



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0101631
(43) 공개일자 2022년07월19일

(51) 국제특허분류(Int. Cl.)
A61K 48/00 (2006.01) A61K 38/17 (2006.01)
A61P 1/16 (2006.01) C07K 14/47 (2006.01)

(52) CPC특허분류
A61K 48/005 (2013.01)
A61K 38/1709 (2020.05)

(21) 출원번호 10-2022-7016178
(22) 출원일자(국제) 2022년10월14일
심사청구일자 없음

(85) 번역문제출일자 2022년05월13일
(86) 국제출원번호 PCT/US2020/055500
(87) 국제공개번호 WO 2021/076566
국제공개일자 2021년04월22일

(30) 우선권주장
62/915,765 2019년10월16일 미국(US)

(71) 출원인
유니버시티 오브 피츠버그 - 오브 더 커먼웰스 시스템 오브 하이어 에듀케이션
미국 펜실베이니아 15260, 피츠버그, 130 새커리 애비뉴, 퍼스트 플로어 가드너 스틸 컨퍼런스 센터

(72) 발명자
소토 구티에레스, 알레한드로
미국 펜실베이니아 15221 피츠버그 디넬 드라이브 11
벨, 애런
미국 펜실베이니아 15122 웨스트 미플린 캠벨 서클 2606
(뒷면에 계속)

(74) 대리인
특허법인다나

전체 청구항 수 : 총 34 항

(54) 발명의 명칭 **간 질환 치료를 위한 조성물 및 방법**

(57) 요약

대상체에서 간세포의 핵으로 전사 인자인 HNF4a의 수송 또는 보유를 증가시켜 대상체에서 간 질환을 치료하기 위한 조성물 및 방법이 개시된다. 일부 구현예에서, 방법은 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군으로부터 선택된 하나 이상의 전사 인자 및 그의 기능적 단편의 발현 또는 기능을 상향 조절하고/하거나 하나 이상의 전사 인자 DNAJB1/HSP40, ATF6, ATF4, 및 PERK 및 그의 기능적 단편의 발현 또는 기능을 하향 조절하는 것을 포함한다.

(52) CPC특허분류

A61P 1/16 (2018.01)

C07K 14/4702 (2013.01)

C12N 2740/16043 (2013.01)

(72) 발명자

프라운호퍼 나바로, 니콜라스

미국 펜실베이니아 15213 피츠버그 카토 스트리트
502

구스만 레페, 조지

미국 펜실베이니아 15221 피츠버그 그린 스트리트
230 아파트먼트 1

하이너, 사라

미국 펜실베이니아 15207 피츠버그 윈터번 애비뉴
4245

미칼로폴로스, 조지 케이.

미국 펜실베이니아 15228 피츠버그 랭커스터 애비뉴
172

오스트로프스카, 알리나

미국 펜실베이니아 15136 맥키스 록스 밴 뷰런 서클
4

폭스, 이라

미국 펜실베이니아 15232 피츠버그 피프스 애비뉴
5237 아파트먼트 비4

타팔렝, 에드거 나오에

미국 펜실베이니아 15232 피츠버그 칼리지 스트리트
630

타케이시, 카즈키

일본 8140006 후쿠오카 사와라구 모모치 3-17-11

명세서

청구범위

청구항 1

PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군에서 선택된 하나 이상의 전사 인자의 양 또는 기능을 증가시키는 조성물을 대상체에 투여하는 단계를 포함하는, 간 질환의 치료를 필요로 하는 대상체에서 간 질환을 치료하는 방법.

청구항 2

제1항에 있어서, 상기 조성물은 벡터이고, 상기 벡터는 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C 중 하나 이상을 암호화하는 하나 이상의 핵산을 포함하는 것인, 방법.

청구항 3

제1항에 있어서, 상기 조성물은 벡터이고, 상기 벡터는 PROX1 및/또는 SREBP1을 암호화하는 하나 이상의 핵산을 포함하는 것인, 방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 하나 이상의 핵산은 DNA 또는 RNA인, 방법.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 조성물의 투여가 대상체의 간세포의 핵에서 HNF4 α 의 양을 증가시키는 것인, 방법.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, 상기 조성물의 투여가 간세포에서 HNF4 α 의 총량을 증가시키지 않는 것인, 방법.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, 상기 조성물의 투여가 간세포에서 HNF4 α 의 총량을 증가시키는 것인, 방법.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 상기 벡터가 HNF4 α 를 암호화하는 핵산을 더 포함하는 것인, 방법.

청구항 9

제1항 내지 제7항 중 어느 한 항에 있어서, HNF4 α 를 암호화하는 핵산을 포함하는 벡터를 대상체에게 투여하는 단계를 더 포함하는, 방법.

청구항 10

제8항 또는 제9항에 있어서, 상기 핵산은 HNF4 α 이소형 2 (P1)을 암호화하는 것인, 방법.

청구항 11

제8항 또는 제9항에 있어서, 상기 HNF4 α 를 암호화하는 핵산은 서열번호 1을 포함하는 것인, 방법.

청구항 12

제1항 내지 제11항 중 어느 한 항에 있어서, 상기 간 질환은 간 섬유증, 간경변, 간암 또는 말기 간 질환인, 방법.

청구항 13

제1항 내지 제12항 중 어느 한 항에 있어서, 상기 간 질환은 간경변인, 방법.

청구항 14

제1항 내지 제13항 중 어느 한 항에 있어서, 상기 대상체는 인간인, 방법.

청구항 15

DNAJB1/HSP40, ATF6, ATF4, 및 PERK으로 이루어진 군으로부터 선택된 하나 이상의 전사 인자의 양을 감소시키거나 기능을 억제시키는 조성물을 대상체에 투여하는 단계를 포함하는, 간 질환의 치료를 필요로 하는 대상체에서 간 질환을 치료하는 방법.

청구항 16

제15항에 있어서, 상기 조성물은 핵산인, 방법.

청구항 17

제16항에 있어서, 상기 핵산은 DNA 또는 RNA인, 방법.

청구항 18

제15항 내지 제17항 중 어느 한 항에 있어서, 상기 조성물의 투여가 대상체의 간세포의 핵에서 HNF4 α의 양을 증가시키는 것인, 방법.

청구항 19

제15항 내지 제18항 중 어느 한 항에 있어서, 상기 조성물의 투여가 간세포에서 HNF4 α의 총량을 증가시키지 않는 것인, 방법.

청구항 20

제15항 내지 제19항 중 어느 한 항에 있어서, 상기 조성물의 투여가 간세포에서 HNF4 α의 총량을 증가시키는 것인, 방법.

청구항 21

제15항 내지 제20항 중 어느 한 항에 있어서, 상기 조성물은 HNF4 α를 암호화하는 핵산을 더 포함하는 것인, 방법.

청구항 22

제15항 내지 제21항 중 어느 한 항에 있어서, HNF4 α를 암호화하는 핵산을 포함하는 벡터를 대상체에게 투여하는 단계를 더 포함하는, 방법.

청구항 23

제21항 또는 제22항에 있어서, 상기 핵산은 HNF4 α 이소형 2를 암호화하는 것인, 방법.

청구항 24

제21항 또는 제22항에 있어서, 상기 HNF4 α를 암호화하는 핵산은 서열번호 1을 포함하는 것인, 방법.

청구항 25

제15항 내지 제24항 중 어느 한 항에 있어서, 상기 간 질환은 간 섬유증, 간경변, 간암 또는 말기 간 질환을 포함하는 것인, 방법.

청구항 26

제15항 내지 제25항 중 어느 한 항에 있어서, 상기 대상체는 인간인, 방법.

청구항 27

PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군으로부터 선택된 하나 이상의 전사 인자, 및 그의 기능적 단편을 암호화하는 하나 이상의 핵산을 포함하는 벡터를 포함하는 조성물.

청구항 28

제27항에 있어서, HNF4 α를 암호화하는 핵산을 포함하는 또 다른 벡터를 더 포함하는 조성물.

청구항 29

제28항에 있어서, 상기 HNF4 α를 암호화하는 핵산은 서열번호 1을 포함하는 것인, 조성물.

청구항 30

HNF4 α 이소형 2를 암호화하는 핵산을 포함하는 벡터를 대상체에 투여하는 단계를 포함하는, 간 질환의 치료를 필요로 하는 대상체에서 간 질환을 치료하는 방법.

청구항 31

제30항에 있어서, 상기 핵산은 서열번호 1을 포함하는 것인, 방법.

청구항 32

제30항 또는 제31항에 있어서, 상기 간 질환은 간 섬유증, 간경변, 간암 또는 말기 간 질환인, 방법.

청구항 33

제32항에 있어서, 상기 간 질환은 간경변인, 방법.

청구항 34

제30항 내지 제33항 중 어느 한 항에 있어서, 상기 대상체는 인간인, 방법.

발명의 설명

기술 분야

[0001] **관련 출원에 대한 상호-참조**

[0002] 본 출원은 2019년 10월 16일자 출원된 미국 가출원 번호 62/915,765의 우선권을 주장하며, 전체로서 본원에 참조로 도입된다.

[0003] **연방 지원 연구 또는 개발에 대한 진술**

[0004] 본 발명은 국립 보건원 (National Institutes of Health)이 부여한 승인 번호 DK099257에 따른 정부 지원으로 이루어졌다. 정부는 발명에 대해 특정 권리를 갖는다.

[0005] 본 개시는 예를 들어 간 질환 및/또는 간 장애를 치료하기 위한 HNF4 α 발현 또는 활성의 조절과 관련이 있다.

배경 기술

[0006] 진행성 간경변(advanced liver cirrhosis)으로 인한 말기 간기능 부전(TLF)은 2015년 전 세계적으로 인구 100,000명당 15.8명이 사망하는 12번째 주요 사망 원인으로 나타났다 (Tsochatzis EA et al., 2014). 미국에서 2015년 만성 간 질환과 간경변으로 인한 사망 등록 건수는 40,326명이었고 사망률은 인구 100,000명당 12.5명으로 연간 3.8% 증가했다 (Murphy SL et al., 2017; Goldman L, et al., 2016). 가장 영향을 받는 연령대는 45세에서 64세 사이이며, 사망률은 인구 100,000명당 26.4명으로, 만성 간 질환 및 간경변이 이 연령대에서 암, 심장 질환, 및 의도치않은 부상에 이어 네 번째 주요 사망 원인이다 (Murphy SL et al., 2017). TLF에 대한 유일한 결정적인 치료법은 간 이식이 필요한 환자의 수와 사용 가능한 장기의 불충분한 수를 고려할 때 TLF를 근본적으로 치료할 수 없는 질병으로 만드는 치환 간 이식(orthotopic liver transplantation)이다 (Lopez PM et

al., 2006).

- [0007] 간염 바이러스에 의한 만성 감염, 알코올-매개 간경변, 및 비-알코올성 지방간염(NASH)을 포함하여 만성 간 질환의 원인은 다양하며 (Archambeaud I et al., 2015; Donato F et al., 2006; Gelatti U et al., 2005; Kuper H et al., 2000), 각각은 간세포 부전을 일으킬 수 있다 (Guzman-Lepe J et al., 2018; Hernaez R et al., 2017; Lee YA et al., 2015; Pessayre D et al., 1978). 인간의 간세포 기능의 악화와 궁극적으로 간 부전의 원인이 되는 메커니즘은 잘 알려져 있지 않다.
- [0008] 만성 간질환, 간경변, 및 최근 TLF의 주요 원인은 B형 및 C형 간염 바이러스에 의한 감염, 알코올 매개 Laennec's 간경변, 비알코올성 지방간염(NASH)/대사 증후군과 관련이 있다 (Archambeaud et al., 2015; Donato F et al., 2006; Gelatti et al., 2005; Kuper et al., 2000). 이러한 병원(etiological agent)은 혈관 구조의 변경으로 정상적인 소엽 구조(lobular architecture)를 방해하는 섬유증을 유발한다 (Goldman L, et al., 2016). 이러한 병리학적 변화는 간세포 부전 및 간세포가 정상적인 기능을 수행하지 못하는 것과 관련이 있으나 (Guzman-Lepe J et al., 2018; Hernaez R et al., 2007; Lee Ya et al., 2015; Pessayre D et al., 1978), 간세포 기능의 악화와 궁극적으로 간 부전에 대한 메커니즘은 인간에게 알려져 있지 않다. 만성 간 손상은 세포 사멸을 유도하는 (Cichoz-Lach H et al., 2014; Malhi H et al., 2011; Zhang XQ et al., 2014; Wang K et al., 2014; Seki E et al., 2015) 산화 스트레스 (Cichoz-Lach H et al., 2014; Simoes ICM et al., 2018) 및 소포체 스트레스 (Malhi H et al., 2011; Zhang XQ et al., 2014)를 생성하며, 궁극적으로 간세포의 증식 능력을 감소시킨다 (Zhang BH et al., 1999; Michalopoulos GK et al., 2015; Dubuquoy L et al., 2015).
- [0009] 말기 간경변이 있는 쥐의 간세포에서 간-강화 전사 인자는 안정적으로 하향조절되고 (Nishikawa T et al., 2014; Guzman-Lepe J et al., 2019), 그 중 하나인 간세포 핵 인자 4 알파 (hepatocyte nuclear factor 4 alpha, HNF4 α)의 강제 재발현은 기능 장애 간세포를 재프로그래밍하여 배양 및 생체 내 모두에서 기능을 회복하는 것으로 확인되었다 (Nishikawa T et al., 2014). 진행성 간 질환 환자의 대규모 코호트 연구는, 병든 간에서 HNF4 α mRNA 발현 수준이 간 기능 장애의 정도 (Childs-Pug 분류)와 상관관계가 있으며, HNF4 α 발현이 핵에 국한되지 않음을 보여준다 (Guzman-Lepe J et al., 2019). 간세포 핵 인자 4 알파 (HNF4 α)의 강제 재발현은 새로운 간세포 또는 줄기 세포의 확장 없이도 말기 간경변 간에서 기능 장애 간세포를 재프로그래밍하여 배양 및 생체 내에서 다시 기능하도록 할 수 있다 (Nishikawa T et al., 2014). HNF4 α 핵 국소화 및 mRNA 발현의 상당한 감소를 포함한 LETF의 하향 조절은 TLF가 있는 인간 간의 대규모 코호트에서 간 기능 장애 정도와 관련이 있다 (Guzman-Lepe J et al., 2018).
- [0010] HNF4 α 는 성인 간에서 간 기관 형성 및 간세포 기능에 중요한 역할을 하는 전사 인자이다 (Nishikawa T et al., 2014; Babeu JP et al., 2014). 주요 HNF4 α 작용은 지질, 포도당, 생체이물(xenobiotic), 및 약물 대사와 관련된 특정 표적 유전자의 조절이다 (Nishikawa T et al., 2014; Babeu JP et al., 2014). 단일 유전자는 인간에서 HNF4 α 를 암호화하며 (Kritis AA et al., 1999), 이는 두 개의 다른 프로모터에 의해 조절된다. 이러한 프로모터는 P1 및 P2의 두 가지 이소형 부류를 생성한다 (Babeu JP et al., 2014). P1 이소형은 주로 성인 간에서 발현되는 반면, P2 이소형은 배아 발달 및 암과 같은 병리학적 조건 하에서 간에서 검출되었다 (Babeu JP et al., 2014; Walesky C et al., 2015; Tanaka T et al., 2006). HNF4 α 의 발현과 기능은 여러 수준에서 조절된다 (Chellappa K et al., 2012; Guo H et al., 2014; Hong YH et al., 2003; Lu H et al., 2016, Song Y et al., 2015; Soutoglou E et al., 2000; Sun K et al., 2007; Xu Z et al., 2007; Yokoyama A et al., 2011; Zhou W et al., 2012).
- [0011] 따라서, HNF4 α 발현을 조절하고/하거나 간 질환 및/또는 간 장애를 치료하기 위한 조성물 및 방법이 필요하다. 본원에 개시된 조성물 및 방법은 이러한 요구 및 기타 요구를 다룬다.

발명의 내용

- [0012] 본원에 개시된 조성물 및 방법은 간 질환 치료에서 충족되지 않은 특정 요구를 다룬다. 일부 측면에서, 간 질환 및/또는 간 장애의 치료용 약제를 위한 조성물 및 이의 용도가 본원에 개시되며, 여기서 상기 조성물은 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및/또는 POM121C로 이루어진 군으로부터 선택된 하나 이상의 전사 인자의 양 또는 기능을 증가시키거나 DNAJB1/HSP40, ATF6, ATF4, 및 PERK으로 이루어진 군으로부터 선택된 하나 이상의 전사 인자의 양을 감소시키거나 기능을 억제한다. 일부 구현예에서, 상기 조성물은 백터이고, 상기 백터는 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및/또는 POM121C 중 하나 이상을 암호화하는 하나 이상의 핵산을 포함한다. 일부 구현예에서, 상기 조성물은 HNF4 α (예를 들어, HNF4 α 이소형 2)를 암호화하는 핵산을 포함하는 백터이다. 본원에 개시된 조성물 및 방법은 간세포에서 HNF4 α 의 양의 놀라운 증가를 초래하여 (예를 들어,

간세포에서 HNF4 α 총량의 증가, 및/또는 간세포의 핵에서 HNF4 α의 양의 증가), 간 질환의 효과적인 치료를 초래한다 (예를 들어, 말기 간 질환).

