420
ggagaatttg gaaactattc tctcttgta aagaacatcc ataattggagt tagtgaaatt	480
gcctgtgacc tggtgtgaa cgaggatcca gttgatagta accttctgt gagcattgca	540
ttccttattg gtcttgctgt catcattgtg atatccttc tgaggctctt gttgagtttg	600
gatgacttta acaattggat ttctaaagcc ataagtctc gagaaactga tcgcctcatc	660
aattctgagc tgggatctcc cagcaggaca gacctctcg atggtgatgt tcagccagca	720
acgtggcgtc taictgcct gccgccccgc ctccgcagcg tggacacctt cagggggatt	780
gctcttatac tcatggtctt tgtcaattat ggaggaggaa aatattgta cttcaaacat	840
gcaagtggga atgggctgac agtggctgac ctctgttcc cgtggtttgt atttattatg	900
ggatcttcca tttttctatc gatgacttct atactgcaac gggggtgttc aaaattcaga	960

ttgctgggga agattgcatg gaggagtttc ctgttaatct gcataggaat tatcattgtg	1020
aatcccaatt attgccttgg tccattgtct tgggacaagg tgcgcattcc tgggtgtgctg	1080
cagcgattgg gagtgcacata ctttgtggtt gctgtgttgg agtcctctct tgctaaacct	1140
gtgcctgaac attgtgcctc ggagaggagc tgcctttctc ttcgagacat cacgtccagc	1200
tggccccagt ggctgctcat cctggtgctg gaaggcctgt ggctgggctt gacattcctc	1260
ctgccagtcc ctgggtgccc tactggttat cttggtcctg ggggcattgg agattttggc	1320
aagtatcaa attgcactgg aggagctgca ggctacatcg accgcctgct gctgggagac	1380
gatacctttt accagcaccc atcttctgct gtactttacc acaccgaggt ggcctatgac	1440
cccgagggca tcctgggcac catcaactcc atcgtgatgg cctttttagg agttcaggca	1500
ggaaaaatac tattgtatta caagctcgg accaaagaca tcctgattcg attcactgct	1560
tgggtgttga ttcttgggct cattttctgt gctctgacga aggtttctga aaatgaaggc	1620
tttatccag taacaaaaa tctctgggtc ctttcgtatg tcaactacgt cagttctttt	1680
gcctttctca tcctgctggt cctgtacca gtgttgatg tgaaggggct gtggacagga	1740
accccttct tttatccagg aatgaattcc attctggtat atgtcgcca cgagggtttt	1800
gagaactact tccccttca gtggaagctg aaggacaacc agtcccaca ggagcacctg	1860
actcagaaca tcgtgccac tgcctctggt gtgctcattg cctacatcct ctatagaaag	1920
aagatttttt ggaaaatctg atggctccca ctgagatgtg ctgctggaag actctagtag	1980
gcctgcaggg aggactgaag cagcctttgt taaaggaag cattcattag gaaattgact	2040
ggctgctgt ttacagactc tgggggaaga cactgatgtc ctcaaactgg ttaactgtga	2100
cacggctcgc cagaactctg cctgtctatt tgtgacttac agatttgaaa tgtaattgtc	2160
tttttctc catcttctgt ggaaatgat gtctttggaa cttcattccg aggagataag	2220
ctttaacttt ccaaaaggga attgccatgg gtgtttttct tctgtggtga gtgaaacaat	2280
ctgaggtctg gtcttctgtg accttgttgc cctgcaaact tcctttccac gtgtacgcgc	2340
acaccaacac gaaatgccat cactcctact gcggctgcta tgaagcttac tggttgtgat	2400
gtgttataat ttagtctgtt tttttgattg aatgcagttt aatgtttcca gaaagccaaa	2460
gtaattttct tttcagatat gcaaggcttt ggtgggtcca aaaaatgtct atcacaagcc	2520
attttttct tttcctctct cgaaaagtta aaatatctat gtgttattcc caaacctct	2580
tacctatgta tcigcctgtc tgtccatcat ctctcttct cctatctct gtgtatctgg	2640
atggcagccg ctgccaggg gagtggctgt ggggagggca ggtactgtct ttgcctgtgg	2700

gtccagctga gccatccctg ctgggtgatg ctgggcaaga cccttggccc gtctgggcct 2760

tggtctctc acttgtgaaa tgagcgggaa gatgactctc agttccttcc acctcttaga 2820

catggtgagg taacagacat caaaagcttt tctgaaatct tcagaagaaa tagttccatt 2880

acagaaaact cticaaata aatagtagtg aaaactttta aaaactctca ttggagtaag 2940

tcttttcaag atgactctcc acaatggagg cagcgttctt acttgtctac acacagctga 3000

agacattgtt tcttaggtgt gaaatcgggg acaaaggaca aacagagaca cacggcattg 3060

ttcatgggag gcatcgtcac cctcctgggt gtctgtggg aatttcctgt gtgaggaaaa 3120

cgtggccaca gggttgtgct gtaccacccc tccccggcg agatggccct cggcctgtgc 3180

cgctgcttc accctcgcca ctccatggca gcttttggtc tgtttccggc tctgccctct 3240

gccctgaact ctcatccggc ttgtacctgc ctgctggacc cctccacctg gaggccagcc 3300

catgtctcag gccagccct agcctcttct cctcaaattc taagtgtttt ctcttttagt 3360

ttccctggct ttgtgaatgg atcatgtgtc tctaggtata aacctgacat catctttcca 3420

cccggcttac ctccaccaga tctccccagt tctgtctcca tcttctacct gcagctgtc 3480

tgttctcatg gtcactgtg catcactgag tctggaccct tgttatcatt ttcaaatgg 3540

cctccttccc tegtcccca cttcttaaag tcacctgtcc attgccacca gattaagctt 3600

tctccagcca gatcacctct ctctgagaaa cctccattga catggaaaca ccattgtctg 3660

gcacacatac tcacatactc accttcccgt ctgatcccc acacatcttt ccagcctccc 3720

ctccactcc actccctgt ctctctcca cctcccac ccttctctc ccttccctc 3780

tgaatccagc ccagcggggc ttctcctgcc tccatccat cacagaagta cctcctgctt 3840

ctggttttaa ttagagcctt ccccgattac atttctctt gaatttttc ctatctacat 3900

ttgatctgtc atgtttaaac cccctacttc taagggaact tctctaact cttatcctca 3960

tccccaaata gtgttttctt cctctgggtt cttataatgt tggatatcaat ctcacagcat 4020

ttagtgcttc ctgcctgggtg tgacagttac ctgtgtgcat gtgcaatttc taatttccca 4080

cgctagactg tgagcttctt aaggcaagaa tcatgccttg ttggtttctg tattctctat 4140

ggtgccaaac acagtgcctt ctacattgca ggcgctgaat aaacatcttt aaagcaaat 4200

gatgtggatt tttaaaataa atatttaagt gctggtaaga tgagcatgta tccggggtgc 4260

ccatgaaatg ttcttggggc cgtgtgggga cagtcgtcat tctcctcct gccacccttt 4320

tctttcagtg agtcactgtg gatgttccca gctgtgtcat cccaaagttc agcagggaaa 4380

gctgagctgg gcctctccag gtgagttttc tagaagcatt tctcaactg tgggttacat 4440

caacttgggt gtcttgagct gtaaggaagg aactccggag tcagctgggc tacaggggag 4500
cttctctaag tcttgcggga ggccagaccc agcctgagct tgctgttagc tagcggaggc 4560
agctgtcgtt ggcccagggt ctcgacacca ggcatccctt ctcctccac gaagggtgtg 4620
ccataatccc cttaacagg aaatgcttcc cagaagcctc tcagcagcct cccctcctgt 4680
cctatcagct agaagcgctt cgcttgctcc aagaccagca gggacaggga actgtccgag 4740
cccgtggctg tgtggaggaa ggcgaccccc agcacaagat tggtttcctt tgggaaggga 4800
agagggagtg tgttggggta aggggtagag cagaggaatg gtcagggggc aacaaccgct 4860

gacagctgca acaggtgcat ggcatctcac agggaggcag ggaggtgcga gctcctaagt 4920
aatggagcaa aaaaattcta ttctgttaga tggggagaga aaatgtgaca ttttaatttt 4980
tttttgattt tatattccta attcctactt aaagtgaata tactgccgct gtagatcata 5040
aaatgtatct ttccatggc caacaagggg catcttttat aaatgcataa taaccagttt 5100
tgtatcaaag ggatcgcact taagtgaat ttcaacatgc tgttactttt tccttttaat 5160
gtaattctgt ttccaaataaat ggggggag acaaatggac tttagtaaa ttcttagca 5220
tatcaaa 5227

<210> 245

<211> 5081

<212> DNA

<213> Homo sapiens

<400> 245

gtctgactcc agccaccggc cttaaggca cggcttttta ttcttcggc tggtcggcct 60
ctcgcccttc agctacctgt gcgtccctcc gtcccgtccc gtcccggggt caccgccgag 120
cctgtccgct atcggtctcc tgctcttagc cccaggtcgg ctccggcggg gcagcccccg 180
ccacctgccc tctgcagcc cagcgtgct actgctggtg ctgggcggct gcctgggggt 240
cttcgggggt gctgcgggaa cccggaggcc caacgtggtg ctgctctca cggacgacca 300
ggacgaagtg ctcgcgga tgacaccgct aaagaaaacc aaagctctca tcggagagat 360

gggatgact tttccagtg cttatgtgcc aagtgtctc tgctgcccc gcagagccag 420
tactctgaca ggaaagtacc cacataatca tcactgttg aacaacactc tggaggggaa 480
ctgcagtagt aagtcttggc agaagatcca agaaccaaat actttccag caattctcag 540
atcaatgtgt ggtatcaga cttttttgc agggaaatat ttaaatgagt acggagcccc 600
agatgcaggt ggactagaac acgttcctct gggttgaggt tactggtatg ctttgaaaa 660
gaattctaag tattataatt acaccctgtc tatcaatggg aaggcacgga agcatggtga 720

aaactatagt gtggactacc tgacagatgt tttggctaata gtctccttgg actttctgga	780
ctacaagtcc aactttgagc ctttcttcat gatgatcgcc actccagcgc ctcatcgcgc	840
ttggacagct gcacctcagt accagaaggc tttccagaat gtctttgcac caagaaacaa	900
gaacttcaac atccatggaa cgaacaagca ctggttaatt aggcaagcca agactccaat	960
gactaattct tcaatacagt ttttagataa tgcatttagg aaaaggtggc aaactctcct	1020
ctcagttgat gaccttgtgg agaaactggt caagaggctg gagttcactg gggagctcaa	1080
caacacttac atcttctata cctcagacaa tggctatcac acaggacagt tttccttgcc	1140
aatagacaag agacagctgt atgagtttga tatcaaagtt ccactgttgg ttcgaggacc	1200
tgggatcaaa ccaaatacaga caagcaagat gctggttgcc aacattgact tgggtcctac	1260
tattttggac attgctggct acgacctaaa taagacacag atggatggga tgccttatt	1320
gccccatttg agaggtgcca gtaacttgac ctggcgatca gatgtcctgg tggaaatacca	1380
aggagaaggc cgtaacgta ctgacccaac atgcccttcc ctgagtcctg gcgtatctca	1440
atgcttccca gactgtgtat gtgaagatgc ttataacaat acctatgcct gtgtgaggac	1500
aatgtcagca ttgtggaatt tgcagtattg cgagttttagt gaccaggagg tgttttaga	1560
agtctataat ctgactgcag acccagacca gatcactaac attgctaaaa ccatagaccc	1620
agagctttta ggaaagatga actatcggtt aatgatgtta cagtctgtt ctgggccaac	1680
ctgtcgcaact ccaggggttt ttgaccccggt atacaggttt gaccccgctc tcatgttcag	1740
caatcgcggc agtgtcagga ctggaagatt ttccaaacat cttctgtagc gacctcacac	1800
agcctctgca gatggatccc tgcacgcctc tttctgatga agtgattgta gtaggtgtct	1860
gtagctagtc ttcaagacca cacctggaag agtttctggg ctggctttaa gtcctgtttg	1920
aaaaagcaac ccagtcagct gacttctcgt tgcaatgtgt taaactgtga actctgcca	1980
tgtgtcagga gtggctgtct ctggctctct ccttttagctg acaaggacac tctgaggtc	2040
tttgttctca ctgtattctt tttatcctgg ggccacagtt cttgattatt cctcttgtgg	2100
ttaaagactg aatttgtaaa cccattcaga taaatggcag tacttttagga cacacacaaa	2160
cacacagaca caccttttga tatgtaaagt tgacctaaag tcaaaggacc tgtgtagcat	2220
ttcagattga gcacttcaat atcaaaaata ctaacatcac atggcttgaa gagtaacat	2280
cagagctgaa tcatccaagt aagaacaagt accattgttg attgataagt agagatacat	2340
tttttatgat gticacaca gtgtggtaag gttgcaaatt caaaacatgt cacccaagct	2400
ctgttcatgt ttttgtgaat tctaggctgg tgctgcactg aaatagagca gtaagcttgt	2460

gataaaggcc aattccaggt agctcttgaa ggtgatagcc atctactttc cagtggctgc	2520
caaccacagg gagtgccagt taacactgga aggattaagg caaggtccct tctcttgaga	2580
ctccctctg agatctgaaa aatgaagtgg cttaggaaca tcagcagtga agaactgcca	2640
agagttgggt aaggttgtct ctccgaggg cttctgaag acagggtct tgaacagaca	2700
agtgaaggg ctgtaccagg gataaaggaa agaagtgcct gtccagcagg gagcttgaat	2760
ttaagtcca tgtatgaagt cattggctct atctgcattt ttctgtcatt ctcttcattt	2820
gttttaaggt ggaaaatttt cttacagttg atgcaaagta tcaactactt taccctacct	2880
tctcccttt tagatgggtt ctctctgagt ttggagttct tgtatgatta tcagtattcc	2940
cctgtcaaaa tcaaatctat tcaggtttct tctactgtga gaacacctaa atgtttttat	3000
ttttgagaag tggggacaga gtctcactat gtcaccagg ctggagtgc atggcatgat	3060
ctcagctcac tgcaaccttc gcctctggg ttcaagcgat tctctgcct ccgcctctg	3120
agtagctggg attataggca cgcaccacca cgcccagcta atttttgta ttttagtag	3180
agacagagtt tcaccatgtt ggccaggctg gtcttgaact cctgaccttg tgatccacc	3240
acctcgccct ccagagtgct tgggattaca ggcatgagcc accacgcttg gctaagaaca	3300
cctaaatttt taigtcttct ggctcaaaaa ccagttccat ttctaattgt gtctcaca	3360
gaaggcta at tgggtgtgag acagcagggg aggaggaaga gctgtggtt gtaacttgtt	3420
caactcaggc aataagcgat tttagcttta tttaaagtct tctgtccagc ttttaagcact	3480
ttgtaagaca tggctgaaag tagcttttct atcagaattg cagatagtca tgttgggcta	3540
acagtcaatt ggatatatt ctttacctca catgacccca gcaactgtgg tggatatctag	3600
aggtgaaaca ggcaagtga atggacacct ctgctgtgaa tgttttagag aaggaaattc	3660
aaaaaatgtt gtaactgaaa gcactgttga atatgggtat cggttttctt tttcactttg	3720
actcttaaca ttatcagtea acttccacat taatgaaagt tgaccatagt tatttccaaa	3780
taaaaagaaa ccaactctta ccaggtcttg gactgtgatg tcatattatt cagttttatg	3840
ctgttctctg agcagaactc ataagagtga catagtcagc tgctgacgc acctcagcca	3900
cgccactctt actcagttca gtgggtgtgc ttgcgtggta ggatgtggg cagccctctc	3960
tacgtcttc tatttttgg atatttctta tctaaccttc aaatagcttc caattctttt	4020
ttcttggac tggcttcatt ctgaatttgt gctaaaataa tctttcataa agagacctca	4080
gtttatagcg taacagacta cacaatgcac tgatgttttc ataatgttta agggaccac	4140
tgcaagaagc ttgctgcctc cttttaattg tattcattta gattttgatt ttccatgtta	4200
agaaggtgag gtccatgttg gtgcccttca gtagtagaaa ccatgtaaac attaggaatg	4260
aacagaggcc ttaggaatga atagagagtt tgccttatac aatttcctgt taaaaagctc	4320

tccctctcat gcaaagtagg gaacaccttt tgagcatctt tgaatttgac aaatggtgct 4380
gttgcaaaca cttttttttt gagatgaagi ctgcggttg tcacccgggc tggagtgcag 4440
tggcgtgac tcggctcact gcaacttcca cctcctgggt tccagcagtt ctcctgcctc 4500
agcctcccaa gtagctgaga ttacaggcgc ctgccacccc acctggctga tttttgtaat 4560

tttagtagag acggggtttc accatgttgg ccaggctgat taactcctga cctcaggatga 4620
tccacctttc tcggcctccc aaagtgcctgg gattacgggt gtgagccacc gtgcccggcc 4680
tgcaaacaca ttttaattga caacactagg gctgtttgac aaaatagtaa tgatagccat 4740
ggaagtttta cttattctg tgagaagtgt tcttaactt attaatgttc taaactaagg 4800
tttagtgctt ttttaaagga aagtgtccc aggattcatc ctaaagaaag caaaagttaa 4860
ttcaactgat ccaccaatgg aattagatgg gtagagttgg gttcttgagt tttaccacca 4920
cttagttccc actgaatttt gtaacttct gtgtttgcat cctctgttcc tattctgccc 4980

ttgctctgtg tcatctcagt catttgactt agaaagtgcc cttcaaaagg accctgttca 5040
ctgctgcact tttcaatgaa ttaaaattta tttctgttct a 5081

<210> 246
<211> 2344
<212> DNA
<213> Homo sapiens
<400> 246

agccggaagg gccggcggac gctcgtcagg tcggctcgt ggccggggct ccgcggctcc 60
cgtggttgcc atggcggcgg ttgtcgcggc gacgaggtgg tggcagctgt tgctggtgct 120
cagcgccgcg gggatggggc cctcgggcgc ccgcagccc ccaacatcc tgctcctgct 180
catggacgac atgggatggg gtgacctcgg ggtgtatgga gagccctcca gagagacccc 240

gaatttgac cggatggctg cagaagggt gcttttcca aacttctatt ctgccaaccc 300
tctgtgctcg ccatcgaggg cggcactgct cacaggacgg ctacccatcc gcaatggctt 360
ctacaccacc aacgcccatt ccagaaacgc ctacacaccg caggagattg tggcgggcat 420
cccagactcg gagcagctcc tgccggagct tctgaagaag gccggctacg tcagcaagat 480
tgtcggaag tggcatctgg gtcacaggcc ccagttccac cccctgaagc acggatttga 540
tgagtgggtt ggatccccc actgccactt tggaccttat gacaacaagg ccaggcccaa 600
catccctgtg tacagggact gggagatggt tggcagatat tatgaagaat ttcctattaa 660

tctgaagacg ggggaagcca acctcaccca gatctacctg caggaagccc tggacttcat 720
taagagacag gcacggcacc accccttttt cctctactgg gctgtcgacg ccacgcacgc 780

accgcgtctat gcctccaaac ctttcttggg caccagtcag cgagggcggt atggagacgc 840
 cgtccgggag attgatgaca gcattgggaa gatactggag ctctctcaag acctgcacgt 900
 cgcggacaac accttcgtct tcttcacgtc ggacaacggc gctgccctca tttccgcccc 960
 cgaacaaggt ggagcaacg gcccttttct gtgtgggaag cagaccacgt ttgaaggagg 1020
 gatgaggag cctgccctcg catggtggcc agggcacgtc actgcaggcc aggtgagcca 1080

ccagctgggc agcatcatgg acctcttcac caccagcctg gcccttgcgg gcctgacgcc 1140
 gccacgcgac agggccattg atggcctcaa ctctctcccc acctcctgc agggccggct 1200
 gatggacagg cctatcttct attaccgtgg cgacacgtg atggcggcca cctcgggca 1260
 gcacaaggct cacttctgga cctggaccaa ctctgggag aacttcagac agggcattga 1320
 tttctgccct gggcagaacg tttcaggggt cacaactcac aatctggaag accacacgaa 1380
 gctgccccct atcttcacc tgggacggga cccaggggag aggttcccc tcagctttgc 1440
 cagcgccgag taccaggagg ccctcagcag gatcacctcg gtcgtccagc agcaccagga 1500

ggccttggtc cccgcgcagc cccagctcaa cgtgtgcaac tgggcggtca tgaactgggc 1560
 acctccgggc tgtgaaaagt tagggaagtg tctgacacct ccagaatcca ttcccaagaa 1620
 gtgcctctgg tcccactagc acctgcgcag actcaggcca ggcctagaat ctccggttgg 1680
 ccctgcaagt gcctggagga aggatggctc tggcctcggt cctccccaa cctgccccaa 1740
 gccagacaga cagcacctgc agacgcaggg ggactgcaca attccacctg cccaggacct 1800
 gacctggcg tgtgcttggc cctctctctc gccacggcg cctcagattt caggacctc 1860
 ctctcgccc acggcgctc agacctcagg acctgcccgt ctacgcctt tgtgaacccc 1920

aaatatctga gaccagtctc agtttatttt gccaaagtta aggatgcacc tgtgacagcc 1980
 tcaggaggtc ctgacaacg gtgcctgagg tgctgggga tacagtttgc ctttatacat 2040
 cttagggaga cacaagatca gtatgtgtat ggcttacatt ggttcagtca gccttccact 2100
 gaatacacga ttgagtctgg cccagtgaat ccgcattttt atgtaaacag taagggaacg 2160
 gggcaatcat ataagcgttt gtctcagggg agccccagag ggatgacttc cagttccgtc 2220
 tgtcctttgt ccacaaggaa tttccctgga cgctaattat gagggaggcg ttagctttct 2280
 tatcattgta actatgttat ttagaaataa aacgggaggc aggtttgcct aattcccagc 2340

ttga 2344

<210> 247

<211> 4852

<212> DNA

<213> Homo sapiens

<400> 247

attctatcag cggtacaagg ggctgggtggc gccacaggcg ctgggaccgc gggcggacaa	60
ggatgggtcc gcgcggcgcg gcgagcttgc cccgaggccc cggacctcgg cggctgctcc	120
tccccgtcgt cctccccgtg ctgctgctgc tgttgttggc gccgccgggc tcgggcgccg	180
gggccagccg gccccccc ac ttggtcttct tgctggcaga cgacctaggc tggaacgacg	240
tcggcttcca cggtccccgc atccgcacgc cgcacctgga cgcgctggcg gccggcgggg	300
tgctcctgga caactactac acgcagccgc tgtgcacgcc gtcgcggagc cagctgctca	360
ctggccgcta ccagatccgt acaggtttac agcaccaaat aatctggccc tgtcagccca	420
gctgtgttcc tctggatgaa aaactcctgc cccagctcct aaaagaagca ggttatacta	480
cccatatggt cggaaaatgg cacctgggaa tgtaccggaa agaatgcctt ccaaccgcc	540
gaggatttga tactactttt ggatatctcc tgggtagtga agattattat tcccatgaac	600
gctgtacatt aattgacgct ctgaatgtca cacgatgtgc tcttgatttt cgagatggcg	660
aagaagtgc aacaggatat aaaaatatgt attcaacaaa catattcacc aaaagggcta	720
tagccctcat aactaacat ccaccagaga agcctctgtt tctctacctt gctctccagt	780
ctgtgcatga gccccttcag gtccctgagg aatacttgaa gccatatgac tttatccaag	840
acaagaacag gcatcactat gcaggaatgg tgtcccttat ggatgaagca gtaggaatg	900
tactgcagc tttaaaaage agtgggctct ggaacaacac ggtgttcac tttctacag	960
ataacggagg gcagactttg gcagggggta ataactggcc ccttcgagga agaaaatgga	1020
gcctgtggga aggaggcgtc cgaggggtgg gctttgtggc aagccccttg ctgaagcaga	1080
agggcgtgaa gaaccgggag ctcatccaca tctctgactg gctgccaaca ctctgaagc	1140
tgccaggggg acacaccaat ggcacaaagc ctctggatgg ctctgacgtg tggaaaacca	1200
tcagtgaagg aagcccatcc cccagaattg agctgctgca taatattgac ccgaacttcg	1260
tggactcttc accgtgtccc aggaacagca tggtccagc aaaggatgac tcttctcttc	1320
cagaatattc agcctttaac acatctgtcc atgctgcaat tagacatgga aattggaaac	1380
tcctcacggg ctacceaggc tgtggttact ggttccctcc accgtctcaa tacaatgttt	1440
ctgagatacc ctcatcagac ccaccaacca agaccctctg gctctttgat attgatcggg	1500
accctgaaga aagacatgac ctgtccagag aatatcctca catcgtcaca aagctcctgt	1560
cccgctaca gtctacat aaacactcag tccccgtgta ctccctgca caggaccccc	1620
gctgtgatcc caaggccact ggggtgtggg gcccttggat gtaggatttc agggaggcta	1680
gaaaaccttt caatiggaag ttggacctca ggccttttct cagactctt gtctcatittg	1740
ttatcccaac ctgggttcac ttggcccttc tcttgctctt aaaccacacc gaggtgtcta	1800

atttcaaccc ctaatgcatt taagaagctg ataaaatctg caacactcct gctgttggct	1860
ggagcatgtg tctagagggtg ggggtggctg ggtttatccc cctttcctaa gccttgggac	1920
agctgggaac ttaacttgaa ataggaagt ttcactgaat cctggaggct ggaacagctg	1980
gctcttttag acitcacaagt cagacgttcg attcccctct gccaatagcc agttttattg	2040
gagtgaatca cattttctac gcaaatgaag ggagcagaca gtgattaatg gttctgttgg	2100
ccaaggcttc tcctgtcgg tgaaggatca tgttcaggca ctccaagtga accaccctc	2160
ttggttcacc ccttactcac ttatctcatc acagagcata aggcccat tttgtgttcag	2220
gtcaacagca aaatgcctgc accatgactg tggcttttaa aataaagaaa tgtgttttta	2280
tcgtaattta tttccccca gccattgtc actctgtcta gacttcctgc cacttccaat	2340
tcttctgtgg cttttcctgc ctttcctttt gacctcagta gtcctatccc tgggaaggcc	2400
actttgcttc tctacctgag caccctgat ttttggaacg ctgctgagcc ctgccttact	2460
tttggcccta gggctgaagc tagaggcctc cccgtaatag gcggtggagt tgctctgtga	2520
ggatgttcat ggtagacact aagagggctg ggtgggagat gcttggctct gtggcatctg	2580
ttcagcgagg cttttcctat attgcatgga gttagtcat tgtattgtag ctttatttca	2640
taatatatta agacttgac tgctatttac tagcagttag aagaaacctc aggaaaggat	2700
atgaaaaagc aagtggccag tgtctgggat actgggcctt ggtaaagcag aggagggcac	2760
accacagtc ctctattct ctgttttact gcttgttttg aggttctggg gtctggcaaa	2820
gaggatgcag ttgacacct gcagcccttt ctcaatccca ctaatgtctt actaatgttg	2880
aacagtcct attagctcca gagagtgtca aaccagaga aatgtgtgca aaaatgatac	2940
tcttttctgc attagcccca ccatttgttt caccaatgct tggaaactg cctgaaggca	3000
ctcatttttt aatttttatt ttatttttaa ttttttatat ctttatgaga cgatctcact	3060
ctgtcaccag gttggagtac agtggtaaaa tcacaactca ccgtagcctc aaactcctgg	3120
gctcaagtga ttctcccacc tcaggcacc aaatagctgg aactacaggc atataccgcc	3180
acaccagct aattttat tttgaaaaga caaggttccc tatgttgcgc agctggtctt	3240
aaactcctgg gctccagcaa ttatcccagc ttgggctcca aaagtgtctg gattacaggc	3300
atgagtcacc atgcctggcc tcatttttta aaacaaatga ataaatggac aaatgagtaa	3360
atgagaaagt ctacaccat gaaagatgct agtccaatga gctgaataga gaggtaatat	3420
aaatgtcttc cagctgttgc ttttctgttc tcaagctgcc cctcctgggg taggagcata	3480
atctacatca ctgggcagtc acaggacact ctatagcaag gttgtagcgt cctctccagt	3540
ggggggagaa aaggaactgt gcctacaaa ggtactctct tgtcagcaat ttccatttct	3600

atactttatg ggacactaga aactaaaagc aacaaataat ctgatataag tccttgtata 3660

gtcatccttc aattcagtag caatatcttc tggctactac taacctgtat tgtattaaaa 3720

tgagactatt ggaaggaaat ggtgctaaaa ctaataacat ctcttaccaa cctttaccea 3780

actcctgggt tggcaaacag ctgaccaaac tgccatcacc tcccacttgg aagtgtatgg 3840

ccgacagcat gaaatagctg agcccagatg ttccttctgc atcctccgaa tcccagggtc 3900

gggtgtaggt agccgttggg ggccatcgtc acagggcacc tatctgttat cgctgctgtc 3960

ctcccaacag ctgtctccag ttctagtctc ttggttttca ggacacagtgg gggatgttct 4020

gcaccacgtg gacttcaaaa gagttttgaa gacttaattt ttgttaaac aagtacttga 4080

gattttggtt taiccataat agaattgtatt tcattagatt ctctgattct atataagaat 4140

gtgaaaagat tgatataattg ttgttagaaa taatgttatt tctttccaat tttttttttt 4200

tttttttttg agatggagtc tcgctctgtc acccaggctg gagtgcagtg gtgtgatctc 4260

ggctcactgc agcctctaac tcccagggtc aagctattct cctgcctcag cctccaagt 4320

agctggatta caggcataca ccaccacgcc tggtatgtt ttgtattttt cgtagagata 4380

gggtttcacc atgttggcca ggctgggtctc aaactcctga cctcaagtga tccaccact 4440

tcagcttccc aaagcactgg gattacaggt gtgagccact gtgcccggca aattttttta 4500

cctttacaga aggttttgc tatttaattg tgagctcatt tttctttgtt acttttgtcc 4560

ccccagattt gggggacaaa ataaaattaa tcttttaaaa tgtgtcagcc atatgtatgg 4620

ggcttcatt tggggtgagg agaaagtctt ggaactagat agtggtcatt gttatacaac 4680

atcataaatg caattactgc cactgaattg tatgttttaa agtggttaaa atgttaagtt 4740

ttatgtttta ttacaatttt taaatgtgtc aaccaacttt ataglacata aattatatct 4800

cagtaaagct gttaataaaa taaatatagt aaaaatttta gaactaaaaa aa 4852

<210> 248

<211> 2199

<212> DNA

<213> Homo sapiens

<400> 248

agacgtggc cgagcggggg accggaagc atggcccggg ggtcggcggt tgcctgggcg 60

gcgctcgggc cgttgttgtg gggctgcgcg ctgggctgc agggcgggat gctgtacccc 120

caggagagcc cgtcgcggga gtgcaaggag ctggacggcc tctggagctt ccgcgccgac 180

ttctctgaca accgacgccg gggcttcgag gagcagtggt accggcggcc gctgtgggag 240

tcaggcccca ccgtggacat gccagttccc tccagttca atgacatcag ccaggactgg 300

cgtctgcggc attttgtcgg ctgggtgtgg tacgaacggg aggtgatcct gccggagcga	360
tggaccagg accctgcgcac aagagtgggtg ctgaggattg gcagtgccca ttctatgcc	420
atcgtgtggg tgaatggggt cgacacgcta gagcatgagg ggggctacct ccccttcgag	480
gccgacatca gcaacctgggt ccaggtgggg ccctgcct cccggctccg aatcactatc	540
gccatcaaca acacactcac cccaccacc ctgccaccag ggaccatcca atacctgact	600
gacacctcca agtateccaa gggttacttt gtccagaaca cataatgtga cttttcaac	660
tacgtggac tgcagcggtc tgtacttctg tacacgacac ccaccaccta catcgatgac	720
atcaccgtca ccaccagcgt ggagcaagac agtgggctgg tgaattacca gatctctgtc	780
aagggcagta acctgttcaa gttggaagtg cgtcttttgg atgcagaaaa caaagtcgtg	840
gcgaatggga ctgggaccca gggccaactt aaggtgccag gtgtcagcct ctggtggccg	900
tacctgatgc acgaacgcc tgcctatctg tattcattgg aggtgcagct gactgcacag	960
acgtcactgg ggctgtgtc tgacttctac acactccctg tggggatccg cactgtggct	1020
gtaccaaga gccagttcct catcaatggg aaacctttct attccacgg tgtcaacaag	1080
catgaggatg cggacatccg aggggaagggc ttgcactggc cgctgctggt gaaggacttc	1140
aacctgcttc gctggcttgg tgccaacgct ttccgtacca gccactacc ctatgcagag	1200
gaagtgatgc agatgtgtga ccgctatggg attgttgtca tcgatgagt tcccggcgtg	1260
ggcctggcgc tgccgcagtt cttcaacaac gtttctctgc atcaccacat gcaggatgatg	1320
gaagaagtgg tgcgtaggga caagaaccac ccgcggtcg tgatgtgttc tgtggccaac	1380
gagcctgctg cccacctaga atctgctggc tactacttga agatggatgat cgctcacacc	1440
aatccttgg acccctccc gcctgtgacc tttgtgagca actctaacta tgcagcagac	1500
aagggggctc cgtatgtgga tgtgatctgt ttgaacagct actactcttg gtatcacgac	1560
tacgggcacc tggagttgat tcagctgcag ctggccacc agtttgagaa ctggtataag	1620
aagtatcaga agccattat tcagagcgag tatggagcag aaacgattgc agggtttcac	1680
caggatccac ctctgatgtt cactgaagag taccagaaaa gtctgctaga gcagtacat	1740
ctgggtctgg atcaaaaacg cagaaaatac gtggttgag agctcatttg gaattttgcc	1800
gatttcatga ctgaacagtc accgacgaga gtgctgggga ataaaaagg gatcttact	1860
cggcagagac aacaaaaag tgcagcgttc cttttcgag agagatactg gaagattgcc	1920
aatgaaacca ggiatcccca ctcagtagcc aagtcacaat gtttgaaaa cagcctgttt	1980
acttagcaa gactgatacc acctgcgtgt ccttcctcc ccgagtcagg gcgacttcca	2040