- [0013] 일부 측면에서, 본 발명은 대상체에게 조성물을 투여하는 것을 포함하는, 간 질환의 치료를 필요로 하는 대상체에서 간 질환을 치료하는 방법을 개시하며, 여기서 상기 조성물은 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자의 양 또는 기능을 증가시킨다.
- [0014] 일부 측면에서, 본 발명은 대상체에게 조성물을 투여하는 것을 포함하는, 간 질환의 치료를 필요로 하는 대상체에서 간 질환 치료용 약제의 제조를 위한 조성물의 용도를 개시하며, 여기서 상기 조성물은 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자의 양 또는 기능을 증가시킨다.
- [0015] 일부 구현예에서, 상기 조성물은 벡터이고, 여기서 벡터는 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C 중 하나 이상을 암호화하는 하나 이상의 핵산을 포함한다. 일부 구현예에서, 상기 벡터는 PROX1 및/또는 SREBP1를 암호화하는 하나 이상의 핵산을 포함한다. 상기 하나 이상의 핵산은 DNA 또는 mRNA일 수 있다.
- [0016] 일부 측면에서, 본 발명은 대상체에게 조성물을 투여하는 것을 포함하는, 간 질환의 치료를 필요로 하는 대상체에서 간 질환을 치료하는 방법을 개시하며, 여기서 상기 조성물은 DNAJB1/HSP40, ATF6, ATF4, 및 PERK로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자의 양을 감소시키거나 기능을 억제한다.
- [0017] 일부 측면에서, 본 발명은 대상체에게 조성물을 투여하는 것을 포함하는, 간 질환의 치료를 필요로 하는 대상체에서 간 질환 치료용 약제의 제조를 위한 조성물의 용도를 개시하며, 여기서 상기 조성물은 DNAJB1/HSP40, ATF6, ATF4, 및 PERK로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자의 양 또는 기능을 감소시킨다.
- [0018] 일부 구현예에서, 상기 조성물의 투여는 대상체의 간세포 핵에서 HNF4 α의 양을 증가시킨다. 일부 구현예에서, 상기 조성물의 투여는 간세포에서 HNF4 α의 총량을 증가시키지 않는다. 일부 구현예에서, 상기 조성물의 투여는 간세포에서 HNF4 α의 총량을 증가시킨다.
- [0019] 일부 구현예에서, 상기 벡터는 HNF4 α를 암호화하는 핵산을 더 포함한다. 일부 구현예에서, 상기 방법은 HNF4 α를 암호화하는 핵산을 포함하는 벡터를 대상체에게 투여하는 것을 더 포함한다. 일 예시에서, 상기 핵산은 HNF4 α 이소형 2를 암호화한다.
- [0020] 일부 측면에서, 본 발명은 벡터를 포함하는 조성물을 개시하며, 여기서 상기 벡터는 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자, 및 이의 기능적 단편을 암호화하는 하나 이상의 핵산을 포함한다.
- [0021] 일부 측면에서, 본 발명은 대상체에게 HNF4 α 이소형 2를 암호화하는 핵산을 포함하는 벡터를 투여하는 것을 포함하는, 간 질환의 치료를 필요로 하는 대상체에서 간 질환을 치료하는 방법을 개시한다.
- [0022] 일부 측면에서, 본 발명은 간 질환의 치료용 약제의 제조를 위한 벡터의 용도를 개시하며, 여기서 상기 벡터는 HNF4 α 이소형 2를 암호화하는 핵산을 포함한다.

도면의 간단한 설명

[0023] 도 1A-1C는 정상 및 간경변 간의 간세포에서 HNF4 α 위치를 나타낸다. 도 1A는 NASH 비대상 간에서 분리된 간세포 및 정상 인간 간세포의 HNF4 α 면역형광의 대표 사진을 나타낸다. 기능적으로 비대상된 간 (NASH 및 알코올-매개 Laennec's 간경변)(n=6)에서 분리된 간세포 및 정상 간 대조군(n=2)에서 분리된 간세포에서 베타-액틴으로 정규화된 HNF4 α의 웨스턴 블롯 분석 및 정량화. 도 1A는 총 HNF4 α를 나타낸다; 정상 인간 간세포 vs 비대상 인간 간세포, P = 0.166. 도 1B는 세포질 HNF4 α를 나타낸다; 정상 인간 간세포 vs 비대상 인간 간세포, P = 0.023. 도 1C는 핵 HNF4 α를 나타낸다; 정상 인간 간세포 vs 비대상 인간 간세포, P = 0.023). A에서 C까지의 그래프는 평균±SD로 표시된다. 통계적 유의미(P < 0.05). 다이아몬드는 Child-Pugh "B"를 나타내고 사각형은 Child-Pugh "C"를 나타낸다.

도 2A-2C는 HNF4 α 번역 후 변형자의 단백질 발현 및 스피어만 순위 상관 테스트를 보여준다. 도 2A 및 2B는 비대상 NASH(n=4) 및 알코올 매개 Laennec's 간경변 간세포(n=2) 및 정상 대조군 간세포(n=2)에서 EGFR(P=0.904), cMET(P=0.023), 총 AMPK α (P > 0.999), p-AMPK α (Thr172)(P=0.547), 총 AKT(P=0.047), p-AKT(Ser473)(P=0.547), p-AKT(Thr308)(P=0.024), p-AMPK α (Thr172)/AMPK α 비율(P=0.5476), p-AKT(Ser473)/AKT 비율(P=0.1667) 및 p-AKT(Thr308)/총 AKT 비율(P=0.0238)의 웨스턴 블롯 분석(2A) 및 정량화

(2B)를 보여준다. 도 2B에서 채워진 사각형은 Child-Pugh "B"를 나타내고 채워진 다이아몬드는 Child-Pugh "C"를 나타낸다. 도 2C는 스피어만 순위 상관 테스트가 표시되고 핵 HNF4 α 가 cMET ($r=0.71$; $P=0.037$), 총 AKT ($r=0.71$; $P=0.037$), 포스포-AKT(Thr308) ($r=0.82$; $P=0.011$), 및 포스포-AKT(Thr308)/총 AKT 비율 ($r=0.73$; $P=0.031$)과 유의한 상관관계가 있음을 보여준다. 세포질 HNF4 α 는 cMET ($r=-0.80$; $P=0.014$), 총 AKT ($r=-0.73$; $P=0.031$), 포스포-AKT(Thr308) ($r=-0.77$; $P=0.021$), 및 포스포-AKT(Thr308)/총 AKT 비율 ($r=-0.72$; $P=0.037$)과 유의한 상관관계를 나타낸다. 도 2B의 막대 그래프는 평균 \pm SD로 표시된다. 통계적 유의미($P < 0.05$).

도 3A-3D는 번역 후 변형자와 HNF4 α 세포 국소화의 관계를 보여준다. 도 3A는 경로 분석이 cMET(0.56; $P=0.004$), 포스포AKT(Thr308)/총 AKT 비율(0.05; $P=0.006$) 및 총 HNF4 α 수준(0.60; $P=0.042$)과 HNF4 α 국소화 사이에 유의한 직접적 관계가 있음을 보여준다. 경로 분석은 cMET과 총 HNF4 α (-0.37; $P=0.024$) 사이에 유의한 관계가 있음을 보여준다. 도 3B는 선형 회귀 분석이 핵 HNF4 α 발현과 간 기능 장애의 정도(Child-Pugh 점수)의 유의한 관계를 보여줌($R^2=0.80$, $P=0.007$)을 도시한다. 도 3C는 주성분 분석(PCA)이 HNF4 α 발현과 상관관계가 있는 단백질 프로파일을 보여주며 정상 인간 간세포($n=2$)의 특성을 명확히 설명하는 반면, 세포질 HNF4 α , 활성 카스파제 3, p-AKT(Ser473)/총 AKT 비율 및 p-AMPK(Ser172)/AMPK는 NASH($n=4$) 및 알코올 매개 Laennec's 간경변 간($n=2$)의 비대상된 인간 간세포의 특성을 설명한다. 도 3D의 그래프는 정상 인간 간세포와 비교하여 비대상 인간 간세포(차일드-피 분류 B 및 C)에서 전체, 핵 또는 세포질 HNF4 α (상단 행의 세 그래프), cMET, p-AMPK(Ser172)/AMPK, p-AKT(Ser473)/총 AKT 및 포스포AKT(Thr308)/총 AKT 비율, p-H3(Ser10) 및 활성화된 카스파제 (두 번째 및 맨 아래 행의 그래프)의 PCA 분석에 사용된 단백질 발현 배수 변화를 보여준다. 그래프는 평균 \pm SD로 표시된다. 통계적 유의미($P < 0.05$).

도 4A-4C는 핵 HNF4 α 의 아세틸화가 NASH 및 알코올 매개 Laennec's 간경변 외식 간(explant liver)의 인간 비대상 간세포에서 변화됨을 나타낸다. 도 4A 및 4B는 비대상 NASH($n=4$) 및 알코올 매개 Laennec's 간경변($n=4$) 외식간($P=0.024$)에서 인간 간세포의 핵 분획에서 HNF4 α (Lys106)의 아세틸화된 형태의 웨스턴 블롯 및 정량화를 보여준다. 도 4C의 선형 회귀 분석은 감소된 아세틸화 형태의 HNF4 α (Lys106)와 간 기능 장애($R^2=0.71$, $P=0.004$)의 유의한 상관 관계를 보여준다. 막대 그래프는 평균 \pm SD로 표시된다. 통계적 유의미($P < 0.05$).

도 5는 HNF4 α 번역 후 변형(PTM)의 인 실리코 분석을 보여준다. 도 5는 HNF4 α -PTM의 목록을 제공한다.

도 6A-6B는 활성화된 AKT 경로와 HNF4 α 번역 후 변형자와 NASH 및 알코올 매개 Laennec's 간경변 외식간에서 인간 비대상 간세포의 p-EGFR 발현과의 관계를 보여준다. 도 6A는 포스포-AKT(Thr308)/AKT에 대한 스피어만 상관관계를 보여준다. 도 6B는 포스포-AKT(Ser473)/AKT에 대한 스피어만 상관관계를 보여준다.

도 7A-7C는 갓 분리된 정상 인간 간세포와 비교하여 mRNA 수준(도 7A)에서, 및 알코올 간염 간세포의 약 40%만이 약한 강도로 핵에서 HNF4 α 를 발현하고 세포의 약 10%가 HNF4 α 세포질 발현을 갖는다는 것을 면역조직화학에 의해 입증하여 단백질 수준(도 7B)에서 알코올성 간염 환자의 외식된 간에서 분리한 인간 간세포에서 간-강화 전사 인자 HNF4 α 의 발현 수준을 보여준다. 도 7C의 경우, 갓 분리된 인간 알코올 간염 간세포를 HNF4 α 를 암호화하는 렌티바이러스(Systems Bioscience, Cat#CS970S-1; CD511B-1의 HNF4 α)로 처리했다. 이 도면은 72시간 후에, HNF4 α 발현이 핵에서 HNF4 α 를 발현하는 간세포의 핵분율을 변경하지 않았음을 보여준다. 그러나 HNF4 α 발현 강도는 기존 세포에서 급격히 증가하였다. 전체적으로 이 데이터는 핵으로의 HNF4 α 수송이 알코올성 간염에 걸린 인간에서 간세포가 기능을 회복하는 데 중요한 역할을 할 수 있음을 보여준다.

도 8A-8F는 1차 인간 간세포에서 MTF1 발현을 보여준다. 도 8A는 NASH 또는 알코올 유발 간경변증으로 간 이식을 받은 환자의 간에서 분리된 1차 인간 간세포가 웨스턴 블롯에 의해 MTF1(MA5-26738 1:1000) 및 HNF4 α (ab41898 1:1000)의 발현에 대해 분석되었음을 보여준다. 도 8B 및 8C는 대조군 간세포, 차일드 B 간세포 및 차일드 C 간세포 중 HNF4 α (도 8B 및 8C) 및 MTF1(도 8D 및 8E)의 상대 강도가 다중 비교를 위해 Brown-Forsythe 및 Welch ANOVA 테스트와 함께 일원 분산 분석(One-Way ANOVA)에 의해 비교되었음을 보여준다. HNF4 α 및 MTF1의 발현은 차일드 B 간세포에 비해 차일드 C 간세포에서 더 낮았다. 도 8B 및 8D, * $p < 0.003$, ** $p < 0.01$, *** $p < 0.0001$, $n=25$. 도 8C $R^2=0.019$, $p=0.06$, $n=19$. 도 8E $R^2=0.015$, $p=0.1$, $n=19$. 도 8F는 차일드-피 점수, HNF4 α 의 단백질 발현 및 MTF1과의 상관 관계 연구가 단순 선형 회귀를 사용하여 수행되었음을 보여준다. 검은색 원은 대조군, 밝은 회색 원은 차일드 B, 어두운 회색 원은 차일드 C를 나타낸다. HNF4 α 의 단백질 발현은 MTF1의 단백질 발현과 상관관계가 있다($R^2=0.28$, $p=0.007$, $n=25$).

도 9A-9D는 1차 인간 간세포에서 NR0B2 발현을 보여준다. 도 9A는 NASH 또는 알코올 유발 간경변증으로 간 이식

을 받은 환자의 간에서 분리된 1차 인간 간세포가 웨스턴 블롯에 의해 NROB2(Abclonal A1836 1:500) 및 HNF4 α (ab41898 1:1000)의 발현에 대해 분석되었음을 보여준다. 도 9B 및 9C는 대조군 간세포, 차일드 B 간세포 및 차일드 C 간세포 중 NROB2의 상대 강도가 다중 비교를 위해 Brown-Forsythe 및 Welch ANOVA 테스트와 함께 일원 분산 분석(One-Way ANOVA)에 의해 비교되었음을 보여준다. 도 9B * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, $n = 25$. 도 9C $R^2 = 0.19$, $p = 0.06$, $n = 19$. NROB2의 발현은 차일드 C, 차일드 B 및 대조군 간세포에서 다르다. 도 9D 차일드-퍼 점수, HNF4 α 의 단백질 발현 및 NROB2와의 상관 관계 연구가 단순 선형 회귀를 사용하여 수행되었음을 보여준다. 검은색 원은 대조군, 밝은 회색 원은 차일드 B, 짙은 회색 원은 차일드 C를 나타낸다($R^2 = 0.12$, $p = 0.1$, $n = 25$).

도 10A-10D는 1차 인간 간세포에서 NR5A2 발현을 보여준다. 도 10A는 NASH 또는 알코올 유발 간경변증으로 간 이식을 받은 환자의 간에서 분리된 1차 인간 간세포가 웨스턴 블롯에 의해 NR5A2(Novus NBP2-27196 1:500) 및 HNF4 α (ab41898 1:1000)의 발현에 대해 분석되었음을 보여준다. 도 10B 및 10C는 대조군 간세포, 차일드 B 간세포 및 차일드 C 간세포 중 NR5A2의 상대 강도가 다중 비교를 위해 Brown-Forsythe 및 Welch ANOVA 테스트와 함께 일원 분산 분석(One-Way ANOVA)에 의해 비교되었음을 보여준다. 도 10B * $p < 0.05$, $n = 25$. 도 10C $R^2 = 0.17$, $p = 0.07$, $n = 19$. NR5A2의 발현은 차일드 B, 차일드 C 및 대조군 간세포에서 다르다. 도 10B는 차일드-퍼 점수, HNF4 α 의 단백질 발현 및 NROB2와의 상관 관계 연구가 단순 선형 회귀를 사용하여 수행되었음을 보여준다. NR5A2의 단백질 발현은 HNF4 α 의 발현과 상관관계가 있다. 검은색 원은 대조군, 밝은 회색 원은 차일드 B, 짙은 회색 원은 차일드 C를 나타낸다($R^2 = 0.17$, $p < 0.05$, $n = 25$).

도 11A-11D는 1차 인간 간세포에서 Prox1 발현을 보여준다. 도 11A는 NASH 또는 알코올 유발 간경변증으로 간 이식을 받은 환자의 간에서 분리된 1차 인간 간세포가 웨스턴 블롯에 의해 PROX1(R&D AF2727 1:500) 및 HNF4 α (ab41898 1:1000)의 발현에 대해 분석되었음을 보여준다. 도 11B 및 11C는 대조군 간세포, 차일드 B 간세포 및 차일드 C 간세포 중 PROX1의 상대 강도가 다중 비교를 위해 Brown-Forsythe 및 Welch ANOVA 테스트와 함께 일원 분산 분석(One-Way ANOVA)에 의해 비교되었음을 보여준다. 도 11B * $p < 0.02$, $n = 25$. 도 11C $R^2 = 0.02$, $p = 0.6$, $n = 19$. PROX1의 발현은 차일드 C 및 대조군 간세포에서 다르다. 도 11D는 차일드-퍼 점수, HNF4 α 의 단백질 발현 및 PROX1과의 상관 관계 연구가 단순 선형 회귀를 사용하여 수행되었음을 보여준다. 검은색 원은 대조군, 밝은 회색 원은 차일드 B, 짙은 회색 원은 차일드 C를 나타낸다($R^2 = 0.02$, $p = 0.46$, $n = 25$).

도 12A-12D는 1차 인간 간세포에서 POM121C 발현을 보여준다. 도 12A는 NASH 또는 알코올 유발 간경변증으로 간 이식을 받은 환자의 간에서 분리된 1차 인간 간세포가 웨스턴 블롯에 의해 POM121C(PA5-85161 1:500) 및 HNF4 α (ab41898 1:1000)의 발현에 대해 분석되었음을 보여준다. 도 12B 및 12C는 대조군 간세포, 차일드 B 간세포 및 차일드 C 간세포 중 POM121C의 상대 강도가 다중 비교를 위해 Brown-Forsythe 및 Welch ANOVA 테스트와 함께 일원 분산 분석(One-Way ANOVA)에 의해 비교되었음을 보여준다. 도 12B $n = 25$. 도 12C $R^2 = 0.08$, $p = 0.24$, $n = 25$. 도 12D는 차일드-퍼 점수, HNF4 α 의 단백질 발현 및 POM121C와의 상관 관계 연구가 단순 선형 회귀를 사용하여 수행되었음을 보여준다. 검은색 원은 대조군, 밝은 회색 원은 차일드 B, 짙은 회색 원은 차일드 C를 나타낸다($R^2 = 0.06$, $p = 0.23$, $n = 25$).

도 13A-13D는 1차 인간 간세포에서 SREBP1 발현을 보여준다. 도 13A는 NASH 또는 알코올 유발 간경변증으로 간 이식을 받은 환자의 간에서 분리된 1차 인간 간세포가 웨스턴 블롯에 의해 SREBP1(Abcam ab28481 1:500) 및 HNF4 α (ab41898 1:1000)의 발현에 대해 분석되었음을 보여준다. 도 13B 및 13C는 대조군 간세포, 차일드 B 간세포 및 차일드 C 간세포 중 SREBP1의 상대 강도가 다중 비교를 위해 Brown-Forsythe 및 Welch ANOVA 테스트와 함께 일원 분산 분석(One-Way ANOVA)에 의해 비교되었음을 보여준다. 도 13B $n = 25$. 도 13C $R^2 = 0.02$, $p = 0.54$, $n = 19$. 도 13D는 차일드-퍼 점수, HNF4 α 의 단백질 발현 및 SREBP1과의 상관 관계 연구가 단순 선형 회귀를 사용하여 수행되었음을 보여준다. 검은색 원은 대조군, 밝은 회색 원은 차일드 B, 짙은 회색 원은 차일드 C를 나타낸다($R^2 = 0.01$, $p = 0.86$, $n = 25$).

도 14A-14D는 1차 인간 간세포에서 EP300 발현을 보여준다. 도 14A는 NASH 또는 알코올 유발 간경변증으로 간 이식을 받은 환자의 간에서 분리된 1차 인간 간세포가 웨스턴 블롯에 의해 EP300(Novus NB100-616 1:500) 및 HNF4 α (ab41898 1:1000)의 발현에 대해 분석되었음을 보여준다. 도 14B 및 14C는 대조군 간세포, 차일드 B 간세포 및 차일드 C 간세포 중 EP300의 상대 강도가 다중 비교를 위해 Brown-Forsythe 및 Welch ANOVA 테스트와 함께 일원 분산 분석(One-Way ANOVA)에 의해 비교되었음을 보여준다. 도 14B $n = 25$. 도 14C $R^2 = 0.32$, $p = 0.01$,

n=19. 도 14D는 차일드-퍼 점수, HNF4 α의 단백질 발현 및 EP300과의 상관 관계 연구가 단순 선형 회귀를 사용하여 수행되었음을 보여준다. 검은색 원은 대조군, 밝은 회색 원은 차일드 B, 짙은 회색 원은 차일드 C를 나타낸다($R^2=0.01$, $p=0.69$, $n=25$).

도 15A 및 B는 HepG2 세포에서 EP300, MTF1, NROB2, NR5A2, POM121C, PROX1 또는 SREBP1의 CRISPR/Cas9 녹아웃이 수행되었고 세포 HNF4 α 위치가 면역형광에서 분석되었음을 보여준다(ab41898 1:500). DAPI 및 HNF4 α에서 양성인 핵의 총 수(도 15A) 및 세포질에서 HNF4 α에 대해 양성인 세포(도 15B)를 카운팅하였다. 다중 비교를 위해 Brown-Forsythe 및 Welch ANOVA 테스트와 함께 일원 분산 분석(One-Way ANOVA)을 사용하여 통계 분석을 수행했다. EP300, MTF1, NROB2, NR5A2, POM121C, PROX1 및 SREBP1의 녹아웃은 HNF4 α의 핵 위치 감소와 HNF4 α의 세포질 위치 증가를 보여주었다. * $p<0.05$.

도 16은 간 이식을 겪은 NASH 환자로부터 분리된 1차 인간 간세포가 105의 MOI를 갖는 AAV-MTF1, NROB2, NR5A2, POM121C, PROX1, SREBP1 또는 GFP 및 AAV-HNF4 α로 형질도입되었음을 나타낸다. HNF4 α에 대해 양성인 핵의 백분율을 계산했다. 다중 비교를 위해 Brown-Forsythe 및 Welch ANOVA 테스트와 함께 일원 분산 분석(One-Way ANOVA)을 사용하여 통계 분석을 수행했다. HNF4 α, MTF1, NROB2, POM121C, PROX1 및 SREBP1과의 동시 형질도입은 GFP/HNF4 α 동시 형질도입 그룹에 비해 증가된 양성 핵의 수를 갖는다(**** $p<0.0001$, *** $p<0.0005$, ** $p<0.001$, * $p<0.05$).

발명을 실시하기 위한 구체적인 내용

- [0024] 본원은 대상체에서 간세포의 핵으로 전사 인자인 HNF4 α의 발현 및/또는 수송 또는 보유를 증가시켜 대상체에서 간 질환을 치료하기 위한 조성물 및 방법을 개시한다. 일부 구현예에서, 상기 방법은 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자 및 그의 기능적 단편의 발현 또는 기능을 상향 조절하고/하거나 하나 이상의 전사 인자 DNAJB1/HSP40, ATF6, ATF4, 및 PERK 및 그의 기능적 단편의 발현 또는 기능을 하향 조절하는 것을 포함한다. 이러한 전사 인자가 HNF4 α의 발현 및/또는 국소화를 조절하여 간 질환 치료에 사용될 수 있다는 것은 놀라운 발견이다.
- [0025] 일부 구현예에서, 상기 방법은 벡터를 투여하는 단계를 포함하며, 여기서 상기 벡터는 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군으로부터 선택된 하나 이상의 전사 인자, 및 그의 기능적 단편을 암호화하는 핵산(예를 들어, DNA, ceDNA, 또는 mRNA)을 포함한다. 일부 측면에서, 상기 방법은 HNF4 α(예를 들어, HNF4 α 이소형 2)를 암호화하는 핵산(예를 들어, DNA, ceDNA 또는 mRNA)을 포함하는 벡터를 투여하는 단계를 포함한다. 다른 또는 추가 구현예에서, 상기 방법은 조성물을 투여하는 단계를 포함하며, 여기서 상기 조성물은 DNAJB1/HSP40, ATF6, ATF4, 및 PERK로 이루어진 군으로부터 선택된 하나 이상의 전사 인자, 및 그의 기능적 단편의 양을 감소시키거나 기능을 억제한다. 다른 구현예에서, 상기 방법은 Lys106에서 HNF4 α의 아세틸화를 증가시키는 단계, cMET의 발현을 증가시키는 단계, 및/또는 Thr308에서 인산화를 통해 AKT의 활성화를 증가시키는 단계를 포함한다. 본원에 개시된 방법은 간세포의 핵에서 HNF4 α의 양을 놀랍게도 증가시키는 것으로 나타났다. HNF4 α의 이러한 조작은 간 질환 환자의 간세포 기능을 향상시킨다.
- [0026] 본 출원 전반에 걸쳐 사용된 용어는 당업자에게 통상적이고 일반적인 의미로 해석되어야 한다. 그러나 출원인은 다음과 같은 용어에 아래 제공된 특정 정의가 부여되기를 바란다.
- [0027] 용어
- [0028] 명세서 및 청구범위에서 사용된 단수 형태 "a", "an" 및 "the"는 문맥상 다르게 지시하지 않는 한 복수의 인물을 포함한다. 예를 들어, 용어 "a cell"은 이의 혼합물을 포함하는 복수의 cells을 포함한다.
- [0029] 양, 백분율 등과 같은 측정 가능한 값을 언급할 때, 본원에서 사용된 용어 "약"은 측정 가능한 값으로부터 $\pm 20\%$, $\pm 10\%$, $\pm 5\%$, 또는 $\pm 1\%$ 의 변동을 포함하는 것을 의미한다.
- [0030] 대상체에 대한 "투여" 또는 "투여하는"은 대상체에게 제제를 도입하거나 전달하는 모든 경로를 포함한다. 투여는 임의의 적합한 경로에 의해 수행될 수 있으며, 정맥내, 복강내 등을 포함한다. 투여는 자가-투여 및 타인에 의한 투여를 포함한다.
- [0031] 본원에서 사용된 용어 "포함하는(comprising)" 및 그의 변형은 용어 "포함하는(including)" 및 그의 변형과 동의어로 사용되며 개방적, 비제한적인 용어이다. 용어 "포함하는(comprising)" 및 "포함하는(including)"이 다양한 구현예를 설명하기 위해 본원에서 사용되었지만, 용어 "~로 필수적으로 구성되는(consisting essentially

of)" 및 "~로 구성되는(consisting of)"는 "포함하는(comprising)" 및 "포함하는(including)" 대신에 보다 구체적인 구현예를 제공하기 위해 사용될 수 있고 또한 개시되어 있다.

[0032] "조성물"은 유익한 생물학적 효과가 있는 임의의 제제를 의미한다. 유익한 생물학적 효과는 치료적 효과 (예: 장애 또는 기타 바람직하지 않은 생리학적 상태의 치료)와 예방 효과 (예: 장애 또는 기타 바람직하지 않은 생리학적 상태 (예: 간 질환)의 예방)를 모두 포함한다. 상기 용어는 또한 본원에 구체적으로 언급된 유익한 제제의 약학적으로 허용가능한, 약리학적 활성 유도체를 포함하며, 벡터, 폴리뉴클레오타이드, 세포, 염, 에스테르, 아마이드, 촉진제(proagent), 활성 대사물, 이성질체, 단편, 유사체 등을 포함하나 이에 제한되는 것은 아니다. "조성물"이라는 용어가 사용되는 경우, 또는 특정 조성물이 구체적으로 확인되는 경우, 그 용어는 조성물 그 자체뿐만 아니라 약학적으로 허용가능한 약리학적 활성 벡터, 폴리뉴클레오타이드, 염, 에스테르, 아마이드, 촉진제, 접합체, 활성 대사물, 이성질체, 단편, 유사체 등을 포함하는 것으로 이해되어야 한다. 일부 측면에서, 본원에 개시된 조성물은 벡터를 포함하며, 여기서 상기 벡터는 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자, 및 이의 기능적 단편을 암호화하는 핵산을 포함한다. 일부 측면에서, 본원에 개시된 조성물은 DNABJ1/HSP40, ATF6, ATF4, 및 PERK로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자, 및 이의 기능적 단편의 양을 감소시키거나 기능을 억제하는 핵산을 포함한다. 일부 측면에서, 본원에 개시된 조성물은 벡터를 포함하며, 여기서 상기 벡터는 HNF α 를 암호화하는 핵산을 포함한다.

[0033] "유효량"은 의학적 상태 EH는 장애 (예: 간 질환)의 증상 또는 징후를 개선, 전환, 완화, 예방 또는 진단할 수 있는 양을 제한 없이 포함한다. 달리 명시되지 않는 한, 명시적으로 또는 문맥에 따라 "유효량"은 상태를 개선하기에 충분한 최소량으로 제한되지 않는다. 질병 또는 장애의 중증도, 및 질병 또는 장애를 예방, 치료 또는 완화하는 치료 능력은 제한 없이 바이오마커 또는 임상 파라미터에 의해 측정될 수 있다. 용어 "벡터의 유효량" 또는 "조성물의 유효량"은 간 질환의 일부 완화 또는 간 기능의 회복을 야기하기에 충분한 벡터 또는 조성물의 양을 지칭한다.

[0034] "단편"은 다른 서열에 부착되었는지 여부에 관계없이, 특정 영역 또는 특정 아미노산 잔기의 삽입, 결실, 치환 또는 기타 선택된 변형을 포함할 수 있다 (단, 단편의 활성이 비변형된 펩타이드 또는 단백질과 비교하여 크게 변경되거나 손상되지 않는다). 이러한 변형은 이황화 결합이 가능한 아미노산의 제거 또는 추가, 생체 수명(bio-longevity) 증가, 분비 특성 변경 등과 같은 일부 추가 속성을 제공할 수 있다. 어떠한 경우든 단편은 표적 유전자의 전사를 조절하는 것과 같은 생리활성 특성을 가져야 한다.

[0035] 용어 "유전자" 또는 "유전자 서열"은 암호화 서열 또는 제어 서열, 또는 이의 단편을 지칭한다. 유전자는 암호화 서열과 조절 서열, 또는 이의 단편의 임의의 조합을 포함할 수 있다. 따라서, 본원에서 언급되는 "유전자"는 천연 유전자의 전부 또는 일부일 수 있다. 본원에서 언급된 폴리뉴클레오타이드 서열은 용어 "유전자"와 상호교환적으로 사용될 수 있거나, 임의의 암호화 서열, 비-암호화 서열 또는 조절 서열, 이의 단편, 및 이의 조합을 포함할 수 있다. 용어 "유전자" 또는 "유전자 서열"은, 예를 들어 암호화 서열의 상류에 있는 제어 서열(예를 들어, 리보솜 결합 부위)을 포함한다.

[0036] 본원에 사용된 "간 질환"은 일반적으로 간에 영향을 미치는 질환, 장애 및 상태를 나타내며, 예를 들어 간세포에 지방의 단순 축적(지방증), 비알코올성 지방간염(NASH), 비알코올성 지방간 질환(NAFLD), 알코올 간 질환(ALD), 알코올-관련 간 질환(지방간, 알코올성 간염, 알코올-관련 간경변을 포함하나 이에 제한되지 않음), 거대세포성 지방증, 문맥주위 및 소엽 염증(지방간염), 간경변, 섬유증, 간 허혈, 간세포 암종을 포함한 간암, A형 간염, B형 간염, C형 간염, 특발성 간 질환, 말기 간 질환 및 간부전을 포함하는 광범위한 중증도를 가질 수 있다. "간경변"은 간 조직 및/또는 재생 결절의 섬유성 비후를 특징으로 하는 간의 만성 질환으로 본원에서 정의된다. "간경변"의 정도 또는 중증도는 차일드-피 점수(Child-Pugh score)에 의해 지정될 수 있으며, 여기서 5가지 임상 측정값인 총 빌리루빈(bilirubin) 수준, 혈청 알부민, 프로트롬빈 시간 연장, 복수(ascite) 및 간성 뇌병증이 1점, 2점, 및 3점 값의 점수 시스템을 사용하여 각 임상 측정값의 다양한 수준에 대해 채점되며 3점 값은 각 측정의 가장 심각한 수준에 할당된다. 차일드-피 점수 및 분류에 도달하기 위해 5개 측정값 모두에 대한 총점이 추가된다. 5-6점은 차일드-피 분류 A, 7-9점은 차일드-피 분류 B, 10-15점은 차일드-피 분류 C이다. 일반적으로 차일드-피 분류 A는 가장 덜 심각한 간 질환을 나타내며, 차일드-피 분류 C는 가장 심각한 간 질환을 나타낸다. 따라서, 일부 구현예에서, 본원에 개시된 방법은 차일드-피 분류 B 또는 차일드-피 분류 C 간 질환을 갖는 대상체를 치료하는 데 사용될 수 있다. 일부 구현예에서, 본원에 개시된 방법은 차일드-피 분류 A 간 질환을 갖는 대상체를 치료하는 데 사용될 수 있다. 다양한 측면에서, 상기 방법은 대상체의 차일드-피 점수를 개선한다. 일부 구현예에서, 상기 간 질환은 알코올성 간염이다. 일부 구현예에서, 본원에 개시된 방법은 생체

의 관류를 위해 허혈성 공여자 간을 치료하는 데 사용될 수 있다. 본 발명은 간 절제 전 또는 후를 포함하는 암 치료 전 또는 후에 간암을 치료하는데 사용될 수 있다.