cagcagcaga acaagtgcct cctggactgt tcacggcaga ccagaacgtt tctggcctgg 2100

gttttgtggt catctattct agcaggggaac actaaagggtg gaaataaaaag attttctatt 2160

atggaaataa agagttagca tgaagtggc tactgaaaa 2199

<210> 249

<211> 3185

<212> DNA

<213> Homo sapiens

<400> 249

ctttccaggg ccggggaacc ccaggaggaa gctgctgagc catgggcgcc tacgcgcggg 60

cttcgggggt ctgcgctcgc ggctgcctgg actcagcagg cccctggacc atgtcccgcg 120

ccctgcggcc accgctcccg cctctctgct ttttctttt gttgctggcg gctgccgggtg 180

ctcgggccgg gggatacagc acatgccccca cagtgcagcc gaacatgctg aacgtgcacc 240

tgctgcctca cacacatgat gacgtgggct ggctcaaaac cgtggaccag tacttttatg 300

gaatcaagaa tgacatccag cagcccggtg tgcagtacat cctggactcg gtcactctcg 360

ccttgctggc agatcccacc cgtcgttca ttacgtgga gattgccttc ttctcccggt 420

ggtggcacca gcagacaaat gccacacagg aagtcgtgcg agaccttggt cgccaggggc 480

gcctggagtt cgccaatggt ggctgggtga tgaacgatga ggacagccacc cactacgggtg 540

ccatcgtgga ccagatgaca cttgggctgc gctttctgga ggacacattt ggcaatgatg 600

ggcgaccccg tgtggcctgg cacattgacc ctttcggcca ctctcgggag caggcctcgc 660

tgtttgcgca gatgggcttc gacggcttct tctttggcg ccttgattat caagataagt 720

gggtacggat gcagaagctg gagatggagc aggtgtggcg ggccagcacc agcctgaagc 780

ccccgaccgc ggacctcttc actggtgtgc ttccaatgg ttacaaccgc ccaaggaatc 840

tgtgctggga tgtgctgtgt gtcgatcagc cgctgggtga ggaccctcgc agccccgagt 900

acaacgcaa ggagctggtc gattacttcc taaatgtggc cactgcccag ggccggtatt 960

accgcaccaa ccacactgtg atgaccatgg gctcggactt ccaatatgag aatgccaaca 1020

tgtggttcaa gaaccttgac aagctcatcc ggctggtaaa tgcgcagcag gcaaaaggaa 1080

gcagtgtcca tgttctctac tccacccccg ctgtttacct ctgggagctg aacaaggcca 1140

acctcacctg gtcagtgaat catgacgact tcttcctta cgcggatggc cccaccagt 1200

tctggaccgg ttacttttcc agtcggccgg cctcaaacg ctacgagcgc ctacgtaca 1260

acttctgca ggtgtgcaac cagctggagg cgtgggtggg cctggcggcc aacgtgggac 1320

cctatggctc cggagacagt gcaccctca atgaggcgat ggctgtgctc cagcatcacg 1380

acgccgtcag cggcacctcc cgccagcacg tggccaacga ctacgcgcgc cagcttgcgg	1440
caggctgggg gccttgcgag gttcttctga gcaacgcgct ggcgcggtc agaggcttca	1500
aagatcactt caccttttgc caacagctaa acatcagcat ctgcccgtc agccagacgg	1560
cggcgcgctt ccaggctatc gtttataatc ccttggggcg gaaggtgaat tggatggtag	1620
ggctgccgtt cagcgaagcg gttttcgtt tgaaggaccc caatggcagg acagtgccca	1680
gcgatgtggt aatatttccc agctcagaca gccaggcgca ccctccggag ctgctgttct	1740
cagcctcact gcccgcctg ggcttcagca cctattcagt agcccagggt cctcgctgga	1800
agccccaggc ccgcgcacca cagcccatcc ccagaagatc ctggtcacct gctttaacca	1860
tcgaaaatga gcacatccgg gcaacgtttg atcctgacac agggctgttg atggagatta	1920
tgaacatgaa tcagcaactc ctgctgcctg ttgccagac cttcttctgg tacaacgcca	1980
gtatagtgta caacgaaagt gaccaggcct cagggtgccta catcttcaga cccaaccaac	2040
agaaaccgct gcctgtgagc cgctgggctc agatccacct ggtgaagaca cccttggtgc	2100
aggaggtgca ccagaacttc tcagcttggg gtgccagggt ggttcgcctg taccaggac	2160
agcggcacct ggagctagag tggtcggtgg ggccgatacc tgtgggcgac acctggggga	2220
aggaggtcat cagccgtttt gacacaccgc tggagacaaa gggacgcttc tacacagaca	2280
gcaatggcgg ggagatcctg gagaggaggc gggattatcg acccacctgg aaactgaacc	2340
agacggagcc cgtggcagga aactactatc cagtcaacac ccggatttac atcacggatg	2400
gaaacatgca gctgactgtg ctgactgacc gctcccaggg gggcagcagc ctgagagatg	2460
gctcgctgga gctcatggtg caccgaagcg tctgaagga cgatggacgc ggagtatcgg	2520
agccactaat ggagaacggg tcggggggcgt ggggtgcgagg gcgccacctg gtgctgctgg	2580
acacagccca ggctgcagcc gccggacacc ggctcctggc ggagcaggag gtctggccc	2640
ctcaggtggt gctggccccg ggtggcggcg ccgcctacaa tctcggggct cctccgcga	2700
cgcagttctc agggctgcgc agggacctgc cgcctcgggt gcacctgctc acgtggcca	2760
gctggggccc cgaatggtg ctgctgcgct tggagcacca gtttccgta ggagaggatt	2820
ccggacgtaa cctgagcgcc ccggttacct tgaacttgag ggacctgttc tccaccttca	2880
ccatcacccg cctgcaggag accacgtggg tggccaacca gctccgcgag gcagcctcca	2940
ggctcaagtg gacaacaaac acaggcccca caccaccca aactccgtac cagctggacc	3000
cggccaacat cagctggaa cccatgaaa tccgcacttt cctggcctca gttcaatgga	3060
aggaggtgga tggttaggtc tgctgggatg ggccctccaa gcccaagcct cctgctccgg	3120

gggcagacca gactctgact ctccctcttg ggctgctgcc attaaaacgc tactactaag 3180

actca 3185

<210> 250

<211> 4001

<212> DNA

<213> Homo sapiens

<400> 250

gagtggcttc accgccggtc ccttgacgct ctgcctttcg atctctccac atctcgggtg 60

cgcgggatct caagatgcgc ctccacctgc tctgctgct cgcgctgtgc ggtgcaggca 120

ccaccgccgc ggagctcagt tacagcttgc gtggcaactg gagcatctgc aatgggaacg 180

gctcgtctga gctccccggg gcggctcctg gctgctgca cagcgcttg ttccagcagg 240

gcctgatcca ggattcttac tacagattta atgacctaa ctacagatgg gtctcttttg 300

ataactggac ctatagcaaa gaatttaaaa tcccccttga aattagcaaa tggcaaaaag 360

taaatttgat tcttgaggga gtggatacgg ttcaaaaaat cctgttcaat gaagtcacta 420

ttggggaaac agacaatatg ttcaatagat atagctttga tattaccaac gtggtcaggg 480

acgtgaactc cattgagctg cgtttccagt cagcgggtgt gtatgcagca cagcagagca 540

aagctcacac tcgtaccag gtccccccag actgcctcc acttgtgcag aagggtgaat 600

gccatgtcaa ctttgttcgg aaggagcaat gtctcttag ttgggactgg gggccttct 660

ttctaccca ggaatctgg aaagatgtta gaattgaagc ctataatatt tgtcacctga 720

actacttcac attttccca atatatgata agagtgccca ggagtggaaat ctggaaatag 780

agtctacatt tgatgtgtc agctcaaagc cagttggtgg tcaagtgtc gtagccatcc 840

ctaagttgca aacacaacag acatacagca ttgaacttca acctgggaaa aggattgttg 900

agctatttgt gaacattagc aagaatatta ctgtagaaac ttggtggcct catggacatg 960

gaaaccagac tgggtacaac atgactgttc tttttgaact ggatggaggc ttaaatttg 1020

aaaaatcagc taaggtttat tttaggacag tggaacttat agaagagcct ataaaagggt 1080

ctcctggttt gattttctat ttcaaaatta atggatttcc catatttcta aaaggctcaa 1140

actggatecc agcagattca ttccaggacc gagtaacctc tgagtgtta cggtccttt 1200

tacagtctgt tgggatgct aatatgaata ctctcgggt ttggggagga ggaatttatg 1260

agcaggatga attctatgaa ctctgtgatg aactaggaat aatggtatgg caggatttta 1320

tgtttgcctg tgccctttat ccaactgatc agggcttctt ggattcagt acagcagaag 1380

ttgcctacca gatcaagaga ctgaaatctc atccttctat catcatatgg agtggcaata 1440

atgaaaatga ggaggcgctg atgatgaatt ggtatcatat cagtttact gaccggccaa	1500
tctacatcaa ggactatgtg acactctatg tgaaaaacat cagagagctc gtactggcag	1560
gagacaagag tcgtcctttt attacgtcca gtcctacaaa tggggctgaa actgttgag	1620
aagcctgggt ctctcaaac cctaatagca attatitttg tgatgtacat tttatgact	1680
atatcagtga ttgctggaac tggaaagttt tcccaaaagc tcgatttgca tctgaatatg	1740
gatatcagtc ctggcctgcc ttcagtacat tagaaaaggt ctcgtctaca gaggactggt	1800
ctttcaatag caagttttca cttcatcgac aacatcacga aggtggtaac aaacaaatgc	1860
tttatcaggc tggacttcat ttcaaaactcc cccaaagcac agatccatta cgcacattta	1920
aagataccat ctaccttact caggatgatgc agggccagtg tgtcaaaaca gaaactgaat	1980
tctaccgccg tagtcgcagc gagatagtgg atcagcaagg gcacacgatg ggggcacttt	2040
attggcagtt gaatgacatc tggcaagctc ctctctgggc ttctcttgag tacggaggaa	2100
agtggaaaat gcttcattac ttgctcaga atttctttgc tccactgttg ccagtaggct	2160
ttgagaatga aaacacgttc tatactatg gtgtgtcaga tcttcactcg gattattcga	2220
tgacactcag tgtgagagtc catacatgga gctccctgga gcccggtgtc tctcgtgtga	2280
ctgaacgttt tgtgatgaaa ggaggagagg ctgtctgcct ttatgaggag ccagtgtctg	2340
aattgctgag gagatgtggg aattgcacac gggaaaagctg tgtggtttcc ttttaccttt	2400
cagctgacca tgaactcctg agcccagcca actaccactt cttgtcctca ccgaaggagg	2460
ccgtggggct ctgcaaggcg cagatcactg ccatcatctc tcagcaaggt gacatatttg	2520
tttttgacct ggagacctca gctgtcgctc cctttgtttg gttggatgta ggaagcatcc	2580
caggagatt tagtgacaat ggtttcctca tgactgagaa gacacgaact atattatatt	2640
accttggga gccaccagc aagaatgagt tggagcaatc ttttcatgtg acctccttaa	2700
cagatattta ctgaaggaat ctaggttgta ttttcagtgg acaatgggaa taaagcattt	2760
ctaaagcacc gactggagag gaaggcaaca gagacaagga gagaagccga gagacatgtc	2820
tgctgtctgc cacgcatttg agcgattgct ctgtgaagag ttgtacactg aacactttca	2880
ggggaggctg ttiaccagg caatgtcctc aaacaagcct gtgccggggt gtcctggaat	2940
ctgtgccagg actgtgtttt tagcccttca cctctcagct ttagcaggac atgaaccagt	3000
tataacaaga tggccctgca gctggttaca agaattgtac atggcaggat ctatggaacc	3060
aaatggaagg ttttgaggtg atgtaggtct ttcacagcta gctttgggga atacggaata	3120
ctcaataaaa gtgctttgtt attatttcag agggaatggc gattgaaatg ttacaacaga	3180

gattttcttgg tggtagctat ttgggtaaag gtatatgggt atttttctgt acatgtgaaa 3240

ttataataaa ataaaagtta tataaattac attgacaact tgatatttca tctgtggctt 3300

atttaaccgt tctataaata ttcaattttg cccaatactg tgctggacce taatttccat 3360

ttagttctcg ttattattct gcttggtaaa ctgggagagt ttgaagaatc taaactaaga 3420

catgtgaaag gttttctctg gaaacagtca tgtaattggg tatagaatgg accagcaaag 3480

agtgtttggc tttaaaatta agatattatt agtggttttg gttttaaaaa tgagaagcat 3540

tatgttcagt gactctaaaa gtaactacag acaatccagt cttcttcagt atctaagagt 3600

atcctcattt tatatgtaca cacacacaca cacacacaca cacacacaca cacacacaca 3660

cacaaagcct tatcaagggg ataaggggac ttagggctgg gactccgtag agtcttactt 3720

ggtaaatatt atgtcatgtt cttatgacta ttgtaaatcc tgttttagca aagcattgct 3780

gcgaggcagg gaggactgct cttcttgaca cctgtgcatg ttctgccagg ggccctgaga 3840

aggggctttc aagctgatga gccagagttt actttgtagt ttgcccttac tagttctgta 3900

accttaatca agttacttaa cttcctgtgc tcagttttct tgtggctaaa tgggagtggg 3960

aacacccacc ttgtaagttt gttttgaaga ataataaaaa t 4001

<210> 251

<211> 2133

<212> DNA

<213> Homo sapiens

<400> 251

cgggccaatc gtiagtca gaaggcgag cgcgccgagg gcacctgcgc gtaagagtg 60

ggccgcgtcg ctgagggtga gcgatgcggg ctccggggat gaggtcgcgg ccggcgggtc 120

ccgcgtgtt gctgctgtg ctcttcctcg gagcgccga gtcggtgcgt cgggccagc 180

ctccgcgcg ctacaccca gactggccga gcctggattc tcggccgtg ccggcctggt 240

tcgacgaagc caagttcggg gtgttcatcc actggggcgt gttctcggtg cccgcctggg 300

gcagcgagtg gttctgggtg cactggcagg gcgaggggag gccgcagtac cagcgttca 360

tgcgcgacaa ctacccgcc ggcttcagct acgccgactt cggaccgcag ttcactgcgc 420

gcttcttcca cccggaggag tgggccgacc tcttcaggc cgcgggcgcc aagtatgtag 480

ttttgacgac aaagcatcac gaaggcttca caaactggcc gagtctgtg tcttggaact 540

ggaactcaa agacgtgggg cctcatcggt atttggttgg tgaattggga acagctctcc 600

ggaagaggaa catccgtat ggactatacc actcactctt agagtgggtc catccactct 660

atctacttga taagaaaaat ggcttcaaaa cacagcattt tgtcagtga aaaacaatgc 720

cagagctgta cgaccttggt aacagctata aacctgatct gatctggtct gatggggagt	780
gggaatgtcc tgatacttac tggaactcca caaatTTTct tTcatggctc tacaatgaca	840
gcccTgtcaa ggatgaggTg gtagtaaatg accgatgggg tcagaactgt tectgtcacc	900
atggaggata ctataactgt gaagataaat tcaagccaca gagcttgcca gatcacaagt	960
gggagatgtg caccagcatt gacaagTTTT cTgggggcta tCgtcgtgac atggcattgt	1020
ctgatgttac agaagaatct gaaatcattt cggaactggT tcagacagta agtttgggag	1080
gcaactatct tctgaacatt ggaccaacta aagatggact gattgttccc atcttccaag	1140
aaaggcttct tCgtgttggg aaatggctga gcatcaatgg ggaggctatc tatgcctcca	1200
aacctggcg ggtgcaatgg gaaaagaaca caacatctgt atgtatacc tcaaaggat	1260
cggtgtttta tgccattttt ctgcactggc cagaaaatgg agtcttaaac ctTgaatccc	1320
ccataactac ctcaactaca aagataacaa tCgtgggaat tcaaggagat ctgaagtggT	1380
ccacagatcc agataaaggT ctcttcatct ctctacccca gttgccaccc tctgtgtcc	1440
ccgcagagtt tCcttggact ataaagctga caggagtga gtaatcattt gagtgaaga	1500
agaaagaggc gctgtctact gttttctgc tTcagttttt ctcttatagt accatcacta	1560
taatcaacga acttctcttc tccaccaga gatggctttt ccaacacatt ttaattaaag	1620
gaactgagta cattaccctg atgtctaaat ggaccaaaga tctgagatcc attgtgatta	1680
tatctgtatc aggtcagcag aagaaggaac tgagcagttg aactctgagt tcatcaattg	1740
taatatTTgg aaattatcta caatggaatc tTccctctgt tctctgataa cctacttgt	1800
tactcaatgc ctttaagcca agtcaccctg tTgcctatgg gaggaggTgg aaggattTgg	1860
caagctcaac cacatgctat ttagttagca tcagttgtca ccaacagtct tTctgcaaag	1920
ggcaggagag ctttggggga aaggaaaagg ctTaccaggc tgctatggTc aactcttcag	1980
aaattttcag agcaatctaa aagcgccaaa attcgctatg tttacagtga tactattaag	2040
aaaatgaatg tgattctgt ctgtcttttt aagtatgatc aaataaaaaa tttgtacatc	2100
acaatcattt ctaccaaaaa aaaaaaaaaa aaa	2133
<210> 252	
<211> 2037	
<212> DNA	
<213> Homo sapiens	
<400> 252	
aattgttcgg cgatcgctgg ctgccgggac tttctcgcg ctggctcttt cggTggtcag	60
ggatggcgcg gaagtccaac ttgcctgtgc tTctcgtgcc gtttctgctc tgccaggccc	120

tagtgcgctg ctccagccct ctgcccttg tgcgtaacac ttggcccttt aagaatgcaa	180
ccgaagcagc gtggaggcca ttagcatctg gaggtctctgc cctggatgca gtggagagcg	240
gctgtgccat gtgtgagaga gagcagtgtg acggctctgt aggcctttgga ggaagtcctg	300
atgaacttgg agaaaccaca ctagatgcca tgatcatgga tggcactact atggatgtag	360
gagcagtagg agatctcaga cgaattaaaa atgctatttg tgtggcacgg aaagtactgg	420
aacatacaac acacacactt ttagtaggag agtcagccac cacatttgc taaagtatgg	480
ggtttatcaa tgaagactta tctaccactg cttctcaagc tcttcattca gattggcttg	540
ctcggaattg ccagccaaat tattggagga atgttatacc agatccctca aaatactgcg	600
gaccctacaa accacctggg atcttaaagc aggatatcc tatccataaa gaaacagaag	660
atgatcgtgg tcatgacact attggcatgg ttgtaatcca taagacagga catattgctg	720
ctggtacatc tacaatggg ataaaaattca aaatacatgg ccgtgtagga gactcaccaa	780
tacctggagc tggagcctat gctgacgata ctgcaggggc agccgcagcc actgggaatg	840
gtgatataat gatgcgcttc ctgccaagct accaagctgt agaatacatg agaagaggag	900
aagatccaac catagcttgc caaaaagtga tttcaagaat ccagaagcat tttccagaat	960
tctttggggc tgttatatgt gccaatgtga ctggaagtta cggctgctgt tgcaataaac	1020
tttcaacatt tactcagttt agtttcatgg ttataattc cgaaaaaat cagccaactg	1080
aggaaaaagt ggactgcatc taatccatct ttactgtcaa catctgtatt taaagaagaa	1140
agaaacaaag gctgaaaagg ctgctcactc tcatcatcta gtgttctca tgtgttctaa	1200
agtctttttg taaaataaac acaaaaataa atttatgttg aattggcact ttctatatct	1260
gaaattgtat tatctatttt tatttaacct ttcttatatt atctgaagat gaatatgtat	1320
gtggctgtat ttgtatata tttaagtttt aagtgcact tggatttctg atcagtgttg	1380
caaaaaattg ctactttgag atttatctcc acagtgatat attaaacca aaaggaatca	1440
tcatatttgg atgcttaaaa tagagtatac aagtttttta gacccaaaga aaaaaacagc	1500
agagtctcag catgatcagt cctgattgaa ttctattatg tcacagtgat gcatgtctta	1560
gcactgatta ttggttttac ttaatatcc ctaaaaactt gaggaaggaa gatgatagaa	1620
catgagatac tctttgtttt gtacatacta gaaaatacag ctatcttctg atgagcaatt	1680
gtgttgattt tctttgcttt ttctctctta tggagctctc agtgtgtgaa agttctaaag	1740
ccacaagaat caatgatgtt ttagctgcta atatttagta ttgatcgttt cttttaaaaa	1800
tgatttgggg aacttgatct gtactgtaat ttactaatt ccttgtgcag aggtgggtat	1860
ggcaattaga agtcaagaaa tgggttggtt agcttggttt gtttcctaa cttctgatta	1920

actctctgta tgacaactct acagaagttg tgcgcgtgct ttctcagcag catttttcct	1980
tcaaaatcat ctttttaatc aacagtcatt aataaatgta catgtacgat attcaaa	2037
<210> 253	
<211> 3353	
<212> DNA	
<213> Homo sapiens	
<400> 253	
ctctgtggag tctagctgcc agggctgcgg cagctgcggg gagagatgac tggggagcga	60
cccagcacgg cgtccccga cagacgctgg gggccgcgga ttctgggctt ctggggaggc	120
tgtagggttt ggggtgtttgc cgcgatcttc ctgctgctgt ctctggcagc ctctgtgtcc	180
aaggctgaga acgacttcgg tctggtgcag ccgtggtga ccatggagca actgctgtgg	240
gtgagcggga gacagatcgg ctgagtgga accttccgca tcccgtcat cacagccact	300
ccgcggggca ctcttctgc ctttctgag gcgaggaaaa tgtctctatc cgatgagggg	360
gccaaagtta tcgcccctgc gaggtccatg gaccagggca gcacatggtc tcctacagcg	420
ttcattgtca atgatggga tgtccccgat gggctgaacc ttggggcagt agtgagcgat	480
gttgagacag gtagtagtatt tcttttctac tccctttgtg ctcaaaaggc cggctgccag	540
gtggcctcta ccatgttgg atggagcaag gatgatggtg tttcttgag cacaccccg	600
aatctctccc tggatatgg cactgaagtg tttgccctg gaccgggctc tggattcag	660
aaacagcggg agccacggaa gggccgcctc atcgtgtgtg gccatgggac gctggagcgg	720
gacggagtct tcgtctcct cagcgatgat catggtgcct cctggcgcta cggaagtggg	780
gtcagcggca tcccctacgg tcagcccaag caggaaaatg atttcaatcc tgatgaatgc	840
cagccctatg agctcccaga tggctcagtc gtcacatg cccgaaacca gaacaactac	900
cactgccact gccgaattgt cctccgcagc tatgatgcct gtgatacact aaggccccgt	960
gatgtgacct tcgacctga gctcgtggac cctgtggtag ctgcaggagc ttagtcacc	1020
agctccggca ttgtcttctt ctccaacca gcacatccag agttccgagt gaacctgacc	1080
ctgcgatgga gcttcagcaa tggtagctca tggcggaaag agacagtcca gctatggcca	1140
ggccccagtg gctattcacc cctggcaacc ctggagggca gcatggatgg agaggagcag	1200
gccccccagc tctacgtcct gtatgagaaa ggccggaacc actacacaga gagcatctcc	1260
gtggccaaaa tcagtgtcta tgggacactc tgagctgtgc cactgccaca ggggtattct	1320
gccttcagga ctctgccttc aggaacacgg gtctgtagag ggtctgctgg agacgcctga	1380
aagacagttc catcttctt tagactccag ccttgcaaaa atcaccttc ctttaccagg	1440

gaaatcactt cctttaggac tgaagctag gcgtcctctc ccacaaaaa gtcctgccct	1500
catctgagaa tactgtcttt ccatatggct aagtgtggcc ccaccacct ctctgccctc	1560
ccgggacatt gattggctct gtcttgggca ggtctagtga gctgtagaat tgaatcaatg	1620
tgaactcagg gaactgggga aggctgagcc tcctcttttg tgttgcgga agataaccga	1680
cagggtggt gaaagtcccc agatggcagg atatttggtt tcagagtaag gactagggtc	1740
accacatga ctgactatca atcaaatgt ttgtaactta aaattttta tgaaggataa	1800
tgaatatattg tagagtctct atggttctgt caatgcacat cttcgtgtct gttttcctca	1860
tgtatccttg tgagcctggg tgagtcttgg ggagagacct gatgtgcga ctgcctgtga	1920
aaatctgact ttggcaaatc aaatcctctt ttccttttga catgccctct tttttgttg	1980
ttgctttttt tgagacaggg ctgcgtctgt caccaggct ggagtcagt tgcacaatca	2040
cggctcactg aagcctcaac ctctgggct caagtgatcc tcacgtctca gcctccggag	2100
tagttgggac tacaggctcag tgacacatg cctggtaaat tttttaatt tttattttca	2160
gtagagacaa ggttgcgta tgttgccag gctggatgg aactcctgtg ctttaagcaat	2220
cctcatgctc cagcttccca aagtctgag gttacagcta tgagccaccg caccagcct	2280
acattccttc ttatcacga gaaacaggtt gatcttcaca ggtglaatga gtatgaaggg	2340
agtgccataa gatatttttt attttttatt tttttttt ttaatttaatt ttttttttt	2400
ttgagatgga gtcttgctct ggcacccagg ctagagtga gtggtgcgat ctcggtcac	2460
tgcaacctct gcctcccagg ttcaagcat tcttctgcct cagcttcccg aatagctggg	2520
attacaggtg cccaccacca caccggcta attttgtat ttttagtaga gacggagttt	2580
cacatgttg gccaggctgg tctcaaactc ctggcctcag gtgatccacc cgcctcggcc	2640
tcccaaagtg ttgtgattac aagcatgagc catggtgccg gcgggctgat ttttttaatt	2700
tttagtagag acaaagtctc actatattgc ctaggttggt ctcaaactgc tgagctcaag	2760
caatctgccg gccatgttct ctcaaagtgc tgggattaca ggcttgagcc cctgcacca	2820
gcctgtaaga ttttttaaaa tccacacttg gccaaagctg gtggtcaca catataatcc	2880
cagcactttg ggaggccaag gtggcgcat cacaaggtca ggagttcgag accagcctgg	2940
ccaatatggt gaaatcccat ctctactaaa aatacaaaaa ctagccaggc gtgttggtt	3000
acacttatag tccagctac tcaggtggct gaggcaggag aatcattga accaggaggc	3060
ggagtttgca gtgagctgag atcacagcac tgcacaccag cctcggaac agagtgagac	3120
tccgtctgaa aaaaaaaaaa tccacgtgtt tttatatattc taacatgggg taggatacca	3180

tcctccatga aataggggtgc aagtgtgtga aatgcagatg ggaatcctgt cccagctagc	3240
ctgacagaga agtgagcact tcctatggca aattatgtaa gaaaatattc cccaggcaga	3300
tgcttatgaa aagaatctgt tcagcctgca ggtaagtggg agagcccttg ctc	3353
<210> 254	
<211> 1858	
<212> DNA	
<213> Homo sapiens	
<400> 254	
agtccccggg cgctcctgg agagcaagga cgcgggggag cagagatgat ccgagccgcg	60
ccgccgccgc tgctcctgct gctgctgctg ctgctgctgc tagtgtcctg ggcgtcccga	120
ggcgaggcag cccccacca ggacgagatc cagcgctcc cgggctggc caagcagccg	180
tccttcgcc agtactccg ctacctaaa ggtccggct ccaagcact ccactactgg	240
tttgtggagt cccagaagga tcccagaaac agccctgtgg tgccttggct caatgggggt	300
cccggtgca gctactaga tgggtctctc acagagcatg gcccttctt ggtccagcca	360
gatggtgtca ccttgagta caaccctat tcttgaatc tgattgcaa tgtgttatac	420
ctggagtccc cagctggggg gggcttctcc tactccgatg acaagtttta tgcaactaat	480
gacactgagg tgcgccagag caattttgag gcccttcaag atttcttccg cctcttccg	540
gagtacaaga acaacaaact tttctgacc ggggagagct atgctggcat ctacatcccc	600
acctggccg tgctggtcat gcaggatccc agcatgaacc ttcaggggct ggctgtgggc	660
aatggactct cctcctatga gcagaatgac aactccctgg tctactttgc ctactacat	720
ggccttctgg ggaacaggt tttgtcttct ctccagacc actgctgctc tcaaaacaag	780
tgtaacttct atgacaaca agacctggaa tgcgtgacca atcttcagga agtggccgc	840
atcgtgggca actctggcct caacatctac aatctctatg cccgtgtgc tggaggggtg	900
cccagccatt ttaggtatga gaaggacact gttgtgttcc aggatttggg caacatcttc	960
actcgctgc cactcaagcg gatgtggcat caggcactgc tgcgtcagg ggataaagtg	1020
cgcatggacc cccctgcac caacacaaca gctgcttcca cctacctcaa caaccgtac	1080
gtgcggaagg cctcaacat cccggagcag ctgccacaat gggacatgtg caactttctg	1140
gtaaaacttac agtaccgccg tctctaccga agcatgaact cccagtatct gaagctgctt	1200
agctcacaga aataccagat cctattatat aatggagatg tagacatggc ctgcaatttc	1260
atgggggatg agtggtttgt ggattccctc aaccagaaga tggaggtgca gcgccggccc	1320
tggttagtga agtacgggga cagcggggag cagattgccg gcttcgtgaa ggagtctcc	1380

cacatcgctt tttcacgat caaggcgcc ggccacatgg ttcccaccga caagccctc	1440
gctgccttca ccatgttctc ccgttctctg aacaagcagc catactgatg accacagcaa	1500
ccagctccac ggctgatgc agccctctcc agcctctccc gctaggagag tctcttctta	1560
agcaaagtgc ccctgcaggc cgggttctgc cgccaggact gccccttcc cagagccctg	1620
tacatcccag actgggcccc gggctctcca tagacagcct gggggcaagt tagcacttta	1680
ttcccgagc agttctgaa tgggttgcc tggcccttc tctgtttaa gaatgccctt	1740
tatgatgcac tgattccatc ccaggaacc aacagagctc aggacagccc acagggaggt	1800
ggtggacgga ctgtaattga tagattgatt atggaattaa attgggtaca gcttcaaa	1858
<210> 255	
<211> 3696	
<212> DNA	
<213> Homo sapiens	
<400> 255	
atcaggttac cggattcgag tcagaagcgg cggcaggtct gaacgcttcg gtaggataag	60
tgcggtggga gggggctctg cggttttggg gagcccaggg cggaaccg gaccggctg	120
gaagtccga gccctggccg agccccgcc ctccgtctt tcttggtgac ctttaagccag	180
tggctgcctt tttctgagc cgggcggggc cgaaggcggc ccgtaggccc tcgggactcc	240
cagcactgca gaggggtgta ggtctgacat ccaagacag ttgtttcgta tttctgaagg	300
aagaactgaa gctccgggaa gtgatggctg gggatggggc gggcaactg gggaccgagt	360
gtacgatcca cgcctaaggt tgaggcgggc cgagctagcc aggcagccgt gacccagtg	420
cttttcagac gtttcttagc ttccagagcc caacacatac agctgataca cgcagaccag	480
atctggtcag gtccctcgaa gctgagtcga gagcgatgct gctgaagaca gtgctcttgc	540
tgggacatgt ggcccagggt ctgatgctgg acaatgggt cctgcagaca ccacccatgg	600
gctggctggc ctgggaacgc ttccgtgca acattaactg tgatgaggac ccaaagaact	660
gcataagtga acagctcttc atggagatgg ctgaccggat ggcacaggat ggatggcggg	720
acatgggcta cacatactc aacattgatg actgctggat cgggtggtcgc gatgccagtg	780
gccgcctgat gccggatccc aagcgttcc ctcatggcat tcttttctg gctgactacg	840
ttcactccct gggcctgaag ttgggtatct acgcggacat gggcaacttc acctgcatgg	900
gttaccagc caccacactg gacaaggtgg tccaggatgc tcagaccttc gccgagtggg	960
aggtagacat gctcaagctg gatggctgct tctccacccc cgaggagcgg gccagggggt	1020
acccaagat ggctgctgcc ctgaatgcca caggccgccc catcgcttc tctgcagct	1080