- [0037] 본원에 사용된 용어 "핵산"은 뉴클레오타이드로 구성된 중합체 (예: 데옥시리보뉴클레오타이드(DNA) 또는 리보뉴클레오타이드(RNA))를 의미한다. 본원에서 사용된 용어 "리보핵산" 및 "RNA"는 리보뉴클레오타이드로 구성된 중합체를 의미한다. 본원에서 사용된 용어 "데옥시리보핵산" 및 "DNA"는 데옥시리보뉴클레오타이드로 구성된 중합체를 의미한다. 일부 구현예에서, 상기 핵산은 DNA(예: ceDNA 또는 cDNA)이다. 일부 구현예에서, 상기 핵산은 mRNA이다.
- [0038] 용어 "폴리뉴클레오타이드"는 뉴클레오타이드 단량체로 구성된 단일 또는 이중 가닥 중합체를 지칭한다.
- [0039] 용어 "폴리펩타이드"는 D- 또는 L-아미노산의 단일 사슬 또는 펩타이드 결합으로 연결된 D- 및 L-아미노산의 혼합물로 구성된 화합물을 지칭한다.
- [0040] 용어 "프로모터" 또는 "조절 요소"는 전사의 시작으로부터 상류 또는 하류에 위치하고 전사를 개시하기 위한 RNA 폴리머레이스 및 기타 단백질의 인식 및 결합에 관여하는 영역 또는 서열 결정자를 지칭한다. 프로모터는 박테리아 기원일 필요는 없으며, 예를 들어 바이러스 또는 다른 유기체로부터 유래된 프로모터는 본원에 기재된 조성물, 시스템 또는 방법에서 사용될 수 있다.
- [0041] "약학적으로 허용가능한 담체"(때때로 "담체"라고도 함)는 일반적으로 안전하고 무독성인 약학적 또는 치료적 조성물을 제조하는 데 유용한 담체 또는 부형제를 의미하며, 수의학 및/또는 인간의 약학적 또는 치료적 용도에 허용되는 담체를 포함한다. 용어 "담체" 또는 "약학적으로 허용가능한 담체"는 인산염 완충 식염수 용액, 물, 에멀전(예: 오일/물 또는 물/오일 에멀전) 및/또는 다양한 유형의 습윤제를 포함할 수 있으나 이에 제한되는 것은 아니다.
- [0042] 본원에 사용된 용어 "담체"는 임의의 부형제, 희석제, 충전제, 염, 완충제, 안정화제, 가용화제, 지질, 안정화제, 또는 약학적 제형에 사용하기 위해 당업계에 널리 공지된 기타 물질을 포함한다. 조성물에 사용하기 위한 담체의 선택은 조성물의 의도된 투여 경로에 따라 달라질 것이다. 이러한 물질을 함유하는 약학적으로 허용가능한 담체 및 제형의 제조는 예를 들어 Remington's Pharmaceutical Sciences, 21st Edition, ed. University of the Sciences in Philadelphia, Lippincott, Williams & Wilkins, Philadelphia, PA, 2005에 기재되어 있다. 생리학적으로 허용가능한 담체의 예로는 식염수, 글리세롤, DMSO, 완충액 (예: 인산염 완충액, 시트르산염 완충액, 및 기타 유기산을 포함하는 완충액); 아스코르브산을 포함하는 항산화제; 저분자량(약 10개 미만의 잔기) 폴리펩타이드; 혈청 알부민, 젤라틴 또는 면역글로불린과 같은 단백질; 폴리비닐피롤리돈과 같은 친수성 고분자; 글리신, 글루타민, 아스파라긴, 아르기닌 또는 리신과 같은 아미노산; 단당류, 이당류, 및 포도당, 만노오스 또는 텍스트린을 포함하는 기타 탄수화물; EDTA와 같은 킬레이트제; 만니톨 또는 소르비톨과 같은 당 알코올; 나트륨과 같은 염-형성 반대이온; 및/또는 TWEEN™(ICI, Inc.; Bridgewater, New Jersey), 폴리에틸렌 글리콜(PEG) 및 PLURONICS™(BASF; Florham Park, NJ)와 같은 비이온성 계면활성제를 포함한다. 원하는 치료적 치료를 위한 이러한 투여량의 투여를 제공하기 위해, 본원에 개시된 조성물은 유리하게는 담체 또는 희석제를 포함하는 총 조성물의 중량을 기준으로 하나 이상의 대상 화합물의 총 중량의 약 0.1% 내지 99%를 포함할 수 있다.
- [0043] 용어 "대상체"는 영장류(예를 들어, 인간), 소, 양, 염소, 말, 개, 고양이, 토끼, 쥐, 마우스 등을 포함하나 이에 제한되지 않는 포유동물과 같은 동물을 포함하는 것으로 본원에서 정의된다. 일부 구현예에서, 상기 대상체는 인간이다.
- [0044] 본 명세서에서 용어 "전사 인자"는 DNA를 RNA로 전사하는 과정에 관여하는 단백질을 지칭한다. 일반적으로, 전사 인자는 특정 유전자의 프로모터 또는 인핸서 영역에 결합하는 영역(domain)을 보유한다. 전사 인자는 또한 RNA 중합효소 및/또는 일부 다른 전사 인자와 상호작용하는 영역을 가질 수 있으며, 이러한 상호작용은 결과적으로 DNA에서 전사된 RNA의 양을 조절한다. 전사 인자는 세포질에 존재하고 활성화 시 핵으로 이동할 수 있다.
- [0045] 본원에서 사용된 용어 "치료하다", "치료하는", "치료" 및 이들의 문법적 변형은 장애 또는 상태의 하나 이상의 수반되는 증상의 강도를 부분적으로 또는 완전히 지연, 경감, 완화 또는 감소시키는 것 및/또는 장애 또는 상태의 하나 이상의 원인을 경감, 완화 또는 방해하는 것을 포함한다. 본 발명에 따른 치료는 예방적으로, 완화적으로 또는 치료적으로 적용될 수 있다. 치료는 발병 전(예: 간 질환의 명백한 징후 전), 조기 발병 동안(예: 간 질환의 초기 징후 및 증상 시), 간 질환의 확립 후, 또는 말기 간부전의 단계에 대상체에게 투여될 수 있다. 예

방적 투여는 간 질환의 증상이 나타나기 이전 몇 일에서 몇 년 동안 발생할 수 있다.

[0046] 일부 예에서, 용어 "치료하다", "치료하는", "치료" 및 이들의 문법적 변형은 대상체의 치료 전과 비교하거나 또는 일반 또는 연구 집단에서 이러한 증상의 발생률과 비교하여 간 질환 완화, 간 기능 회복 및/또는 대상체의 간세포 핵에서 HNF α 양의 증가를 포함한다.

[0047] 본원에 사용된 "벡터"는 분리된 핵산을 포함하고 분리된 핵산을 세포 내부로 전달하는 데 사용될 수 있는 물질의 조성물이다. 벡터는 자가-복제, 염색체의 벡터 또는 숙주 게놈에 통합되는 벡터일 수 있다. 또는, 벡터는 또한 전술한 핵산 서열을 포함하는 비히클일 수 있다. 벡터는 플라스미드, 박테리오파지, 바이러스 벡터(분리, 약독화, 재조합, 바이러스 입자로 캡슐화 등), 리포솜, 엑소솜, 세포외 소포, 마이크로입자 및/또는 나노입자일 수 있다. 벡터는 이중-가닥 또는 단일-가닥 DNA, RNA, 또는 이중-가닥 및/또는 단일-가닥 뉴클레오타이드를 포함하는 하이브리드 DNA/RNA 서열을 포함할 수 있다. 일부 구현예에서, 상기 벡터는 하나 또는 복수의 폴리펩타이드를 암호화하는 하나 또는 복수의 핵산 서열을 패키징하는 역할을 하는 바이러스 패키징 서열인 핵산 서열을 포함하는 바이러스 벡터이다. 일부 구현예에서, 벡터는 상기 플라스미드이다. 일부 구현예에서 상기 벡터는 엑소솜이다. 일부 구현예에서, 상기 벡터는 바이러스 입자이다. 일부 구현예에서, 상기 바이러스 입자는 렌티바이러스 입자이다. 일부 구현예에서, 상기 벡터는 천연 및/또는 조작된 캡시드를 갖는 바이러스 벡터이다. 일부 구현예에서, 상기 벡터는 조절 서열에 작동 가능하게 연결된 핵산 서열을 포함하는 바이러스 입자를 포함하고, 여기서 상기 핵산 서열은 하나 또는 복수의 AAV 바이러스 입자 폴리펩타이드 또는 그의 단편을 포함하는 융합 단백질 암호화한다. 일부 구현예에서, 상기 벡터는 핵산 또는 폴리펩타이드를 포함하는 나노입자이다. 다양한 구현예에서, 상기 벡터는 지질계 나노입자이다.

[0048] **조성물**

[0049] 상기 언급된 바와 같이, 본원에서는 대상체에서 간세포의 핵으로 전사 인자인 HNF4 α 의 발현 및/또는 수송 및/또는 보유를 증가시켜 간 질환을 치료하는 방법을 개시한다. 일부 구현예에서, 상기 방법은 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자의 발현 또는 기능을 상향 조절하고/하거나 DNAJB1/HSP40, ATF6, ATF4 및 PERK로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자의 발현 또는 기능을 하향 조절하는 단계를 포함한다. 일부 구현예에서, 상기 방법은 선택적으로 내인성 HNF α 의 발현을 증가시키거나 외인성 HNF α 를 도입하여 세포 (예를 들어, 간 세포) 내부, 바람직하게는 세포의 핵 내부에 HNF4 α 의 발현 또는 기능을 상향 조절하는 단계를 더 포함한다. PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300 및 POM121C의 상향 조절 및/또는 DNAJB1/HSP40, ATF6, ATF4 및 PERK의 하향 조절이 간세포의 핵에서 HNF4 α 의 수준을 증가시켜 간 기능의 회복 및 간 질환 완화를 유도한다는 것이 본원에 설명되어 있다. 다른 구현예에서, 상기 방법은 Lys106에서 HNF4 α 의 아세틸화를 증가시키는 단계, cMET의 발현을 증가시키는 단계, 및/또는 Thr308에서 인산화를 통해 AKT의 활성화를 증가시키는 단계를 포함한다. 이들 방법은 전체로서 참조로 도입되는 미국 특허 출원 공개 2014/0249209에 기재된 방법과 함께 사용될 수 있다.

[0050] 일부 구현예에서, 상기 방법은 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군으로부터 선택된 하나 이상의 전사 인자와 함께 HNF4 α 의 발현 또는 기능을 상향 조절하는 단계를 추가로 포함한다. 일부 구현예에서, 상기 조성물은 HNF4 α 및 PROX1의 발현 또는 기능을 상향 조절하는 것을 추가로 포함한다. 일부 구현예에서, 상기 방법은 HNF4 α 및 SREBP1의 발현 또는 기능을 상향 조절하는 단계를 추가로 포함한다. 일부 구현예에서, 상기 방법은 HNF4 α , PROX1 및 SREBP1의 발현 또는 기능을 상향 조절하는 단계를 추가로 포함한다.

[0051] 따라서, PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300 및 POM121C로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자의 발현 또는 기능을 증가시키고/시키거나 DNAJB1/HSP40, ATF6, ATF4 및 PERK로 이루어진 군에서 선택되는 하나 이상의 전사 인자의 발현 또는 기능을 감소시키는 조성물이 본원에 포함된다. 일부 구현예에서, 상기 조성물은 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군으로부터 선택된 하나 이상의 전사 인자와 함께 HNF4 α 의 발현 또는 기능을 상향 조절한다. 일부 구현예에서, 상기 조성물은 HNF4 α 및 PROX1의 발현 또는 기능을 상향 조절한다. 일부 구현예에서, 상기 조성물은 HNF4 α 및 SREBP1의 발현 또는 기능을 상향 조절한다. 일부 구현예에서, 상기 조성물은 HNF4 α , PROX1 및 SREBP1의 발현 또는 기능을 상향 조절한다.

[0052] 벡터를 포함하는 조성물이 본원에 개시되며, 여기서 상기 벡터는 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및/또는 POM121C로 이루어진 군으로부터 선택된 하나 이상의 전사 인자, 및 그의 기능적 단편을 암호화하는 핵산을 포함한다. 일부 구현예에서, 상기 벡터는 PROX1 또는 그의 기능적 단편을 암호화하는 핵산을 포함한다. 다른 구현예에서, 상기 벡터는 NR5A2 또는 그의 기능적 단편을 암호화하는 핵산을 포함한다. 다른 구현예에서, 상기 벡터는 NROB2 또는 그의 기능적 단편을 암호화하는 핵산을 포함한다. 다른 구현예에서, 상기 벡터는 MTF1 또

는 그의 기능적 단편을 암호화하는 핵산을 포함한다. 다른 구현예에서, 상기 벡터는 SREBP1 또는 그의 기능적 단편을 암호화하는 핵산을 포함한다. 다른 구현예에서, 상기 벡터는 EP300 또는 그의 기능적 단편을 암호화하는 핵산을 포함한다. 다른 구현예에서, 상기 벡터는 POM121C 또는 그의 기능적 단편을 암호화하는 핵산을 포함한다. 일부 구현예에서, 상기 벡터는 HNF4 α 를 암호화하는 핵산을 추가로 포함한다. 일부 구현예에서, 상기 벡터는 PROX1 및 SREBP1을 암호화하는 핵산을 포함한다. 일부 구현예에서, 상기 벡터는 HNF4 α , PROX1, 및 SREBP1을 암호화하는 핵산을 포함한다. 일부 구현예에서, 상기 벡터는 HNF4 α 및 PROX1을 암호화하는 핵산을 포함한다. 일부 구현예에서, 상기 벡터는 HNF4 α 및 SREBP1을 암호화하는 핵산을 포함한다.

[0053] HNF4 α 는 간 외에도 신장, 소장, 결장 및 췌장에서도 많이 발현되며, 중요한 역할을 한다. HNF4 α 유전자의 다형성 돌연변이는 성숙기발병당뇨병 (MODY), 크론병 및 염증성 장 증후군을 포함하는 광범위한 질병과 관련이 있다. 대체 스플라이싱과 결합된 P1 또는 P2 프로모터의 전사는 12개의 다른 전사체를 생성할 수 있다. 상대적인 이소형 발현은 조직에 따라 다르다. 12개의 이소형은 각각 전사 활성화 및 억제를 담당하는 N 및 C 말단에서만 상이하다 (Ko et al., Cell Rep. 2019 Mar 5;26(10):2549-2557.e3 참조, 전체로서 본원에 참조로 도입됨). 각 이소형이 조직- 의존적 방식으로 유전자의 특정 하위 집합을 조절하는 별개의 기능을 수행한다는 것이 점점 인식되고 있다. 예를 들어, HNF4 α 이소형 2는 보고된 바에 따르면 간에서 풍부하고 본 출원에 기재된 바와 같이 손실이 간암(hepatocarcinoma) 또는 간부전과 관련된 중앙 억제인자로 작용하는 반면, HNF4 α 이소형 8은 결장에서 고도로 발현되고 성장-촉진 유전자의 발현을 조절한다.

[0054] 따라서, 일부 구현예에서, 본원에 개시된 벡터는 HNF4 α 이소형 2 폴리펩타이드를 암호화하는 핵산을 추가로 포함한다. 일부 구현예에서, HNF4 α 이소형 2 폴리펩타이드는 서열번호 1 또는 그의 단편과 적어도 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 상동성인 서열을 포함한다. 일부 구현예에서, 상기 핵산은 서열번호 31 또는 그의 단편과 적어도 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 상동성이 있다. 일부 측면에서, HNF4 α 이소형 2 폴리펩타이드는 프로모터 1(P1)에 의해 구동되거나, 다시 말해서 이의 발현은 HNF4 α 의 P1 프로모터에 의해 구동된다. 이는 본원에서 HNF4 α 이소형 2(P1)로 지정된다. 따라서, 일부 구현예에서, HNF4 α 이소형 2 폴리뉴클레오타이드 또는 핵산은 P1 프로모터에 작동 가능하게 연결된다.

[0055] 벡터는 발현 가능한 유전자의 복제를 제어하는 조절 핵산 서열을 포함하는 핵산 서열일 수 있다. 일부 구현예에서, 제2 핵산에 작동 가능하게 연결된 프로모터(예를 들어, 전사 인자를 암호화하는 폴리뉴클레오타이드)를 포함하는 벡터는 인간 조작(예: Sambrook et al., *Molecular Cloning—A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., (1989) or *Current Protocols in Molecular Biology Volumes 1-3*, John Wiley & Sons, Inc. (1994-1998)에 기재된 방법에 의한)의 결과로서 제2 핵산(예를 들어, 전사 인자를 암호화하는 폴리뉴클레오타이드)과 이중인 프로모터를 포함할 수 있다. 본원에 기재된 임의의 측면의 벡터는 프로모터, 인핸서, 항생제 내성 유전자 및/또는 원점을 추가로 포함할 수 있으며, 이는 상기 언급된 전사 인자 중 하나 이상에 작동 가능하게 연결될 수 있음을 이해해야 한다.