ggccagccta tgaaggcggc ctcccccaaa gggatgaacta cagtctgctg gcggacatct	1140
gcaacctctg gcgtaatat gatgacatcc aggactcctg gtggagcgtg ctctccatcc	1200
tgaattggtt cgtggagcac caggacatac tgcagccagt ggccggccct gggcactgga	1260
atgacctga catgctgctc attgggaact ttggtctcag cttagagcaa tcccgggccc	1320
agatggccct gtggacggtg ctggcagccc cctctttgat gtccacagac ctgcgtacca	1380
tctccgccc gaacatggac attctgcaga atccactcat gatcaaaatc aaccaggatc	1440
ccttaggcat ccaggagcgc aggattcaca aggaaaaatc tctcatcgaa gtgtacatgc	1500
ggcctctgtc caacaaggct agcgcttag tcttcttcag ctgcaggacc gatatgcctt	1560
atcgctacca ctctccctt ggccagctga acttcaccgg gtctgtgata tatgaggccc	1620
aggacgtcta ctcaggtgac atcatcagt gcctccgaga tgaaccaac ttcacagtga	1680
tcatcaacc ttcaggggta gtgatgtgtt acctgtatcc catcaagaac ctggagatgt	1740
cccagcagtg aggagctggg acatgtgaca ggctgtggtg gcaccactga gcctagacca	1800
tggagccttg gcatgccag ggcaagtggg gaggttctct gctcccagg cctgctcggt	1860
gactgacccc atcataccca aagtgaatc tcacggccag gttctatgcc ctgtccaagc	1920
glaaaccctc ttggaaactt cttttggggc aattttctg tggccttctt ggctctact	1980
tccatgtgcg cagccccaca gacgttctg agcaactgc cagcctctg agctccatgc	2040
ccatcaggac tctagcctt gaccttctg ttgactctga aatcaggatt tggagtttt	2100
cgaattagga gtagagagat ctgacctctt gccaggaatg cccatggatc atgtgattgg	2160
cttttctacc catagagggc cttgcagcct gataccacct gggagtgagg gtcacaaagg	2220
agaccttggc tccctcaggt caccaataaa cctgttcttt aatcaagtag ttgtaatatg	2280
gacctggggt acctttagtt agggactctg gtggcactac acggggtgct ggattcaaat	2340
ctcacctcta tggatgatgt ctggcatcgt ctggccctc cttttccaaa gtgaccagaa	2400
gattgtactt cctagagatg ccagacttgc ctcacacca ggagctgccc ctgctgtgcc	2460
tggatgtctg gcagtgtga tggtgacaac acaatgaggc caagtagcaa tgctcccagg	2520
tcctgcagct tigtctctgc ccagcagtgc ccttgGCCa gcacctcagc ctgctcttct	2580
ctctggtggc tccacccaa cccctacttc ttttcttta tttctctccc aacctggctg	2640
caaaggaggt tcatattccc cagatgtgtc cacaacctg tccaggagt gtgtcttaaa	2700
attgatttcc tgtccctctt cctaaggaca ggggtaacag tcaaagtga ggaatgatgg	2760
gttcttgctt gtcttgcttt ggtccgtatt tcccaagccc tctctttta gccaaaggcta	2820
cagggcactt gtcttcctg gtggctctcc ctttggaga gagctagtgg cagagcaacc	2880

tagctaggtc ctccctgggc ttggaatgca ggagcaccca gtgccccagg ccttgcactg	2940
tgctgcccac ggagtatttt ctgggtggat ctcccatact cctcagggtt ctgaggtcct	3000
tgtatttagg gtatggtgta aaaacttaca tatgcacatt ttgttacatc atttaagatg	3060
attgaagctt tactggttaa gtataggggtc accagagaaa gttcctttcc ccagttgatg	3120
tcttgaggct gttgagaaac ttaggacaga agcctacctc atcccaggcc tgcccctccg	3180
tgcatagctc tgcctttggg tcttgttcat agtgcgtgtc aggcaaaatg gaatagtctc	3240
aggtgggaac accttctctt gctagctcca ggaaggcttt gtgggaaatg agttgagacc	3300
atggacttgg gagtggggg ccagagttga gcctggactt gaccaattaa ttccttacc	3360
tttcagccca agatagctac atgtctttac ttgctgtata agttttcctt ttgtcctggg	3420
ggtgcctgat tgatcctatc tcttcacctc tcattctttc aacaacatg ccatctatct	3480
cccaggacat ttatccctt ctattcaagg ccagtttagga atgcaattat ttttttttt	3540
cagttaaata cagacttgtt tggacgcaag gtaccctttc tctttttaa aactttttat	3600
catgaaaact ttcaaataca aaaataaaat ggtagagtga acctccatgt acccagcatc	3660
cagctccagt aaccatcaac gtgtaccatt atttca	3696
<210> 256	
<211> 5720	
<212> DNA	
<213> Homo sapiens	
<400> 256	
gcagccgcct ctgcccggc cagaccggg cggcgccggc cgcctcgcc gccgcaggtc	60
cgcctccgc gcgccaggg cgtgggctcc cggctcccgg aagcggcggc cgcggcggc	120
agccgagcgg gcgtccgtc cggagctgc aatgagcggc gcccgaggc tgtgacctgc	180
gcgcggcggc ccgaccggg cccctgaatg gcggctcgt gagcggcgg cggcgcggc	240
ggcggtcag gtcctcggg gcgtggcgt gcggtgaagg ggtgatgtg ttcaagctc	300
tgcagagaca gacctatcc tgctgtccc acaggtatgg gctctacgt tgcttcttg	360
gcgtcgttgt caccatcgt tccgccttc agttcgaga ggtggttctg gaatggagc	420
gagatcaata ccatgtttg ttgtattcct atagagacaa tattgctgga aagtccttc	480
agaatcggct ttgtctgcc atgccattg acgttggtta cacctgggtg aatggcacag	540
atcttgaact actgaaggaa ctacagcagg tcagagaaca gatggaggag gagcagaaag	600
caatgagaga aatccttggg aaaaacacaa cggaacctac taagaagagt gagaagcagt	660
tagagtgttt gctaacacac tgcattaagg tgccaatgct tgtcctggac ccagccctgc	720

cagccaacat caccctgaag gacctgccat ctctttatcc tttttttcat tctgccagt	780
acattttcaa tgttgcaaaa ccaaaaaacc ctctaccaa tgtctcagtt gttgtttttg	840
acagtactaa ggatgttgaa gatgccact ctggactgct taaaggaaat agcagacaga	900
cagtatggag gggctacttg acaacagata aagaagtccc tggattagt ctaatgcaag	960
atttggtttt cctgagtgga ttccaccaa cattcaagga aacaaatcaa ctaaaaacaa	1020
aattgccaga aaatctttcc tctaaagtca aactgttgca gttgtattca gaggccagt	1080
tagcgcttct aaaactgaat aacccaagg attttcaaga attgaataag caaactaaga	1140
agaacatgac cattgatgga aaagaactga ccataagtcc tgcatattta ttatgggatc	1200
tgagcgccat cagccagttc aagcaggatg aagacatctc tgccagtcgt ttgaaagata	1260
acgaagaact gaggtactca ttgcgatcta tcgagaggca tgcaccatgg gttcggaata	1320
ttttcattgt caccaacggg cagattccat cctggctgaa ccttgacaat cctcagtgga	1380
caatagtaac acaccaggat gtttttcgaa atttgagcca ctgcctacc tttagttcac	1440
ctgctattga aagtcacatt catcgcatcg aagggtgtc ccagaagttt atttacctaa	1500
atgatgatgt catgtttggg aaggatgtct ggccagatga tttttacagt cactccaaag	1560
gccagaaggt ttatttgaca tggcctgtgc caaactgtgc cgagggtgc ccaggttcct	1620
ggattaagga tggctattgt gacaaggctt gtaataattc agcctgcgat tgggatggtg	1680
gggattgctc tggaacaggt ggagggagtc gctatattgc aggaggtgga ggtactggga	1740
gtattggagt tggacagccc tggcagtttg gtggaggaat aaacagtgtc tcttactgta	1800
atcagggatg tgcgaattcc tggctcgtg ataagttctg tgaccaagca tgcaatgtct	1860
tgtcctgtgg gtttgatgct ggcgactgtg ggcaagatca ttticatgaa ttgtataaag	1920
tgatccttct cccaaccag actcactata ttattccaaa aggtgaatgc ctgccttatt	1980
tcagctttgc agaagtagcc aaaagaggag ttgaaggtgc ctatagtgac aatccaataa	2040
ttcgacatgc ttctattgcc aacaagtgga aaaccatcca cctcataatg cacagtggaa	2100
tgaatgccac cacaatacat tttaatctca cgtttcaaaa tacaacgat gaagagttca	2160
aaatgcagat aacagtggag gtggacacaa gggagggacc aaaactgaat tctacagccc	2220
agaagggtta cgaaaattta gttagtcca taacattct tccagaggcg gaaatccttt	2280
ttgaggatat tcccaaagaa aaacgcttcc cgaagttaa gagacatgat gtttaactca	2340
caaggagagc ccaggaagag gtgaaaattc ccctggtaaa tatttcactc cttccaaaag	2400
acgccagtt gagtctcaat accttgatt tgcaactgga acatggagac atcactttga	2460

aaggatacaa tttgtccaag tcagccttgc tgagatcatt tctgatgaac tcacagcatg 2520

ctaaaaataa aaatcaagct ataataacag atgaaacaaa tgacagtttg gtggctccac 2580

aggaaaaaca ggttcataaa agcatcttgc caaacagctt aggagtgtct gaaagattgc 2640

agaggttgac ttttctgca gtgagtgtaa aagtgaatgg tcatgaccag ggtcagaatc 2700

cacccttgga ctgggagacc acagcaagat ttagagtgga aactcacacc caaaaaacca 2760

taggcggaaa tgtgacaaaa gaaaagcccc catctctgat tgttccactg gaaagccaga 2820

tgacaaaaga aaagaaaatc acagggaag aaaaagagaa cagtagaatg gaggaaaatg 2880

ctgaaaatca cataggcggt actgaagtgt tacttggaag aaagctgcag cattacacag 2940

atagtactt gggctttttg ccatgggaga aaaaaaagta tttccaagat cttctcgacg 3000

aagaagagtc attgaagaca caattggcat acttactga tagcaaaaat actgggaggc 3060

aactaaaaga tacatttgca gattccctca gatatgtaaa taaaattcta aatagcaagt 3120

ttggattcac atcgcgaaa gtccctgctc acatgcctca catgattgac cggattgtta 3180

tgcaagaact gcaagatatg ttcctgaag aatttgacaa gacgtcattt cacaaagtgc 3240

gccattctga ggatatgcag ttgccttct cttattttta ttatctcatg agtgcagtgc 3300

agccactgaa tatatctcaa gtctttgatg aagttgatac agatcaatct ggtgtcttgt 3360

ctgacagaga aatccgaaca ctggctacca gaattcacga actgccgtta agtttgcagg 3420

atttgacagg tctggaacac atgctaataa attgctcaaa aatgcttctt gctgatatca 3480

cgcagctaaa taatattcca ccaactcagg aatcctacta tgatcccaac ctgccaccgg 3540

tcactaaaag tctagtaaca aactgtaaac cagtaactga caaaatccac aaagcatata 3600

aggacaaaaa caaataatagg ttgaaatca tgggagaaga agaaatcgct tttaaaatga 3660

ttcgtaccaa cgtttctcat gtggttggcc agttggatga cataagaaaa aaccctagga 3720

agtttgtttg cctgaatgac aacattgacc acaatcataa agatgctcag acagtgaagg 3780

ctgttctcag ggacttctat gaatccatgt tccccatacc ttccaattt gaactgcaa 3840

gagagtatcg aaaccgtttc cttcatatgc atgagctgca ggaatggagg gcttatcgag 3900

acaaattgaa gttttggacc catttgttac tagcaacatt gattatgttt actatattct 3960

cattttttgc tgagcagtta attgcactta agcgggaagat atttcccaga aggaggatac 4020

acaaagaagc tagtcccaat cgaatcagag tatagaagat cttcatttga aaaccatcta 4080

cctcagcatt tactgagcat tttaaaactc agcttcacag agatgtcttt gtgatgtgat 4140

gcttagcagt ttggcccgaa gaaggaaaat atccagtacc atgctgtttt gtggcatgaa 4200

tatagccac tgaccaggaa ttatttaacc aaccactga aaacttgtgt gttgagcagc 4260
 tctgaactga ttttactttt aaagaatttg ctcatggacc tgtcatcctt ttataaaaa 4320
 ggctcactga caagagacag ctgttaattt cccacagcaa tcattgcaga ctaactttat 4380
 taggagaagc ctatgccagc tgggagtgat tgctaagagg ctccagtctt tgcattccaa 4440
 agccttttgc taaagttttg cacttttttt ttttcatttc ccatttttaa gtagttacta 4500
 agttaactag ttattcttgc ttctgagtat aacgaattgg gatgtctaaa cctattttta 4560
 tagatgttat ttaaataatg cagcaataac acctcttatt gacaatacct aaattatgag 4620

ttttattaat atttaagact gtaaatggtc ttaaaccact aactactgaa gagtcaatg 4680
 attgacatct gaaatgcttt gtaattattg acttcagccc ctaagaatgc tatgatttca 4740
 cgtgcaggtc taatttcaaa gggctagagt tagtactact taccagatgt aattatgttt 4800
 tggaaatgta catattcaaa cagaagtgcc tcattttaga aatgagtagt gctgatggca 4860
 ctggcacatt acagtgggtg ctgttttaat actcattggg atattccagt agctatctct 4920
 ctgattgggt tttgataga acagaggcca gcaaactttc tttgtaaaag gctggtagt 4980
 aaattattgc aggccacctg tgtctttgtc atacattctt ctgtctgttg tttagtttgt 5040

ttttttcaa acaaccctct aaaaatgtaa aaacatggt tagcttgagc ctgtacaaaa 5100
 actgcccacc agccagatgt gacctcagg ccattcttg ccaatcactg agaattagtt 5160
 tttgttgttg ttgttgttgt tgtttttgag acagagtcct tctctgttgc ccaggtgga 5220
 gtgcagtggc gcaatctcag ctactgcaa cctccgcctc ccgggttcaa gcagtctgt 5280
 ctgagccttc tgagtagctg ggactacagg tgcatgccac cacaccctgc taatttttgt 5340
 atttttagta gagacggggg ttccaccata ttggtcaggc ttatcttgaa ctctgacct 5400
 caggtgatcc acctgcctct gcctcccaaa gtgctgagat tacaggcata agccagtga 5460

cccagccgag aattagtatt tttatgtatg gttaaacctt ggctcttagc catattttat 5520
 gtcataatac aatggatttg tgaagagcag attccatgag taactctgac aggtatttta 5580
 gatcatgac tcaacaatat ttttccaaaa tggcatacat cttttgtaca aagaacttga 5640
 aatgtaata ctgtgtttgt gctgtaagag ttgtgtattt caaaaactga aatctcataa 5700
 aaagttaaat tttgtctga 5720

<210> 257

<211> 4139

<212> DNA

<213> Homo sapiens

<400> 257

gcatttcctga ccggcacctg gcgaggctca tgcgtcccgt gagggcgggt cctcgagcct 60

gggggcgctc agattgcttt ggagacgctg agagaacctt tgcgagagcg ccggttgacg 120

tgcggagtgc ggggctccgg gggactgagc agcacgagac cccatcctcc cctccgggtt 180

ttcacactgg gcgaaggag gactcctgag cctgcctct tccagtaaca ttgaggatta 240

ctgtgttttg tgagagctcg ctaggcgcgc taagcaacag agttctgaga aatcgagaaa 300

catgataagg aattggctga ctatctttat cctttttccc ctgaagctcg tagagaaatg 360

tgagtcaagc gtcagcctca ctgttccctc tgcgttaaag ctggagaacg gcagctcgac 420

caacgtcagc ctcacctgc ggccaccatt aaatgcaacc ctggtgatca cttttgaaat 480

cacatttcgt tccaaaaata ttactatcct tgagctcccc gatgaagttg tggcgcctcc 540

tggagtgcac aactcctctt ttcaagtgc atctcaaaat gttggacaac ttactgttta 600

tctacatgga aatcactcca atcagaccgg ccgaggata cgctttcttg tgatccgcag 660

cagcgccatt agcatcataa accaggtgat tggctggatc tactttgtgg cctgggtccat 720

ctccttctac cctcaggtga tcatgaattg gaggcggaag agtgtcattg gtctgagctt 780

cgacttcgtg gctctgaacc tgacgggctt cgtggcctac agtgtattca acatcgccct 840

cctctgggtg ccctacatca aggagcagtt tctcctcaaa taccccaacg gagtgaaccc 900

cgtgaacagc aacgacgtct tcttcagcct gcacgcgggt gtcttcacgc tgatcatcat 960

cgtgcagtgc tgctgtatg agcgcggtgg ccagcgctg tcctggcctg ccatcggtt 1020

cctgggtctc gcgtggctct tcgcatttct caccatgatc gtggctgcag tgggagtgc 1080

cacgtggctg cagtttctct tctgtctctc ctacatcaag ctgcagtc caactgtcaa 1140

gtattttcca caggccctca tgaactttta ctacaaaagc actgagggtt ggagcattgg 1200

caacgtgctc ctggacttca ccgggggcag cttcagcctc ctgcagatgt tctccagtc 1260

ctacaacaac gaccagtga cgctgatctt cggagacca accaagtttg gactcgggt 1320

cttctccatc gtcttcgacg tcgtcttctt catccagcac ttctgtttgt acagaaagag 1380

accggggtat gaccagctga actagcacc agggaccagc tgtaccagc ctctggcctc 1440

gtgccctgct ggggaaggcc tcaccagcg aaggccggag aagcggttgg gccctggcac 1500

acagggtctg ctcagtgtgc ggacagagga gaccactctg ctctggggc cagaggccat 1560

tcaatagcct gccttcgctc gggccctcc tgggctccc cggccaggca cgtggcaccg 1620

tcgccttgac accgcatct cttttcttta aggcttcagg cagcgcgac aggcctctggc 1680

agccgtctca ggcaggactg ggcaccaagc ttgcagccga aggccttgcc ccaaactacc 1740

agcgtttctg caagcagctt gaagggtga ccttgcagcc gggtagacca agggcacttt	1800
gctgccaccg ctgcattccc agagatcaag cagcccgtg ccgtggccag tgaactcaga	1860
ggtagctggg gacgggctag gactttgggg ttaggccatg gggctctttc tctgaaggcc	1920
actttcctga cgtactctct gtacataact cagcgtccgt gactgcagta acagccagcc	1980
ctaccagag tatttctgag ccatgagggg cccaccagat tggttctgaa ttggattcat	2040
gcccagcgca ttagcatagt aactcctttc agattttttg gagggacgtt tgggaagtggc	2100
ttactctctt ctgccctctc tctacctcc accttctcag attagcccca tctgagcaca	2160
tccagctgct ccttaccag catctggagt acaggacata gctctctcct gctaccagtc	2220
tgtgccttag aggtctgtta ggcctgccaa acggcgacca gctcccctgg agcgagggca	2280
ggcccccttc ctctctttcc ccagacacct acttgagact caccaatttc tggcctgttc	2340
aggagcctca gataagtatt tgtacttgag accacctcac acaatctgta tgggccaac	2400
cctgatctca aacctccttc cctctgccca aagctgtcct tcctatggca ggaggggtgg	2460
gggtcccagg acgtgcctca tacatgactt gagcttgta gtccactgag ttctcttcta	2520
cgagatcaac gcgaggggcc tgtatcttga attaaaacct actcgcttcc ttcttgcaact	2580
ctgtgtgctc gatacagaag gcaggtgtgc tggccttcag agccaggtgc cctgctcccc	2640
tcgtgtaagg aagcgataa ggttgtgaag gtactgacc cgagagcgct gcgtgtgac	2700
cccacacttc actggggagc cctcttgggt agcgcgagac gcctaccgc ccagggccag	2760
gtcagctctc caggtccgt gtcttcacgg ttaccagggc acgccaactca atctggatgt	2820
gagccggggt gcatgaggtt ctgaaggga caccagcact gggagcgggg gcagtggggg	2880
tcctattctc cagacgttc ctctatctag ttgtaacaaa cgtcaaagaa tgtcaggtct	2940
tgtttctgca gacagactt tggcctgacg ggtccacag acctgtttca ctgcaacac	3000
ccagtaaatg tgttactggg tgccctactt taaaaatcag agatttcaa tgattgactt	3060
ttccggctga gtgcggtggc tcacgcctgt aatcccagca ctttgggagg ccaaggcaga	3120
tggatcacct gacatcagga gttcgagacc agcctagcca acattgcaa accccatctc	3180
tactaaagat acaaaaatta gctgggcatg gtggcgggca cctgtaatcc cagctacttg	3240
gttggctgag gcaggagaat tacttgaacc caggaggcgg aggttgcagt gagcgaagat	3300
cacgccattg cactccagcc tgggcaacaa gaacgaaact ccatctcaa aaaaaaaaaa	3360
aattattgac ttttcttta aatctgattt ggcagtgcta ggtaggtcca ggagtgctaa	3420
caccttccca gaggtggtac gcaggtagtc aggtgctctc tgactcacc ccatctggcc	3480
agatcacggc cccagcaac acgtgggggt gtgtccctcg ctgggataca cggagtgcc	3540
ttataggcag ggagtcttat aggcaggggt cctcccagg cacctctggg tgtccctacg	3600

gccccctcca ggtgggagct gaggctaggt tgcaaaaagg aggttgggag caagtagggg	3660
cttcacagta acaccaaga aagcacacgc accccagggt cccacccag tgctcccaga	3720
cactcccaca gatgtggttc tggtcaggcg cagtgcagct gctgcggctc caccagcccc	3780
gggctttcag gcagggtgggc ctgggagtag ttgagctttg tcccagaaca gtgggggatg	3840
gggccatcat taaaaccaca gctctggccc caatatccca ccgccacccc caccgcgata	3900
tcccagccca tgcctcagca aagccccctg tgaattccag catttttatt gagcgcacca	3960
catcggggag gggcggcagt ggtttccacg gtaaccaaag taaggcttgt gcacttgctg	4020
agcttcaga tgcgaattag taaaagctaa attcaaaagt agaatgggca tctccaaaga	4080
gtacgataaa aacattttgt atcaaacttg tgattaaaaa aatacaaaag ttaaaccga	4139
<210> 258	
<211> 3292	
<212> DNA	
<213> Homo sapiens	
<400> 258	
gcccccgccc cgcccgttcc agccagctcg gcccgggggc ttcgggctgt cgggccggcg	60
ctcccttctc tgccagggtg cgagtacacc tgctcacgta ggcgtcatga ggtctccggt	120
tcgagacctg gcccggaacg atggcgagga gacacggac cgcacgcctc ttctaccggg	180
cgccccacgg gccgaagccg ctccagtgtg ctgctctgct cgttacaact tagcaatfff	240
ggcctttfff ggtttcttca ttgtgtatgc attacgtgtg aatctgagt ttgcgttagt	300
ggatatggta gattcaata caactttaga agataataga acttccaagg cgtgtccaga	360
gcattctgct ccataaaag ttcatcataa tcaaacgggt aagaagtacc aatgggatgc	420
agaaactcaa ggatggattc tcggttctt tttttatggc tacatcatca cacagattcc	480
tggaggatat gttgccagca aaataggggg gaaaatgctg ctaggatttg ggatccttgg	540
cactgtctgc ctacacctgt tcactcccat tgctgcagat ttaggagttg gaccactcat	600
tgtactcaga gactagaag gactaggaga ggggtgttaca tttccagcca tgcattccat	660
gtggtcttct tgggctcccc ctcttgaaag aagcaaaact cttagcattt catatgcagg	720
agcacagctt gggacagtaa tttctcttcc tctttctgga ataatttgc actatatgaa	780
ttggacttat gtcttctact tttttgttac tattggaata ttttggttc ttttgtgat	840
ctggttagtt agtgacacac caaaaaaca caagagaatt tcccattatg aaaaggaata	900
cattctttca tcattaaaga atcagctttc ttcacagaag tcagtgccgt gggtagccat	960
tttaaaatcc ctgccacttt gggctatcgt agttgcacac ttttcttaca actggacttt	1020

ttatacttta ttgacattat tgcctactta tatgaaggag atcctaaggt tcaatgttca	1080
agagaatggg tttttatctt cattgcctta tttaggctct tggttatgta tgatcctgtc	1140
tggatcaagct gctgacaatt taagggcaaa atggaatfff tcaactttat gtgttcgcag	1200
aatfffftagc cttataggaa tgattggacc tgcagtattc ctggtagctg ctggcttcat	1260
tggctgtgat tattctttgg ccgttgcttt cctaactata tcaacaacac tgggaggctt	1320
ttgctcttct ggatttagca tcaaccatct ggatattgct ccttcgtatg ctggtatcct	1380
cctgggcac cacaatacat ttgccactat tccaggatg gttgggcccgc tcattgctaa	1440
aagtctgacc cctgataaca ctgttggaga atggcaaac gtgttctata ttgctctgc	1500
tattaatgtt ttgggtgcca tttcttttac actattcgcc aaaggtgaag tacaaaactg	1560
ggctctcaat gatcaccatg gacacagaca ctgaaggac caataataa tctgcctct	1620
attaatgtat tttatttat catgtaacct caaagtgcct tctgtattgt gtaagcattc	1680
tatgtctttt ttaattgta ctgtatttag atttttaagg cctataatca tgaaatatca	1740
ctagttgcca gaataataaa atgaactgtg ttaattatg aataatatgt aagctaggac	1800
ttctacttta gggtcacata cctgcctgct agtcgggcaa catgaagtag gacagttctg	1860
ttgatttttt agggccatag taaagggaat gagctgaaac agacctcctg atacctttgc	1920
ttaattaaac tagatgataa ttctcaggta ctgataaaca cctgttgttg ttcactttcc	1980
tcataaaaaa tgcagctct ctctgacact tagacctcaa actttagcat ctctgtggag	2040
ctgccatcca ctgtataatt tcgcctggca actggactga ggggagtgtg cccaggcagc	2100
tgccaagcac tccctccctg gcttcagggt cagagtgcgc agcgtttatc agaggcagca	2160
tcgaagccca gagccagtgt cgactcttcg gctgggtgcct ttctctctgag gggctatcaa	2220
tgtgtagata aagccctgag taggcaagag cagttagatc cactgctatg gtcttgatac	2280
atctcaaac tttcccttcc cagcacagag gaatatggc tggcatgcaa cctgcaaaag	2340
aaaaatgcga agcggccggg cacggtggct catgcctgta atcccagcac tttggggggc	2400
tgaggtgggc gaatcatgag atcaggagt tgcagaccagc ctggccagca tgggtgaaac	2460
ccatctctac taaaaatata aaaaattagc tgggcgtggt gacgggcgcc tgtaatccca	2520
gatactcagg aggctgaggt aggagaatca ctgtaacctg ggaggtggaa gttgcagtga	2580
accaagatca cgccactgca ctccagcctg ggcatggag cgagactcca actcaaaaaa	2640
aaaaaaagaa taaagaaaga aaagtgcgat gccagtcac tcacaaataa gatcatcctg	2700
gtttaaatct actctcacat ggatcacagt ataaatttct atgtgctgtg tttgttctgt	2760
ttgtattttg tagagatggg gtctcgtttt gtgcgccagg ctggttttga actcctggct	2820
tcaagcgatc ctctgtctc ggcctcaca agtgttgaga ctacaggcat gagccactgt	2880

gcccagcctg ttctatgttt ttaagctaca cgagaatfff ttttttaatt aattctcact	2940
gtttgttcag tctgtcttca tctaagtttg tgttgcagtt taaagttaaa gtgactttta	3000
aaggccacat cacctgagac tagggtaatc atctttactt ctggttccctg aaatcatatt	3060
tttccagtgg accatcctcc agtggctgtg gttgttgagc atgctttcag aacacctatg	3120
tggcttaaaa cttagtttat gttttgtgtt caacactacg tgtaatatft taaaactgtt	3180
taatgtgatg tgaatacatt tatgtacatt tatftttaaa tttgtaaata gctttaaatt	3240
gctatggcaa tgtttctftt ataatcatc aaaataaacc tttgtgaatt ga	3292
<210> 259	
<211> 1879	
<212> DNA	
<213> Homo sapiens	
<400> 259	
gtgctgtcac gtgatccgac aaacggcctc tgcatagtgc agaacattct gctgctctta	60
aaggtacagg cctcagggtc cctgctgtag acggggcggg ggagagtacg atgggtgggg	120
cgtgggtgggt cgtagggcgc tccagatgga gccccagct tccctgatgg atcgcggggc	180
gcgagtgcce tagacaagcc ggagctggga ccggcaatcg ggcgttgatc cttgtcacct	240
gtcgcagacc ctcatcctc ccgtgggagc cccctttgga cactctatga ccttggaacc	300
tcgggggacc tgaacttgat gcgatgggag gctgtgcagg ctgcgcggcg cgcttttcgg	360
attccgaggg ggaggagacc gtcccggagc cccggctccc tctgttgga catcagggcg	420
cgcattggaa gaacgcggtg ggcttctggc tctgggcct ttgcaacaac ttctcttatg	480
tggtgatgct gagtgccgc cagcacatcc ttagccacaa gaggacatcg ggaaaccaga	540
gccatgtgga cccaggccca acgccgatcc cccacaacag ctcatcacga ttgtactgca	600
actctgtctc tacggctgct gtgctcctgg cggacatcct cccacactc gtcacaaat	660
tgttgctcc tcttggcctt cacctgctgc cctacagccc ccgggttctc gtcagtggga	720
tttgtgctgc tggaagcttc gtcttggtt ccttttctca ttctgtgggg accagcctgt	780
gtggtgtggt ctctgctagc atctcatcag gccttgggga ggtcaccttc ctctccctca	840
ctgccttcta cccagggcc gtgatctct ggtggtctc agggactggg ggagctgggc	900
tgctgggggc cctgtcctac ctgggcctca cccaggccgg cctctcccct cagcagaccc	960
tgctgtccat gctgggtatc cctgcctgc tcttgccag ctatttcttg ttgtcacat	1020
ctctgaggc ccaggacct ggagggggaag aagaagcaga gagcgagcc cggcagcccc	1080
tcataagaac cgaggccccc gagtcaagc caggctccag ctccagctc tcccttcggg	1140

aaaggtggac agtgttcaag ggtctgctgt ggtacattgt tcccttggtc gtagtttact	1200
ttgccagta ttctattaac cagggaacttt ttgaactcct ctttttctgg aacacttccc	1260
tgagtacgc tcagcaatac cgctgggtacc agatgctgta ccaggctggc gtctttgcct	1320
cccgtcttc tcctcgctgc tgtcgcatcc gtttcacctg ggccctggcc ctgctgcagt	1380
gcctcaacct ggtgttctcg ctggcagacg tgtggttcgg ctttctgcca agcatctacc	1440
tcgtcttctt gatcattctg tatgaggggc tcttgggagg cgcagcctac gtgaacacct	1500
tccacaacat cgccctggag accagtgatg agcaccggga gtttgcaatg gcggccacct	1560
gcatctctga cacactgggg atctccctgt cggggctcct ggctttgcct ctgcatgact	1620
tcctctgcca gcctctctga tactcgggat cctcaggacg caggtcacat tcacctgtgg	1680
gcagagggac aggtcagaca cccaggccca ccccagagac cctccatgaa ctgtgctccc	1740
agccttcccg gcaggtctgg gagtagggaa gggctgaagc cttgtttccc ttgcaggggg	1800
gccagccatt gtctccact tggggagttt ctctctggca tcatgccttc tgaataaatg	1860
ccgattttat ccatggaaa	1879
<210> 260	
<211> 2504	
<212> DNA	
<213> Homo sapiens	
<400> 260	
ttattttgat tcaccgcaga gggcggctta cgagagcgca gagccccact cggccagcgg	60
ggtctggcgg gggacctgtc gcgctgaaag ctccagggtta gggccgacgc ccatcaggct	120
gggcatccgt tcgggatgcg caggttgcca tctgcaaccg gcggcgccac gcccaggcgg	180
gcggagcgcg gttcccgag tctcgcgccc gcggtcatgt gacacagcga agatggcgtc	240
gcccggctgc ctgtggctct tggctgtggc tctctgcca tggacctgcg cttctcgggc	300
gctgcagcat ctggaccgc cggcgccgct gccgttggtg atctggcatg ggatgggaga	360
cagctgttgc aatcccttaa gcatgggtgc tattaaaaaa atggtggaga agaaaatacc	420
tggaaattac gtcttatctt tagagattgg gaagacctg atggaggacg tggagaacag	480
cttctctttg aatgtcaatt cccaagtaac aacagtgtgt caggcacttg ctaaggatcc	540
taaattgcag caaggctaca atgctatggg attctcccag ggaggccaat ttctgagggc	600
agtggctcag agatgccctt cacctcccat gatcaatctg atctcggttg ggggacaaca	660
tcaagtggtt ttggactcc ctcatgccc aggagagagc tctcacatct gtgacttcat	720
ccgaaaaaca ctgaatgctg gggcgctact caaagtgtt caggaacgcc tcgtgcaagc	780

cgaatactgg catgacccca taaaggagga tgtgtatcgc aaccacagca tcttcttggc 840
agataataat caggagcggg gtatcaatga gtccatacaag aaaaacctga tggccctgaa 900
gaagtttgtg atggtgaaat tcttcaatga ttccattgtg gacctgtag attcggagtg 960
gtttggattt tacagaagtg gccaaagcaa ggaaaccatt cccttacagg agacctccct 1020

gtacacacag gaccgcctgg ggctaaagga aatggacaat gcaggacagc tagtgtttct 1080
ggctacagaa ggggaccatc ttcagttgtc tgaagaatgg ttttatgccc acatcatacc 1140
attccttggg tgaaacccgt atagttcaca atagagctca gggagcccct aactcttcca 1200
aaccacatgg gagacagtgt ccttcattgc caagcctgag ctgagatcca gcttgcaact 1260
aatccttcta tcatctaaca tgccctactt ggaaagatct aagatctgaa tcttatcctt 1320
tgccatcttc tgitaccata tgggtgtgaa tgcaagtta attacatgg agattgtttt 1380
acaaactttt gatgtggtca agttcagttt tagaaaaggg agtctgttcc agatcagtgc 1440

cagaactgtg cccaggccca aaggagacaa ctaactaaag tagtgagata gattctaagg 1500
gcaaacattt ttccaagtct tgccatattt caagcaaaga ggtgcccagg cctgaggtag 1560
tcacataaat gctttgtttt gctggtgatt taaccagtgc ttggaaaaat ctgtcttggc 1620
tatttctgca tcatcttcta aggtctcctt cctctctcag tacgttgccc tctgtctat 1680
catcttatca tcaattatta gacaaatccc actggcctac agtcttgctt ctgcagcacc 1740
cactttgtct cctcaggtag tgatgaatta gttgctgtca caaaaggagg gaagtagcac 1800
ccaaattaag ttgcttaaga gaggaatgt acatcttgta taacttaggg agcgaagaaa 1860

atgtaggcgc gaaagtgaag agtgaggcag ctagttcttc ctattccatt ctgaccaac 1920
ctgccctttc ttaatatgac tagtggtctt gatgctagag tcaacttact ctgttgctgg 1980
ctttagcaga gaataggagg aaccatatga aaaagatcag gctttctgac ttccatcccc 2040
aaaacacatt taccagcata ctccaaactg tttctgatgt gttccatgag aaaaggattg 2100
tttgctcaaa aagcttgga aatactacac actcccttc tccttctgga gatcaacca 2160
cattagagtg tctaaggact cctgagaatt cctgttacag taaacaaaac taacgtaac 2220
taccatttcc tacactattt gagcatggaa atcatagtcc ccactctgtg aaaacttaac 2280

gctttttgga agacatttct gtagcatgtc agtttgaga aatgatgagc tacgccttga 2340
tgaaagaacc gtgttggtgc tgctaagttt agccattatg gtttttctt tctctctctt 2400
aagccttatt cticaactaa aagatgagga ttaagagcaa gaagttgggg gggatgtgaa 2460
aataatttta tgaggttgtc taaaataaag agtagtttct tate 2504