[0056] 일부 구현예에서, 상기 벡터는 바이러스 벡터일 수 있다. 본원에 개시된 "바이러스 벡터"는 비히클과 관련하여, 바이러스, 바이러스-유사 입자, 비리온, 바이러스 입자. 또는 유사형 바이러스(pseudotyped virus) 내 핵산 서열의 패키징을 지시하는 핵산 서열을 포함하는 임의의 바이러스, 바이러스-유사 입자, 비리온, 바이러스 입자, 또는 유사형 바이러스를 의미한다. 일부 구현예에서, 상기 바이러스, 바이러스-유사 입자, 비리온, 바이러스 입자 또는 유사형 바이러스는 벡터(예: 핵산 벡터)를 숙주 세포 내로 및/또는 숙주 세포 간에 전달할 수 있다. 일부 구현예에서, 상기 바이러스, 바이러스-유사 입자, 비리온, 바이러스 입자, 또는 유사형 바이러스는 벡터(예: 핵산 벡터)를 대상체의 간 내 간세포와 같은 표적 세포 내로 및/또는 표적 세포 간에 전달할 수 있다. 중요하게는, 일부 구현예에서, 상기 바이러스, 바이러스-유사 입자, 비리온, 바이러스 입자 또는 유사형 바이러스는 표적 세포(예를 들어, 간세포)의 핵 내로 수송될 수 있다. 용어 "바이러스 벡터"는 또한 모든 목적을 위해 본원에 참조로 도입된 미국 특허 출원 공개 U.S. 2018/0057839에 보다 완전하게 기술된 형태를 지칭하는 것을 의미한다. 적합한 바이러스 벡터는 예를 들어, 아데노바이러스, 아데노연관바이러스(AAV), 백시니아 바이러스, 헤르페스바이러스, 배칼로바이러스 및 레트로바이러스, 파보바이러스, 및 렌티바이러스를 포함한다. 일부 구현예에서, 상기 바이러스 벡터는 렌티바이러스 벡터 또는 아데노연관바이러스 벡터이다.

[0057] 복제-결함 아데노바이러스(replication-defective adenoviruse)의 구성이 (Berkner et al., J. Virology 61:1213-1220 (1987); Massie et al., Mol. Cell. Biol. 6:2872-2883 (1986); Haj-Ahmad et al., J. Virology 57:267-274 (1986); Davidson et al., J. Virology 61:1226-1239 (1987); Zhang "Generation and identification of recombinant adenovirus by liposome-mediated transfection and PCR analysis" *BioTechniques* 15:868-872 (1993))에 기재되어 있다. 이러한 바이러스를 벡터로 사용하는 이점은 초기 감염된

세포 내에서 복제할 수 있지만 새로운 감염성 바이러스 입자를 형성할 수 없기 때문에 다른 세포 유형으로 퍼질 수 있는 범위가 제한된다는 것이다. 재조합 아데노바이러스는 기도 상피, 간세포, 혈관 내피, CNS 실질(parenchyma) 및 다수의 기타 조직 부위로 직접 생체내 전달된 후 고효율 유전자 전달을 달성하는 것으로 나타났다 (Morsy, J. Clin. Invest. 92:1580-1586 (1993); Kirshenbaum, J. Clin. Invest. 92:381-387 (1993); Roessler, J. Clin. Invest. 92:1085-1092 (1993); Moullier, Nature Genetics 4:154-159 (1993); La Salle, Science 259:988-990 (1993); Gomez-Foix, J. Biol. Chem. 267:25129-25134 (1992); Rich, Human Gene Therapy 4:461-476 (1993); Zabner, Nature Genetics 6:75-83 (1994); Guzman, Circulation Research 73:1201-1207 (1993); Bout, Human Gene Therapy 5:3-10 (1994); Zabner, Cell 75:207-216 (1993); Caillaud, Eur. J. Neuroscience 5:1287-1291 (1993); and Ragot, J. Gen. Virology 74:501-507 (1993)). 재조합 아데노바이러스는 특정 세포 표면 수용체에 결합하여 유전자 도입(transduction)을 달성하고, 그 후 바이러스는 야생형 또는 복제-결합 아데노바이러스와 동일한 방식으로 수용체-매개 내포작용(endocytosis)에 의해 내재화된다 (Chardonnet and Dales, Virology 40:462-477 (1970); Brown and Burlingham, J. Virology 12:386-396 (1973); Svensson and Persson, J. Virology 55:442-449 (1985); Seth, et al., J. Virol. 51:650-655 (1984); Seth, et al., Mol. Cell. Biol. 4:1528-1533 (1984); Varga et al., J. Virology 65:6061-6070 (1991); Wickham et al., Cell 73:309-319 (1993)).

[0058] 다른 유형의 바이러스 벡터는 아데노연관바이러스(AAV)를 기반으로 한다. 이 결손(defective) 과보바이러스는 많은 세포 유형을 감염시킬 수 있으며 인간에게는 비병원성이다. AAV 유형 벡터는 약 4 내지 5kb를 수송할 수 있고 야생형 AAV는 19번 염색체에 안정적으로 삽입되는 것으로 알려져 있다. AAV 역 말단 반복부(ITR) 또는 이의 변형은 감염성 및 부위 특이적 도입을 부여하지만, 세포 독성은 부여하지 않으며, 프로모터는 세포 특이적 발현을 지시한다. 미국 특허 번호 6,261,834는 AAV 벡터와 관련된 물질에 대해 참조로 본원에 도입된다. 생체내에서 간 세포를 형질도입하기 위해 AAV 벡터를 사용하는 방법은 당업계에 공지되어 있다. 전체로서 본원에 참조로 도입된 미국 특허 번호 9,981,048 참조.

[0059] 바이러스 벡터, 특히 아데노바이러스 벡터는 양이온성 지질, 폴리-L-리신(PLL) 및 디에틸아미노에틸텍스트란(DELAE-텍스트란)과 같은 양이온성 양친매성 물질과 복합될 수 있으며, 이는 표적 세포의 바이러스 감염 효율을 증가시킨다 (예: Nov. 20, 1997 출원된 PCT/US97/21496 참조, 본원에 참조로 도입됨). 예를 들어 본원에 도입된 Zhong et al., J. Genet Syndr Gene Therapy 2012 Jan. 10; S1. pii: 008, U.S. Pat. Nos. 5,139,941, 5,252,479 및 5,753,500 및 PCT 공보 WO 97/09441에 개시된 AAV 벡터는, 또한 이러한 벡터가 벡터의 반복 투여에 대한 최소한의 필요로 숙주 염색체에 도입되기 때문에 유용하다. 유전자 요법에서 바이러스 벡터에 대한 리뷰는 McConnell et al., 2004, Hum Gene Ther. 15(11):1022-33; Mccarty et al., 2004, Annu Rev Genet. 38:819-45; Mah et al., 2002, Clin. Pharmacokin. 41(12):901-11; Scott et al., 2002, Neuromuscul. Disord. 12(Suppl 1):S23-9를 참고할 수 있다.

[0060] 일부 구현예에서, 상기 벡터는 나노입자이다. 본원에서 사용되는 나노입자는 핵산의 전달에 유용한 임의의 나노입자일 수 있다. 본원에 사용된 용어 "나노입자"는 생체적합성이 있고 그러한 사용 환경에 의한 화학적 및/또는 물리적 파괴에 충분히 내성이 있어 충분한 수의 나노입자가 적용 또는 치료 부위에 전달된 후 실질적으로 손상되지 않은 상태로 유지되고, 크기가 나노미터 범위에 있는 입자 또는 구조를 지칭한다. 일부 구현예에서, 상기 나노입자는 지질-유사 나노입자를 포함한다. 예를 들어, 본원에 참조로 도입된 WO WO/2016/187531A1, WO/2017/176974, WO/2019/027999, 또는 Li, B et al., An Orthogonal array optimization of lipid-like nanoparticles for mRNA delivery in vivo. Nano Lett. 2015, 15, 8099-8107를 참조할 수 있다. 일부 구현예에서, 상기 나노입자는 지질 이중층 또는 리포솜을 포함할 수 있다. 일부 구현예에서, 상기 벡터는 mRNA 지질 나노입자이다.

[0061] 일부 구현예에서, 개시된 나노입자는 생물학적 개체, 예를 들어 표적 세포 상의 특정 막 성분 또는 세포 표면 수용체(예: 간 세포로의 전달을 촉진하는 수용체 또는 간 세포 상의 수용체)에 효율적으로 결합하거나 회합될 수 있다. 예를 들어, 개시된 나노입자는 정상 또는 병든 간 세포 상에서 발현되는 수용체 (예: 간 아시알로당단백질 수용체(ASGPR) 또는 저밀도 지단백질(LDLR) 수용체)에 결합하는 리간드를 포함하도록 조작성될 수 있다.

[0062] 일부 측면에서, 본원에 개시된 나노입자는 간 세포 내로 핵산의 전달을 촉진하는 보충 성분을 포함할 수 있다. 상기 나노입자는 양이온성 지질, 보조 지질, 콜레스테롤, 및 폴리에틸렌 글리콜 (PEG)을 포함할 수 있다. 일부 구현예에서, 상기 나노입자는 5A2-SC8, 1,2-디올레오일-sn-글리세로-3-포스포에탄올아민 (DOPE), 콜레스테롤, 및/또는 1,2-디미리스토일-rac-글리세롤-메톡시(폴리(에틸렌 글리콜)), 또는 이들의 임의의 조합을 포함한다. 일부 구현예에서, 상기 나노입자는 5A2-SC8, 1,2-디올레오일-sn-글리세로-3-포스포에탄올아민 (DOPE), 콜레스테

를, 및 1,2-디미리스토일-rac-글리세롤-메톡시(폴리(에틸렌 글리콜))을 더 포함한다. 일부 구현예에서, 상기 나노입자는 1,2-디올레오일-3-트리메틸암모늄-프로판 (DOTAP)을 더 포함한다. 일부 구현예에서, 상기 나노입자 5A2-SC8, 1,2-디올레오일-sn-글리세로-3-포스포에탄올아민 (DOPE), 콜레스테롤, 및 1,2-디미리스토일-rac-글리세롤-메톡시(폴리(에틸렌 글리콜))의 몰비(molar ratio)는 약 15/15/30/3이다.

[0063] 일부 구현예에서, 상기 나노입자는 DLin-MC3-DMA, 1,2-디스테아로일-sn-글리세로-3-포스포콜린 (DSPC), 콜레스테롤, 및 1,2-디미리스토일-rac-글리세롤-메톡시(폴리(에틸렌 글리콜))을 포함한다. 일부 구현예에서, DLin-MC3-DMA, 1,2-디스테아로일-sn-글리세로-3-포스포콜린 (DSPC), 콜레스테롤, 및 1,2-디미리스토일-rac-글리세롤-메톡시(폴리(에틸렌 글리콜))의 몰비는 약 50/10/38.5/1.5이다.

[0064] 일부 구현예에서, 상기 나노입자는 C12-200, 1,2-디올레오일-sn-글리세로-3-포스포에탄올아민 (DOPE), 콜레스테롤, 및 1,2-디미리스토일-rac-글리세롤-메톡시(폴리(에틸렌 글리콜))을 포함한다. 일부 구현예에서, C12-200, 1,2-디올레오일-sn-글리세로-3-포스포에탄올아민 (DOPE), 콜레스테롤, 및 1,2-디미리스토일-rac-글리세롤-메톡시(폴리(에틸렌 글리콜))의 몰비는 약 35/16/46.5/2.5이다.

[0065] 일부 구현예에서, 본원에 개시된 상기 나노입자는 5A2-SC8, 1,2-디올레오일-sn-글리세로-3-포스포에탄올아민 (DOPE), 콜레스테롤, 1,2-디미리스토일-rac-글리세롤-메톡시(폴리(에틸렌 글리콜)), 및 1,2-디올레오일-3-트리메틸암모늄-프로판 (DOTAP)을 포함한다. 상기 나노입자는 DOTAP을 약 0.1% 내지 약 30% 몰/몰(mol/mol)을 포함할 수 있다. 예를 들어, 나노입자에 존재하는 DOTAP의 양은 나노입자의 약 0.1%, 약 0.2%, 약 0.3%, 약 0.4%, 약 0.5%, 약 0.6% 몰/몰, 약 0.7% 몰/몰, 약 0.8% 몰/몰, 약 0.9% 몰/몰, 약 1% 몰/몰, 약 2% 몰/몰, 약 2.5% 몰/몰, 약 3% 몰/몰, 약 3.5% 몰/몰, 약 4% 몰/몰, 약 4.5% 몰/몰, 약 5% 몰/몰, 약 5.5% 몰/몰, 약 6% 몰/몰, 약 6.5% 몰/몰, 약 7% 몰/몰, 약 7.5% 몰/몰, 약 8% 몰/몰, 약 8.5% 몰/몰, 약 9% 몰/몰, 약 9.5% 몰/몰, 약 10% 몰/몰, 약 10.5% 몰/몰, 약 11% 몰/몰, 약 11.5% 몰/몰, 약 12% 몰/몰, 약 12.5% 몰/몰, 약 13% 몰/몰, 약 13.5% 몰/몰, 약 14% 몰/몰, 약 15% 몰/몰, 약 16% 몰/몰, 약 17% 몰/몰, 약 18% 몰/몰, 약 19% 몰/몰, 약 20% 몰/몰, 약 22% 몰/몰, 약 24% 몰/몰, 약 26% 몰/몰, 약 28% 몰/몰, 약 30% 몰/몰 일 수 있다. 일부 구현예에서, 나노입자에 존재하는 DOTAP의 양은 나노입자의 약 20% 몰/몰 이다.

[0066] 일부 구현예에서, 본원에 개시된 나노입자 및 간-특이적 전달 방법은 당업계, 예를 들어 Cheng et al., Nat Nanotechnol. 2020 Apr;15(4):313-320. Epub 2020 Apr 6; Trepotec et al., Mol Ther. 2019 Apr 10;27(4):794-802. Epub 2018 Dec 22; Truong, et al., Proc Natl Acad Sci USA. 2019 Oct 15;116(42):21150-21159. Epub 2019 Sep 9;에 기재되어 있으며, 전체로서 본원에 참조로 도입된다.

[0067] 추가 구현예에서, 본원에 개시된 벡터는 폴리(아미도-아민), 폴리-베타 아미노-에스터 (PBAEs), 및/또는 폴리에틸렌이민 (PEI)을 포함한다. 일부 구현예에서, 상기 벡터는 폴리아크리딘 PEG를 포함한다. 일부 구현예에서, 본원에 개시된 벡터는 외부 PEG 셸 및 나노입자-기반 코어를 포함한다.

[0068] 지질-기반 나노 입자는 간으로 치료용 페이로드를 성공적으로 전달한다. 예를 들어, Witzigmann et al., Adv Drug Deliv Rev. 2020 Jul, doi: 10.1016/j.addr.2020.06.026 참고. 리포솜은 여러 유형의 지질로부터 만들어질 수 있지만, 인지질은 약물 운반체로 지질-기반 나노 입자를 생성하는 데 가장 일반적으로 사용된다. 본 발명에서 사용하기 위한 지질 입자는 리포솜-형성 지질 및 인지질, 및 막 활성 스테롤(예: 콜레스테롤)을 포함하도록 제조될 수 있다. 리포솜은 리포솜 형성 지질이 아닌 다른 지질 및 인지질을 포함할 수 있다.

[0069] 인지질은 예를 들어 레시틴(예: 계란 또는 대두 레시틴); 포스파티딜콜린(예: 계란 포스파티딜콜린); 수소화된 포스파티딜콜린; 리소포스파티딜 콜린; 디팔미토일포스파티딜콜린; 디스테아로일 포스파티딜콜린; 디미리스토일 포스파티딜콜린; 디라우로일포스파티딜콜린; 글리세로인지질(예: 포스파티딜글리세롤, 포스파티딜세린, 포스파티딜에탄올아민, 리소포스파티딜에탄올아민, 포스파티딜이노시톨, 포스파티딜이노시톨 포스페이트, 포스파티딜이노시톨 비스포스페이트 및 포스파티딜이노시톨 트리스포스페이트); 스펅고미엘린; 카디오리핀; 포스파티드산; 플라즈마로겐; 또는 이들의 혼합물로부터 선택될 수 있다. 각 경우는 본 발명의 각각의 구현예를 나타낸다. 사용될 수 있는 다른 지질의 예는 당지질(예를 들어, 글리세로당지질(예: 갈락토리피드 및 설포리피드), 글리코스핑고리피드(예: 세레브로사이드, 글루코세레브로사이드 및 갈락토세레브로사이드), 및 글리코실포스파티딜이노시톨); 포스포스핑고리피드(예: 세라마이드 포스포릴콜린, 세라마이드 포스포릴에탄올아민, 및 세라마이드 포스포릴글리세롤 등); 또는 이들의 혼합물을 포함한다. 각 경우는 본 발명의 각각의 구현예를 나타낸다. 음전하 또는 양전하를 띤 지질 나노입자는 예를 들어 음이온성 또는 양이온성 인지질 또는 지질을 사용하여 얻을 수 있다. 이러한 음이온성/양이온성 인지질 또는 지질은 일반적으로 스테롤, 아실 또는 디아실 사슬과 같은 친유성 모이어티를 가지며, 여기서 지질은 전체 알짜 음전하/양전하를 갖는다.

[0070] 일부 구현예에서, 본원에 개시된 나노입자는 1개, 2개, 3개 또는 그 이상의 생체적합성 및/또는 생분해성 중합체를 포함한다. 예를 들어, 고려되는 나노입자는 생분해성 중합체 및 폴리에틸렌 글리콜을 포함하는 하나 이상의 블록 공중합체를 약 10 내지 약 99 중량%로 포함할 수 있으며, 생분해성 단독중합체를 약 0 내지 약 50 중량%로 포함할 수 있다. 중합체는 예를 들어 미세결정질 셀룰로오스, 히드록시프로필 셀룰로오스, 히드록시프로필 메틸셀룰로오스, 폴리알킬렌 옥사이드, 예를 들어 폴리에틸렌 옥사이드(PEG), 폴리무수물, 폴리(에스테르 무수물), 폴리하이드록시산, 예를 들어 폴리락타이드(PLA), 폴리글리콜라이드(PGA), 폴리(락타이드-코-글리콜라이드)(PLGA), 폴리-3-하이드록시부티레이트(PHB) 및 이들의 공중합체, 폴리-4-하이드록시부티레이트(P4HB) 및 이들의 공중합체, 폴리카프로락톤 및 이들의 공중합체, 및 이들의 조합과 같은 생안정성 및 생분해성 중합체 모두를 포함할 수 있다.

[0071] 일부 구현예에서, 상기 나노입자는 약 1 nm 내지 약 1000 nm의 직경을 갖는다. 일부 구현예에서, 상기 나노입자는 예를 들어 약 1000 nm, 약 950 nm, 약 900 nm, 약 850 nm, 약 800 nm, 약 750 nm, 약 700 nm, 약 650 nm, 약 600 nm, 약 550 nm, 약 500 nm, 약 450 nm, 약 400 nm, 약 350 nm, 약 300 nm, 약 290 nm, 약 280 nm, 약 270 nm, 약 260 nm, 약 250 nm, 약 240 nm, 약 230 nm, 약 220 nm, 약 210 nm, 약 200 nm, 약 190 nm, 약 180 nm, 약 170 nm, 약 160 nm, 약 150 nm, 약 140 nm, 약 130 nm, 약 120 nm, 약 110 nm, 약 100 nm, 약 90 nm, 약 80 nm, 약 70 nm, 약 60 nm, 약 50 nm, 약 40 nm, 약 30 nm, 약 20 nm, 또는 약 10 nm 미만의 직경을 갖는다. 일부 구현예에서, 상기 나노입자는 예를 들어 약 20 nm 내지 약 1000 nm, 약 20 nm 내지 약 800 nm, 약 20 nm 내지 약 700 nm, 약 30 nm 내지 약 600 nm, 약 30 nm 내지 약 500 nm, 약 40 nm 내지 약 400 nm, 약 40 nm 내지 약 300 nm, 약 40 nm 내지 약 250 nm, 약 50 nm 내지 약 250 nm, 약 50 nm 내지 약 200 nm, 약 50 nm 내지 약 150 nm, 약 60 nm 내지 약 150 nm, 약 70 nm 내지 약 150 nm, 약 80 nm 내지 약 150 nm, 약 90 nm 내지 약 150 nm, 약 100 nm 내지 약 150 nm, 약 110 nm 내지 약 150 nm, 약 120 nm 내지 약 150 nm, 약 90 nm 내지 약 140 nm, 약 90 nm 내지 약 130 nm, 약 90 nm 내지 약 120 nm, from 100 nm 내지 약 140 nm, 약 100 nm 내지 약 130 nm, 약 100 nm 내지 약 120 nm, 약 100 nm 내지 약 110 nm, 약 110 nm 내지 약 120 nm, 약 110 nm 내지 약 130 nm, 약 110 nm 내지 약 140 nm, 약 90 nm 내지 약 200 nm, 약 100 nm 내지 약 195 nm, 약 110 nm 내지 약 190 nm, 약 120 nm 내지 약 185 nm, 약 130 nm 내지 약 180 nm, 약 140 nm 내지 약 175 nm, from 150 nm 내지 175nm, 또는 약 150 nm 내지 약 170 nm의 직경을 갖는다. 일부 구현예에서, 상기 나노입자는 약 100 nm 내지 약 250 nm의 직경을 갖는다. 일부 구현예에서, 상기 나노입자는 약 150 nm 내지 약 175 nm의 직경을 갖는다. 일부 구현예에서, 상기 나노입자는 약 135 nm 내지 약 175 nm의 직경을 갖는다. 상기 입자는 임의의 모양을 가질 수 있으나, 일반적으로 구형이다.

[0072] 일부 구현예에서, 본원에 사용된 벡터는 엑소좀이다. 본원에 사용된 용어 "미세소포(microvesicle)" 및 "엑소좀"은 약 10 nm 내지 약 5000 nm, 보다 일반적으로 30 nm 내지 1000 nm의, 및 가장 일반적으로 약 50 nm 내지 750 nm의 직경을 갖는 막형 입자를 지칭하고 (또는 입자가 회전 타원체(spheroid)가 아닌 경우 최대 치수), 여기서 엑소좀 막의 적어도 일부는 세포로부터 직접 수득된다. 가장 일반적으로 엑소좀은 도너 세포 크기의 최대 5%인 크기(평균 직경)를 갖는다. 따라서, 특히 고려되는 엑소좀은 세포로부터 배출된 엑소좀을 포함한다. 엑소좀의 제조 방법은 당업계에 공지되어 있다. 예를 들어, 전체로서 본원에 참조로 도입된 미국 공개 번호 2018/0177727을 참조할 수 있다. 폴리뉴클레오타이드 및 폴리펩타이드를 전달하기 위한 엑소좀 및 이의 용도는 당업계에 공지되어 있다. 본원에 참조로 도입된 미국 특허 번호 10,577,630을 참조할 수 있다.

[0073] 또한 RNA 활성화(RNAa)를 통해 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군으로부터 선택된 하나 이상의 전사 인자의 발현 또는 기능을 증가시키는 조성물이 본원에 포함된다. 따라서, 일부 구현예에서, 상기 조성물은 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C 중 하나 이상을 활성화하는 짧은 헤어핀 RNA(shRNA)를 포함한다.

[0074] 본원에서 "HNF4 a"는 인간에서 HNF4A 유전자에 의해 암호화되는 폴리펩타이드를 지칭한다. 일부 구현예에서, HNF4 a 폴리펩타이드는 HGNC: 5024, Entrez Gene: 3172, Ensembl: ENSG00000101076, OMIM: 600281, UniProtKB: P41235와 같이 하나 이상의 공개적으로 이용가능한 데이터베이스에서 확인된 것이다. 일부 구현예에서, HNF4 a 폴리펩타이드는 서열번호 1의 서열(HNF4 a 이소형 2), 또는 서열번호 1과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리펩타이드 서열, 또는 서열번호 1의 부분을 포함하는 폴리펩타이드를 포함한다. 서열번호 1의 HNF4 a 폴리펩타이드는 성숙한 HNF4 a의 미성숙 또는 전처리된 형태를 나타낼 수 있고, 이에 따라 본원에 포함되는 것은 서열번호 1의 HNF4 a 폴리펩타이드의 성숙 또는 가공된 부분이다. 일부 구현예에서, 상기 HNF4 a 폴리펩타이드는 모든 목적을 위해 본원에 참고로 포함된 미국 특허 출원 공개 US 2014/0249209에 기재된 것이다. 일부 구현예에서, 상기 HNF4 a 폴리뉴클레오타이드는 서열번호 31, 서열번호

32, 서열번호 33, 또는 서열번호 34의 서열, 또는 서열번호 31, 서열번호 32, 서열번호 33, 또는 서열번호 34의 서열과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리뉴클레오타이드 서열, 또는 서열번호 31, 서열번호 32, 서열번호 33, 또는 서열번호 34의 부분을 포함하는 폴리뉴클레오타이드를 포함한다.

[0075] PROX1(Prospero-related homeobox 1)은 보통 간 기관 형성 동안 생쥐의 내배엽 세포에서 배아 8.5일(E8.5)에 처음 발현된다. 성인 간에서, PROX1의 역할은 간세포의 에너지 대사를 제어할 수 있다. 더 중요하게는 Prox1이 특정 DNA 요소에 대한 호메오도메인의 직접 결합에 의해 유전자 전사의 활성제로 기능할 수 있다고 보고되었다. 본원에서 "PROX1"은 인간에서 PROX1 유전자에 의해 암호되는 폴리펩타이드를 지칭한다. 일부 구현예에서, PROX1 폴리펩타이드는 HGNC: 9459, Entrez Gene: 5629, Ensembl: ENSG00000117707, OMIM: 601546, UniProtKB: Q92786와 같이 하나 이상의 공개적으로 이용가능한 데이터베이스에서 확인된 것이다. 일부 구현예에서, PROX1 폴리펩타이드는 서열번호 2의 서열, 또는 서열번호 2와 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리펩타이드 서열, 또는 서열번호 2의 부분을 포함하는 폴리펩타이드를 포함한다. 서열번호 2의 PROX1 폴리펩타이드는 성숙한 PROX1의 미성숙 또는 전처리된 형태를 나타낼 수 있고, 이에 따라 본원에 포함되는 것은 서열번호 2의 PROX1 폴리펩타이드의 성숙 또는 가공된 부분이다. 일부 구현예에서, 상기 PROX1 폴리뉴클레오타이드는 서열번호 13의 서열, 또는 서열번호 13의 서열과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리뉴클레오타이드 서열, 또는 서열번호 13의 부분을 포함하는 폴리뉴클레오타이드를 포함한다.

[0076] NR5A2(Nuclear receptor 5A2; Liver receptor homologue-1; LRH-1)는 표적 유전자의 프로모터 및 조절 영역 내의 특정 반응 요소에 단량체로 결합하는 핵 수용체이다. NR5A2는 또한 담즙산 생산 효소, 지방산 대사 및 미토콘드리아 기능을 암호화하는 유전자를 양성 조절할 수 있다. 본원에서 "NR5A2"는 인간에서 NR5A2 유전자에 의해 암호화되는 폴리펩타이드를 지칭한다. 일부 구현예에서, NR5A2 폴리펩타이드는 HGNC: 7984, Entrez Gene: 2494, Ensembl: ENSG00000116833, OMIM: 604453, UniProtKB: 000482와 같은 하나 이상의 공개적으로 이용가능한 데이터베이스에서 확인된 것이다. 일부 구현예에서, NR5A2 폴리펩타이드는 서열번호 3의 서열, 또는 서열번호 3과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리펩타이드 서열, 또는 서열번호 3의 부분을 포함하는 폴리펩타이드를 포함한다. 서열번호 3의 NR5A2 폴리펩타이드는 성숙한 NR5A2의 미성숙 또는 전처리된 형태를 나타낼 수 있고, 이에 따라 본원에 포함되는 것은 서열번호 3의 NR5A2 폴리펩타이드의 성숙 또는 가공된 부분이다. 일부 구현예에서, 상기 NR5A2 폴리뉴클레오타이드는 서열번호 14, 서열번호 15, 또는 서열번호 16의 서열, 또는 서열번호 14, 서열번호 15, 또는 서열번호 16의 서열과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리뉴클레오타이드 서열, 또는 서열번호 14, 서열번호 15, 또는 서열번호 16의 부분을 포함하는 폴리뉴클레오타이드를 포함한다.

[0077] NROB2(Nuclear receptor small heterodimer partner; SHP)는 보통 정상 간세포에서 높게 발현되며 담즙산, 포도당 및 지질 대사에 대한 중요한 전사 조절자 역할을 한다. NROB2의 수모일화(SUMOylation)는 핵 수송 및 담즙산 항상성을 유지하고 간 독성으로부터 보호하는 데 중요한 담즙산 생합성의 피드백 억제에서 SHP의 유전자 억제 기능에 필요할 수 있다. (Kim DH et al., 2016). 본원에서 "NROB2"는 인간에서 NROB2 유전자에 의해 암호화되는 폴리펩타이드를 지칭한다. 일부 구현예에서, NROB2 폴리펩타이드는 HGNC: 7961, Entrez Gene: 8431, Ensembl: ENSG00000131910, OMIM: 604630, UniProtKB: Q15466과 같이 하나 이상의 공개적으로 이용가능한 데이터베이스에서 확인된 것이다. 일부 구현예에서, NROB2 폴리펩타이드는 서열번호 4의 서열, 또는 서열번호 4와 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리펩타이드 서열, 또는 서열번호 4의 부분을 포함하는 폴리펩타이드를 포함한다. 서열번호 4의 NROB2 폴리펩타이드는 성숙한 NROB2의 미성숙 또는 전처리된 형태를 나타낼 수 있고, 이에 따라 본원에 포함되는 것은 서열번호 4의 NROB2 폴리펩타이드의 성숙 또는 가공된 부분이다. 일부 구현예에서, 상기 NROB2 폴리뉴클레오타이드는 서열번호 17의 서열, 또는 서열번호 17의 서열과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리뉴클레오타이드 서열, 또는 서열번호 17의 부분을 포함하는 폴리뉴클레오타이드를 포함한다.

[0078] MTF1(Metal-responsive transcription factor 1)은 메탈로티오네인(metallothionein) 유전자의 기초 및 중금속 유도 전사를 매개할 수 있으며 또한 세포 스트레스 반응 및 금속 항상성에 관련된 다른 유전자를 조절할 수 있다. MTF1은 또한 아연 수송체 1과 같은 다른 금속 반응성 유전자의 전사 조절에 관여할 수 있다. MTF1은 간세포의 아연 수준을 조절할 수 있다. 본원에서 "MTF1"은 신호전달 림프구 활성화 분자 패밀리의 자가-리간드 수용체인 폴리펩타이드를 지칭하며, 인간에서 MTF1 유전자에 의해 암호화된다. 일부 구현예에서, MTF1 폴리펩타이드는 HGNC: 7428, Entrez Gene: 4520, Ensembl: ENSG00000188786, OMIM: 600172, UniProtKB: Q14872와 같이 하나 이상의 공개적으로 이용가능한 데이터베이스에서 확인된 것이다. 일부 구현예에서, MTF1 폴리펩타이드는 서열번호

호 5의 서열, 또는 서열번호 5와 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리펩타이드 서열, 또는 서열번호 5의 부분을 포함하는 폴리펩타이드를 포함한다. 서열번호 5의 MTF1 폴리펩타이드는 성숙한 MTF1의 미성숙 또는 전처리된 형태를 나타낼 수 있고, 이에 따라 본원에 포함되는 것은 서열번호 5의 MTF1 폴리펩타이드의 성숙 또는 가공된 부분이다. 일부 구현예에서, 상기 MTF1 폴리뉴클레오타이드는 서열번호 18의 서열, 또는 서열번호 18의 서열과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리뉴클레오타이드 서열, 또는 서열번호 18의 부분을 포함하는 폴리뉴클레오타이드를 포함한다.

[0079] SREBP1(Sterol regulator element binding proteins 1)은 콜레스테롤, 지방산 및 트리글리세리드의 생합성에 관여하는 전사 인자이다. SREBP1은 AKT/PI3K 신호전달 경로의 발현과 활성을 제어할 수 있으며 그 반대의 경우도 마찬가지이다 (Shi Q et al., 2016; Porstmann T et al., 2008). 본원에서 "SREBF1"은 신호전달 림프구 활성화 분자 패밀리의 자가-리간드 수용체인 폴리펩타이드를 지칭하며, 인간에서 SREBF1 유전자에 의해 암호화된다. 일부 구현예에서, SREBF1 폴리펩타이드는 HGNC: 11289, Entrez Gene: 6720, Ensembl: ENSG00000072310, OMIM: 184756, UniProtKB: P36956과 같이 하나 이상의 공개적으로 이용가능한 데이터베이스에서 확인된 것이다. 일부 구현예에서, SREBF1 폴리펩타이드는 서열번호 6의 서열, 또는 서열번호 6과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리펩타이드 서열, 또는 서열번호 6의 부분을 포함하는 폴리펩타이드를 포함한다. 서열번호 6의 SREBF1 폴리펩타이드는 성숙한 SREBF1의 미성숙 또는 전처리된 형태를 나타낼 수 있고, 이에 따라 본원에 포함되는 것은 서열번호 6의 SREBF1 폴리펩타이드의 성숙 또는 가공된 부분이다. 일부 구현예에서, 상기 SREBP1 폴리뉴클레오타이드는 서열번호 19, 서열번호 20, 또는 서열번호 21의 서열, 또는 서열번호 19, 서열번호 20, 또는 서열번호 21의 서열과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리뉴클레오타이드 서열, 또는 서열번호 19, 서열번호 20, 또는 서열번호 21의 부분을 포함하는 폴리뉴클레오타이드를 포함한다.