<210> 261

<211> 2084

<212> DNA

<213> Homo sapiens

<400> 261

acagatcagc tgatgccgga gggtttgaag ccgcgccgcg agggagcagag gtcgcagtga	60
cagcggcggg cgatcggacc caggctgccc cgccgtaccc gcctgcgtcc cgcgctccc	120
ccccagcatg acagccccgg cgggtccgcg cggtcagag accgagcggc ttctgacccc	180
caaccccggg tatgggaccc aggcggggcc ttaccggcc cctccgacac cccagaaga	240
ggaagacctt cgcgctgctc tcaatactt ttcatgagt ccctgcgaca agtttcgagc	300
caaggccgc aagccctgca agctgatgct gcaagtggc aagatcctgg tggtcacggt	360
gcagctcatc ctgtttgggc tcagtaatca gctggctgtg acattccggg aagagaacac	420
catcgcttc cgacacctt tctgctggg ctactcggac ggagcggatg acaccttcgc	480
agcctacacg cgggagcagc tgtaccaggc catcttccat gctgtggacc agtacctggc	540
gttgctgac gtgtactgg gccgtatgc gtatgtccgt ggtgggggtg acccttggac	600
caatggctca gggcttgctc tctgccagcg gtactaccac cgaggccacg tggaccggc	660
caacgacaca ttgacattg atccgatggg ggttactgac tgcattcagg tggatcccc	720
cgagcggccc cctccgccc ccagcgacga tctaccctc ttgaaagca gctccagtta	780
caagaacctc acgtcaaat tccacaagct ggtcaatgc accatccact tccggctgaa	840
gaccattaac ctccagagcc tcatcaataa tgagatcccg gactgctata cttcagcgt	900
cctgatcacg ttgacaaca aagcacacag tgggcggatc cccatcagcc tggagacca	960
ggcccacatc caggagtga agcaccccag tgtcttcag cacggagaca acagcttccg	1020
gctctgttt gacgtgggtg tcatctcac ctgctccctg tcttctctc tctgcgccc	1080
ctcactcctt cgaggttcc tgctgcagaa cgagtttgtg gggttcatgt ggcggcagcg	1140
gggacgggtc atcagcctgt gggagcggct ggaatttgtc aatggctggt acatcctgct	1200
cgtcaccagc gatgtgctca ccatctcggg caccatcatg aagatcgga tcgaggccaa	1260
gaacttggcg agctacgacg tctgcagcat cctcctgggc acctcgacgc tgctggtgtg	1320
ggtgggcgtg atccgtacc tgaccttctt ccacaactac aatacctca tcgccacact	1380
gcgggtggcc ctgccagcg tcatgcgtt ctgctgctgc gtggtgtca tctacctggg	1440
ctactgcttc tgggctgga tcgtgctggg gccctatcat gtgaagtcc gctactctc	1500
catggtgtct gagtgctgt tctcgtcat caatggggac gacatgttg tgacgttcgc	1560
cgccatgcag gcgcagcagg gccgcagcag cctggtgtgg ctcttctccc agctctacct	1620
ttactccttc atcagcctt tcatctacat ggtgctcagc ctcttcatcg cgctcatcac	1680

cggcgcctac gacaccatca agcatcccgg cggcgcaggc gcagaggaga gcgagctgca 1740
 ggccctacatc gcacagtgcc aggacagccc cacctccggc aagtccgcc gcgggagcgg 1800

ctcggcctgc agccttctct gctgctgcgg aagggacccc tcggaggagc attcgtctct 1860
 ggtgaattga ttgcacctga ctgccgttgg accgtaggcc ctggactgca gagacccccg 1920
 ccccgaccc cgcttattta ttgttaggt ttgttttaaa ggatcggtc cctgtcgcgc 1980
 ccgaggaggg cctggacctt tcgtgtcgga cccttggggg cggggagact gggtagggag 2040
 ggtgttgaat aaaagggaataaataatgtgt cgttttcatt tttta 2084

<210> 262

<211> 2748

<212> DNA

<213> Homo sapiens

<400> 262

gcattgcaga ctgcggagtc agacggcgct atgtacgcc tcttcctcct gccagcctc 60

ctgggcgcgg ctctagccgg cccggtcctt ggactgaaag aatgcaccag gggctcgga 120
 gtgtggtgcc agaattgtga gacggcgctc gactgcgggg cagtgaagca ctgctgcag 180
 accgtttgga acaagccaac agtgaatcc ctccctgcg acatatgcaa agacgttgc 240
 accgcagctg gtgatatgt gaaggacaat gccactgagg aggagatcct tgtttacttg 300
 gagaagacct gtgactggt tccgaaccg aacatgtctg cttcatgcaa ggagatagt 360
 gactcctacc tccctgtcat cctggacatc attaaaggag aatgagccg tcctggggag 420
 gtgtgctctg ctctcaacct ctgcgagtct ctccagaagc acctagcaga gctgaatcac 480

cagaagcagc tggagtccaa taagatcca gagctggaca tgactgaggt ggtggcccc 540
 ttcatggcca acatccctct cctcctctac cctcaggacg gccccgcag caagccccag 600
 ccaaaggata atggggacgt ttgccaggac tgcatcaga tggtagtga catccagact 660
 gctgtacgga ccaactccac ctttgtccag gccttggtgg aacatgtcaa ggaggagtgt 720
 gaccgcctgg gccctggcat ggccgacata tgcaagaact atatcagcca gtattctgaa 780
 attgctatcc agatgatgat gcacatgcaa cccaaggaga tctgtgcgct ggttgggttc 840
 tgtgatgagg tgaaagagat gcccatgcag actctggtcc ccgcaaaagt ggcctccaag 900

aatgtcatcc ctgcccctgga actggtggag cccattaaga agcacgaggt ccagcaaag 960
 tctgatgttt actgtgaggt gtgtgaattc ctggtgaagg aggtgaccaa gctgattgac 1020
 aacaacaaga ctgagaaga aatactcgac gcttttgaca aaatgtgctc gaagctgccg 1080
 aagtcctgt cggaagagt ccaggagggt gtggacacgt acggcagctc catcctgtcc 1140

atcctgctgg aggaggtcag ccctgagctg gtgtgcagca tgctgcacct ctgctctggc 1200
acgcggctgc ctgcactgac cgttcacgtg actcagccaa aggacggtgg cttctgcgaa 1260
gtgtgcaaga agctggtggg ttatttggat cgcaacctgg agaaaaacag caccaagcag 1320

gagatcctgg ctgctcttga gaaaggctgc agcttcctgc cagaccctta ccagaagcag 1380
tgtgatcagt ttgtggcaga gtacgagccc gtgctgatcg agatcctggt ggaggtgatg 1440
gatccttccct tcgtgtgctt gaaaattgga gcctgcccct cggcccataa gcccttgttg 1500
ggaactgaga agtgtatatg gggcccaagc tactggtgcc agaacacaga gacagcagcc 1560
cagtgcaatg ctgtcgagca ttgcaaagc catgtgtgga actaggagga ggaatattcc 1620
atcttggcag aaaccacagc attggttttt ttctacttgt gtgtctgggg gaatgaacgc 1680
acagatctgt ttgactttgt tataaaaata gggtccccc acctccccca tttctgtgtc 1740

ctttattgta gcattgctgt ctgcaaggga gccctagcc cctggcagac atagctgctt 1800
cagtgccctt ttctctctg ctatagtgat gttgatgcac tggaggtctt ttagcctgcc 1860
cttgcattggc gcctgctgga ggaggagaga gctctgctgg catgagccac agtttcttga 1920
ctggaggcca tcaacctctt tggttgaggc cttgttctga gccctgacat gtgcttgggc 1980
actgttgggc ctgggcttct gagggtggcct cctgccctga tcagggaccc tccccgttt 2040
cctgggcctc tcagtgaac aaagcagcaa aacaaaggca gttttatatg aaagattaga 2100
agcctggaat aatcaggctt tttaaatgat gtaattccca ctgtaatagc atagggattt 2160

tggaagcagc tgctggtggc ttgggacatc agtggggcca agggttctct gtccctggtt 2220
caactgtgat ttggctttcc cgtgtctttc ctgggtgatgc cttgtttggg gtctctgtggg 2280
tttgggtggg aagagggcca tctgcctgaa tgtaacctgc tagctctccg aaggccctgc 2340
gggcctggct tgtgtgagcg tgtggacagt ggtggccgcg ctgtgcctgc tcgtgttgcc 2400
tacatgtccc tggtgttga ggcgtgctt cagcctgcac cctcccttg tctcatagat 2460
gtccttttg accttttcaa ataaatatgg atggcgagct cctaggcctc tggttctctg 2520
gtagaggcg gcatgccgaa gggctctgctg ggtgttgatt ggatgctggg gtgtgggggt 2580

tggaagctgt ctgtggccca cttgggcacc cacgttctg tccacttctg gttgccagga 2640
gacagcaagc aaagccagca ggacatgaag ttgctattaa atggacttcg tgattttgt 2700
tttgactaa agtttctgtg atttaacaat aaaattctgt tagccaga 2748

<210> 263

<211> 2976

<212> DNA

<213> Homo sapiens

<400> 263

gtcgcgaagt ttccagccct gcgagcgccg ccgggtcggc cgatcgtccc ccatacctcg	60
gccatgcggc ccttctgtct actggccctg ctgggctggc tgctgtctggc cgaagcgaag	120
ggcgacgcca agccggagga caacctttta gtccctcacgg tggccactaa ggagaccgag	180
ggattccgtc gcttcaagcg ctcagctcag ttcttcaact acaagatcca ggcgcttggc	240
ctaggggagg actggaatgt ggagaagggg acgtcggcag gtggagggca gaaggctcgg	300
ctgctgaaga aagctctgga gaagcacgca gacaaggagg atctggatcat tctcttcgca	360
gacagctatg acgtgctgtt tgcacggggg ccccgggagc tctgaagaa gttccggcag	420
gccaggagcc aggtgggtctt ctctgtctgag gagctcatct acccagaccg caggctggag	480
accaagtatc cgggtggtgc cgatggcaag aggttcctgg gctctggagg cttcatcggt	540
tatgccccca acctcagcaa actggtggcc gagtgggagg gccaggacag cgacagcgat	600
cagctgtttt acaccaagat cttcttggac ccggagaaga gggagcagat caataccacc	660
ctggaccacc gctgccgtat cttccagaac ctggatggag ccttggatga ggtcgtgctc	720
aagtttgaaa tgggcatgt gagagcgagg aacctggcct atgacaccct cccggctctg	780
atccatggca acgggccaac caagctgcag ttgaactacc tgggcaacta catcccgcgc	840
ttctggacct tcgaaacagg ctgcaccgtg tgtgacgaag gcttgcgcag cctcaagggc	900
attggggatg aagctctgcc cacggtcctg gtccgctgtt tcatcgaaca gcccacgccg	960
tttgtgtccc tgttcttcca gcggctcctg cggctccact acccccagaa acacatgcga	1020
cttttcatcc acaaccacga gcagcaccac aaggctcagg tggaagagtt cctggcacag	1080
catggcagcg agtaccagtc tgtgaagctg gtgggcccctg aggtgcggat ggcgaatgca	1140
gatgccagga acatgggcgc agacctgtgc cggcaggacc gcagctgcac ctactacttc	1200
agcgtggatg ctgacgtggc cctgaccgag cccaacagcc tgcggctgct gatccaacag	1260
aacaagaacg tcattgcccc gctgatgacc cggcatggga ggctgtggtc gaacttctgg	1320
ggggctctca gtgcagatgg ctactatgcc cgttccgagg actacgtgga cattgtgcag	1380
gggcggcgtg ttggtgtctg gaatgtgcc tatatttcaa acatctactt gatcaagggc	1440
agtgccctgc ggggtgagct gcagtcctca gatctcttcc accacagcaa gctggacccc	1500
gacatggcct tctgtgccaa catccggcag caggatgtgt tcatgttctt gaccaaccgg	1560
cacacccttg gccatctgct ctccctagac agctaccgca ccaccacct gcacaacgac	1620
ctctgggagg tgttcagcaa ccccaggac tggaaggaga agtacatcca ccagaactac	1680
accaaagccc tggcagggaa gctggtggag acgccctgcc cggatgtcta ttggttcccc	1740
atcttcacgg aggtggcctg tgatgagctg gtggaggaga tggagcactt tggccagtgg	1800

tctctgggca acaacaagga caaccgcatc cagggtggct acgagaacgt gccgactatt 1860

gacatccaca tgaaccagat cggctttgag cgggagtggc acaaattcct gctggagtac 1920

attgcgcccc tgacggagaa gctctacccc ggctactaca ccagggccca gtttgacctg 1980

gcctttgtcg tccgtacaa gcctgatgag cagccctcac tgatgccaca ccatgatgcc 2040

tccaccttca ccatcaacat cgccctgaac cgagtcgggg tggattacga gggcgggggc 2100

tgctggttcc tgcgctacaa ctgttccatc cgagccccaa ggaagggtg gaccctcatg 2160

caccttggac gactcacgca ttacatgag gggtcccca ccaccagggg caccgctac 2220

atcgagtct cttctgtcga tccctaattg gccaggcctg accctcttgg acctttcttc 2280

tttgccgaca accactgccc agcagcctct gggacctcgg ggtcccaggg aaccagttcc 2340

agcctcctgg ctgttgactt cccattgtct ttggagccac caatcaaaga gattcaaaga 2400

gattcctgca ggccagagcg ggaacacacc tttatggctg gggtctctcg tgggtgtctg 2460

gaccagccc ctggagacac cattcacttt tactgtttg tagtgactcg tgctctcaa 2520

cctgtcttcc tgaaaaacca aggccccctt ccccccctc ttccatgggg tgagacttga 2580

gcagaacagg ggcttcccca agttgcccag aaagactgtc tgggtgagaa gccatggcca 2640

gagcttctcc caggcacagg tggtgcacca gggacttctg cttcaagttt tggggtaaag 2700

acacctggat cagactccaa gggctgcctt gactctggga cttctgcctc catggctggt 2760

catgagagca aaccgtatgc ccctggagac agcgactcca gagaacctct tgggagacag 2820

aagaggcatc tgtgcacagc tcgatcttct acttgctgt ggggagggga gtgacaggtc 2880

cacacaccac actgggtcac cctgtcctgg atgcctctga agagagggac agaccgtcag 2940

aaactggaga gtttctatta aaggtcattt aaacca 2976

<210> 264

<211> 4164

<212> DNA

<213> Homo sapiens

<400> 264

agcagagcaa tcaccacaa gcctggaata actgcaaggg ctctgctgac atcttctga 60

ggtgccaagg aaatgaggat ggaggaagga atgaatgtt tccatgactt tgggatccag 120

tcaacacatt acctccagg gaattacaa gactcccagg actggttcat cttgggtgcc 180

gtgatcgag acctcaggaa tgccttctac gtctcttcc ccatctggtt ccatcttcag 240

gaagctgtgg gcattaaact ctttgggta gctgtgattg gagactggct caacctctc 300

tttaagtgga ttctcttgg acagcgcca tactggtggg ttttgatac tgactactac 360

agcaacactt ccgtgccct gataaagcag ttcctgttaa cctgtgagac tggaccaggg	420
agccccctctg gccatgccat gggcacagca ggtgtatact acgtgatggg cacatctact	480
ctttccatct ttcagggaaa gataaagccg acctacagat ttcgggtgctt gaatgtcatt	540
ttgtgggttg gattctgggc tgtgcagctg aatgtctgtc tgtcacgaat ctaccttgct	600
gctcatcttc ctcatcaagt tgttgctgga gtctgtcag gcattgctgt tgcagaaact	660
ttcagccaca tccacagcat ctataatgcc agcctcaaga aatatcttct cattaccttc	720
ttcctgttca gcttcgcat cggaatttat ctgctgtca agggactggg ttagacctc	780
ctgtggactc tggagaaagc ccagaggtgg tgcgagcagc cagaatgggt ccacattgac	840
accacaccct ttgccagcct cctcaagaac ctgggcacgc tctttggcct ggggctggct	900
ctcaactcca gcatgtacag ggagagctgc aaggggaaac tcagcaagt gctccattc	960
cgcctcagct ctattgtage ctccctctc ctctgcacg tctttgactc cttgaaacct	1020
ccatcccaag tcgagctggt cttctacgtc ttgtccttct gcaagagtgc ggtagtgccc	1080
ctggcatccg tcagtgtcat cccctactgc ctgcccagg tcctgggcca gccgcacaag	1140
aagtcgttgt aagagatgtg gagtcttcgg tgtttaaagt caacaacct gccagggatt	1200
gaggaggact actatttgaa gcaatgggca ctggtatttg gagcaagtga catgccatcc	1260
attctgccgt cgtggaatta aatcacggat ggcagattgg agggctgcct ggcttattcc	1320
catgtgtgac tccagcctgc cctcagcaca gactctttca gatggagggt ccataacacg	1380
tacaccatat gcaagtcttc cgccaggagg tcctctctc tctacttgaa tactctcaca	1440
agtagggagc tcactccac tggaacagcc catcttatct ttgaatggtc ttctgccagc	1500
ccattttgag gccagaggtg ctgtcagctc aggtggctct cttttacaat cctaataata	1560
ttgggtaagt tttttgaaa gctaataag ctattgagaa agacctgttg ctagaagttg	1620
ggttgttctg gattttcccc tgaagactta cttattcttc cgtcacatat acaaaagcaa	1680
gacttccagg tagggccagc tcacaagccc aggctggaga tcctaactga gaattttcta	1740
cctgtgttca ttcctaccga gaaaaggaga aaggagctct gaatctgata ggaaaagaag	1800
gctgcctaag gaggagtttt tagtatgtgg cgtatcatgc aagtgctatg ccaagccatg	1860
tctaaatggc ttttaattata tagtaatgca ctctcagtaa tgggggacca gcttaagtat	1920
aattaataga tggttagttg ggtaattctg cttctagtat ttttttact gtgcatacat	1980
gttcatcgta tttccttgga tttctgaatg gctgcagtga cccagatatt gcactaggtc	2040
aaaacattca ggtatagctg acatctctc tatcacatta catcatctc cttataagcc	2100

cagctctgct tttccagat tcttcactg gctccacatc caccctactg gatcttcaga 2160

aggctagagg gcgactctgg tgggtctttt gtatgtttca attaggctct gaaatcttgg 2220

gcaaaatgac aaggggaggg ccaggattcc tctctcaggt cactccagtg ttacttttaa 2280

ttcctagagg gtaaataatga ctcttttctc tatcccaagc caaccaagag cacattctta 2340

aagggaaaagt caacatcttc tctctttttt ttttttttg agacagggtc tcactatgtt 2400

gcccaggctg ctcttgaatt cctgggctca agcagtcctc ccacctacc acagcgtccc 2460

gcgtagctgg gactacaggt gcaagccact atgtccagct agccaactcc tccttgctg 2520

cttttctttt tttttctttt tttagagagg cgcacctatc acccaggctg gagtggagtg 2580

gcacgatctt ggctcactgc aacctcttcc tcttgggtca agcgattctc atgtctcagc 2640

ctctcagta gctaggacta ccggcgtgca ccacctgcc aggcataattt ttatatTTTT 2700

agaattttag aagagatggg atttcatcat gttggccagg ctggtctcga actcctgacc 2760

tcaagtgatc cactgcctt ggcttcccaa ggtgctagga ttacaggcat gagccaccgc 2820

accgggccct ccttgctgt tttcaatct catctgatat gcagagtatt tctgccccac 2880

ccacctacce ccaaaaaaa gctgaagcct atttattga aagtccttgt ttttctact 2940

aattatatag tataccatc attatcttc aaaacaacca tctgctcat aacatctttg 3000

aaaagaaaaa tatatatgtg cagtatttta ttaaagcaac attttattta agaataaagt 3060

cttgtaatt actatatttt agatgcaatg tgatctgaag tttctaattc tggccaact 3120

aaatttctag ctctgtttcc ctaaacaaat aatttggttt ctctgtgcct gcattttccc 3180

tttgagaag aaaagtctc tctcttgagt tgaccgagag tccattagg gatagggaga 3240

cttaaatgca tccacagggg cacaggcaga gttgagcaca taaacggagg ccaaaatca 3300

gcatagaacc agaaagattc agagtggcc aagaatgaac attggctacc agaccacaag 3360

tcagcatgag ttgctctatg gcatcaaatt gcaacttgag agtagatggg cagggtcact 3420

atcaaattaa gcaatcaggg cacacaagt gtagtaacac aacaagacta ggccagctct 3480

ggaatccagt aactcagtg cagcaagggt ttgggttata gttcaagaaa gtctaaacag 3540

agccagtcac agcaccaagg aatgtcaag ggagctattg caggtttctc tgctaagaga 3600

tttatttcat cctgggtgca gggttcgacc tccaaaggcc tcaaatcatc accgtatcaa 3660

tggatttctt gagggtaagc tccgtattt cacacctgaa ctccggagtc tgtatattca 3720

gggaagattg cattctcta ctggatttgg gctctcagag ggcttgttg gaaccaggcc 3780

cctcacagaa tcaaatggtc ccaaccaggg agaaagaaaa tagtctttt tttttttta 3840

atagagatgg gggctctcact atgctgcccc ggctggctctt gaactcctgg gttcaagtga 3900
 tcctcctgcc tcagcctccc aaagtgtctgg gattacagtg tgagccactg cgcttggcca 3960
 gaaatggttt tgatctgtct gaactgaacc ctactgttta ggcatagccc catccttgat 4020
 aatctatittg ctcccaagga ccaagtccaa gatccttaca agaaaggtct gccagaaagt 4080
 aaatactgcc cccactccct gaagtttatg aggttgataa gaaaacataa cagataaagt 4140
 ttattgagtg ctaactttat gcca 4164

<210> 265

<211> 2312

<212> DNA

<213> Homo sapiens

<400> 265

gtcctattcg ggctcccgcc tctgttcagg aactgggtc cccttgagc ctccccaggc 60
 ttaatgattg tccagaaggc ggctataaag ggagcctggg aggcctgggtg gaggaggag 120
 cagaaaaaac ccaactcagc agatctggga actgtgagag cggcaagcag gaactgtggt 180
 cagaggctgt gcgtcttggc tggtagggcc tgccttttc taccatggca gcccagggt 240
 atggctatta tcgactgtg atcttctcag ccatgtttgg gggctacagc ctgtattact 300
 tcaatcgaa gaccttctcc ttgtcatgc catcattggt ggaagagatc cctttggaca 360
 aggatgattt ggggttcatc accagcagcc agtcggcagc ttatgctatc agcaagtttg 420

tcagtgggt gctgtctgac cagatgagtg ctgcctggct cttctcttct gggtgctcc 480
 tggttggcct ggtaacata ttctttgcct ggagctccac agtacctgtc ttgtctgcc 540
 tctggttcct taatggcctg gcccaggggc tgggtggcc ccatgtggg aaggtcctgc 600
 ggaagtgggt tgagccatct cagtttggca ctgtgtggc catcctgtca accagcatga 660
 acctggctgg agggctgggc cctatcctgg caaccatcct tgcccagagc tacagctggc 720
 gcagcacgt gccctatct ggggcactgt gtgtggttgt ctccttctc tgtctcctgc 780
 tcatccaaa tgaacctgct gatgttggac tccgaacct ggaccccatg ccctctgagg 840

gcaagaaggc ctcttgaag gaggagagca cctgcagga gctgctgctg tccccctacc 900
 tgtgggtgct ctccactggt taccttgtgg tgtttggagt aaagacctgc tgtactgact 960
 ggggccagtt cticcttacc caggagaaag gacagtcagc cctttaggt agctcctaca 1020
 tgagtgcctt ggaagtggg gccctttag gcagcatcgc agctggctac ctgtcagacc 1080
 gggccatggc aaaggcggga ctgtccaaact acgggaaccc tcgccatggc ctgttctgt 1140
 tcatgatggc tggcatgaca gtgtccatgt acctcttccg ggtaacagt accagtact 1200

cccccaagct ctggatcctg gtattgggag ctgtatttgg tttctcctcg tatggcccca 1260

 ttgccctgtt tggagtcata gccaacgaga gtgccctcc caacttgtgt ggcacctccc 1320
 acgccattgt gggactcatg gccaatgtgg gcggctttct ggctgggctg ccttcageca 1380
 ccattgccaa gcactacagt tggagcacag ccttctgggt ggctgaagtg atttgtcgcg 1440
 ccagcacggc tgccttcttc ctctacgaa acatccgcac caagatgggc cgagtgtcca 1500
 agaaggctga gtgaagagag tccaggttcc ggagcaccat cccacggtgg cttccccct 1560
 gcacgctctg cggggagaaa aggaggggcc tgcctggcta gccctgaacc tttcactttc 1620
 ctttctgcg cttttctgt caccgggtg gcgctggaag ttatcagtgg ctagtagggt 1680

 cccagctccc tgatcciatg ctctatttaa aagataacct ttggccttag actccgttag 1740
 ctctatttc ctgccttcag acaaacagga aacttctgca gtcaggaagg ctctgtacc 1800
 cttctctttt tctaggccc tgtctgccc gcatcctacc ccatccccac ctgaagttag 1860
 gctatccctg cagctgcagg gcaataatga ccttgactt ctgctgggtc ctaagtcctc 1920
 tcagcagtgg gtgactgctg ttgccaatac ctgagactcc agggaaagag aggaggccat 1980
 cattctcact gtaccactag gcgcagttgg atataggtgg gaagaaaagg tgacttgta 2040
 tagaagatta aaactagatt tgatactgaa cactgtcagt gattcatttt ttcaaagtga 2100

 gacagcttcc ttgggaaata ttgtcaatac ctgctcttcc caccceaaaa tggaaagact 2160
 tcatttccct ggaatgggga agctgaagtg tagatggact cttttaaacc tcatgcctcc 2220
 tctgccttca cctgaacatt gaatagttca atgattattt tagagataaa gctattgggt 2280
 tgtggacaga ttaaataact tgatggaggg aa 2312

 <210> 266
 <211> 3325
 <212> DNA
 <213> Homo sapiens
 <400> 266

 gcgcattgcg cttcttggga cccacggcag gcgcgaatcc cagcgccgg cggcgccgg 60
 ggatacttct acatagacat aatcaagttt tgactatttg gaaaccaagc atcattaaaa 120

 ttctctcaaa ctctaattg cgaagaatcg ataacatttc aagaagtgat aacatttctc 180
 tgaacaagaa aagaagtgat tgaccacgtt ttaaaagtac tctggcactg gtgctgtgtt 240
 ttcttccct ccctaaattt gaagaactat ggagaaatgg tacttgatga cagtagtggt 300
 tttaatagga ctaacagtac gatggacagt gtctcttaat tcttattcag gtgctggtaa 360
 accgcctatg ttgggtgatt atgaagctca gagacactgg caagaaataa cttttaattt 420

accggtcaaa caatggtatt ttaacagcag tgataacaat ttacagtatt ggggattgga	480
ttaccacact cttacagctt atcatagtct cctatgtgca tatgtggcaa agtttataaa	540
tccagactgg attgctctcc atacatcacg tggatatgag agtcaggcac ataagctctt	600
catgcgtaca acagttttaa ttgctgatct gctgatttac atacctgcag tggttttgta	660
ctgtttgtgc ttaaaagaaa tctcaactaa gaaaaagatt gctaatgcat tatgcatctt	720
gctgtatcca ggccttattc ttatagacta tggacatttt caatataatt ctgtgagtct	780
tggctttgct ttgtggggtg ttcttggaa atcttgtgac tgcgacctcc tagggtcact	840
ggcattttgc ttagctataa attataaaca gatggaactt taccacgcct tgccattttt	900
ttgcttttta ctggcaagt gttttaaaaa aggcctcaaa ggaaaggggt ttgtgttgc	960
agttaagcta gcttgtattg ttgtggcttc ctctgtctc tgctggctgc cattctttac	1020
agaaaggga caaacctgc aggttctaag aagactcttc ccggttgatc gtggattatt	1080
tgaggataaa gtagccaata ttgtgtgcag ctccaatgct tttctgaaga ttaaggatat	1140
tttgccacgt cacatccaat taataatgag cttttgttct acgtttttga gcctgcttcc	1200
tgcattgcata aaattaatac ttacgccctc ttccaaagga ttcaaattta cactggttag	1260
ctgtgcgcta tcattctttt tattttcttt ccaagtacat gaaaaatcca ttctcttgg	1320
gtcactacca gtctgcttag ttttaagtga aattcctttt atgtctactt ggtttttact	1380
tgtgtcaaca ttiagtatgc tacctcttct attgaaggat gaactcctaa tgccctctgt	1440
tgtgacaaca atggcatttt ttatagcttg tgtaacttcc ttttcaatat ttgaaaagac	1500
ttctgaagaa gaactgcagt tgaaatcctt ttccatttct gtgaggaaat atcttccatg	1560
ttttacattt ctttccagaa ttatacaata ttgttttctt atctcagtca tcactatggt	1620
gcttctgacg ttgatgactg tcacactgga tctctctcag aaactaccgg acttgtttcc	1680
tgtattggtg tgttttgtat ctgtcttgaa ctctctgttc ttcttggtat actttaacat	1740
tattattatg tgggattcca aaagtgaag aatcagaag aaaatcagct agctgtattc	1800
ctaaacaaat tgtttccctaa acaaatgtga aaatgtgaac agtgctgaaa ggttttgtga	1860
actttttgct atgtataaat gaaattacca ttttgagaac catggaacca caggaaagga	1920
aatggtgaaa agtcattgtt gtctacacaa aataaatgta tatggagacc aaagaccaat	1980
ggaggcttgt ctaagtactg ctggccaaa acatgtggtt ttattattca cagctattca	2040
agtggggaaa aaagagaac atgtccttaa tccattgtt tactcaggaa atgcactttt	2100
taaatgtctt taaaaaactg agaaatattt ctgatattt gctttgtcta aacctacttc	2160

cttcatcaga ccatatagtg agcttcatgg gagaattgag ccccttcttt ttaatggaaa 2220

tcactaattt tggatttcaa acttacattt tagaatgcaa gtagtgtgta tttcactgtt 2280

tagaagtcac atcttaggga attaaatttg tgataatctt ttttaattact gagaaggcaa 2340

tttaactgc ttgaagtttg aatatgcctt attacacatg caaatttgat tgttttaaca 2400

aggcaataaa accataatct tttttatttc caagaatgca tttttaata caattaaaaa 2460

tgaattttcc tgaatttaaa tgtaagcttt aacttaactt taagtgggtt gagtgaagtc 2520

ttctaaatta cacataatca ctggttctaa acttcagaa ccattctgtt taattaaaca 2580

ttatattata tatatattat tattgacatg gttcagggtt accactgccca gagcattgat 2640

tcattgttaac ttcacacctat taatactgia tcaccctgta ttagcagaag ccaatattta 2700

ttatagtatt caactggtaa aaataataca tgatatccat gttgcagttt ctcataataa 2760

gatatgtttt ctttagcctc tgtaataaag ataatcattt tcccatcaac attttttaag 2820

gaatatattt ctcttttcat aagatagcca ttaaactata ttaagtgggtg gtatatattta 2880

atatagtgca aatcaaacca catacaaaaa aaaactgggt ctagacacgg acagtcacag 2940

gctgatataa gaaggatatg gtttctaagt ctggtaaagg aagtgggtt agacaaagca 3000

aaaatctaga aaggtttctt tctcaaagca aatgtgcata aagactcaaa gggcagggtca 3060

acgcagaata gcagccttga atcactgccg aatcagcaag aacattgctg tccatgggtca 3120

ccatctgtaa tcaactgtgcc caaatgtcag catcttggct ggctgggtga gctttatttt 3180

taaaaatatt ttttgaata gtacacttgt gtttaacata gaatgttgac ttagataaat 3240

gatgaggaag ttataataaa ttggagtagc ataactttat tatttggaaa taaaagtatt 3300

atgtacatat attagtaaaa tgaaa 3325

<210> 267

<211> 1552

<212> DNA

<213> Homo sapiens

<400> 267

acacaagcgg cgcaccgtta agatggcggc tgggctgcgg aaacgcggcc ggtccggttc 60

cgcgccccag gcagagggac tctgcaagca atggctgcag cgcgcttggc aagagcggcg 120

cctgctgctg cgggagccgc gctacacgct gctgggtggc gccctgctct gcctggcggc 180

ggtgggcatac accttctggg tcattcacag ggtggcatac acagagattg actggaaggc 240

ctacatggcc gaggtagaag gcgtcatcaa tggtaacctat gactataccc aactgcaggg 300

tgacaccgga ccacttgtgt acccagctgg tttcgtgtac atctttatgg ggttgtacta 360

tgccaccagc cgaggcactg acatccgcat ggcccagaac atctttgctg tgctctacct 420
ggctaccttg ctgcttgctt tcttgatcta tcaccagacc tgcaaggtac ctcccttcgt 480

ctttttcttc atgtgctgcg cctcttaccg tgccactcc atctttgtgc tgcggtcttt 540
caatgacca gtggccatgg tgctgctctt cctcagatc aacctcctgc tggcccagcg 600
ctggggctgg ggttgctgct ttttcagcct ggcagtctct gtgaagatga atgtgctgct 660
cttcgcccc gggttactgt tttctctect cacacagttt ggcttcctg gggccctccc 720
caagctggga atctgtgctg gccttcaggt ggtgctgggg ctgcccttcc tgctggagaa 780
ccccagcggc tacctgtccc gctcctttga ccttgccgc cagtctctgt tccactggac 840
agtgaactgg cgttctctcc cagaggcgct ctctctgcat cgagccttcc acctggccct 900

gttgactgcc cacctacccc tgctcctgct gtttgccctc tgcaggtggc acaggacagg 960
ggaaagtatc ttgtcgtgc tgagggatcc ctccaaaagg aaggttccac cccagcccc 1020
tacaccaac cagatcgttt ctacctctt cacctccaac ttcatggca tctgcttcag 1080
ccgtccctc cactaccagt tctacgtctg gtatttccac acactgccct acctcctgtg 1140
ggcatgctt gcacgtggc tcacacacct gctcaggttg ttggtgctgg ggctcatcga 1200
gctctcctgg aacacatacc ctccacatc ctgcagctct gctgccctgc acatatgcca 1260
tgccgtcatc ctgctgcagc tctggctggg cccgcagcct tcccccaaga gcaccaaca 1320

cagcaagaaa gccactgaa gtccaccct ttcctcagg acctgagtct acctcagga 1380
cctggggttg gttggactct gcccttccaa ataaaccttg ctaagtccaa ctctgtgcaa 1440
cctacatgga ggtgggggca gccgatgctt ggtccaggct gtgagggaca cgtatggagc 1500
agataaagaa ttactcaag cagtcttaag aggcacttta ttgcatagga tt 1552

<210> 268

<211> 2683

<212> DNA

<213> Homo sapiens

<400> 268

gcagttgcgg gttgtcataa cgggtccccgc cggagtgagg cgaggccgcg tcgctcagtt 60
ctggccgtct agggcccctg taaggatgag agcgcagagg acgcagggcc gctggaggcg 120

caggtaacga agctagggtg cggttgggac cgcggctgag ctttttccgg gacctgtgt 180
gctgaatgga gaggacggag acgaagccga gccgcggctc ctagcggcgg gccgatgct 240
cgagctgtag ctgccaggcg aggatgtgtg gagcgaggc ggcgcggggt aaatgagagg 300
tctcggggcc caggaccccc gggggccggg atgagttagc gagggcagcc gcgggggcca 360

gttccgaccg tgacaggcca aggcgacggc cgccgccgc cgcgcccttc cgtgcagaag	420
cagctgctcc ttcccgcc cgcccctg cgtcccgcc cctggagacc atgaggttc	480
gcatctacaa acggaagggt ctaatcctga cgtcgtggt ggccgcctgc ggcttcgtcc	540
tctggagcag caatggcgca caaaggaaga acgaggccct cgcgccaccg ttgctggacg	600
ccgaaccgc gcgggggtgc ggccggccgc gtggggacca cccctctgtg gctgtgggca	660
tccgcagggt ctccaacgtg tcggcggtt ccttggtccc ggcggtcccc cagcccagg	720
cggacaacct gacgtcgcggt taccggtccc tgggtgacca gctgaacttt gatcagacc	780
tgaggaatgt agataaggct ggcacctggg cccccggga gctggtgctg gtggtccagg	840
tgcataaccg gccgaatac ctgagactgc tctggactc acttcgaaa gcccgaggaa	900
ttgacaacgt cctcgtcatc tttagccatg acttctggtc gaccgagatc aatcagctga	960
tcgccgggt gaatttctgt ccggttctgc aggtgttctt tctttcagc attcagttgt	1020
accctaacga gttccagggt agtgacctta gagattgtcc cagagacctg ccgaagaatg	1080
ccgctttgaa attgggggtgc atcaatgctg agtatccga ctccttcggc cattatagag	1140
aggccaaatt ctccagacc aaacatcact ggtggtggaa gctgcatttt gtgtgggaaa	1200
gagtgaataa tcttcgagat tatgtggcc ttatactttt cctagaagag gatcactact	1260
tagccccaga cttttacat gtcttcaaaa agatgtggaa actgaagcag caagagtgc	1320
ctgaatgtga tgttctctcc ctggggacct atagtgcag tcgcagtttc tatggcatgg	1380
ctgacaagggt agatgtgaaa acttggaat ccacagagca caatatgggt ctgacctga	1440
cccggaatgc ctatcagaag ctgacgagt gcacagacac tttctgtact tatgatgatt	1500
ataactggga ctggactctt caatacttga ctgtatcttg tcttcaaaa ttctggaaag	1560
tgctggttcc tcaaattcct aggatcttct atgctggaga ctgtggtatg catcacaaga	1620
aaacctgtag accatccact cagagtgcc aaattgagtc actcttaaat aatacaaac	1680
aatacatgtt tccagaaact ctaactatca gtgaaaagtt tactgtgta gccatttccc	1740
cacctagaaa aaatggaggg tggggagata ttagggacca tgaactctgt aaaagttata	1800
gaagactgca gtgaaaatca cagtiacaaa agcgacagtc ttctatttt gatatttgc	1860
caaacaggac atacaattga ataaaagagt ttaggaactg gtttctgctt taatacaaaa	1920
acaaaatctt gtaaaagggt tccaaataca tagtaatctt ttccagttat gtctgattaa	1980
gatttaaac tgaaggtttc attttgggag tagggtttta aagctcaatc tgttatctgc	2040
taaaattgat tattgttgat atgagagaag aggggaaatt ttatttaaat tgcatttatt	2100
aatcttttta tctgaaactt tgtacacttt tccactttca aaacctattt taagtacagc	2160

aaaatttatt taaaactgtc atagcagtaa aaagtattac gatgaaattg ttaggttatt 2220

 aatggaacaa acccagtttc actctcttga cacacttatt aggaagggat tgcttcactg 2280
 gttaataat ttaaaagtta tgtttgtaa acaccctgtc agaacagtca ttttcagtat 2340
 tagattcctg tactattgtg ttttgagtgt gttttggaac cttcatagaa cacactttct 2400
 tttggaatgt atttgattga taagaaagt taaacattgt tttcacctca atgtagaaat 2460
 acagtgggtt tgttttttt tttcttttag tgcagacaaa ataaaatact catTTTTgca 2520
 taaaaaggtt cctaactctt ttgcagaata agttttgttt actctttata ccaaaattca 2580
 gtgaaggcat tctacaagt ttgagttagc attacatttt aatatttact attgctacat 2640

 tgtataattg agtttgaaat aaaaccagc ttatgacaat gca 2683
 <210> 269
 <211> 2867
 <212> DNA
 <213> Homo sapiens
 <400> 269

 gcaggcgctg gctggcaggt gtcgctaacc ggacggtggg cgccaggcg agaggcgga 60
 gccggagagg tgaggcagga cccgggctcc actgccct ctcgagctc ttgtgacgcg 120
 gacctcagt ccaggatggc tcggggcgag cggcgcgcc gcgcagtgcc ggcagaggga 180
 gtgcggacag ccgagagggc ggctcgggga ggccccgggc gacgggacgg ccggggcggc 240
 gggccgcgta gcacggctgg aggagtggct ctggccgtcg tggctcctgtc tttggccctg 300

 ggtatgtcgg ggcgctgggt gctggcgtgg taccgtgcgc ggcgggcggt cacgctgcac 360
 tccgcgcctc ctgtgttggc tgccgactcc tccagccccg ccgtggcccc ggacctcttc 420
 tggggaacct accgacctca cgtctacttc ggcatgaaga cccgcagccc gaagccctc 480
 ctaccggac tgatgtgggc gcagcagggc accaccccg ggactcctaa gtcaggcac 540
 acgtgtgagc agggggacgg tgtgggtccc tatggctggg agttccacga cggcctctcc 600
 ttcggcgcc aacacatcca ggatggggcc ttaaggetca cactgagtt cgtcaagagg 660
 cctgggggtc agcacggagg ggactggagc tggagagtga ctgtagagcc tcaggactca 720

 ggtacttctg cctcccttt ggtctcctg ttcttctatg tggtagaca tggcaaggaa 780
 gtccactac cagaggttgg ggccaagggg cagttgaagt ttatcagtgg gcacaccagt 840
 gaacttgggt acttccgtt tacacttttg ccaccaacca gtccaggga tacagcccc 900
 aagtatggca gctacaatgt cttctggacc tccaaccag gactgccct gctgacagag 960
 atggtaaaga gtcgcctaaa tagctgggtt cagcatcggc cccaggggc cccctgaa 1020

cgctacctcg gcttgccagg atccctgaag tgggaggaca gaggtccaag tgggcaaggg 1080
caggggcagt tcttgatata gcaggtgacc ctgaaaattc ccatttccat agagtttgtg 1140

tttgaatcag gcagtgccca ggcaggagga aatcaagccc tgccaagact ggcaggcagt 1200
ctactgacct aggccctgga gagccatgct gaaggcttta gagagcgctt tgagaagacc 1260
ttccagctga aggagaaggg cctgagctct ggcgagcagg ttttgggtca ggctgccctc 1320
agcggcctcc ttggtggaat tggctacttc tacggacaag ggctgggtatt gccagacatc 1380
ggggtggaag ggtctgagca gaaggtggac ccagccctct ttccaccctg acctcttttt 1440
acagcagtgc cctcccgtc attcttccca cgaggcttcc tttgggatga aggttttcac 1500
cagctgggtg ttacagcgtg ggatccctcc ctacccggg aagcccttgg ccactggctg 1560

gggctgctaa atgctgatgg ctggattggg agggagcaga tactggggga tgaggcccga 1620
gccccgggtg ctccagaatt cctagtataa cgagcagtc acgccaaccc cccaacccta 1680
cttttgctg tagcccatat gctagagggt ggtgacctg acgacttggc tttcttccga 1740
aaggccttgc ccgcctgca tgcctggttt tcttggctcc atcagagcca ggcaggccca 1800
ctgccactat cttaccgtg gcggggacgg gacctgcct taccaacctt actgaacccc 1860
aagacctac ccttgggtg ggatgactac ccccggtct cacaccttc agtaaccgag 1920
cggcacctgg acctgcgatg ttgggtggca ctgggtgccc gtgtgctgac gcggctggca 1980

gagcatctgg gtgaggctga ggtagctgct gagctgggcc cactggctgc ctactggag 2040
gcagcagaga gcctggaatg gctgactggt gccccagagc taggagtctt tgcagacttt 2100
gggaaccaca caaaagcagt acagctgaag cccaggcccc ctacggggct cgttcgggtg 2160
gtgggtcggc cccaacctca actgcagtat gtagatgctc ttggctatgt cagtcttttt 2220
cccttctgct tgcgactgct ggaccccacc tcattccgcc ttggggccct gctggacatt 2280
ctagccgaca gccgcatct ctggagcccc tttggtttac gctcccttgc agcctccagc 2340
tccttttatg gccagcgcaa ttcagagcat gatccccct actggcgggg tgctgtgtgg 2400

ctcaatgta actacctggc tttgggagca ctccaccact atgggcatct ggagggtcct 2460
caccaggctc gggctgccaa actccacggt gagctccgtg ccaacgtggt aggcaatgta 2520
tggcgccagt accaggctac aggttttctt tgggagcagt acagtaccg cgatgggcga 2580
ggcatgggt gcccctttt ccacggctgg accagccttg tcttactggc catggctgaa 2640
gactactgaa gggagggaga ggaggggagc caagacactc atgccactct ggctctgaag 2700
ggacaaaggc ttttggcttt tgccccagc cccttgata ccagtaattc aaaccttct 2760

catttcatct cagggtgtctc cttgctgtca tcccacatag ccctgggggtg aatgtgaatc 2820

cagagtctat ttttctaaat aaattggaaa aaacattttg aactcta 2867

<210> 270

<211> 280

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 270

ccctggagag tccttttagca gggcaaagtg caacataggc agaccttaag ggatgactca 60

gtaacagata agcttttgtgt gcctgcaggg ggatgggaag aggggtggggc aggagaggga 120

cataaaaggg ctctgaggca ttgtactgtg aattccttca gtctcctgtt tggagaagac 180

agagccaatg aggccctcgt tccagggaag cagaatatgc tcagcatgac gcagcactcc 240

ctgaactttc cggttacatc acccaatagc tgagatcaga 280

<210> 271

<211> 241

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 271

ccctggagag tccttttagca gggcaaagtg caacataggc agaccttaag ggatgactca 60

gtaacagata agcttttgtgt gcctgcaggg ggatgggaag aggggtggggc aggagaggga 120

cataaaaggg ctctgaggca ttgtactgtg aattccttca gtctcctgct ctgctcagcc 180

agtcagccct gcctcccttg tttaggacca cacagcactg ctgggtgtct gcctttcctt 240

g 241

<210> 272

<400> 272

000

<210> 273

<211> 212

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 273

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggtaataat ctccaggagca caaacattcc tttggagaag acagagccaa	120
tgaggccctc gtccaggga aacagaatat gctcagcatg acgcagcact ccctgaactt	180
tccggttaca tcaccaata gctgagatca ga	212

<210> 274

<211> 284

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 274

cactgaaccc ttgaccctg ccctgcagcc cccgcagctt gctgtttgcc cactctattt	60
gcccagcccc agaggttaat ttttaaaaag cagtcaaaag tccaagtggc ccttggcagc	120
atttactctc tctgtttgct ctggtaata atctcaggag cacaacatt ctttggaga	180
agacagagcc aatgaggccc tcgtccagg gaaacagaat atgctcagca tgacgcagca	240
ctccctgaac tttccggtta catcacccaa tagctgagat caga	284

<210> 275

<211> 240

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 275

cactgaaccc ttgaccctg ccctgcagcc cccgcagctt gctgtttgcc cactctattt	60
gcccagcccc agccctggag agtccttttag cagggcaaag tgcaacatag gcagacctta	120
agggatgact cagtaacaga taagctttgt gtgcctgcag ggggatggga agagggtggg	180
gcaggagagg gacataaaag ggctctgagg cattgtactg tgaattcctt cagtctcctg	240

<210> 276

<211> 241

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 276

ccctggagag tccttttagca gggcaaagtg caacataggc agaccttaag ggatgactca	60
gtaacagata agcttttgtgt gcctgcaggg ggatgggaag aggggtggggc aggagaggga	120
cataaaaggg ctctgaggca ttgtactgtg aattccttca gtctcctgct ctgctcagcc	180
agtcagccct gcctcccttg tttaggacca cacagcactg ctgggtgtct gcctttcctt	240
g	241

<210> 277

<211> 168

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 277

ccctggagag tccttttagca gggcaaagtg caacataggc agaccttaag ggatgactca	60
gtaacagata agcttttgtgt gcctgcaggg ggatgggaag aggggtggggc aggagaggga	120
cataaaaggg ctctgaggca ttgtactgtg aattccttca gtctcctg	168

<210> 278

<211> 174

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 278

ccctggagag tccttttagca gggcaaagtg caacataggc agaccttaag ggatgactca	60
gtaacagata agcttttgtgt gcctgcaccc tggagagtcc tttagcaggg caaagtgcaa	120
cataggcaga ccttaaggga tgactcagta acagataagc tttgtgtgcc tgca	174

<210> 279

<211> 159

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 279

cactgaaccc ttgaccctg ccctgcagcc cccgcagctt gctgtttgcc cactctat	60
ttt gccagcccc agccctggag agtccttttag cagggcaaag tgcaacatag gcagac	120
ctta agggatgact cagtaacaga taagctttgt gtgcctgca	159

<210> 280

<211> 207

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 280

agaatgaaca ttgaactttg gactatacct gaggggtgag gtaacaaca ggactataaa	60
tgcccgaggga ggcgcccttt ggaccttttg caatcctggc gcactgaacc ctgaccct	120
gccctgcagc cccgcagct tgctgtttgc ccactctatt tgcccagccc caggacaga	180
gctgacctt gaactcttaa gttccac	207

<210> 281

<211> 313

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 281

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactcttc	60
tgtttctct ggtaataat ctgaggagca caaacattcc tgtaccagaa tgaacattga	120
actttggact atacctgagg ggtgaggtaa acaacaggac tataaatggc ccgggaggcg	180
ccctttggac cttttgcaat cctggcgcac tgaacccttg acccctgccc tgcagcccc	240
gcagcttgct gtttgeccac tctatttggc cagccccagg gacagagctg atccttgaac	300
tcttaagttc cac	313

<210> 282

<211> 269

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 282

aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccctgta	60
ccagaatgaa cattgaactt tggactatac ctgaggggtg aggtaaacaa caggactata	120
aatggcccgg gaggcgcctt ttggaccttt tgcaatcctg gcgcactgaa cccttgaccc	180
ctgccctgca gccccgcag cttgtctgtt gccactcta ttgtcccagc cccagggaca	240
gagctgatcc ttgaactctt aagttccac	269

<210> 283

<211> 258

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 283

ggcgcccttt ggaccttttg caatcctgga gcaaacagca aacctgtac cagaatgaac	60
attgaacttt ggactatacc tgaggggtga ggtaaacac aggactataa atggcccggg	120
aggcgccctt tggacctttt gcaatcctgg cgactgaac ccttgacccc tgcctgcag	180
ccccgcagc ttgtgtttg cccactctat ttgccagcc ccaggacag agctgatcct	240
tgaactctta agttccac	258

<210> 284

<211> 194

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 284

agaatgaaca ttgaactttg gactatacct gaggggtgag gtaacaaca ggactataaa	60
tagaatgaac attgaacttt ggactatacc tgaggggtga ggtaaacac aggactataa	120

atcactgaac ccttgacccc tgccctgcag cccccgcagc ttgctgtttg cccactctat 180
 ttgccagcc ccag 194
 <210> 285
 <211> 199
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<400> 285
 ggccccgggag gcgccccttg gaccttttgc aatcctggcg cactgaaccc ttgaccctg 60
 ccctgcagcc cccgcagctt gctgtttgcc cactctatctt gccagcccc agccctggag 120
 agtccttttag cagggcaaag tgcaacatag gcagacctta agggatgact cagtaacaga 180
 taagctttgt gtgcctgca 199
 <210> 286
 <211> 260
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide
 <400> 286
 agaatgaaca ttgaactttg gactatacct gaggggtgag gtaacaaca ggactataaa 60
 tggccccgga ggcgcccttt ggaccttttg caatcctggc gactgaacc cttgaccct 120
 gccctgcagc cccgcagctt tgctgtttgc ccactctatt tgcccagccc cagccctgga 180
 gagtccttta gcagggcaaa gtgcaacata ggcagacctt aaggatgac tcagtaacag 240
 ataagctttg tgtgcctgca 260
 <210> 287
 <211> 101
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide
 <400> 287

agaatgaaca ttgaactttg gactatacct gaggggtgag gtaacaaca ggactataaa 60

tggcccggga ggcgcccttt ggaccttttg caatcctggc g 101

<210> 288

<211> 201

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 288

gctggtttct tataaaactg atggaagata caaacactat taaagaactg ttgcatgtt 60

gcaaagtatg tccaaagtcc aaacattgtt aataattaat actccaataa acatcatgtc 120

agaatttctg ttttcttttc ctttgaacc ttgacaggat tgccacatca tcaggaccac 180

accttcatca ggaatgaata t 201

<210> 289

<211> 232

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 289

attggcatct tctattgctt ttctgggtga cttcattttt cactcttggc taaaaatggg 60

tctctgatga tttattctat cctgggtgtt gacaagtga agaagttgtg tggggcctgc 120

tgccagtaac cctgggtgac gaagcgtgac tcaccactcc gaggtcagtg gggggatgga 180

aggcagggga gtcagctgac aagatctgct gctttgtcac caggccttct gc 232

<210> 290

<211> 200

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 290

tgaaatgcct gccatatatt agtgcctga agtccaaagg tagaggaacc gagtgtttaa 60

aaattactgt ggctgtggag tcaacatgat gtaaaaaaac aaacatttgg ataacaccaa 120

gaagccagat atggttgaaa tgttgactgg ttgacaaaaa taatttgggt tgcttaatgg 180
 tgcacaaagg taatgcaaaa 200
 <210> 291
 <211> 103
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<400> 291
 caggctttca ctttctgcc aacttacaag gcctttctgt gtaaacaata cctgaacctt 60
 taccgcgttg cccggcaacg gccaggtctg tgccaagtgt ttg 103
 <210> 292
 <211> 402
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<400> 292
 caggctttca ctttctgcc aacttacaag gcctttctgt gtaaacaata cctgaacctt 60
 taccgcgttg cccggcaacg gccaggtctg tgccaagtgt ttgaggttaa tttttaaaaa 120

gcagtcaaaa gtccaagtgg cccttggcag catttactct ctctgtttgc tctggttaat 180
 aatctcagga gcacaaacat tccgccccgg gaggcgcct ttggaccttt tgcaatcctg 240
 gcgcactgaa cccttgacct ctgccctgca gccccgcag cttgctgttt gccactcta 300
 ttgcccagc ccagccctg gagagtcctt tagcagggca aagtgaaca taggcagacc 360
 ttaagggatg actcagtaac agataagctt tgtgtgcctg ca 402
 <210> 293
 <211> 103
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<400> 293

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60
tacccegttg cccggcaacg gccaggtctg tgccaagtgt ttg 103
<210> 294
<211> 299
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 294
aggttaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
tgtttgctct ggTTaataat ctCaggagca caaacattcc ggcccgggag gcgccctttg 120
gaccttttgc aatcctggcg cactgaaccc ttgacccttg cctgcagcc cccgcagctt 180

gctgtttgcc cactctatTT gccagcccc agccctggag agtcctttag cagggcaaag 240
tgcaacatag gcagacctta agggatgact cagtaacaga taagctttgt gtgcctgca 299
<210> 295
<211> 212
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 295
aggttaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
tgtttgctct ggTTaataat ctCaggagca caaacattcc ggcccgggag gcgccctttg 120
gaccttttgc aatcctggcg cactgaaccc ttgacccttg cctgcagcc cccgcagctt 180

gctgtttgcc cactctatTT gccagcccc ag 212
<210> 296
<211> 499
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 296
tgaaatgcct gccatatatt agtgcctga agtccaaagg tagaggaacc gagtgtttaa 60

aaattactgt ggctgtggag tcaacatgat gtaaaaaaac aaacatttgg ataacaccaa 120
gaagccagat atggttgaaa tgttgactgg ttgacaaaaa taatttgggt tgcttaatgg 180
tgacaaaagg taatgcaaaa aggttaattt ttaaaaagca gtcaaaagtc caagtggccc 240

ttggcagcat ttactctctc tgtttgctct ggtaataat ctgaggagca caaacattcc 300
ggcccgggag gcgccctttg gaccttttgc aatcctggcg cactgaaccc ttgaccttg 360
ccctgcagcc cccgcagctt gctgtttgcc cactctatit gccagcccc agccctggag 420
agtcctttag cagggcaaaag tgcaacatag gcagacctta agggatgact cagtaacaga 480
taagctttgt gtgcctgca 499

<210> 297

<211> 297

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 297

ctttggacct ttgcaatcc tggcgactg aaccttgac ccctgccctg cagccccgc 60
agcttgctgt ttgcccactc tatttgccca gcccagtg aatgcctgcc atatattagt 120
gccctgaagt ccaaaggtag aggaaccgag tgtttaaaaa ttactgtggc tgtggagtca 180
acatgatgta aaaaaacaaa catttgata acaccaagaa gccagatatg gttgaaatgt 240
tgactggttg acaaaaataa tttgggttgc ttaatggtgc acaaagtaa tgcaaaa 297

<210> 298

<211> 212

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 298

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
tgtttgctct ggtaataat ctgaggagca caaacattcc ggcccgggag gcgccctttg 120
gaccttttgc aatcctggcg cactgaaccc ttgaccttg ccctgcagcc cccgcagctt 180
gctgtttgcc cactctatit gccagcccc ag 212

<210> 299

<211> 304

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 299

gctggtttct tataaaactg atggaagata caaacactat taaagaactg ttgcatgtt	60
gcaaatgatg tccaaagtcc aaacattgtt aataattaat actccaataa acatcatgtc	120
agaatttctg ttttcttttc cctttgaacc ttgacaggat tgccacatca tcaggaccac	180
accttcatca ggaatgaata tcaggctttc actttctcgc caacttaca ggcccttctg	240
tgtaaacaat acctgaacct ttaccccggt gcccggaac ggccaggtct gtgccaagt	300
tttg	304

<210> 300

<211> 315

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 300

aggttaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggtaataat ctcaggagca caaacattcc ggcccgggag gcgccctttg	120
gaccttttgc aatcctggcg cactgaacct ttgaccttg ccctgcagcc cccgcagctt	180
gctgtttgcc cactctatTT gcccgacccc agcaggtttt cactttctcg ccaacttaca	240
aggcctttct gtgtaaacaa tacctgaacc ttaccccggt tgcccggcaa cggccaggtc	300
tgtgccaagt gtttg	315

<210> 301

<211> 315

<212> DNA

<213>

> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 301

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
 tgtttctctt ggtaataat ctccaggagca caaacattcc ggcccgggag gcgccctttg 120
 gaccttttgc aatcctggcg cactgaacct ttgacctctg ccctgcagcc cccgcagctt 180
 gctgtttgcc cactctattt gccagcccc agcaggtttt cactttctcg ccaacttaca 240
 aggcctttct gtgtaaacaa tacctgaacc ttaccctgt tgcccggcaa cggccaggtc 300
 tgtgccaagt gtttg 315

<210> 302

<211> 413

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 302

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
 tgtttctctt ggtaataat ctccaggagca caaacattcc ggcccgggag gcgccctttg 120
 gaccttttgc aatcctggcg cactgaacct ttgacctctg ccctgcagcc cccgcagctt 180
 gctgtttgcc cactctattt gccagcccc agcgtgtttt cttataaaac tgatggaaga 240
 tacaaacact attaaagaac tgtttgcatt ttgcaaatga tgtccaaagt ccaaacttg 300

ttaataatta atactccaat aaacatcatg tcagaatttc tgttttcttt tccctttgaa 360

cctttgcagg attgccacat catcaggacc acaccttcat caggaatgaa tat 413

<210> 303

<211> 268

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 303

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
 tgtttctctt ggtaataat ctccaggagca caaacattcc ggcccgggag gcgccctttg 120
 gaccttttgc aatcctggcg cactgaacct ttgacctctg ccctgcagcc cccgcagctt 180

gctgtttgcc cactctattt gccagcccc agaagcaaat atttgtggtt atggattaac 240

tcgaactgtt tgcccactct atttgccc 268

<210> 304

<211> 159

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 304

aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccccagg	60
ctttcacttt ctgcceaact tacaaggcct ttctgtgtaa acaatactg aacctttacc	120
ccgttgcccg gcaacggcca ggtctgtgcc aagtgtttg	159

<210> 305

<211> 159

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 305

aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccccagg	60
ctttcacttt ctgcceaact tacaaggcct ttctgtgtaa acaatactg aacctttacc	120
ccgttgcccg gcaacggcca ggtctgtgcc aagtgtttg	159

<210> 306

<211> 243

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 306

aggttaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggtaataat ctcaggagca caaacattcc ggcccgggag gcgccctttg	120
gaccttttgc aatcctggcg caggctttca ctttctcgcc aacttacaag gcctttctgt	180
gtaaacaata cctgaacctt taccctgttg cccggcaacg gccaggtctg tgccaagtgt	240
ttg	243

<210> 307

<211> 243

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 307

aggTTAattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggTTaataat ctcaggagca caaacattcc ggccccggag gcgccctttg	120
gaccttttgc aatcctggcg caggctttca ctttctcgcc aacttacaag gcctttctgt	180
gtaaacaata cctgaacctt taccctgttg cccggcaacg gccaggtctg tgccaagtgt	240
ttg	243

<210> 308

<211> 335

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 308

attggcatct tctattgctt ttcttggtga cttcattttt cactcttggc taaaaatggg	60
tctctgatga tttattctat cctgggtgtt gacaagctga agaagttgtg tggggcctgc	120
tgccagtaac cctgggtgac gaagcgtgac tcaccactcc gaggtcagtg gggggatgga	180
aggcagggga gtcagctgac aagatctgct gctttgtcac caggccttct gccagctttt	240
cactttctcg ccaacttaca aggcttttct gtgtaaaca tacctgaacc tttaccccg	300
tgccccgcaa cggccaggtc tgtgccaagt gtttg	335

<210> 309

<211>

> 304

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 309

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaaacaata cctgaacctt	60

taccccgcttg cccggcaacg gccaggtctg tgccaagtgt ttggctgggt tcttataaaa 120
ctgatggaag atacaaacac tattaaagaa ctgtttgcat gttgcaaatg atgtccaaag 180
tccaaacatt gttataaatt aatactccaa taaacatcat gtcagaattt ctgttttctt 240
ttccctttga acctttgcag gattgccaca tcatcaggac cacaccttca tcaggaatga 300

atat 304

<210> 310

<211> 334

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 310

ctgtttgctg cttgcaatgt ttgccattt tagggagggt aatttttaaa aagcagtcaa 60
aagtccaagt ggcccttgge agcatttact ctctctgttt gctctggtta ataatctcag 120
gagcacaac attccggccc gggaggcgcc ctttgacat tttgcaatcc tggcgactg 180
aacccttgac ccttgccctg cagcccccgc agcttgctgt ttgccactc tatttgccca 240

gccccagccc tggagagtcc tttagcaggg caaagtgcaa cataggcaga ccttaaggga 300
tgactcagta acagataagc tttgtgtgcc tgca 334