[0080] EP300(The histone acetyltransferase p300) 은 C/EBP 단백질과 복합체를 형성할 수 있고, 간 지방증 발달, 포도당 대사 및 간에서 높게 발현되는 Foxo1 및 파네소이드 X 수용체(farnesoid X receptor, FXR)와 같은 여러 전사 인자의 조절 동안 트리글리세리드 합성에 관여하는 유전자의 프로모터를 활성화할 수 있다. 본원에서 "EP300"은 신호전달 림프구 활성화 분자 패밀리의 자가-리간드 수용체인 폴리펩타이드를 나타내며, 인간에서 EP300 유전자에 의해 암호화된다. 일부 구현예에서, EP300 폴리펩타이드는 HGNC: 3373, Entrez Gene: 2033, Ensembl: ENSG00000100393, OMIM: 602700, UniProtKB: Q09472와 같이 하나 이상의 공개적으로 이용가능한 데이터베이스에서 확인된 것이다. 일부 구현예에서, EP300 폴리펩타이드는 서열번호 7의 서열, 또는 서열번호 7과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리펩타이드 서열, 또는 서열번호 7의 부분을 포함하는 폴리펩타이드를 포함한다. 서열번호 7의 EP300 폴리펩타이드는 성숙한 EP300의 미성숙 또는 전처리된 형태를 나타낼 수 있고, 이에 따라 본원에 포함되는 것은 서열번호 7의 EP300 폴리펩타이드의 성숙 또는 가공된 부분이다. 일부 구현예에서, 상기 EP300 폴리뉴클레오타이드는 서열번호 22 또는 서열번호 23의 서열, 또는 서열번호 22 또는 서열번호 23의 서열과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리뉴클레오타이드 서열, 또는 서열번호 22 또는 서열번호 23의 부분을 포함하는 폴리뉴클레오타이드를 포함한다.

[0081] POM121C(Nuclear envelope pore membrane protein POM 121)는 핵공 생합성에 관여하는 것으로 여겨지는 공극막 단백질이라고 하는 단백질 그룹의 구성원인 막 단백질이다. 본원에서 "POM121C"는 신호전달 림프구 활성화 분자 패밀리의 자가-리간드 수용체인 폴리펩타이드를 나타내며, 인간에서 POM121C 유전자에 의해 암호화된다. 일부 구현예에서, POM121C 폴리펩타이드는 HGNC: 34005, Entrez Gene: 100101267, Ensembl: ENSG00000272391, OMIM: 615754, UniProtKB: A8CG34와 같이 하나 이상의 공개적으로 이용가능한 데이터베이스에서 확인된 것이다. 일부 구현예에서, POM121C 폴리펩타이드는 서열번호 8의 서열, 또는 서열번호 8과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리펩타이드 서열, 또는 서열번호 8의 부분을 포함하는 폴리펩타이드를 포함한다. 서열번호 8의 POM121C 폴리펩타이드는 성숙한 POM121C의 미성숙 또는 전처리된 형태를 나타낼 수 있고, 이에 따라 본원에 포함되는 것은 서열번호 8의 성숙 또는 가공된 부분이다. 일부 구현예에서, 상기 POM121C 폴리뉴클레오타이드는 서열번호 24의 서열, 또는 서열번호 24의 서열과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리뉴클레오타이드 서열, 또는 서열번호 24의 부분을 포함하는 폴리뉴클레오타이드를 포함한다.

[0082] 또한 DNAJB1/HSP40, ATF6, ATF4, 및 PERK로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자의 양을 감소시키거나 기능을 억제하는 조성물이 본원에 개시된다. 따라서, DNAJB1/HSP40, ATF6, ATF4, 및 PERK 폴리뉴클레오타이드에 상관되고/되거나 이에 작용하는 작은 간섭 RNA(siRNA) 및 마이크로RNA(miRNA)와 같은 작은 활성화

RNA(saRNA), 또는 crisgrRNA 또는 tracr/mate RNA와 같은 CRISPR RNA를 포함하는 조성물이 본원에 포함된다. 단백질의 기능을 감소 또는 억제하기 위해 saRNA를 사용하는 방법은 당업계에 공지되어 있다. 예를 들어, 전체로서 본원에 참조로 도입된 국제 공개 번호 WO2019/048632를 참고할 수 있다. 따라서, DNAJB1/HSP40, ATF6, ATF4 및 PERK로 이루어진 군으로부터 선택된 하나 이상의 전사 인자의 양을 감소시키거나 기능을 억제하기 위해 saRNA를 사용하여 HNF4 α의 발현을 증가시키는 방법이 본원에 포함된다. 일부 구현예에서, 상기 조성물은 DNAJB1/HSP40의 양을 감소시키거나 기능을 억제하는 핵산을 포함한다. 일부 구현예에서, 상기 조성물은 ATF6의 양을 감소시키거나 기능을 억제하는 핵산을 포함한다. 일부 구현예에서, 상기 조성물은 ATF4의 양을 감소시키거나 기능을 억제하는 핵산을 포함한다. 일부 구현예에서, 상기 조성물은 PERK의 양을 감소시키거나 기능을 억제하는 핵산을 포함한다. 일부 구현예에서, 상기 조성물은 HNF4 α를 암호화하는 핵산을 추가로 포함한다.

[0083] DNAJB1/HSP40(Heat shock protein 40)은 유전자 발현 및 번역 개시, 접힘 및 비-접힘, 및 단백질의 전위 및 분해에 필수적인 역할을 할 수 있는 분자 샤페론 단백질이다. DNAJ/HSP40의 활성은 여러 번역 후 변형에 의해 조절된다. 많은 경우에, DNAJ/HSP40은 인단백질(예: DnaJA1, DnaJB4, DnaJC1, DnaJC29)이며, 그 발현과 기능은 각각 아세틸화(예: DnaJA1, DnaJB2, DnaJB12, DnaJC5, DnaJC8, DnaJC13), 당화(DnaJB11, DnaJC10, DnaJC16), 팔미토일화(DnaJC5, DnaJC5B, DnaJC5G), 메틸화(DnaJA1-4), 프레닐화(DnaJA1, DnaJA2, DnaJA4), 및 분자내 이황화 결합(DnaJB11, DnaJC3, DnaJC10)에 의해 번역 동시 및 번역 후 추가로 조절될 수 있다. 본원에서 "DNAJB1/HSP40"은 신호전달 림프구 활성화 분자 패밀리의 자가-리간드 수용체인 폴리펩타이드를 지칭하며, 인간에서 DNAJB1 유전자에 의해 암호화된다. 일부 구현예에서, DNAJB1 폴리펩타이드는 HGNC: 5270, Entrez Gene: 3337, Ensembl: ENSG00000132002, OMIM: 604572, UniProtKB: P25685와 같이 하나 이상의 공개적으로 이용가능한 데이터베이스에서 확인된 것이다. 일부 구현예에서, DNAJB1 폴리펩타이드는 서열번호 9의 서열, 또는 서열번호 9와 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리펩타이드 서열, 또는 서열번호 9의 부분을 포함하는 폴리펩타이드를 포함한다. 서열번호 9의 DNAJB1 폴리펩타이드는 성숙한 DNAJB1의 미성숙 또는 전처리된 형태를 나타낼 수 있고, 이에 따라 본원에 포함되는 것은 서열번호 9의 DNAJB1 폴리펩타이드의 성숙 또는 가공된 부분이다. 일부 구현예에서, 상기 DNAJB1 폴리뉴클레오타이드는 서열번호 25 또는 서열번호 26의 서열, 또는 서열번호 25 또는 서열번호 26의 서열과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리뉴클레오타이드 서열, 또는 서열번호 25 또는 서열번호 26의 부분을 포함하는 폴리뉴클레오타이드를 포함한다.

[0084] ATF6은 비-접힘 단백질 반응(UPR)의 센서일 수 있고 전사 발현을 조절하는 기능을 할 수 있다. 어떤 경우에는 ER 스트레스 시 ATF6이 ER에서 골지로 이동하여 단백질 분해로 절단되어 단백질의 접힘 및 이동에 관여하는 유전자의 전사 인자인 N-말단 ATF6 조각을 방출하는 것으로 나타났다. 본원에서 "ATF6"은 신호전달 림프구 활성화 분자 패밀리의 자가-리간드 수용체인 폴리펩타이드를 나타내며, 인간에서 ATF6 유전자에 의해 암호화된다. 일부 구현예에서, ATF6 폴리펩타이드는 HGNC: 791, Entrez Gene: 22926, Ensembl: ENSG00000118217, OMIM: 605537, UniProtKB: P18850와 같이 하나 이상의 공개적으로 이용가능한 데이터베이스에서 확인된 것이다. 일부 구현예에서, ATF6 폴리펩타이드는 서열번호 10의 서열, 또는 서열번호 10과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리펩타이드 서열, 또는 서열번호 10의 부분을 포함하는 폴리펩타이드를 포함한다. 서열번호 10의 ATF6 폴리펩타이드는 성숙한 ATF6의 미성숙 또는 전처리된 형태를 나타낼 수 있고, 이에 따라 본원에 포함되는 것은 서열번호 10의 ATF6 폴리펩타이드의 성숙 또는 가공된 부분이다. 일부 구현예에서, 상기 ATF6 폴리뉴클레오타이드는 서열번호 27의 서열, 또는 서열번호 27의 서열과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리뉴클레오타이드 서열, 또는 서열번호 27의 부분을 포함하는 폴리뉴클레오타이드를 포함한다.

[0085] ATF4는 UPR 표적 유전자의 전사 활성제로, 아미노산 대사 및 산화 스트레스에 대한 내성에 관여하는 유전자의 전사 발현을 향상시키는 역할을 할 수 있다. (Fusakio ME et al., 2016) 본원에서 "ATF4"는 신호전달 림프구 활성화 분자 패밀리의 자가-리간드 수용체인 폴리펩타이드를 나타내며, 인간에서 ATF4 유전자에 의해 암호화된다. 일부 구현예에서, ATF4 폴리펩타이드는 HGNC: 786, Entrez Gene: 468, Ensembl: ENSG00000128272, OMIM: 604064, UniProtKB: P18848과 같은 하나 이상의 공개적으로 이용가능한 데이터베이스에서 확인된 것이다. 일부 구현예에서, ATF4 폴리펩타이드는 서열번호 11의 서열, 또는 서열번호 11과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리펩타이드 서열, 또는 서열번호 11의 부분을 포함하는 폴리펩타이드를 포함한다. 서열번호 11의 ATF4 폴리펩타이드는 성숙한 ATF4의 미성숙 또는 전처리된 형태를 나타낼 수 있고, 이에 따라 본원에 포함되는 것은 서열번호 11의 ATF4 폴리펩타이드의 성숙 또는 가공된 부분이다. 일부 구현예에서, 상기 ATF4 폴리뉴클레오타이드는 서열번호 28의 서열, 또는 서열번호 28의 서열과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리뉴클레오타이드 서열, 또는 서열번호 28의 부분을 포함하는 폴리뉴

클레오타이드를 포함한다.

- [0086] PERK(protein kinase RNA-like endoplasmic reticulum kinase)는 PERK에서 떨어진 샤페론 모집(recruitment)을 통해 일반적으로 활성화되어 세포질 키나아제 도메인의 올리고머화 및 활성화를 유도하는 유형 1 막횡단 단백질이다. PERK 및 ATF4의 하류에서 세포자멸사 신호 전달을 매개하는 중요한 단백질은 간 질환의 진행과 관련된 CCAAT 인핸서-결합 단백질(C/EBP) 상동 단백질(CHOP)이다. (Malhi H et al., 2011) "EIF2AK3"으로도 알려진 "PERK"는 본원에서 신호전달 림프구 활성화 분자 패밀리의 자가-리간드 수용체인 폴리펩타이드를 지칭하며, 인간에서 EIF2AK3 유전자에 의해 암호화된다. 일부 구현예에서, PERK 폴리펩타이드는 NC: 3255, Entrez Gene: 9451, Ensembl: ENSG00000172071, OMIM: 604032, UniProtKB: Q9NZJ5와 같이 하나 이상의 공개적으로 이용가능한 데이터베이스에서 확인된 것이다. 일부 구현예에서, PERK 폴리펩타이드는 서열번호 12의 서열, 또는 서열번호 12와 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리펩타이드 서열, 또는 서열번호 12의 부분을 포함하는 폴리펩타이드를 포함한다. 서열번호 12의 PERK 폴리펩타이드는 성숙한 PERK의 미성숙 또는 전처리된 형태를 나타낼 수 있고, 이에 따라 본원에 포함되는 것은 서열번호 12의 PERK 폴리펩타이드의 성숙 또는 가공된 부분이다. 일부 구현예에서, 상기 PERK 또는 EIF2AK3 폴리펩타이드는 서열번호 29 또는 서열번호 30의 서열, 또는 서열번호 29 또는 서열번호 30의 서열과 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98% 이상의 상동성을 갖는 폴리뉴클레오타이드 서열, 또는 서열번호 29 또는 서열번호 30의 부분을 포함하는 폴리뉴클레오타이드를 포함한다.
- [0087] 임의의 전술한 측면의 조성물은 HNF4 α 작용제를 추가로 포함할 수 있고, 여기서 HNF4 α 작용제는 모든 목적을 위해 본원에 참고로 도입되는 미국 특허 출원 공개 US2014/0249209에 보다 완전히 기재된 조성물을 지칭하는 것을 의미한다.
- [0088] 일부 구현예에서, 임의의 전술한 측면의 조성물 및/또는 벡터는 생물학적으로 허용되는 담체와 함께 제제화될 수 있다. 일부 구현예에서, 생물학적으로 허용되는 담체는 조성물 및/또는 벡터를 숙주 세포 내로 및/또는 숙주 세포 간에 전달할 수 있다. 일부 구현예에서, 상기 생물학적으로 허용되는 담체는 대상체의 간에 있는 간세포와 같은 표적 세포 내로 및/또는 표적 세포 사이에서 조성물 및/또는 벡터를 전달할 수 있다. 중요하게도, 일부 구현예에서, 생물학적으로 허용되는 담체와 함께 조성물 및/또는 벡터는 DNA 및 RNA와 같은 기능적 거대분자를 표적 세포(예를 들어, 간세포)의 핵으로 수송할 수 있다.
- [0089] **치료 방법**
- [0090] 본원에서는 대상체에서 간세포의 핵으로 전사 인자인 HNF4 α 의 발현 및/또는 수송 및/또는 보유를 증가시켜 간 질환을 치료하는 방법을 제공한다. 일부 구현예에서, 상기 방법은 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및 POM121C로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자의 발현 또는 기능을 상향 조절하고/하거나 DNAJB1/HSP40, ATF6, ATF4, 및 PERK로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자의 발현 또는 기능을 하향 조절하는 것을 포함한다. 일부 측면에서, 본원은 벡터를 대상체에게 투여하는 것을 포함하는, 간 질환의 치료를 필요로 하는 대상체에서 간 질환을 치료하는 방법이 개시하며, 여기서 상기 벡터는 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및/또는 POM121C로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자 및 그의 기능적 단편을 암호화하는 핵산을 포함한다. 일부 측면에서, 본원은 DNAJB1/HSP40, ATF6, ATF4 및 PERK로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자의 발현 또는 기능을 하향 조절하는 조성물을 대상체에게 투여하는 것을 포함하는, 간 질환의 치료를 필요로 하는 대상체에서 간 질환을 치료하는 방법이 개시한다. 일부 측면에서, 상기 조성물은 siRNA, miRNA, sgRNA 또는 tracr/mate RNA를 포함한다. 다른 구현예에서, 상기 방법은 Lys106에서 HNF4 α 의 아세틸화를 증가시키는 단계, cMET의 발현을 증가시키는 단계, 및/또는 Thr308에서 인산화를 통해 AKT의 활성화를 증가시키는 단계를 포함한다.
- [0091] 일부 구현예에서, 본원은 PROX1을 암호화하는 핵산을 포함하는 벡터를 대상체에게 투여하는 것을 포함하는 간 질환의 치료를 필요로 하는 대상체에서 간 질환을 치료하는 방법을 개시한다. 일부 구현예에서, SREBP1을 암호화하는 핵산을 포함하는 벡터를 대상체에게 투여하는 것을 포함하는, 간 질환의 치료를 필요로 하는 대상체에서 간 질환을 치료하는 방법이 개시된다. 일부 구현예에서, 상기 벡터는 HNF4 α 를 암호화하는 핵산을 추가로 포함한다. 일부 구현예에서, 상기 벡터는 HNF4 α , PROX1, 및 SREBP1을 암호화하는 하나 이상의 핵산을 포함한다. 일부 구현예에서, 상기 방법은 HNF4 α 를 암호화하는 핵산을 포함하는 벡터를 투여하는 단계를 추가로 포함한다.
- [0092] 앞서 언급한 바와 같이, PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300 및 POM121C는 모두 예를 들어 HNF4 α 의 아세틸화, 세포 대사 경로, 또는 핵공 복합체의 형성을 포함하는 직접 EH는 간접 메커니즘을 통해 HNF4 α 핵 수송을 조절하는 전사 인자 및/또는 조절자이다. HNF4 α 에 대한 이러한 효과는 간 질환이 있는 환자의 간 세포 기능

을 회복시킬 수 있다.