<210> 311

<211> 236

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 311

ctgtttgctg cttgcaatgt ttgccattt taggggctgg tttcttataa aactgatgga 60
agatacaaac actattaaag aactgtttgc atgttgcaaa tgatgtccaa agtccaaaca 120
ttgttaataa ttaatactcc aataaacatc atgtcagaat ttctgttttc ttttcctttt 180

gaacctttgc aggattgcca catcatcagg accacacctt catcaggaat gaatat 236

<210> 312

<211> 103

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 312

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60

taccccgttg cccggcaacg gccaggtctg tgccaagtgt ttg 103

<210> 313

<211> 305

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 313

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60

tgtttgctct ggtaataat ctccaggagca caaacattcc tgtaccggcc cgggaggcgc 120

cctttggacc ttttgcaatc ctggcgcaact gaacccttga cccctgcct gcagcccccg 180

cagcttgctg ttgcccact ctatttggcc agccccagcc ctggagagtc ctttagcagg 240

gcaaagtga acataggcag accttaaggg atgactcagt aacagataag ctttgtgtgc 300

ctgca 305

<210> 314

<211> 261

<212> DNA

<213>

> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 314

aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccctgta 60

ccggccccggg aggcgccctt tggacctttt gcaatcctgg cgcactgaac ccttgacccc 120

tgccctgcag cccccgcagc ttgtgttttg cccactctat ttgccagacc ccagccctgg 180

agagtcttt agcagggcaa agtgcaacat aggagacct taagggatga ctgagtaaca 240

gataagcttt gtgtgcctgc a 261

<210> 315

<211> 250

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 315

ggcgcccttt ggaccttttg caatcctgga gcaaacagca aacactgtac cggcccggga	60
ggcgcccttt ggaccttttg caatcctggc gactgaacc cttgaccct gccctgcagc	120
ccccgcagct tctgtttgc ccactctatt tgcccagccc cagccctgga gagtcttta	180
gcagggcaaa gtgcaacata ggcagacctt aagggatgac tcagtaacag ataagctttg	240
tgtgcctgca	250

<210> 316

<211> 366

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 316

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggtaataat ctcaggagca caaacattcc tgtaccagaa tgaacattga	120
actttggact atacctgagg ggtgaggtaa acaacaggac tataaatggc cggggaggcg	180
ccctttggac cttttgcaat cctggcgac tgaacccttg acccctgccc tgcagcccc	240
gcagcttgct gtttgeccac tctatttgcc cagccccage cctggagagt cctttagcag	300
ggcaaagtgc aacataggca gaccttaagg gatgactcag taacagataa gctttgtgtg	360
cctgca	366

<210> 317

<211> 322

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 317

aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccctgta	60

ccagaatgaa cattgaactt tggactatac ctgaggggtg aggtaaacaa caggactata	120
aatggcccgaggaggcgccct ttggaccttt tgcaatcctg gcgcactgaa cccttgaccc	180
ctgccctgca gccccgcag cttgctgttt gccactcta tttgcccagc ccagccctg	240
gagagtcctt tagcaggga aagtgaaca taggcagacc ttaagggatg actcagtaac	300
agataagctt tgtgtgcctg ca	322
<210> 318	
<211> 311	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 318	
ggcgcccttt ggaccttttg caatcctgga gcaaacagca aacctgtac cagaatgaac	60
attgaacttt ggactatacc tgaggggtga ggtaaacaac aggactataa atggcccgagg	120
aggcgccctt tggacctttt gcaatcctgg cgactgaac ccttgacccc tgccctgcag	180
ccccgcagc ttgctgtttg cccactctat ttgcccagcc ccagccctgg agagtccttt	240
agcaggggcaa agtgaacat aggagacct taagggatga ctcagtaaca gataagcttt	300
gtgtgcctgc a	311
<210> 319	
<211> 207	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 319	
aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgcctt ggtaataat ctcaggagca caaacattcc tgtaccagaa tgaacattga	120
actttggact atacctgagg ggtgaggtaa acaacaggac tataaatggc ccgggaggcg	180
ccctttggac cttttgcaat cctggcg	207
<210> 320	
<211> 163	
<212> DNA	

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 320

aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccctgta 60

ccagaatgaa cattgaactt tggactatac ctgaggggtg aggtaaacaa caggactata 120

aatggcccgg gaggcgcctt ttggaccttt tgcaatcctg gcg 163

<210> 321

<211> 152

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 321

ggcgcccttt ggaccttttg caatcctgga gcaaacagca aacactgtac cagaatgaac 60

attgaacttt ggactatacc tgaggggtga ggtaaacaac aggactataa atggcccggg 120

aggcgccctt tggacctttt gcaatcctgg cg 152

<210> 322

<211> 100

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 322

aggttaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60

tgtttgctct ggtaataat ctcaggagca caaacattcc 100

<210> 323

<211> 146

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 323

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
 tgtttctct ggtaataat ctcaggagca caaacattcc tgtaccggcc cgggaggcgc 120

cctttggacc ttttgaatc ctggcg 146

<210> 324
 <211> 172
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 polynucleotide
 <400> 324
 tgctctctga caaagatacg gtgggtccca ctgatgaact gtgctgccac agtaaatgta 60
 gccactatgc ctatctccat tctgaagatg tgcctgttc aaacatgtcc taatactctg 120
 tctctgcaag ggtcatcagt agttttccat cttactcaac atcctcccag tg 172

<210> 325
 <211> 180
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 polynucleotide
 <400> 325
 cctccccgtg ttctgtctt ttgtccctct gtccactta gactaatatt tgccttgggt 60
 actgcaaaca ggaaatgggg gagggacagg agtagggcgg ccctgttcaa acatgtccta 120
 atactctgtc tctgcaaggg tcatcagtag ttttccatct tactcaacat cctcccagtg 180

<210> 326
 <211> 159
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 polynucleotide
 <400> 326
 ccgccccac tgaaccttg acccctgcc tgcagccccc gcagettgt gtttggccac 60
 tctatttggc cagccccagc cctgttcaaa catgtcctaa tactctgtct ctgcaagggt 120

catcagtagt ttccatctt actcaacatc ctcccagtg 159

<210> 327

<211> 165

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 327

aagctttctg aacagccaaa cagagattcc aaagttcagg caccaaagtt cagaccctaa 60

cagttattta caagggtcag ttaacccctg ttcaaacatg tcctaatact ctgtctctgc 120

aagggtcatc agtagttttc catcttactc aacatcctcc cagtg 165

<210> 328

<211> 140

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 328

aggttaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60

tgtttgctct ggTTaataat ctCaggagca caaacattcc ggcccgggag gcgcctttg 120

gaccttttgc aatcctggcg 140

<210> 329

<211> 100

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 329

aggttaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60

tgtttgctct ggTTaataat ctCaggagca caaacattcc 100

<210> 330

<211> 233

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 330

aggTTaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggTTaataat ctCaggagca caaacattcc tgtaccggcc cgggaggcgc	120
cctttggacc ttttgcaatc ctggcgccct ggagagtcct ttagcagggc aaagtgaac	180

ataggcagac cttaagggat gactcagtaa cagataagct ttgtgtgcct gca	233

<210> 331

<211> 265

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 331

aggTTaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggTTaataat ctCaggagca caaacattcc tgtaccact gaaccttga	120
cccctgccct gcagcccccg cagcttgctg ttgcccact ctatttgccc agccccagcc	180
ctggagagtc ctttagcagg gcaaagtga acataggcag accttaagg atgactcagt	240

aacagataag ctttgtgtgc ctgca	265

<210> 332

<211> 469

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 332

caggctttca ctttctgcc aacttacaag gcctttctgt gtaacaata cctgaacctt	60
taccccgttg cccggcaacg gccaggctctg tgccaagtgt ttgaggTTaa tttttaaaaa	120
gcagtcaaaa gtccaagtgg cccctggcag catTTactct ctctgtttgc tctggttaat	180
aatctcagga gcacaaacat tcctgtacca gaatgaacat tgaactttgg actataacctg	240

aggggtgagg taaacaacag gactataaat ggcccgggag gcgccctttg gaccttttgc	300

aatcctggcg cactgaaccc ttgacccctg ccctgcagcc cccgcagctt gctgtttgcc 360
cactctatatt gccagcccc agccctggag agtccttttag cagggcaaag tgcaacatag 420
gcagacctta agggatgact cagtaacaga taagctttgt gtgcctgca 469

<210> 333

<211> 61

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 333

agaatgaaca ttgaactttg gactatacct gaggggtgag gtaaacaaca ggactataaa 60

t 61

<210> 334

<211> 364

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 334

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaaacaata cctgaacctt 60

taccccgctt gccggcaacg gccaggtctg tgccaagtgt ttgaagcaa ttttgttgtt 120

tatggattaa ctggaactgt ttgcccactc tttttgcct gtaccggccc gggaggcgcc 180

ctttggacct ttgcaatcc tggcgcactg aacccttgac ccctgcctg cagccccgcg 240

agcttctgt ttgcccactc ttttgccca gcccagccc tggagagtcc tttagcaggg 300

caaagtgcaa cataggcaga ccttaaggga tgactcagta acagataagc tttgtgtgcc 360

tgca 364

<210> 335

<211> 425

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 335

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60
taccgccgttg cccggcaacg gccaggtctg tgccaagtgt ttgaagcaaa ttttgtggt 120

tatggattaa ctggaactgt ttgcccactc ttttgcct gtaccagaat gaacattgaa 180
ctttggacta tacctgaggg gtgaggtaaa caacaggact ataatggcc cgggaggcgc 240
cctttggacc ttttgaatc ctggcgact gaaccttga cccctgcct gcagcccccg 300
cagcttgctg ttgcccact ctattgccc agccccagcc ctggagagtc ctttagcagg 360
gcaaagtga acataggcag accttaaggg atgactcagt aacagataag ctttgtgtgc 420
ctgca 425

<210> 336
<211> 338
<212> DNA
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 336

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60
taccgccgttg cccggcaacg gccaggtctg tgccaagtgt ttgaagcaaa ttttgtggt 120
tatggattaa ctggaactgt ttgcccactc ttttgcct gtaccagaat gaacattgaa 180
ctttggacta tacctgaggg gtgaggtaaa caacaggact ataatggcc cgggaggcgc 240
cctttggacc ttttgaatc ctggcgact gaaccttga cccctgcct gcagcccccg 300
cagcttgctg ttgcccact ctattgccc agccccag 338

<210> 337
<211> 140
<212> DNA
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 337

aggtaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
tgtttgctct ggtaataat ctgaggagca caaacattcc ggcccgagg gcgcctttg 120
gaccttttgc aatcctggcg 140

<210> 338

<211> 203

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 338

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggtaataat ctcaggagca caaacattcc caggctttca ctttctcgcc	120
aacttacaag gcctttctgt gtaaacaata cctgaacctt taccccggtg cccggcaacg	180
gccaggtctg tgccaagtgt ttg	203

<210> 339

<211> 180

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 339

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggtaataat ctcaggagca caaacattcc ccctgttcaa acatgtccta	120
atactctgtc tctgcaaggg tcatcagtag ttttccatct tactcaacat cctcccagtg	180

<210> 340

<211> 80

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 340

ccctgttcaa acatgtccta atactctgtc tctgcaaggg tcatcagtag ttttccatct	60
tactcaacat cctcccagtg	80

<210> 341

<211> 146

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 341

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggtaataat ctcaggagca caaacattcc tgtaccggcc cgggaggcgc	120
cctttggacc ttttgaatc ctggcg	146

<210> 342

<211> 356

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 342

ccctggagag tcctttagca gggcaaagtg caacataggc agaccttaag ggatgactca	60
gtaacagata agcttttgtg gcctgcaggg ggatgggaag aggggtggggc aggagaggga	120
cataaaaggg ctctgaggca ttgtactgtg aattccttca gtctcctgtt tggagaagac	180
agagccaatg aggccctcgt tccagggaag cagaatatgc tcagcatgac gcagcactcc	240
ctgaactttc cgtttacatc acccaatagc tgagatcaga gggcatataa aacaggggca	300
aggcacagac tcatagcaga gcaatcacca ccaagcctgg aataactgca gccacc	356

<210> 343

<211> 453

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 343

cctggagagt cctttagcag ggcaaagtgc aacataggca gaccttaagg gatgactcag	60
taacagataa gctttgtgtg cctgcagggg gatgggaaga ggggtggggca ggagaggac	120
ataaaagggc tctgaggcat tgtactgtga attccttcag tctcctgctc tgctcagcca	180
gtcagccctg cctcccttgt ttaggaccac acagcactgc tgggtgtctg cctttccttg	240
ggtaattttt ttttctgggt aatatttagc aagaattctg cagagtgatc aaaaaaatca	300

aatactcagt atttcagaaa tagattaaat aggttacttt tttactgata atgtgaaaga 360
atgatataaa aacttgattt tctcaacaa cattactttc ttttgtaaat gtggtttcta 420

caaagatgaa actactaaaa cttacaggcc acc 453

<210> 344
<211> 288
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 344
aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
tgtttgctct ggtaataat ctccaggagca caaacattcc tttggagaag acagagccaa 120
tgaggccctc gttccaggga aacagaatat gctcagcatg acgcagcact ccttgaactt 180
tccgggttaca tcaccaata gctgagatca gagggcatat aaaacagggg caaggcacag 240

actcatagca gagcaatcac caccaagcct ggaataactg cagccacc 288

<210> 345
<211> 360
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 345
cactgaaccc ttgacccttg ccttcagacc cccgcagctt gctgtttgcc cactctattt 60
gcccagcccc agaggttaat ttttaaaaag cagtcaaaag tccaagtggc ccttggcagc 120
atttactctc tctgtttgct ctggttaata atctcaggag cacaaacatt cctttggaga 180
agacagagcc aatgaggccc tcgttccagg gaaacagaat atgctcagca tgacgcagca 240

ctccctgaac tttccggtta catcacccaa tagctgagat cagagggcat ataaaacagg 300
ggcaaggcac agactcatag cagagcaatc accaccaagc ctggaataac tgcagccacc 360

<210> 346
<211> 316
<212> DNA
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 346

cactgaaccc ttgaccctg ccctgcagcc cccgcagctt gctgtttgcc cactctat	60
gcccagcccc agccctggag agtccttttag cagggcaaag tgcaacatag gcagacctta	120
agggatgact cagtaacaga taagctttgt gtgcctgcag ggggatggga agagggtggg	180
gcaggagagg gacataaaag ggctctgagg cattgtactg tgaattcctt cagtctcctg	240
gggcataata aacaggggca aggcacagac tcatagcaga gcaatcacca ccaagcctgg	300
aataactgca gccacc	316

<210> 347

<211> 411

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 347

ccctggagag tccttttagca gggcaaagtg caacataggc agaccttaag ggatgactca	60
gtaacagata agcttttgtgt gcctgcaggg ggatgggaag aggggtgggc aggagaggga	120
cataaaaggg ctctgaggca ttgtactgtg aattccttca gtctcctgct ctgctcagcc	180
agtcagccct gcctcccttg tttaggacca cacagcactg ctgggtgtct gcctttcctt	240
gtggacttag cccctgtttg ctctccgat aactgggtg accttggtta atattacca	300
gcagcctccc ccgttgcccc tctggatcca ctgcttaaat acggacgagg acaggccct	360
gtctcctcag cttcaggcac caccactgac ctgggacagt gaatcgccac c	411

<210> 348

<211> 338

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 348

ccctggagag tccttttagca gggcaaagtg caacataggc agaccttaag ggatgactca	60
gtaacagata agcttttgtgt gcctgcaggg ggatgggaag aggggtgggc aggagaggga	120

cataaaaggg ctctgaggca ttgtactgtg aattccttca gtctcctgtg gacttagccc 180
ctgtttgctc ctccgataac tggggtgacc ttggtaata ttcaccagca gcctcccccg 240
ttgccctct ggatccactg cttaaatacg gacgaggaca gggccctgtc tcctcagctt 300
caggcaccac cactgacctg ggacagtga tgcgccacc 338

<210> 349

<211> 250

<212> DNA

<213>

> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 349

ccctggagag tccttttagca gggcaaagtg caacataggc agaccttaag ggatgactca 60
gtaacagata agcttttgtg gcctgcaccc tggagagtcc tttagcaggg caaagtgcaa 120
cataggcaga ccttaaggga tgactcagta acagataagc tttgtgtgcc tgcagggcat 180
ataaaacagg ggcaaggcac agatcatag cagagcaatc accaccaagc ctggaataac 240
tgcagccacc 250

<210> 350

<211> 385

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 350

cactgaacc ttgaccctg ccctgcagcc cccgcagctt gctgtttgcc cactctattt 60
gccagcccc agccctggag agtccttttag cagggaag tgcaacatag gcagacctta 120
agggatgact cagtaacaga taagctttgt gtgcctgcac tctttgttt tacatgaagg 180
gtctggcagc caaagcaatc actcaaagt caaaccttat catTTTTTgc tttgttcctc 240
ttggccttgg ttttgtacat cagctttgaa aataccatcc cagggttaat gctggggtta 300

attataact aagagtgtc tagttttgca atacaggaca tgctataaaa atggaaagat 360

gttgctttct gagagatgcg ccacc 385

<210> 351

<211> 283

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 351

agaatgaaca ttgaactttg gactatacct gaggggtgag gtaacaaca ggactataaa	60
tggcccgga ggcgcccttt ggaccttttg caatcctggc gactgaacc cttgaccct	120
gcctgcage cccgcagct tgctgtttgc ccactctatt tgccagccc caggacaga	180

gtgatcctt gaactcttaa gttccacggg catataaac aggggcaagg cacagactca	240
tagcagagca atcaccacca agcctggaat aactgcagcc acc	283

<210> 352

<211> 389

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 352

aggtaattt taaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggtaataat ctcaggagca caaacattcc tgtaccagaa tgaacattga	120
actttggact atacctgagg ggtgaggtaa acaacaggac tataaatggc ccgggaggcg	180

ccctttggac cttttgcaat cctggcgcac tgaacccttg acccctgccc tgcagcccc	240
gcagcttgct gtttgccac tctatttgcc cagccccagg gacagagctg atccttgaac	300
tcttaagttc cacgggcata taaaacaggg gcaaggcaca gactcatagc agagcaatca	360
ccaccaagcc tggaataact gcagccacc	389

<210> 353

<211> 345

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 353

aagcaaatat ttgtggttat ggattaactc gaactgtttg ccactctat ttgccctgta	60

ccagaatgaa cattgaactt tggactatac ctgaggggtg aggtaaacaa caggactata 120
aatggcccgaggcgccct ttggaccttt tgcaatcctg gcgcactgaa cccttgaccc 180
ctgccctgca gccccgcag cttgctgttt gccactcta ttgcccagc cccagggaca 240
gagctgatcc ttgaactctt aagtccacg ggcataataa acaggggcaa ggcacagact 300
catagcagag caatcaccac caagcctgga ataactgcag ccacc 345

<210> 354

<211> 334

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 354

ggcgcccttt ggaccttttg caatcctgga gcaaacagca aacactgtac cagaatgaac 60
attgaacttt ggactataacc tgaggggtga ggtaaacac aggactataa atggcccgagg 120
aggcgccctt tggacctttt gcaatcctgg cgactgaac ccttgacccc tgccctgcag 180
ccccgcagc ttgctgtttg cccactctat ttgccagcc ccagggacag agctgatcct 240
tgaactctta agttccacg gcatataaaa caggggcaag gcacagactc atagcagagc 300
aatcaccacc aagcctggaa taactgcagc cacc 334

<210> 355

<211> 270

<212> DNA

<213>

> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 355

agaatgaaca ttgaactttg gactataact gaggggtgag gtaacaaca ggactataaa 60
tagaatgaac attgaacttt ggactataacc tgaggggtga ggtaaacac aggactataa 120
atcactgaac ccttgacccc tgccctgcag cccccgcagc ttgctgtttg cccactctat 180
ttgccagcc ccaggggcat ataaacagg ggcaaggcac agactcatag cagagcaatc 240
accaccaagc ctggaataac tgcagccacc 270

<210> 356

<211> 275

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 356

ggccccgggag ggcgcccttg gaccttttgc aatcctggcg cactgaaccc ttgaccctg	60
ccctgcagcc cccgcagctt gctgtttgcc cactctatct gcccgacccc agccctggag	120
agtcctttag caggggcaaag tgcaacatag gcagacctta agggatgact cagtaacaga	180
taagctttgt gtgcctgcag ggcatataaa acaggggcaa ggcacagact catagcagag	240
caatcaccac caagcctgga ataactgcag ccacc	275

<210> 357

<211> 336

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 357

agaatgaaca ttgaactttg gactataacct gaggggtgag gtaacaaca ggactataaa	60
tggccccgga ggcgcccttt ggaccttttg caatcctggc gcactgaacc cttgaccct	120
ggcctgcagc ccccgcagct tgctgtttgc ccactctatt tgcccagccc cagccctgga	180
gagtccttta gcagggcaaa gtgcaacata ggcagacctt aagggatgac tcagtaacag	240
ataagctttg tgtgcctgca gggcatataa aacaggggca aggcacagac tcatagcaga	300
gcaatcacca ccaagcctgg aataactgca gccacc	336

<210> 358

<211> 327

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 358

agaatgaaca ttgaactttg gactataacct gaggggtgag gtaacaaca ggactataaa	60
tggccccgga ggcgcccttt ggaccttttg caatcctggc gctcttttgt ttacatgaa	120
gggtctggca gccaaagcaa tcaactcaaag ttcaaactt atcatTTTTT gctttgttcc	180

tcttggcctt ggttttgtac atcagctttg aaaataccat cccagggtta atgctggggt	240
taatttataa ctaagagtgc tctagttttg caatacagga catgctataa aaatggaaag	300
atgttgcttt ctgagagatg cgccacc	327
<210> 359	
<211> 480	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 359	
gctggtttct tataaaactg atggaagata caaacactat taaagaactg ttgcatgtt	60
gcaaatgatg tccaaagtcc aaacattgtt aataattaat actccaataa acatcatgtc	120
agaatttctg ttttcttttc ctttgaacc ttgacaggat tgccacatca tcaggaccac	180
accttcatca ggaatgaata tccgatgacc taatgattct gagcttggca aaggtcttat	240
ctccagctc gccaggccc agtgttccag gaatgtgacc ttgctgcag cagccgctgg	300
agggggcaga ggggatgggc tggaggttga gcaaacagag cagcagaaaa ggcagttcct	360
cttctccagt gccctcctc cctgtctctg cctctccctc ccttctcag gcatcagagc	420
ggagacttca gggagaccag agcccagctt gccaggcact gagctagaag cctgccatg	480
<210> 360	
<211> 511	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 360	
attggcatct tctattgctt ttccgtgtga cttcattttt cactcttggc taaaaatggg	60
tctctgatga tttattctat cctgggtgtt gacaagctga agaagttgtg tggggcctgc	120
tgccagtaac cctgggtgac gaagcgtgac tcaccactcc gaggtcagtg gggggatgga	180
aggcagggga gtcagctgac aagatctgct gctttgtcac caggccttct gcccgatgac	240
ctaattgattc tgagcttggc aaaggtctta tctcccagct cgcccaggcc cagtgttcca	300
ggaatgtgac ctttctgca gcagccgctg gagggggcag aggggatggg ctggaggttg	360

agcaaacaga gcagcagaaa aggcagttcc tcttctccag tgcctcctt ccctgtctct 420

gcctctccct cccttctca ggcatcagag cggagacttc agggagacca gagcccagct 480

tgccaggcac tgagctagaa gccctgccat g 511

<210> 361

<211> 479

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 361

tgaaatgcct gccatatatt agtgcctga agtccaaagg tagaggaacc gagtgtttaa 60

aaattactgt ggctgtggag tcaacatgat gtaaaaaaac aaacatttgg ataacaccaa 120

gaagccagat atggttgaaa tgttgactgg ttgacaaaaa taatttgggt tgcttaatgg 180

tgcacaaagg taatgcaaaa ccgatgacct aatgattctg agcttgcaa aggtcttatac 240

tcccagctcg cccaggccca gtgttccagg aatgtgacct ttgctgcagc agccgctgga 300

gggggcagag gggatgggct ggaggttgag caaacagagc agcagaaaag gcagttcctc 360

ttctccagtg cctccttcc ctgtctctgc ctctcctcc cttctcagg catcagagcg 420

gagacttcag ggagaccaga gccagcttg ccaggcactg agctagaagc cctgccatg 479

<210> 362

<211> 340

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 362

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60

taccctgttg cccggcaacg gccaggtctg tgccaagtgt ttgcctttgc aacagcttat 120

cggaagcaaa caagctgagg ggaattgagc aagaatttct gggataccaa cagcatagga 180

ggaacaaagg acgtagaggg aggggtgact gtctacacag gacaaagcca atgattaacc 240

aaacctcttg cagattttaa taggatggga actaggagtg gcagcaatcc tttctttcag 300

ctggagtgtc cctcaggagc cagccccacc ctagaaaaag 340

<210> 363

<211> 575

<212> DNA

<213>

> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 363

```

aggTTAATTT TAAAAAGCA GTCAAAGTC CAAGTGGCCC TTGGCAGCAT TTACTCTCTC      60
TGTTTGCTCT GGTTAATAAT CTCAGGAGCA CAAACATTCC GGCCCGGGAG GCGCCCTTTG      120
GACCTTTTGC AATCCTGGCG CACTGAACCC TTGACCCCTG CCCTGCAGCC CCCGCAGCTT      180
GCTGTTTGCC CACTCTATTT GCCCAGCCCC AGCCCTGGAG AGTCCTTTAG CAGGGCAAAG      240
TGCAACATAG GCAGACCTTA AGGGATGACT CAGTAACAGA TAAGCTTTGT GTGCCTGCAC      300
CGATGACCTA ATGATTCTGA GCTTGGCAAA GGTCTTATCT CCCAGCTCGC CCAGGCCAG      360

TGTTCCAGGA ATGTGACCTT TGCTGCAGCA GCCGCTGGAG GGGGCAGAGG GGATGGGCTG      420
GAGGTTGAGC AAACAGAGCA GCAGAAAAGG CAGTTCTCTT TCTCCAGTGC CTCCTTCCC      480
TGTCTCTGCC TCTCCCTCCC TTCTCAGGC ATCAGAGCGG AGACTTCAGG GAGACCAGAG      540
CCCAGCTTGC CAGGCCTGA GCTAGAAGCC CTGCC      575

```

<210> 364

<211> 488

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 364

```

aggTTAATTT TAAAAAGCA GTCAAAGTC CAAGTGGCCC TTGGCAGCAT TTACTCTCTC      60

TGTTTGCTCT GGTTAATAAT CTCAGGAGCA CAAACATTCC GGCCCGGGAG GCGCCCTTTG      120
GACCTTTTGC AATCCTGGCG CACTGAACCC TTGACCCCTG CCCTGCAGCC CCCGCAGCTT      180
GCTGTTTGCC CACTCTATTT GCCCAGCCCC AGCCGATGAC CTAATGATTC TGAGCTTGCC      240
AAAGGTCTTA TCTCCAGCT CGCCAGGCC CAGTGTTCCT GGAATGTGAC CTTTGTGCA      300
GCAGCCGCTG GAGGGGGCAG AGGGGATGGG CTGGAGGTTG AGCAAACAGA GCAGCAGAAA      360
AGGCAGTTCC TCTTCTCCAG TGCCCTCTCT CCCTGTCTCT GCCTCTCCCT CCTTCTCTCA      420
GGCATCAGAG CGGAGACTTC AGGGAGACCA GAGCCCAGCT TGCCAGGCAC TGAGCTAGAA      480

```

gccctgcc	488
<210> 365	
<211> 575	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 365	
tgaaatgcct gccatatatt agtgcctga agtccaaagg tagaggaacc gagtgtttaa	60
aaattactgt ggctgtggag tcaacatgat gtaaaaaaac aaacatttgg ataacaccaa	120
gaagccagat atggttgaaa tgttgactgg ttgacaaaaa taatttgggt tgcttaatgg	180
tgacaaaagg taatgcaaaa aggttaattt ttaaaaagca gtcaaaagtc caagtggccc	240
ttggcagcat ttactctctc tgtttgcctt ggtaataat ctcaggagca caaacattcc	300
ggcccgggag gcgcccttg gaccttttgc aatcctggcg cactgaaccc ttgaccctg	360
ccctgcagcc cccgcagctt gctgtttgcc cactctattt gccagcccc agccctggag	420
agtcctttag cagggcaaag tgcaacatag gcagacctta agggatgact cagtaacaga	480
taagctttgt gtgcctgcag ggcatataaa acaggggcaa ggcacagact catagcagag	540
caatcaccac caagcctgga ataactgcag ccacc	575
<210> 366	
<211> 688	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 366	
aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgcctt ggtaataat ctcaggagca caaacattcc ggcccgggag gcgcccttg	120
gaccttttgc aatcctggcg cactgaaccc ttgaccctg ccctgcagcc cccgcagctt	180
gctgtttgcc cactctattt gccagcccc agtgaaatgc ctgcatata ttagtgcctt	240
gaagtccaaa ggtagaggaa ccgagtgttt aaaaattact gtggctgtgg agtcaacatg	300
atgtaaaaaa acaaacattt ggataacacc aagaagccag atatggttga aatgttgact	360
ggttgacaaa aataatttgg gttgcttaat ggtgcacaaa ggtaatgcaa aaccgatgac	420

ctaagtattc tgagcttggc aaaggtctta tctcccagct cgcccaggcc cagtgttcca 480
 ggaatgtgac ctttgcgtca gcagccgctg gagggggcag aggggatggg ctggaggttg 540
 agcaaacaga gcagcagaaa aggcagttcc tcttctccag tgcctcctt ccctgtctct 600
 gcctctccct cccttctca ggcatcagag cggagacttc agggagacca gagcccagct 660
 tgccaggcac tgagctagaa gccctgcc 688

<210> 367

<211> 580

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 367

gctggtttct tataaaactg atggaagata caaacactat taaagaactg ttgcatgtt 60
 gcaaagtatg tccaaagtcc aaacattgtt aataattaat actccaataa acatcatgtc 120
 agaatttctg ttttcttttc ctttgaacc ttgacagat tgccacatca tcaggaccac 180
 accttcatca ggaatgaata tcaggtcttc actttctcgc caacttaca gccctttctg 240
 tgtaaacat acctgaacct ttaccccggt gcccggaac ggccaggtct gtgccaagtg 300
 ttgcccagat acctaagtat tctgagcttg gcaaaggtct tatctcccag ctgcccagg 360

cccagtgttc caggaatgtg acctttgtg cagcagccgc tggagggggc agaggggatg 420
 ggctggaggt tgagcaaca gagcagcaga aaaggcagtt cctcttctcc agtgccttcc 480
 ttccctgtct ctgccttccc ctccttctc caggtatcag agcggagact tcaggagac 540
 cagagcccag cttgccaggc actgagctag aagccctgcc 580

<210> 368

<211> 391

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 368

aggttaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
 tgtttctct ggtaataat ctgaggagca caaacattcc ggcccgggag gcgccctttg 120
 gaccttttg aatcctggcg cactgaacc ttgaccttg ccctgcagcc cccgcagctt 180

gctgtttgcc cactctatTT gcccagcccc agcaggcttt cactttctcg ccaacttaca 240
aggcctttct gtgtaaacaa tacctgaacc ttTaccctgt tgcccggcaa cggccaggtc 300
tgtgccaagt gtttggggca tataaaacag gggcaaggca cagactcata gcagagcaat 360
caccaccaag cctggaataa ctgcagccac c 391

<210> 369

<211> 507

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 369

aggTTaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
tgTTTgtctt ggTTaataat ctCaggagca caaacattcc ggcccgggag gcgccctttg 120
gaccttttgc aatcctggcg cactgaaccc ttgacctctg cctgcagcc cccgcagctt 180
gctgtttgcc cactctatTT gcccagcccc agcaggcttt cactttctcg ccaacttaca 240
aggcctttct gtgtaaacaa tacctgaacc ttTaccctgt tgcccggcaa cggccaggtc 300
tgtgccaagt gtttggggcg actcagatcc cagccagtgg acttagcccc tgTTTgtctc 360

tccgataact ggggtgacct tggTTaataT tcaccagcag cctccccctg tgcccctctg 420
gatccactgc ttaaatacgg acgaggacag ggccctgtct cctcagcttc aggcaccacc 480
actgacctgg gacagtgaat cgccacc 507

<210> 370

<211> 689

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 370

aggTTaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
tgTTTgtctt ggTTaataat ctCaggagca caaacattcc ggcccgggag gcgccctttg 120

gaccttttgc aatcctggcg cactgaaccc ttgacctctg cctgcagcc cccgcagctt 180
gctgtttgcc cactctatTT gcccagcccc aggctggTTT cttataaaac tgatggaaga 240
tacaacact attaaagaac tgTTTgcattg ttgcaaatga tgtccaaagt ccaaacttg 300

ttaataatta atactccaat aaacatcatg tcagaatttc tgttttcttt tccctttgaa 360
cctttgcagg attgccacat catcaggacc acaccttcat caggaatgaa tatccgatga 420
cctaattgatt ctgagcttgg caaaggtctt atctcccagc tgcgccaggc ccagtgttcc 480
aggaatgtga cctttgctgc agcagccgct ggagggggca gaggggatgg gctggaggtt 540

gagcaaacag agcagcagaa aaggcagttc ctcttctcca gtgccctcct tccctgtctc 600
tgctctccc tcccttctc aggcatcaga gcggagactt caggagacc agagcccagc 660
ttgccaggca ctgagctaga agccctgcc 689

<210> 371

<211> 344

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 371

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
tgtttctctt ggtaataat ctgaggagca caaacattcc ggcccgggag gcgccctttg 120

gaccttttgc aatcctggcg cactgaacc ttgaccttg cctgcagcc cccgcagctt 180
gctgtttgcc cactctattt gccagcccc agaagcaaatt atttgtggtt atggattaac 240
tcgaactgtt tgcccactct atttcccgg gcatataaaa caggggcaag gcacagactc 300
atagcagagc aatcaccacc aagcctggaa taactgcagc cacc 344

<210> 372

<211> 401

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 372

aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccccagg 60

ctttcacttt ctgcceaact tacaaggcct tctgtgttaa acaatactg aacctttacc 120
ccgttgcccc gcaacggcca ggtctgtgcc aagtgtttgc ctttgcaaca gcttatcgga 180
agcaacaag ctgaggggaa ttgagcaaga atttctggga taccaacagc ataggaggaa 240
caaaggacgt agagggaggg ttgactgtct acacaggaca aagccaatga ttaaccaaac 300

ctcttcgaga tttaaatagg atgggaacta ggagtggcag caatccttc tttcagctgg 360
 agtgctctc aggagccagc cccaccctta gaaaagccac c 401
 <210> 373
 <211> 435
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<400> 373
 aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccccagg 60
 ctttcacttt ctgcceaact tacaaggcct tttgtgttaa acaatactg aacctttacc 120
 ccgttgcccg gcaacggcca ggtctgtgcc aagtgtttgc cgatgaccta atgattctga 180
 gcttggcaaa ggtcttatct cccagctcgc ccaggcccag tgttccagga atgtgacctt 240
 tgctgcagca gccgctggag ggggcagagg ggatgggctg gaggttgagc aaacagagca 300
 gcagaaaagg cagtctctct tctccagtgc cctccttccc tgtctctgcc tctccctccc 360
 ttctcaggc atcagagcgg agacttcagg gagaccagag cccagcttgc caggcactga 420
 gctagaagcc ctgcc 435

<210> 374

<211> 485

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<400> 374
 aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
 tgtttgctct ggtaataat ctgaggagca caaacattcc ggcccgggag gcgccctttg 120
 gaccttttgc aatcctggcg caggctttca ctttctcgcc aacttacaag gcctttctgt 180
 gtaacaata cctgaacctt taccccgttg cccggcaacg gccaggtctg tgccaagtgt 240
 ttgcctttgc aacagcttat cggaagcaaa caagctgagg ggaattgagc aagaatttct 300
 gggataccaa cagcatagga ggaacaaagg acgtagaggg agggttgact gtctacacag 360
 gacaaagcca atgattaacc aaacctttg cagatttaaa taggatggga actaggagtg 420
 gcagcaatcc tttctttcag ctggagtgt cctcaggagc cagccccacc cttagaaaag 480

ccacc	485
<210> 375	
<211> 519	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 375	
aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggtaataat ctgaggagca caaacattcc ggcccgggag gcgccctttg	120
gaccttttgc aatcctggcg caggctttca cttctcgc aacttacaag gcctttctgt	180
gtaaacaata cctgaacctt taccccggtg cccggcaacg gccaggtctg tgccaagtgt	240
ttgccgatga cctaagtatt ctgagcttgg caaaggtctt atctcccagc tcgcccaggc	300
ccagtgttcc aggaatgtga cctttgctgc agcagccgct ggagggggca gaggggatgg	360
gctggagggt gagcaaacag agcagcagaa aaggcagttc ctcttctcca gtgccctcct	420
tcctgtctc tgcctctccc tcccttctc aggcattcaga gcggagactt caggagagacc	480
agagcccagc ttgccaggca ctgagctaga agccctgcc	519
<210> 376	
<211> 611	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 376	
attggcatct tctattgctt ttcttgggtga ctctatttt cactcttggc taaaaatggg	60
tctctgatga ttattctat cctgggtgtt gacaagctga agaagttgtg tggggcctgc	120
tgccagtaac cctgggtgac gaagcgtgac tcaccactcc gaggtcagtg gggggatgga	180
aggcagggga gtcagctgac aagatctgct gctttgtcac caggccttct gccaggcttt	240
cactttctcg ccaacttaca aggcctttct gtgtaacaa tacctgaacc ttaccccg	300
tgcccggcaa cggccagtc tgtgccaagt gtttgccgat gacctaatga ttctgagctt	360
ggcaaaggtc ttatctccca gctcggccag gccagtggt ccaggaatgt gacctttgct	420
gcagcagccg ctggaggggg cagaggggat gggttgagg ttgagcaaac agagcagcag	480

aaaaggcagt tcctcttctc cagtgccctc cttccctgtc tctgcctctc cctcccttcc 540

tcaggcatca gagcggagac ttcagggaga ccagagccca gcttgccagg cactgagcta 600

gaagccctgc c 611

<210> 377

<211> 580

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 377

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60

taccccgttg cccggcaacg gccaggtctg tgccaagtgt ttggctgggt tcttataaaa 120

ctgatggaag atacaaacac tattaaagaa ctgtttgcat gttgcaaag atgtccaaag 180

tccaaacatt gttaataatt aatactccaa taaacatcat gtcagaattt ctgttttctt 240

ttccctttga acctttgcag gattgccaca tcatacaggac cacaccttca tcaggaatga 300

atatccgatg acctaatgat tctgagcttg gcaaaggctc tatctcccag ctgcccagg 360

cccagtgttc caggaatgtg acctttgctg cagcagccgc tggagggggc agaggggatg 420

ggctggagggt tgagcaaaca gagcagcaga aaaggcagtt cctcttctcc agtgccctcc 480

ttccctgtct ctgcctctcc ctcccttctc caggcatcag agcggagact tcaggagagac 540

cagagcccag cttgccaggc actgagctag aagccctgcc 580

<210> 378

<211> 410

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 378

ctgtttgctg cttgcaatgt ttgccattt tagggagggt aatttttaaa aagcagtcaa 60

aagtccaagt ggcccttggc agcatttact ctctctgttt gctctggtta ataatctcag 120

gagcacaaac attccggccc gggaggcgcc ctttgacct tttgcaatcc tggcgactg 180

aacccttgac ccctgccttg cagcccccgc agcttgctgt ttgccactc tatttgccca 240

gccccagccc tggagagtcc tttagcaggg caaagtgcaa cataggcaga ccttaaggga	300
tgactcagta acagataagc tttgtgtgcc tgcagggcat ataaaacagg ggcaaggcac	360
agactcatag cagagcaatc accaccaagc ctggaataac tgcagccacc	410
<210> 379	
<211> 512	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 379	
ctgtttgctg cttgcaatgt ttgccattt taggggtgg tttcttataa aactgatgga	60
agatacaaac actattaaag aactgtttgc atgttgcaaa tgatgtccaa agtccaaaca	120
ttgttaataa ttaatactcc aataaacatc atgtcagaat ttctgttttc ttttccttt	180
gaacctttgc aggattgcca catcatcagg accacacctt catcaggaat gaatatccga	240
tgacctaatg attctgagct tggcaaaggt cttatctccc agctcgcca ggcccagtgt	300
tccaggaatg tgacctttgc tgcagcagcc gctggagggg gcagagggga tgggctggag	360
gttgagcaaa cagagcagca gaaaaggcag ttctcttctt ccagtgccct ccttcctgt	420
ctctgcctct cctcccttc ctcaggcatc agagcggaga cttcagggag accagagccc	480
agcttgccag gcactgagct agaagccctg cc	512
<210> 380	
<211> 345	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 380	
caggctttca ctttctgcc aacttacaag gcctttctgt gtaaacaata cctgaacctt	60
taccccgttg cccggcaacg gccaggtctg tgccaagtgt ttgcctttgc aacagettat	120
cggaagcaaa caagctgagg ggaattgagc aagaatttct gggataccaa cagcatagga	180
ggaacaaagg acgtagaggg agggttgact gtctacacag gacaaagcca atgattaacc	240
aaacctcttg cagatttaaa taggatggga actaggagtg gcagcaatcc tttctttcag	300

ctggagtgcct cctcaggagc cagccccacc cttagaaaag ccacc 345

<210> 381

<211> 337

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 381

aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccctgta 60

ccggcccggg aggcgccttt tggacctttt gcaatcctgg cgactgaac ccttgacccc 120

tgccctgcag cccccgcagc ttgctgtttg cccactctat ttgccagcc ccagccctgg 180

agagtccttt agcagggcaa agtgcaacat aggcagacct taagggatga ctcagtaaca 240

gataagcttt gtgtgcctgc agggcatata aacaggggc aaggcacaga tcatagcag 300

agcaatcacc accaagcctg gaataactgc agccacc 337

<210> 382

<211> 326

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 382

ggcgcccttt ggaccttttg caatcctgga gcaaacagca aacactgtac cggcccggga 60

ggcgcccttt ggaccttttg caatcctggc gcaactgaacc cttgaccctt gccctgcagc 120

ccccgcagct tgcgttttgc ccaactctatt tgcccagccc cagccctgga gaggccttta 180

gcagggcaaa gtgcaacata ggcagacctt aagggatgac tcagtaacag ataagctttg 240

tgtgcctgca gggcatataa aacaggggca aggcacagac tcatagcaga gcaatcacca 300

ccaagcctgg aataactgca gccacc 326

<210> 383

<211> 442

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 383

aggTTAATTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60

tgttttctct ggTTAATAAT ctCaggagca caaacattcc tgtaccagaa tgaacattga 120

actttggact atacctgagg ggtgaggtaa acaacaggac tataaatggc ccgggaggcg 180

ccctttggac cttttgcaat cctggcgac tgaaccttg acccctgccc tgcagccccc 240

gcagcttgct gtttgccac tctatttgcc cagccccagc cctggagagt ctttagcag 300

ggcaaagtgc aacataggca gaccttaagg gatgactcag taacagataa gctttgtgtg 360

cctgcagggc atataaaaca ggggcaaggc acagactcat agcagagcaa tcaccaccaa 420

gcctggaata actgcagcca cc 442

<210> 384

<211> 398

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 384

aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccctgta 60

ccagaatgaa cattgaactt tggactatac ctgaggggtg aggtaaacaa caggactata 120

aatggcccgaggaggccct ttggaccttt tgcaatcctg gcgcactgaa cccttgacct 180

ctgccctgca gccccgcag cttgtgttt gccactcta ttgcccagc cccagccctg 240

gagagtcctt tagcagggca aagtgaaca taggcagacc ttaagggatg actcagtaac 300

agataagctt tgtgtgcctg cagggcataat aaaacagggg caaggcacag actcatagca 360

gagcaatcac caccaagcct ggaataactg cagccacc 398

<210> 385

<211> 387

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 385

ggcgcccttt ggaccttttg caatcctgga gcaaacagca aacactgtac cagaatgaac 60

attgaacttt ggactatacc tgaggggtga ggtaacaac aggactataa atggcccggg 120
aggcgccctt tggacctttt gcaatcctgg cgactgaac ccttgacccc tgccttcag 180
ccccgcagc ttgctgtttg cccactctat ttgccagcc ccagccctgg agagtccttt 240
agcagggcaa agtgaacat aggcagacct taaggatga ctacagtaaca gataagcttt 300

gtgtgcctgc agggcatata aaacaggggc aaggcacaga ctcatagcag agcaatcacc 360
accaagcctg gaataactgc agccacc 387

<210> 386

<211> 433

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 386

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60
tgtttgctct ggtaataat ctacaggagca caaacattcc tgtaccagaa tgaacattga 120
actttggact atacctgagg ggtgaggtaa acaacaggac tataaatggc ccgggaggcg 180

ccctttggac cttttgcaat cctggcgctc tttgtttta catgaagggt ctggcagcca 240
aagcaatcac tcaaagtca aacctatca ttttttgctt tgttcctctt ggccttggtt 300
ttgtacatca gctttgaaaa taccatcca gggttaatgc tgggttaat ttataactaa 360
gagtgtctta gttttgcaat acaggacatg ctataaaaat ggaaagatgt tgctttctga 420
gagatgcgcc acc 433

<210> 387

<211> 389

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 387

aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccctgta 60
ccagaatgaa cattgaactt tggactatac ctgaggggtg aggtaaacaa caggactata 120
aatggcccgg gaggcgcctt ttggaccttt tgcaatcctg gcgctctttt gttttacatg 180
aagggtctgg cagccaaagc aatcactcaa agttcaaacc ttatcathtt ttgctttggt 240

cctcttggcc ttggttttgt acatcagctt tgaataacc atcccagggt taatgctggg 300
gttaatttat aactaagagt gctctagttt tgcaatacag gacatgctat aaaaatggaa 360
agatgttgct ttctgagaga tgcgccacc 389

<210> 388

<211> 378

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 388

ggcgcccttt ggaccttttg caatcctgga gcaaacagca aacactgtac cagaatgaac 60
attgaacttt ggactatacc tgagggtga ggtaacaac aggactataa atggcccggg 120
aggcgccctt tggacctttt gcaatcctgg cgctcttttg tttacatga agggctctggc 180
agccaaagca atcactcaaa gtccaacct tatcattttt tgctttgttc ctcttggcct 240
tggttttgta catcagcttt gaaaatacca tcccagggtt aatgctgggg ttaatttata 300

actaagagtg ctctagtttt gcaatacagg acatgctata aaaatggaaa gatgttgctt 360
tctgagagat gcgccacc 378

<210> 389

<211> 611

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 389

attggcatct tctattgctt ttcttgggtga ctctattttt cactcttggc taaaaatggg 60
tctctgatga ttattctat cctgggtgtt gacaagctga agaagttgtg tggggcctgc 120
tgccagtaac cctgggtgac gaagcgtgac tcaccactcc gaggtcagtg gggggatgga 180

aggcagggga gtcagctgac aagatctgct gctttgtcac caggccttct gccaggcttt 240
cactttctcg ccaacttaca aggcttttct gtgtaacaa tacctgaacc ttaccccgt 300
tgcccggcaa cggccaggtc tgtgccaagt gtttgccgat gacctaatga ttctgagctt 360
ggcaaaggtc ttatctccca gctcgcccag gccagtggtt ccaggaatgt gacctttgct 420
gcagcagccg ctggaggggg cagaggggat gggttgaggg ttgagcaaac agagcagcag 480

aaaaggcagt tctcttctc cagtgccctc cttccctgtc tctgcctctc cctcccttcc	540
tcaggcatca gagcggagac ttcagggaga ccagagccca gcttgccagg cactgagcta	600
gaagccctgc c	611
<210> 390	
<211> 580	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 390	
caggctttca ctttctcgcc aacttacaag gcctttctgt gtaaacaata cctgaacctt	60
taccccgttg cccggcaacg gccaggtctg tgccaagtgt ttggctggtt tcttataaaa	120
ctgatggaag atacaaacac tattaaagaa ctgtttgcat gttgcaaag atgtccaaag	180
tcctaacatt gttaataatt aatactccaa taaacatcat gtcagaattt ctgttttctt	240
ttccctttga acctttgcag gattgccaca tcctcaggac cacaccttca tcaggaatga	300
atatccgatg acctaatgat tctgagcttg gcaaaggctt tatctcccag ctgcccagg	360
cccagtgttc caggaatgtg acctttgctg cagcagccgc tggagggggc agaggggatg	420
ggctggagggt tgagcaaaca gagcagcaga aaaggcagtt cctcttctcc agtgccctcc	480
ttccctgtct ctgcctctcc ctcccttctt caggcatcag agcggagact tcaggagagac	540
cagagcccag cttgccaggc actgagctag aagccctgcc	580
<210> 391	
<211> 410	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 391	
ctgtttgctg cttgcaatgt ttgccattt tagggagggt aatttttaaa aagcagtcaa	60
aagtccaagt ggcccttggc agcatttact ctctctgttt gctctggtta ataatctcag	120
gagcacaaac attccggccc gggaggcgcc ctttgacct tttgcaatcc tggcgactg	180
aacccttgac cctgccttg cagccccgc agcttgctgt ttgccactc tatttgccca	240
gccccagccc tggagagtcc ttagcaggg caaagtgcaa cataggcaga ccttaaggga	300

tgactcagta acagataagc tttgtgtgcc tgcagggcat ataaaacagg ggcaaggcac 360

agactcatag cagagcaatc accaccaagc ctggaataac tgcagccacc 410

<210> 392

<211> 512

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 392

ctgtttgctg ctigcaatgt ttgcccatit taggggctgg tttcttataa aactgatgga 60

agatacaaac actattaaag aactgtttgc atgttgcaaa tgatgtccaa agtccaaaca 120

ttgttaataa ttaatactcc aataaacatc atgtcagaat ttctgttttc ttttcctttt 180

gaacctttgc aggattgcca catcatcagg accacacctt catcaggaat gaatatccga 240

tgacctaatg attctgagct tggcaaaggt cttatctccc agctcgccca ggcccagtgt 300

tccaggaatg tgacctttgc tgcagcagcc gctggagggg gcagagggga tgggctggag 360

gttgagcaaa cagagcagca gaaaaggcag ttcctcttct ccagtgcctt ctttcctgt 420

ctctgcctct cctcccttc ctcaggcatc agagcggaga cttcaggag accagagccc 480

agcttgccag gcactgagct agaagccctg cc 512

<210> 393

<211> 345

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 393

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60

taccctgttg cccggcaacg gccaggtctg tgccaagtgt ttgcctttgc aacagcttat 120

cggaagcaaa caagctgagg ggaattgagc aagaatttct gggataccaa cagcatagga 180

ggaacaaagg acgtagaggg aggggtgact gtctacacag gacaaagcca atgattaacc 240

aaacctcttg cagatttaaa taggatggga actaggagtg gcagcaatcc tttctttcag 300

ctggagtgtt cctcaggagc cagcccccacc ctagaaaag ccacc 345

<210> 394

<211> 381

<212> DNA

<213>

> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 394

aggTTaatTT tTaaaaagca gTcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgttttctct ggTTaataat ctCaggagca caaacattcc tgtaccggcc cgggaggcgc	120
cctttggacc ttttgaatc ctggcgcact gaacccttga cccctgcct gcagcccccg	180
cagcttgctg ttgcccact ctatttggcc agcccagcc ctggagagtc ctttagcagg	240
gcaaagtga acataggcag accttaaggg atgactcagt aacagataag ctttgtgtgc	300
ctgcagggca tataaaacag gggcaaggca cagactcata gcagagcaat caccaccaag	360

cctggaataa ctgcagccac c	381

<210> 395

<211> 337

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 395

aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccctgta	60
ccggccccggg aggcgccctt tggacctttt gcaatcctgg cgactgaac ccttgacccc	120
tgccctgcag cccccgcagc ttgctgtttg cccactctat ttgccagacc ccagccctgg	180
agagtcttt agcagggcaa agtgcaacat aggcagacct taagggatga ctcagtaaca	240

gataagcttt gtgtgcctgc agggcatata aaacaggggc aaggcacaga ctcatagcag	300
agcaatcacc accaagcctg gaataactgc agccacc	337

<210> 396

<211> 326

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 396

ggcgcccttt ggaccttttg caatcctgga gcaaacagca aacactgtac cggcccggga	60
ggcgcccttt ggaccttttg caatcctggc gcaactgaacc cttgaccctt gccctgcagc	120
ccccgcagct tgctgtttgc ccactctatt tgcccagccc cagccctgga gattccttta	180

gcagggcaaa gtgcaacata ggcagacctt aaggatgac tcagtaacag ataagctttg	240
tgtgcctgca gggcatataa aacaggggca aggcacagac tcatagcaga gcaatcacca	300
ccaagcctgg aataactgca gccacc	326

<210> 397

<211> 442

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 397

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggtaataat ctcaggagca caaacattcc tgtaccagaa tgaacattga	120

actttggact atacctgagg ggtgaggtaa acaacaggac tataaatggc ccgggaggcg	180
ccctttggac cttttgcaat cctggcgcac tgaacccttg acccctgccc tgcagccccc	240
gcagcttgct gtttgccac tctatttgcc cagccccagc cctggagagt cctttagcag	300
ggcaaagtgc aacataggca gaccttaagg gatgactcag taacagataa gctttgtgtg	360
cctgcagggc atataaaaca ggggcaaggc acagactcat agcagagcaa tcaccaccaa	420
gcctggaata actgcagcca cc	442

<210> 398

<211> 398

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 398

aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccctgta	60
ccagaatgaa cattgaactt tggactatac ctgaggggtg aggtaaacaa caggactata	120
aatggcccgg gaggcgcctt ttggaccttt tgcaatcctg gcgcactgaa cccttgacct	180

ctgccctgca gccccgcag cttgctgttt gccactcta tttgccagc ccagccctg 240
gagagtcctt tagcaggga aagtgaaca taggcagacc ttaaggatg actcagtaac 300
agataagctt tgtgtgcctg caggcatat aaaacagggg caaggcacag actcatagca 360

gagcaatcac caccaagcct ggaataactg cagccacc 398

<210> 399

<211> 387

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 399

ggcgcccttt ggacctttt caatcctgga gcaaacagca aacctgtac cagaatgaac 60
attgaacttt ggactatacc tgagggtga ggtaaacac aggactataa atggccggg 120
aggcgccctt tggacctttt gcaatcctgg cgactgaac cttgacccc tgccctgcag 180
ccccgcagc ttgtgtttt gccactctat ttgccagcc ccagccctgg agagtccttt 240

agcagggcaa agtgaacat aggcagacct taaggatga ctgagtaaca gataagcttt 300
tgtgtgcctg agggcatata aaacaggggc aaggcacaga ctcatagcag agcaatcacc 360
accaagcctg gaataactgc agccacc 387

<210> 400

<211> 433

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 400

aggtaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactcttc 60
tgtttctct ggtaataat ctgaggagca caaacattcc tgtaccagaa tgaacattga 120

actttggact atacctgagg ggtgaggtaa acaacaggac tataaatggc ccgggaggcg 180
ccctttggac cttttgcaat cctggcgctc tttgtttta catgaagggt ctggcagcca 240
aagcaatcac tcaaagtta aaccttatca tttttgctt tgttcctctt ggccttggtt 300
ttgtacatca gctttgaaa taccatccca gggttaatgc tgggttaat ttataactaa 360
gagtgtctta gttttgcaat acaggacatg ctataaaaat ggaaagatgt tgctttctga 420

gagatgcgcc acc 433

<210> 401

<211> 389

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 401

aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccctgta	60
ccagaatgaa cattgaactt tggactatac ctgaggggtg aggtaaacaa caggactata	120
aatggcccgg gaggcgccct ttggaccttt tgcaatcctg gcgctctttt gttttacatg	180
aagggtctgg cagccaaagc aatcactcaa agttcaaacc ttatcatttt ttgctttgtt	240
cctcttggcc ttggttttgt acatcagctt tgaaaatacc atcccagggt taatgctggg	300
gttaatttat aactaagagt gctctagttt tgcaatacag gacatgctat aaaaatggaa	360

agatgttgct ttctgagaga tgcgccacc 389

<210> 402

<211> 378

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 402

ggcgcccttt ggaccttttg caatcctgga gcaaacagca aacactgtac cagaatgaac	60
attgaacttt ggactatacc tgaggggtga ggtaaacac aggactataa atggcccggg	120
aggcgccctt tggacctttt gcaatcctgg cgctcttttg tttacatga agggctctggc	180
agccaaagca atcactcaaa gttcaaacct tatcattttt tgctttgttc ctcttgcct	240

tggttttgta catcagcttt gaaaatacca tcccagggtt aatgctgggg ttaatttata	300
actaagagtg ctctagtttt gcaatacagg acatgctata aaaatggaaa gatgttgctt	360
tctgagagat gcgccacc	378

<210> 403

<211> 203

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 403

aggTTAATTT TAAAAAGCA GTCAAAGTC CAAGTGGCCC TTGGCAGCAT TTACTCTCTC 60

TGTTTGCTCT GGTTAATAAT CTCAGGAGCA CAAACATTCC AGTCATATGT TTGTCCTCTG 120

AAGGTTACTA GTTAACAGGC ATCCCTTAAA CAGGATATAA AAGGACTTCA GCAGGACTGC 180

TCGAAACATC CCCTCAGCC ACC 203

<210> 404

<211> 222

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 404

aggTTAATTT TAAAAAGCA GTCAAAGTC CAAGTGGCCC TTGGCAGCAT TTACTCTCTC 60

TGTTTGCTCT GGTTAATAAT CTCAGGAGCA CAAACATTCC TGTACCGGCC CGGGAGGCGC 120

CCTTTGGACC TTTTGCAATC CTGGCGGGGC ATATAAACA GGGGCAAGGC ACAGACTCAT 180

AGCAGAGCAA TCACCACCAA GCCTGGAATA ACTGCAGCCA CC 222

<210> 405

<211> 275

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 405

TGCTCTCTGA CAAAGATACG GTGGGTCCA CTGATGAAC TGTCTGCCAC AGTAAATGTA 60

GCCACTATGC CTATCTCCAT TCTGAAGATG TGCCCTGTTC AAACATGTCC TAATACTCTG 120

TCTCTGCAAG GGTCATCAGT AGTTTTCCAT CTACTCAAC ATCCTCCAG TGAGTCATAT 180

GTTTGCTCAC TGAAGGTTAC TAGTTAACAG GCATCCCTTA AACAGGATAT AAAAGGACTT 240

CAGCAGGACT GCTCGAAACA TCCCACTCAG CCACC 275

<210> 406

<211> 283

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 406

cctccccgtg ttctgtctct ttgtccctct gtcctactta gactaatatt tgccttgggt	60
actgcaaaca ggaaatgggg gagggacagg agtagggcgg ccctgttcaa acatgtccta	120
atactctgtc tctgcaaggg tcatcagtag ttttccatct tactcaacat cctcccagtg	180
agtcatatgt ttgtcactg aagggtacta gttaacaggc atcccttaaa caggatataa	240
aaggacttca gcaggactgc tcgaaacatc ccactcagcc acc	283

<210> 407

<211> 262

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 407

ccgccccac tgaacccttg acccctgccc tgcagccccc gcagcttgct gtttgccac	60
tctatttgcc cagccccage cctgttcaaa catgtcctaa tactctgtct ctgcaagggt	120
catcagtagt tttccatctt actcaacatc ctcccagtg gtcataatgt ttgtcactga	180
aggttactag ttaacaggca tcccttaaac aggatataaa aggacttcag caggactgct	240
cgaaacatcc cactcagcca cc	262

<210> 408

<211> 268

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 408

aagctttctg aacagccaaa cagagattcc aaagttcagg caccaaagtt cagaccctaa	60
cagttattta caagggtcag ttaacccctg ttcaaacatg tcctaatact ctgtctctgc	120
aagggtcatc agtagttttc catcttactc aacatcctcc cagtgaagtca tatgtttgct	180

cactgaaggt tactagttaa caggcatccc ttaaacagga tataaaagga cttcagcagg	240
actgctcgaa acatcccact cagccacc	268
<210> 409	
<211> 389	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 409	
aagcaaatat ttgtggttat ggattaactc gaactgtttg cccactctat ttgccctgta	60
ccagaatgaa cattgaactt tggactatac ctgaggggtg aggtaaacaa caggactata	120
aatggcccgg gaggcgcctt ttggaccttt tgcaatcctg gcgctctttt gttttacatg	180
aagggtctgg cagccaaagc aatcactcaa agttcaaacc ttatcatttt ttgctttgtt	240
cctcttggcc ttggttttgt acatcagctt tgaaaatacc atcccagggt taatgctggg	300
gttaatttat aactaagagt gctctagttt tgcaatacag gacatgctat aaaaatggaa	360
agatgttgct ttctgagaga tgcgccacc	389
<210> 410	
<211> 378	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 410	
ggcgcccttt ggaccttttg caatcctgga gcaaacagca aacactgtac cagaatgaac	60
attgaacttt ggactatacc tgaggggtga ggtaaacac aggactataa atggcccggg	120
aggcgccctt tggacctttt gcaatcctgg cgctcttttg ttttacctga agggctctggc	180
agccaaagca atcactcaaa gttcaaacct tatcattttt tgctttgttc ctcttggcct	240
tggttttgta catcagcttt gaaaatacca tcccagggtt aatgctgggg ttaatttata	300
actaagagtg ctctagtttt gcaatacagg acatgctata aaaatggaaa gatgttgctt	360
tctgagagat gcgccacc	378
<210> 411	
<211> 203	

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 411

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggtaataat ctccaggagca caaacattcc agtcataatg ttgctcactg	120
aaggttacta gttaacaggc atcccttaaa caggatataa aaggacttca gcaggactgc	180
tcgaaacatc ccactcagcc acc	203

<210> 412

<211> 222

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 412

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttgctct ggtaataat ctccaggagca caaacattcc tgtaccggcc cgggaggcgc	120
cctttggacc ttttgcaatc ctggcggggc atataaaaca ggggcaaggc acagactcat	180
agcagagcaa tcaccaccaa gcctggaata actgcagcca cc	222

<210> 413

<211> 275

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 413

tgctctctga caagatacg gtgggtccca ctgatgaact gtgctgccac agtaaatgta	60
gccactatgc ctatctccat tctgaagatg tgcctgttc aaacatgtcc taatactctg	120
tctctgcaag ggtcatcagt agttttccat cttactcaac atcctcccag tgagtcatat	180
gtttgtcac tgaaggttac tagttaacag gcattccctta aacaggatat aaaaggactt	240
cagcaggact gctcgaataa tccactcag ccacc	275

<210> 414

<211> 283

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 414

cctccccgtg ttctgtctt ttgtccctct gtcctactta gactaatatt tgccttgggt 60

actgcaaaca ggaaatgggg gagggacagg agtagggcgg ccctgttcaa acatgtccta 120

atactctgtc tctgcaaggg tcatcagtag ttttccatct tactcaacat cctcccagtg 180

agtcatatgt ttgtcactg aaggttacta gttaacaggc atcccttaaa caggatataa 240

aaggacttca gcaggactgc tcgaaacatc ccactcagcc acc 283

<210> 415

<211> 262

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 415

ccgccccac tgaaccttg acccctgcc tgcagccccc gcagcttgct gtttgccac 60

tctatttgcc cagccccagc cctgttcaaa catgtcctaa tactctgtct ctgcaagggt 120

catcagtagt tttccatctt actcaacatc ctcccagtg gtcataatgt tgctcactga 180

aggttactag ttaacaggca tcccttaaac aggatataaa aggacttcag caggactgct 240

cgaaacatcc cactcagcca cc 262

<210> 416

<211> 268

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 416

aagctttctg aacagccaaa cagagattcc aaagttcagg caccaaagtt cagaccctaa 60

cagttattta caaggtcag ttaacccctg ttcaaactg tcctaatact ctgtctctgc 120

aagggtcacc agtagttttc catcttactc aacatcctcc cagtgagtca tatgtttgct 180

cactgaaggt tactagttaa caggcatccc ttaaacagga tataaaagga cttcagcagg 240

actgctcgaa acatcccact cagccacc 268

<210> 417

<211> 422

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 417

aggttaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60

tgtttgctct ggtaataat ctcaggagca caaacattcc ggcccgggag gcgccctttg 120

gaccttttgc aatcctggcg ccgatgacct aatgattctg agcttggcaa aggtcttacc 180

tcccagctcg cccaggccca gtgttcagg aatgtgacct ttgctgcagc agccgctgga 240

gggggcagag gggatgggct ggaggttgag caaacagagc agcagaaaag gcagttcctc 300

ttctccagtg cctccttcc ctgtctctgc ctctcctcc cttctcagg catcagagcg 360

gagacttcag ggagaccaga gccagcttg ccaggcactg agctagaagc cctgccgcca 420

cc 422

<210> 418

<211> 176

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 418

aggttaatTT ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60

tgtttgctct ggtaataat ctcaggagca caaacattcc gggcatataa aacaggggca 120

aggcacagac tcatagcaga gcaatcacca ccaagcctgg aataactgca gccacc 176

<210> 419

<211> 309

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 419

aggTTAATTT TAAAAAGCA GTCAAAGTC CAAGTGGCCC TTGGCAGCAT TTACTCTCTC 60

TGTTTGCTCT GGTTAATAAT CTCAGGAGCA CAAACATTCC TGTACCGGCC CGGGAGGCGC 120

CCTTTGGACC TTTTGCAATC CTGGCGCCCT GGAGAGTCCT TTAGCAGGGC AAAGTGCAAC 180

ATAGGCAGAC CTTAAGGGAT GACTCAGTAA CAGATAAGCT TTGTGTGCCT GCAGGGCATA 240

TAAAACAGGG GCAAGGCACA GACTCATAGC AGAGCAATCA CCACCAAGCC TGGAAATAACT 300

GCAGCCACC 309

<210> 420

<211> 341

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 420

aggTTAATTT TAAAAAGCA GTCAAAGTC CAAGTGGCCC TTGGCAGCAT TTACTCTCTC 60

TGTTTGCTCT GGTTAATAAT CTCAGGAGCA CAAACATTCC TGTACCCACT GAACCCCTGA 120

CCCCTGCCCT GCAGCCCCCG CAGCTTGCTG TTGCCCCT CTATTGCCC AGCCCCAGCC 180

CTGGAGAGTC CTTAGCAGG GCAAAGTGCA ACATAGGCAG ACCTTAAGGG ATGACTCAGT 240

AACAGATAAG CTTTGTGTGC CTGCAGGGCA TATAAACAG GGGCAAGGCA CAGACTCATA 300

GCAGAGCAAT CACCACCAAG CCTGGAATAA CTGCAGCCAC C 341

<210> 421

<211> 545

<212> DNA

<213>

> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 421

CAGGCTTTCA CTTTCTGCC AACTTACAAG GCCTTTCTGT GTAACAATA CCTGAACCTT 60

TACCCCGTTG CCCGGCAACG GCCAGGTCTG TGCCAAGTGT TTGAGGTTAA TTTTAAAAA 120

gcagtcaaaa gtccaagtgg cccttggcag ctttactct ctctgtttgc tctggttaat 180
aatctcagga gcacaaacat tcctgtacca gaatgaacat tgaactttgg actatacctg 240
aggggtgagg taaacaacag gactataaat ggcccgggag gcgccctttg gaccttttgc 300
aatcctggcg cactgaaccc ttgacccctg ccctgcagcc cccgcagctt gctgtttgcc 360

cactctattht gccagcccc agccctggag agtcccttag cagggcaaag tgcaacatag 420
gcagacctta agggatgact cagtaacaga taagctttgt gtgcctgcag ggcatataaa 480
acaggggcaa ggcacagact catagcagag caatcaccac caagcctgga ataactgcag 540
ccacc 545

<210> 422
<211> 137
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 422
agaatgaaca ttgaactttg gactatacct gaggggtgag gtaacaaca ggactataaa 60

tgggcatata aaacaggggc aaggcacaga ctcatagcag agcaatcacc accaagcctg 120
gaataactgc agccacc 137

<210> 423
<211> 440
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 423
caggtttta ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60
taccctgtt cccggcaacg gccaggtctg tgccaagtgt ttgaagcaa ttttgttgt 120
tatggattaa ctgaaactgt ttgcccactc ttttgcct gtaccggccc gggaggcgcc 180

ctttggacct ttgcaatcc tggcgcactg aacccttgac ccctgccctg cagccccgcg 240
agcttgctgt ttgcccactc ttttgcceca gccccagccc tggagagtcc tttagcaggg 300
caaagtgcaa cataggcaga ccttaaggga tgactcagta acagataagc tttgtgtgcc 360
tgcaggcat ataaaacagg ggcaaggcac agactcatag cagagcaatc accaccaagc 420

ctggaataac tgcagccacc 440

<210> 424

<211> 501

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 424

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60

taccccgttg cccggcaacg gccaggtctg tgccaagtgt ttgaagcaa ttttgtggt 120

tatggattaa ctggaactgt ttgccactc ttttgcct gtaccagaat gaacattgaa 180

ctttggacta tacctgaggg gtgaggtaaa caacaggact ataatggcc cgggaggcgc 240

cctttggacc ttttgaatc ctggcgact gaaccctga cccctgcct gcagccccg 300

cagcttgctg tttgccact ctatttgcgc agccccagcc ctggagagtc ctttagcagg 360

gcaaagtga acataggcag accttaaggg atgactcagt aacagataag ctttgtgtgc 420

ctgcagggca tataaacag gggcaaggca cagactcata gcagagcaat caccaccaag 480

cctggaataa ctgcagccac c 501

<210> 425

<211> 414

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 425

caggctttca ctttctcgcc aacttacaag gcctttctgt gtaacaata cctgaacctt 60

taccccgttg cccggcaacg gccaggtctg tgccaagtgt ttgaagcaa ttttgtggt 120

tatggattaa ctggaactgt ttgccactc ttttgcct gtaccagaat gaacattgaa 180

ctttggacta tacctgaggg gtgaggtaaa caacaggact ataatggcc cgggaggcgc 240

cctttggacc ttttgaatc ctggcgact gaaccctga cccctgcct gcagccccg 300

cagcttgctg tttgccact ctatttgcgc agccccaggg gcatataaaa caggggcaag 360

gcacagactc atagcagagc aatcaccacc aagcctggaa taactgcagc cacc 414

<210> 426

<211> 243

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 426

aggTTAATTT TAAAAAGCA GTCAAAAGTC CAAGTGGCCC TTGGCAGCAT TTACTCTCTC 60

TGTTTGCTCT GGTTAATAAT CTCAGGAGCA CAAACATTCC GGCCCGGGAG GCGCCCTTTG 120

GACCTTTTGC AATCCTGGCG AGTCATATGT TTGCTCACTG AAGGTTACTA GTTAACAGGC 180

ATCCCTTAAA CAGGATATAA AAGGACTTCA GCAGGACTGC TCGAAACATC CCCTCAGCC 240

ACC 243

<210> 427

<211> 279

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 427

aggTTAATTT TAAAAAGCA GTCAAAAGTC CAAGTGGCCC TTGGCAGCAT TTACTCTCTC 60

TGTTTGCTCT GGTTAATAAT CTCAGGAGCA CAAACATTCC CAGGCTTTCA CTTTCTCGCC 120

AACCTACAAG GCCTTTCTGT GTAAACAATA CCTGAACCTT TACCCGTTG CCCGGCAACG 180

GCCAGGTCTG TGCCAAGTGT TTGGGGCATA TAAACAGGG GCAAGGCACA GACTCATAGC 240

AGAGCAATCA CCACCAAGCC TGGAATAACT GCAGCCACC 279

<210> 428

<211> 256

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 428

aggTTAATTT TAAAAAGCA GTCAAAAGTC CAAGTGGCCC TTGGCAGCAT TTACTCTCTC 60

TGTTTGCTCT GGTTAATAAT CTCAGGAGCA CAAACATTCC CCCTGTTCAA ACATGTCCTA 120

atactctgtc tctgcaaggg tcatcagtag ttttccatct tactcaacat cctcccagtg 180
gggcatataa aacaggggca aggcacagac tcatagcaga gcaatcacca ccaagcctgg 240
aataactgca gccacc 256

<210> 429

<211> 183

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 429

ccctgttcaa acatgtccta atactctgtc tctgcaaggg tcatcagtag ttttccatct 60

tactcaacat cctcccagtg agtcatatgt ttgctcactg aaggttacta gttaacaggc 120

atcccttaaa caggatataa aaggacttca gcaggactgc tcgaaacatc ccactcagcc 180

acc 183

<210> 430

<211> 249

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 430

aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc 60

tgtttgctct ggtaataat ctcaggagca caaacattcc tgtaccggcc cgggaggcgc 120

cctttggacc ttttgcaatc ctggcgagtc atatgtttgc tactgaagg ttactagtta 180

acaggcatcc cttaaacagg atataaaagg acttcagcag gactgctcga aacatcccac 240

tcagccacc 249

<210> 431

<211> 349

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 431

tcgagcttgg gctgcaggtc gagggcactg ggaggatgtt gagtaagatg gaaaactact	60
gatgaccctt gcagagacag agtattagga catgtttgaa caggggccgg gcgatcagca	120
ggtagctcta gaggatcccc gtctgtctgc acatttcgta gagcgagtgt tccgatactc	180
taatctccct aggcaagggt catatttgtag taggttactt attctccttt tgttgactaa	240
gtcaataatc agaatacagca ggtttggagt cagcttgga gggatcagca gcctgggttg	300
gaaggagggg gtataaaagc cccttcacca ggagaagccg tcacacaga	349
<210> 432	
<211> 558	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 432	
ccctaaaatg ggcaaacatt gcaagcagca aacagcaaac acacagccct ccctgcctgc	60
tgaccttga gctggggcag aggtcagaga cctctctggg cccatgccac ctccaacatc	120
cactcgacct cttggaattt cgggtggagag gagcagaggt tgtcctggcg tggtttaggt	180
agtgtgagag gggaatgact cctttcggta agtgagtggt aagctgtaca ctgccaggc	240
aaagcgtccg ggcagcgtag gcgggcgact cagatcccag ccagtggtact tagccccgtg	300
ttgtctctcc gataactggg gtgaccttgg ttaatatcca ccagcagcct cccccgttgc	360
ccctctggat ccaactgcta aatacggacg aggacagggc cctgtctcct cagcttcagg	420
caccaccact gacctgggac agtgaatccg gactctaagg taaatataaa atttttaagt	480
gtataatgtg ttaaaactact gattctaatt gtttctctct tttagattcc aacctttgga	540
actgaattct agaccacc	558
<210> 433	
<211> 580	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 433	
atagtaatca attacggggg cattagttca tagcccatat atggagttcc gcgttacata	60
acttacggta aatggcccg ctagctgacc gccaacgac ccccgcccat tgacgtcaat	120

aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga 180

gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc 240

ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt acatgacctt 300

atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtgat 360

gcggttttgg cagtacatca atgggcgtgg atagcggttt gactcacggg gatttccaag 420

tctccacccc attgacgtca atgggagttt gttttggcac caaaatcaac gggactttcc 480

aaaatgtcgt aacaactccg cccattgac gcaaattgggc ggtaggcgtg tacggtggga 540

ggtctatata agcagagctg gtttagtgaa ccgtcagatc 580

<210> 434

<211> 1775

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 434

agatctgaat tcggtaccta gttattaata gtaatcaatt acggggtcat tagttcatag 60

cccatatatg gagttccgga ttacataact tacggtaaat ggcccgcctg gctgaccgcc 120

caacgacccc cgccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg 180

gactttccat tgacgtcaat ggggtggagta ttacggtaa actgccact tggcagtaca 240

tcaagtgtat catatgcaa gtacgcccc tattgacgtc aatgacggt aatggccgc 300

ctggcattat gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt 360

attagtcac gctattacca tggctgaggt gagccccacg ttctgcttca ctctcccat 420

ctccccccc tccccaccc caattttgta ttattttatt ttttaattat tttgtgcagc 480

gatgggggcg gggggggggg gggggcgcg gccaggcggg gcggggcggg gcgagggcg 540

gggcggggcg aggcggagag gtgcggcggc agccaatcag agcggcgcg tccgaaagtt 600

tccttttatg gcgaggcggc ggcggcggc gccctataaa aagcgaagcg gcggcgggc 660

gggagtcgt gcgcgtgcc ttcccccgt gcccgcctcc gccgccct cgcgccgcc 720

gccccggctc tgactgaccg cgttactccc acaggtgagc gggcgggacg gcccttctcc 780

tccgggctgt aattagcgt tggtttaatg acggcttgtt tcttttctgt ggctgcgtga 840

aagccttag gggctccggg agggcccttt gtgcgggggg agcggctcgg ggggtgcgtg 900

cgtgtgtgtg tgcgtgggga gcgcccgctg cggtcccg cgtcccggcg gctgtgagcg 960

ctgcgggcgc ggcgcggggc tttgtgcgct ccgcagtggt cgcgaggga gcgcggccgg	1020
gggcggtgcc ccgcggtgcg gggggggctg cgagggaac aaaggctgcg tgcgggtgt	1080
gtgcgtgggg gggtagcag ggggtgtggg cgcgtcggtc gggctgcaac cccccctgca	1140
ccccctccc cgagttgctg agcacggccc ggcttcgggt gcggggctcc gtacggggcg	1200
tggcgcgggg ctgccgtgc gggcggggg gtggcggcag gtgggggtgc cggcggggc	1260
ggggccgcct cggcggggg agggctcggg ggaggggcg gcggccccc ggagcgccgg	1320
cggctgtcga ggcgcggcga gccgcagcca ttgcctttta tggtaatcgt gcgagagggc	1380
gcagggaatt ctttgtccc aaatctgtgc ggagccgaaa tctgggaggc gccgccgac	1440
cccctctagc gggcgcgggg cgaagcgggt cggcgccggc aggaaggaaa tggcgggga	1500
gggccttcgt gcgtcgccgc gccgccgtcc cttctccct ctccagctc ggggctgtcc	1560
gcggggggac ggctgccttc gggggggacg gggcaggcg gggttcggt tctggcgtgt	1620
gaccggcggc tctagacct ctgtaacca tgttcagcc ttctctttt tctacagct	1680
cctgggcaac gtgctggtta ttgtctgtc tcacatctt ggcaaagaat tctcgaaga	1740
tctaggcaac gcgtctcgag gcggccgccg ccacc	1775
<210> 435	
<211> 890	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
polynucleotide	
<400> 435	
aggttaattt ttaaaaagca gtcaaaagtc caagtggccc ttggcagcat ttactctctc	60
tgtttctct ggtaataat ctgaggagca caaacattcc agatccaggt taatttttaa	120
aaagcagtea aaagtccaag tggcccttgg cagcatttac tctctctgtt tgctctggtt	180
aataatctca ggagcacaaa cattccagat ccggcgcgcc agggctggaa gctacctttg	240
acatcatttc ctctcgcaat gcatgtataa ttctacaga acctattaga aaggatcacc	300
cagcctctgc ttttgtacaa tttccctta aaaaactgcc aattccactg ctgtttggcc	360
caatagttag aactttttcc tgctgcctct tgggtgcttt gcctatggcc cctattctgc	420
ctgctgaaga cactcttgcc agcatggact taaacccctc cagctctgac aatcctcttt	480
ctcttttgtt ttacatgaag ggtctggcag ccaaagcaat cactcaaagt tcaaacctta	540
tcattttttg ctttgttctt cttggccttg gttttgtaca tcagctttga aaataccatc	600

ccagggttaa tgctgggggtt aatttataac taagagtgt ctagttttgc aatacaggac 660
atgctataaa aatggaaaga tgttgctttc tgagagactg cagaagttgg tcgtgaggca 720
ctgggcaggt aagtatcaag gttacaagac aggtttaagg agaccaatag aaactgggct 780
tgtcgagaca gagaagactc ttgcgtttct gataggcacc tattggtctt actgacatcc 840
actttgcctt tctctccaca gggcaatccg gtactgttgg taaagccacc 890

<210> 436

<400> 436

000

<210> 437

<400> 437

000

<210> 438

<211> 290

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 438

tcagatgcc tggagacgcc atccacgtg ttttgacctc catagaagac accgggaccg 60
atccagcctc cgcggccggg aacggtgcat tggaacgcgg attccccgtg ccaagagtga 120
cgtaagtacc gcctatagac tctataggca caccctttg gctcttatgc atgaacggtg 180
gagggcagtg tagtctgagc agtactcgtt gctgccgcgc gcgccaccag acataatagc 240

tgacagacta acagactgtt cctttccatg ggtcttttct gcaggccacc 290

<210> 439

<211> 12

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 439

tgataactcg ag 12

<210> 440

<211> 9817

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 440

ctgcgcgctc gctcgtcac tgaggccgcc cgggcaaagc cggggtcg ggcgaccttt	60
ggtcgcccgg cctcagttag cgagcgagcg cgagagagg gaggggcaa ctccatcact	120
aggggttctt ggaggggtgg agtcgtgacg tgaattacgt catagggtta gggaggtcgg	180
ccgctctagg agttaatttt taaaaagcag taaaagtcc aagtgcctt gcgagcattt	240
actctctctg ttgctctgg ttaataatct caggagcaca aacattcctt actagtctta	300
ggagttaatt tttaaaaagc agtcaaaagt ccaagtcccc ttgcgagcat ttactctctc	360
tgtttgctct ggtaataat ctacaggagca caaacattcc ttactagttc tagagcggcc	420
gccagtgtgc tggaattcgg ctttttagg gctggaagct accttgaca tcatttcctc	480
tgcgaatgca tgtataattt ctacagaacc tattagaaag gatcacccag cctctgcttt	540
tgtacaactt tcccttaaaa aactgccaat cccactgctg tttggcccaa tagtgagaac	600
tttttctgc tgcctcttgg tgttttgcc tatggccctt attctgctg ctgaagacac	660
tcttgccagc atggacttaa accctccag ctctgacaat cctctttctc tttgtttta	720
catgaagggt ctggcagcca aagcaatcac tcaaagtca aaccttatca tttttgctt	780
tgttcctctt ggccttggtt ttgtacatca gctttgaaa taccatcca gggttaatgc	840
tggggttaat ttataactga gagtgtctta gtctgcaat acaggacatg ctataaaaat	900
ggaaagatgt tgctttctga gagatcagct tacatgtgct agcgtagcca accatgggag	960
tgaggcaccg gccctgctcc caccggctcc tggccgtctg cgcctcgtg tccctggcaa	1020
ccgctgcact cctggggcac atcctactcc atgatttctt gctggttccc cgagagctga	1080
gtggctcttc cccagtctg gaggagactc acccagctca ccagcagga gccagcagac	1140
cagggccccg ggacgccag gcacaccccg gccgtcccag agcagtgcc acacagtgcg	1200
acgtcccccc caacagccg ttcgattgcg ccctgacaa ggccatcacc caggaacagt	1260
gcgagggccc cggtgctgc tacatccctg caaagcaggg gctgcagga gccagatgg	1320
ggcagccctg gtgcttcttc ccaccagct acccagcta caagctggag aacctgagct	1380
cctctgaaat gggtacacg gccaccctga cccgtaccac cccaccttc tccccagg	1440
acatctgac cctgcggtg gacgtgatga tggagactga gaaccgctc cacttcacga	1500
tcaaagatcc agctaacagg cgctacgagg tgcccttga gaccccgct gtccacagcc	1560

gggcaccgtc cccactctac agcgtggagt tctccgagga gcccttcggg gtgatcgtgc 1620
accgcgagct ggacggccgc gtgctgctga acacgacggt ggcgccccctg ttctttgcgg 1680
accagttcct tcagctgtcc acctcgctgc cctcgagta tatcacaggc ctgcccagac 1740

acctcagtcc ccgatgctc agcaccagct ggaccaggat caccctgtgg aaccgggacc 1800
ttgcgcccac gcccggtgcg aacctctacg ggtctcacc tttctacctg gcgctggagg 1860
acggcgggtc ggacacacggg gtgttcctgc taaacagcaa cgccatggac gtggtcctgc 1920
agccgagccc tgcctttagc tggaggtcga cagggtggat cctggacgtc tacatcttcc 1980
tgggcccaga gcccagagc gtggtgcagc agtacctgga cgtggtggga taccggttca 2040
tgccgccata ctggggcctg ggcttccacc tgtgccgtg gggtactcc tccaccgcta 2100
tcaccgcca ggtggtggag aacatgacca ggcccactt cccctggac gtccagtga 2160

acgatctgga ctacatggac tcccggaggg acttcacgtt caacaaggat ggcttcggg 2220
acttcccgc catggtgcag gagctgcacc agggcgccg gcgctacatg atgatcgtgg 2280
atcctgccat cagcagctcg ggccctgccg ggagctacag gccctacgac gagggctcgc 2340
ggaggggggt ttcatcacc aacgagaccg gccagccgt gattgggaag gtgtggcccg 2400
ggtccactgc ctccccgac ttcaccaacc ccacagccct ggctggtgg gaggacatgg 2460
tggtgagtt ccatgaccag gtgcccttcg acggcatgtg gattgacatg aacgagcctt 2520
ccaacttcat cagaggctct gaggacggct gcccacaaca cgagctggag aaccaccct 2580

acgtgcctgg ggtggttggg gggaccctcc aggcggccac catctgtgcc tccagccacc 2640
agtttctctc cacacactac aacctgcaca acctctacgg cctgaccgaa gccatcgct 2700
cccacagggc gctggtgaag gctcggggga cagccatt tgtgatctcc cgctcgacct 2760
ttgctggcca cgccgatac gccggccact ggacggggga cgtgtggagc tcctgggagc 2820
agctcgctc ctccgtgcca gaaatcctgc agtttaacct gctgggggtg cctctggtcg 2880
gggccgacgt ctgcgcttc ctgggcaaca cctcagagga gctgtgtgtg cgctggaccc 2940
agctgggggc ctctacccc ttcatgcgga accacaacag cctgtcagt ctgcccagg 3000

agccgtacag cticagcag ccggcccagc aggccatgag gaaggccctc accctgcgt 3060
acgcactcct cccccactc tacacactgt tccaccaggc ccacgtcgcg ggggagaccg 3120
tgcccggcc cctcttctg gatttccca aggaactctag cacttggaact gtggaccacc 3180
agctcctgtg gggggaggcc ctgctcatca cccagtgtc ccaggccggg aaggccgaag 3240
tgactggcta ctccccctg ggcacatggt acgacctgca gacggtgcca atagaggccc 3300
ttggcagcct cccacccca cctgcagctc cccgtgagcc agccatccac agcgaggggc 3360

agtgggtgac gctgccggcc cccctggaca ccatcaacgt gcacctccgg gctgggtaca 3420

 tcatccccct gcagggccct ggcctcacia ccacagagtc ccgccagcag cccatggccc 3480
 tggctgtggc cctgaccaag ggtggagagg cccgagggga gctgttctgg gacgacggag 3540
 agagcctgga agtgctggag cgaggggcct acacacaggt catcttcctg gccaggaata 3600
 acacgatcgt gaatgagctg gtacgtgtga ccagttaggg agctggcctg cagctgcaga 3660
 aggtgactgt cctgggcgtg gccacggcgc cccagcaggt cctctccaac ggtgtccctg 3720
 tctccaactt cacctacagc cccgacacca aggtcctgga catctgtgtc tcgctgttga 3780
 tgggagagca gtttctcgtc agctggtgtt gataactcga gccgggcgga gtgtgttagt 3840

 ctctccagag ggaggctggt tccccaggga agcagagcct gtgtgcgggc agcagctgtg 3900
 tgcgggcctg ggggttggtt agtgcaatta tttttaataa aaggggcatt tggaaaaaaa 3960
 aaaaaagggt agcagtcgac agatgaattc tgcagatctg tggtctctag ctgcccggt 4020
 ggcatccctg tgaccctcc ccagtgcctc tcctggcctt ggaagttgcc actccagtgc 4080
 ccaccagcct tgcctaata aaattaagtt gcatcatttt gtctgactag gtgtccttct 4140
 ataataattat ggggtggagg ggggtggtat ggagcaaggg gcaagttggg aagacaacct 4200
 gagtgttgg gattccaggc atcgagtaga taagtagcat ggccgggttaa tcattaacta 4260

 caaggaacct ctagtgatgg agttggccac tccctctctg cgcgtctgct cgctcactga 4320
 ggccgggcga ccaaaggtcg cccgacgccc gggctttgcc cgggcggcct cagtgcgca 4380
 gcgagcgcgc agagagggac agatccgaag ccgggcaaata cagcgcctgg cagcagtggc 4440
 gtctggcgga aaacctcagt gtgacgctcc ccgccgctc ccacgcttgt tcccggatct 4500
 gaccaccagc gaaatccgat ttttgaccg agctgggtta taagcgttgg caatttaacc 4560
 gccagtccag ctttcttca cagtgtggat tggcgataaa aaacaactgc tgacgccgt 4620
 gcgcatcag ttcaccggtt caccgctgga taacgacttg gcgtaagtga agcgaccgt 4680

 aagaccctaa cgcttgggtc gaacgctgga aggcggcggg ccaaaccagg ccgaagcagc 4740
 gttgttcgag ttcacggcag atacacttgc tgttgcggtg ctgattacga ccgtcactc 4800
 gtggcagcaa caggggaaaa cttatattat cagccggaata acctaccgga ttgttggtag 4860
 tggatcaatag gcgattaccg ttgtgttgaa gtggcgagcg atacaccgct tccggcgagg 4920
 attggcctga actgccaaat ggcgaggtga gcagagcggg taaactggct cggattaggg 4980
 ccgcaagaaa aciatcccga ccgccttact gccgcctgtt ttgaccgctg ggatctgcc 5040
 agtcagacag tatagccgt acgtcttccc gagcgaaaac ggtctgcgt gcgggacgcg 5100

cgaattgaat ttggcccaca ccagttggcgc ggcgacttcc agttcaatat cagccgctac 5160
 agtgaacagc aactgtttga aaccagcctt cgccaactgc tgcacgcgga agaaggcact 5220
 ggctgaatat cgacggtttc cagttgggga ttggtggcga cgactcctgg agcccgctcag 5280
 tatcggcgga cticcaactg agcgccggtc gctaccttac cagttggtct ggtgtcaaaa 5340
 agcgtccgct tgagtctagc gatcgcgcg agatctgtca tgtgagcaaa aggccagcaa 5400
 aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgccccct 5460
 gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa 5520

 agataccagg cgtttccccc tggaagctcc ctctgtcgct ctctgtttcc gaccctgccg 5580
 ctaccggat acctgtccgc cttttccct tcgggaagcg tggcgctttc tcatagctca 5640
 cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa 5700
 cccccgttc agcccgaccg ctgcgcctta tccgtaact atcgtcttga gtccaacccg 5760
 glaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg 5820
 tatgtaggcg gtgctacaga gttcttgaag tggtagccta actacggcta cactagaaga 5880
 acagtatttg gtatctgcgc tctgtgaag ccagttacct tcggaaaaag agttggtagc 5940

 tcttgatccg gcaaacaaac caccgtggt agcgggtggt ttttgtttg caagcagcag 6000
 attacgcga gaaaaaagg atctcaagaa gatccttga tcttttctac ggggtctgac 6060
 gctcagtga acgaaaactc acgttaaggg attttggta tgagattatc aaaaaggatc 6120
 ttcacctaga tcttttctac gtagaaagcc agtccgcaga aacggtgctg acccggatg 6180
 aatgtcagct actgggctat ctggacaagg gaaaacgcaa gcgcaaagag aaagcaggta 6240
 gcttgacgtg ggcttacatg gcgatagcta gactgggcgg ttttatggac agcaagcgaa 6300
 ccggaattgc cagctggggc gccctctggt aaggttggga agccctgcaa agtaactgg 6360

 atggctttct tgccgccaag gatctgatgg cgcaggggat caagatctga tcaagagaca 6420
 ggatgaggat cgtttcgcat gattgaacaa gatggattgc acgcaggttc tccggccgct 6480
 tgggtggaga ggctattcgg ctatgactgg gcacaacaga caatcggtg ctctgatgcc 6540
 gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt ttgtcaagac cgacctgtcc 6600
 ggtgccctga atgaactgca agacaggca gcgcggctat cgtggctggc cagcagggc 6660
 gttccttgcc cagctgtgct gcagttgtc actgaagcgg gaagggactg gctgctattg 6720
 ggcaagtgc cggggcagga tctctgtca tctcaccttg ctctgccga gaaagtatcc 6780

 atcatggctg atgcaatgcg gcggctgcat acgcttgatc cggctacctg cccattcgac 6840
 caccaagcga aacatcgcat cgagcgagca cgtactcgga tggaagccgg tcttgtcgat 6900
 caggatgac tggaacgaaga gcatcagggg ctgcgccag ccgaactgtt cgccaggctc 6960

aaggcgagca tgccccgacgg cgaggatctc gtcgtgaccc atggcgatgc ctgcttgccg 7020
aatatcatgg tggaaaatgg ccgcttttct ggattcatcg actgtggccg gctgggtgtg 7080
gcggaccgct atcaggacat agcgttggct acccgtgata ttgctgaaga gcttggcggc 7140
gaatgggctg accgcttcct cgtgctttac ggtatcgccg ctcccgattc gcagcgcatc 7200

gccttctatc gccttcttga cgagtcttc tgaatttaaa gccaataacg caaacgcct 7260
ctccccgcgc gttggccgaa gcggttcaat attttgtaa aattcgctt aaattttgt 7320
taaatcagct attttttaac caataggccg aaatcgcaa aatcccttgt aaatcaaaag 7380
aatagaccga gatagggttg agtgtgttc cagtttgga caagagtcca ctattaaaga 7440
acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgttggc ccactacgtg 7500
aaccttcacc ctaatcaagt ttttgggggt cgaggtgccg taaagcacta aatcggaacc 7560
ctaaaggag ccccgattt agagcttgac ggggaaaccg gcgaacgtgg cgagaaagga 7620

agggaagaaa gcgaaaggag cggcgcttag ggcgtggca agtglagcgg tcacgtgcg 7680
cgtaaccacc acaccgccc cgctaagcgc cgctacaggg cgcgtccctt cgccttcagg 7740
ctgcgtcgag tactgtactg tgagccagag ttgcccgccg ctctccggt gcggtagtgc 7800
aggcagttca atcaactgtt tacctgttg agcgactcca gaggcacttc accgcttgcc 7860
agcggcttac gatccagcgc cagcatccag tgcaggagat cgttatcgct atacggaaca 7920
ggtattcgct ggtcacttcg ataaggtttg cccggataaa cggaactgga aaaactgctg 7980
ctggtgtttt gcttccgtca gtgctggatc ggcgtgcggt cggcaaagac cagaccgttc 8040

taacagaact ggcgattgtt cggcgatatc caaaaatcac cgccgtaagc cgaccacggg 8100
ttgccgtttt cagcaggatt taatcagcga ctgatccacc cagtcccaga cgaagccgcc 8160
ctgtaaacgg ggatactgac gaaacgcctg ccagtattta gcgaaaccgc caagactgtt 8220
acccaagcgt gggcgatttc gcaaaggatc agcgggcgcg tctctccagg tagcgaaagc 8280
cttttttgat cgaccttgcg gcacagccgg gaagggttg tcttcaacca cgcgcgcgta 8340
caacgggcaa ataatatcgg tggccgtggt gtggtctccg ccgccttcaa ctgcaccggg 8400
cgggaaggat cgacagattt gatccagcga tacagcgcgt cgtgattagc gccgtggcct 8460

gattcaattc ccacgcgacc agtagatcac actcgggtga ttacgattgc gctgcaccag 8520
tcgcttacg gttcgtctt cgccggtagc cagcgcggat cacggtcaga cgattcgttg 8580
gcacgatccg tgggtttcaa tactggcttc aaaccaccac taacaggccg tagcggtcgc 8640
acagcgtgta ccacagcggg tggttcggat aatcgaaacag cgcacggcgt taaagtgtt 8700
ctgcttcaac agcaggatat tctgcactt cgtctgtct tctaacctg accaagcaga 8760
ggatctgctc gtgacgggta atcctcgaat cagcaacggc ttgccgttca gcagcagcag 8820

accaagtcca atccgcacct cgcggaaacc gacaacgcag gcttctgctt caatcagcgt 8880

gccgtcggcg gtgtgcagtt caaccaccgc acgatataga ttcgggattt cggcgtcca 8940

cagtttcggg ttttcgacgt tcagacgtag tgtgacgga tctgcaaacc accacgtca 9000

acgataattt caccgccgaa aggcgcgggt ccgctggcga cctgcgtttc accctgccag 9060

aaagaaactg ttaccctgag gtagtcacgc aactcgccgc aactgaact tcagcctcca 9120

gtacagcgcg gctgaaatcg tcttaagcg agtggcaact ggaaatcgct gatttgtgta 9180

gtcggtttag cagcaacgag acttcacgga aaatccgcta atccgccaca gatcctgac 9240

ttccagataa ctgccgtcac tccaacgcag caccttcacc gcgaggcgggt tttctccggc 9300

gcgtaaaaat cgctcaggtc aaattcagac ggcaaacgac tgtcctggcc gtaaccgacc 9360

cagcgcccggt tgcaccacag attgaaacgc cgagtttacg cctcaaaaat aattcgcgtc 9420

tggccttctt gtagccagct ttcacaacta taatagttag cgagtaacaa cccgtcggat 9480

tctccgtggg acaaaacggc ggattgaccg tatagggata gggtacgttg gtgtagtagg 9540

gcgctccgta accgtgctac tgccagtttg aggggacgac gacagtatcg gcctcaggaa 9600

gacgcactc cagccagctt tccggcaccg cttctggtac tggaaaccag gcaaagcgcc 9660

tatcgcttat caggctgcac aactgttggg aaggcgatc tgtgcgggcc tcttcgtat 9720

tacgccagct tgcgaaaggg ggtagtgctg caaggcgatt aagtgggta acgccagggt 9780

tttccagtc acgacgtgga tctgggccac tcctct 9817

<210> 441

<211> 130

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 441

ctgcgcgtc gctcgtcac tgaggccgc cgggcaaagc cgggcgtcg ggcgacctt 60

ggtcgcccgg cctcagttag cgagcgagcg cgagagagg gaggggcaa ctccatcact 120

aggggttct 130

<210> 442

<211> 130

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 442

aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcaactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcgccctca gtgagcgagc 120

gagcgcgag 130

<210> 443

<211> 145

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 443

aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcaactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcgccctca gtgagcgagc 120

gagcgcgag agagggagtg gccaa 145

<210> 444

<211> 136

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 444

aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcaactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcgccctca gtgagcgagc 120

gagcgcgag agaggg 136

<210> 445

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 445

acgctgggct acttgatc	18
<210> 446	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 446	
cgaggaggag ctattcttg	19
<210> 447	
<211> 18	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic probe	
<400> 447	
tttcgggtcg tgctcatg	18