- [0093] 따라서, 일부 구현예에서, 벡터 또는 벡터들의 투여는 대상체의 간세포의 핵에서 HNF4 α 의 양을 증가시킨다. 일부 구현예에서, 벡터(들)의 투여는 간세포에서 HNF4 α 의 총량을 증가시키지 않는다. 일부 구현예에서, 벡터(들)의 투여는 간세포에서 HNF4 α 의 총량을 증가시킨다. 일부 구현예에서, 임의의 선행하는 측면의 벡터는 HNF4 α 를 암호화하는 핵산을 추가로 포함한다. 일부 구현예에서, 본원에 개시된 벡터는 HNF4 α 이소형 2 폴리펩타이드를 암호화하는 핵산을 추가로 포함한다. 일부 구현예에서, HNF4 α 이소형 2 폴리펩타이드는 서열번호 1 또는 그의 단편과 적어도 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98%의 상동성을 갖는 서열을 포함한다. 일부 구현예에서, 핵산은 서열번호 31 또는 그의 단편과 적어도 약 80%, 약 85%, 약 90%, 약 95%, 또는 약 98%의 상동성을 갖는다.
- [0094] 따라서, 일부 측면에서, 본원은 HNF4 α 이소형 2를 암호화하는 핵산을 포함하는 벡터를 대상체에게 투여하는 것을 포함하는, 간 질환의 치료를 필요로 하는 대상체에서 간 질환을 치료하는 방법을 개시한다. 또한 본원은 HNF4 α 이소형 2를 암호화하는 핵산을 포함하는 조성물을 대상체에게 투여하는 것을 포함하는 간 질환의 치료를 필요로 하는 대상체에서 간 질환을 치료하기 위한 약제의 제조를 위한 조성물의 용도를 포함한다.
- [0095] 상기 방법에 사용된 벡터는 플라스미드, 박테리오파지, 바이로소 입자(분리, 약독화, 재조합 등), 엑소솜, 세포 외 소포 및/또는 나노입자를 포함하는 본원에 기재된 임의의 것일 수 있다. 일부 구현예에서, 상기 벡터는 플라스미드이다. 일부 구현예에서, 상기 벡터는 바이러스 입자이다. 일부 구현예에서, 상기 벡터는 천연 및/또는 조작된 캡시드를 갖는 바이러스 벡터이다. 일부 구현예에서, 상기 벡터는 엑소솜이다. 일부 구현예에서 상기 벡터는 나노입자이다. 일부 구현예에서, 상기 벡터는 mRNA 지질 나노입자이다. 일부 구현예에서, 상기 핵산은 DNA (예를 들어, 폐쇄형 DNA(ceDNA)) 또는 RNA이다. ceDNA 및 ceDNA의 제조 및 사용 방법은 당업계에 공지되어 있다. 예를 들어, 전체로서 본원에 참조로 도입된 국제 공개 번호 WO2019/169233 및 WO2017152149를 참조할 수 있다. 방법, 재료, 나노 입자의 전달 및 이의 제조 및 사용을 포함하는(양 및 제형 포함) 나노 입자, 이의 구성요소 및 이러한 구성요소의 전달과 관련하여, 모두 본 발명의 실시예 유용하며, Wu et al., J. Biol. Chem. 262, 4429, 1987, U.S. 특허 출원 공개 2011/0274706, 및 WO2018/170405는 모든 목적을 위해 본원에 참조로 도입된다.
- [0096] 일부 측면에서, 본원에 대상체에게 조성물을 투여하는 것을 포함하는 간 질환의 치료를 필요로 하는 대상체에서 간 질환을 치료하는 방법이 개시되며, 여기서 상기 조성물은 DNAJB1/HSP40, ATF6, ATF4, 및 PERK로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자의 양을 감소시키거나 기능을 억제한다.
- [0097] 앞서 언급한 바와 같이, DNAJB1/HSP40, ATF6, ATF4 및 PERK는 모두 소포체(ER) 스트레스의 전사 조절자이다. ER은 단백질의 적절한 접힘, 변형 및 수송에 중요한 진핵 세포의 일종의 막성 소기관이다. 단백질을 접는 ER의 능력이 포화될 때 발생하는 ER 스트레스는 세포 사멸 및/또는 염증과 같은 반응으로 이어질 수 있다. 이러한 전사 조절자는 ER 스트레스와 관련된 경로를 통해 HNF4 α 핵 수송을 조절한다. 이러한 전사 조절제 중 하나 이상의 양의 감소 또는 기능의 억제가 간 질환이 있는 환자에서 간 세포 기능을 회복시킬 수 있음이 본원에서 제시된다. 따라서, 일부 구현예에서, 조성물의 투여는 대상체의 간세포의 핵에서 HNF4 α 의 양을 증가시킨다. 일부 구현예에서, 조성물의 투여는 간세포에서 HNF4 α 의 총량을 증가시키지 않는다. 일부 구현예에서, 조성물의 투여는 간세포에서 HNF4 α 의 총량을 증가시킨다. 일부 구현예에서, 조성물은 HNF4 α 를 암호화하는 핵산을 추가로 포함한다.
- [0098] 일부 구현예에서, 조성물은 이들 유전자의 녹다운에 의해 DNAJB1/HSP40, ATF6, ATF4, 및/또는 PERK로 이루어진 군으로부터 선택된 하나 이상의 전사 인자의 양을 감소시킨다. DNAJB1/HSP40, ATF6, ATF4 및/또는 PERK의 녹다운은 상보적 RNA 분자에 의한 DNAJB1/HSP40, ATF6, ATF4 및/또는 PERK를 암호화하는 mRNA 또는 DNAJB1/HSP40, ATF6, ATF4 및/또는 PERK의 활성화에 필요한 효소와 같은 관련 mRNA의 인식에 의해 야기될 수 있으며, RNA 간섭에 의해 매개될 수 있다. 예를 들어, 간섭 RNA(RNAi)를 암호화하는 분자는 형질감염(transfection) 또는 형질도입(transduction)과 같은 방법을 통해 렌티바이러스, 레트로바이러스 벡터 또는 나노입자와 같은 적합한 벡터에 의해 간세포 또는 간세포 전구체에 도입될 수 있다.
- [0099] 일부 구현예에서, DNAJB1/HSP40, ATF6, ATF4 및/또는 PERK로 이루어진 군으로부터 선택된 하나 이상의 전사 인자를 녹아웃시키기 위한 임의의 하나 또는 복수의 CRISPR 복합체 성분은 본원에 개시된 바이러스 입자, 비리온(virion), 또는 바이러스 벡터와 함께 또는 그 안에 투여될 수 있다. 일부 구현예에서, sgRNA 또는 tracr/mate RNA는 하나 이상의 재프로그래밍 인자와 함께 패키징될 수 있다. 일부 구현예에서, 바이러스 입자, 비리온 또는 바이러스 벡터에 의해 캡슐화된 sgRNA 분자는 하나 이상의 재프로그래밍 인자와 함께 패키징될 수 있다.

CRISPRCas 시스템에 관한 일반 정보, 그의 구성 요소 및 방법, 재료, 전달 비히클, 벡터, 입자, AAV를 포함하는 이러한 구성 요소의 전달 및 양 및 제형을 포함하여 이의 제조 및 사용과 관련하여, 본 발명의 실시예에 유용한 모든 목적을 위해 본 명세서에 참고로 도입된 미국 공개 2018/0057839를 참조한다.

[0100] 상기 언급한 바와 같이, 본원에서 사용된 "간 질환"은 일반적으로 간에 영향을 미치는 질환, 장애 및 상태를 지칭하며, 예를 들어 간세포에 단순 지방 축적(지방증), 거대낭포성 지방증, 문맥주위 및 소엽 염증(지방간염), 간경변, 섬유증, 간 허혈, 간세포 암종을 포함하는 간암, 초기 질환 단계의 간 질환, 말기 간 질환, 및 간 부전을 포함하는 광범위한 중증도를 가질 수 있다. 따라서, 지방증, 거대낭포성 지방증, 지방간염, 간경변, 섬유증, 간암, 간세포 암종, 말기 간 질환, 만성 간 질환, 및 간 부전은 각각 "간 질환"의 정의에 포함된다. "간 경변"의 정도 또는 중증도는 차일드-피 점수에 의해 지정될 수 있으며, 여기서 5가지 임상 측정값인 총 빌리루빈(bilirubin) 수준, 혈청 알부민, 프로트롬빈 시간 연장, 복수(ascite) 및 간성 뇌병증이 1점, 2점, 및 3점 값의 점수 시스템을 사용하여 각 임상 측정값의 다양한 수준에 대해 채점되며 3점 값은 각 측정의 가장 심각한 수준에 할당된다. 차일드-피 점수 및 분류에 도달하기 위해 5개 측정값 모두에 대한 총점이 추가된다. 5-6점은 차일드-피 분류 A, 7-9점은 차일드-피 분류 B, 10-15점은 차일드-피 분류 C이다. 일반적으로 차일드-피 분류 A는 가장 덜 심각한 간 질환을 나타내며, 차일드-피 분류 C는 가장 심각한 간 질환을 나타낸다. 일부 구현예에서, 본원에 개시된 방법은 차일드-피 분류 B 또는 차일드-피 분류 C 간 질환을 갖는 대상체를 치료하는 데 사용될 수 있다. 일부 구현예에서, 본원에 개시된 방법은 차일드-피 분류 A 간 질환을 갖는 대상체를 치료하는 데 사용될 수 있다. 일부 구현예에서, 상기 간 질환은 알코올성 간염이다. 일부 구현예에서, 본원에 개시된 방법은 생체의 관류를 갖는 허혈성 공여자 간을 치료하는 데 사용될 수 있다. 본 발명은 간 절제 전 또는 후를 포함하는 암 치료 전 또는 후에 간암을 치료하는데 사용될 수 있다. 초기 질환 단계의 간 질환은 비알코올성 지방간 질환(NAFLD), 비알코올성 지방간염(NASH), 알코올-관련 간 질환(지방간, 알코올성 간염 및 알코올-관련 간경변을 포함하나 이에 제한되지 않음)일 수 있음을 이해하고 본원에서 고려한다. 본원에 개시된 말기 간 질환은 예를 들어 바이러스성, 알코올성, 비-알코올성 및 잠복성을 포함하여 당업계에서 공지된 모든 원인에 기인할 수 있음을 또한 이해해야 한다.

[0101] 따라서, 본원에 개시된 방법은 임의의 상기 측면의 간 질환을 갖는 대상체에서 간 기능을 개선하는데 사용될 수 있다. 이러한 간 기능의 개선은 예를 들어 혈청 알부민의 증가, 혈청 암모니아 수준의 감소, 총 빌리루빈의 감소, 뇌병증 점수의 증가 및/또는 프로트롬빈 시간 연장의 감소로 나타날 수 있다. 따라서, 본원에 개시된 방법은 혈청 알부민 수준을 증가시키고, 혈청 암모니아 수준을 감소시키고, 총 빌리루빈 수준을 감소시키고, 뇌병증 점수를 증가시키고 및/또는 프로트롬빈 시간 연장을 감소시키는데 사용될 수 있다.

[0102] 간 질환은 염증, 섬유증, 간경변, 말기 간 질환, 및 간암을 포함하는 여러 단계로 진행될 수 있다. 간염 바이러스에 의한 만성 감염, 알코올-매개 간경변, 및/또는 비알코올성 지방간염(NASH)을 포함하여 만성 간 질환의 다양한 원인이 있음을 알아야 한다. 간 질환의 시기를 예측할 수 없는 경우가 많기 때문에, 간 질환을 치료, 예방, 감소 및/또는 억제하는 개시된 방법은 염증, 섬유증, 간경변, 말기 간 질환, 및/또는 간암의 발병 전 또는 후에, 심지어 간염 바이러스 감염, 알코올 매개 간경변, 및/또는 비알코올성 지방간염의 전 또는 도중에 간 질환의 임의의 단계를 치료, 예방, 억제 및/또는 완화하기 위해 사용될 수 있음을 이해해야 한다. 개시된 방법은 염증, 섬유증, 간경변, 말기 간 질환 및/또는 간암이 발병하기 전에 언제든지 수행할 수 있다. 한 측면에서, 개시된 방법은 염증, 섬유증, 간경변, 말기 간 질환, 및/또는 간암의 발병 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 또는 1년; 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 또는 1개월; 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 또는 3일; 60, 48, 36, 30, 24, 18, 15, 12, 10, 9, 8, 7, 6, 5, 4, 3, 또는 2시간 전; 또는 염증, 섬유증, 간경변, 말기 간 질환, 및/또는 간암의 발병 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 75, 90, 105, 120분; 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 18, 24, 30, 36, 48, 60시간; 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, 45, 60, 90일 또는 그 이상; 4, 5, 6, 7, 8, 9, 10, 11, 12개월 또는 그 이상; 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1년 후에 사용될 수 있다.

[0103] 간 절제는 간 질환(예: 간경변, 말기 간 질환, 및/또는 간암)이 있는 대상체의 간 전체 또는 일부를 외과적으로 제거하는 것이다. 개시된 방법은 간 절제 전 또는 후에 언제든지 대상체에게 수행될 수 있다. 한 측면에서, 개

시된 방법은 간 절제 수술 5, 4, 3, 2, 또는 1년; 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 또는 1개월; 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 또는 3일; 60, 48, 36, 30, 24, 18, 15, 12, 10, 9, 8, 7, 6, 5, 4, 3, 또는 2시간 전; 또는 간 절제 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 75, 90, 105, 120분; 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 18, 24, 30, 36, 48, 60시간; 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, 45, 60, 90일 또는 그 이상; 4, 5, 6, 7, 8, 9, 10, 11, 12개월 또는 그 이상; 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1년 후에 사용될 수 있다.

[0104] 본원에 기재된 벡터 또는 조성물은 경구, 국소, 정맥내, 피하, 경피, 경피, 근육내, 관절내, 비경구, 동맥내, 피내, 뇌실내, 두개내, 복강내, 병변내, 비강내, 직장, 질, 흡입 또는 이식된 저장소를 통하는 것을 포함하는 임의의 경로를 통해 대상체에게 투여될 수 있다. 용어 "비경구"는 피하, 정맥내, 근육내, 관절내, 활막내, 흉골내, 척추강내, 간내, 병변내, 및 두개내 주사 또는 주입 기술을 포함한다. 일부 구현예에서, 벡터 또는 조성물의 투여는 정맥내이다.

[0105] 본 개시의 또 다른 측면은 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, 및/또는 POM121C로 이루어진 군으로부터 선택되는 하나 이상의 전사 인자, 및 이의 기능적 단편의 양 또는 기능을 증가시키는 상기 기재된 조성물 및 DNAJB1/HSP40, ATF6, ATF4 및 PERK로 이루어진 군으로부터 선택된 하나 이상의 전사 인자의 양을 감소시키거나 기능을 억제하는 조성물 모두를 투여하는 것에 관한 것이다. 일부 구현예에서, 두 조성물은 동시에 투여된다. 다른 구현예에서, 하나의 조성물은 다른 것보다 먼저 투여된다. 일부 구현예에서, 임의의 전술한 측면의 방법은 HNF4a의 발현 또는 기능을 상향조절하는 것을 추가로 포함한다. 일부 구현예에서, 임의의 전술한 측면의 방법은 HNF4a 작용제를 추가로 투여하는 것을 포함하며, 용어 작용제는 모든 목적을 위해 본원에 참조로 포함된 미국 특허 출원 공개 US2014/0249209에 보다 완전히 기재된 조성물을 지칭한다.

[0106] 벡터 또는 임의의 전술한 측면의 조성물에 대한 투여 빈도는 적어도 1년에 1회, 2년에 1회, 3년에 1회, 4년에 1회, 5년에 1회, 6년에 1회, 7년에 1회, 8년에 1회, 9년에 1회, 10년에 1회, 적어도 2개월에 1회, 3개월에 1회, 4개월에 1회, 5개월에 1회, 6개월에 1회, 7개월에 1회, 8개월에 1회, 9개월에 1회, 10개월에 1회, 11개월에 1회, 적어도 1개월에 1회, 3주에 1회, 2주에 1회, 1주에 1회, 1주에 2회, 1주에 3회, 1주에 4회, 1주에 5회, 1주에 6회, 또는 매일을 포함하나 이에 제한되는 것은 아니다. 투여는 또한 연속적일 수 있고 임의의 원하는 특정 범위 내에서 화합물의 수준을 유지하도록 조정될 수 있다. 임의의 전술한 측면의 벡터 및/또는 조성물을 사용하여 간 질환을 치료하기 위해 본원에 사용된 용어 "투여" 또는 "투여하는"은 모든 목적을 위해 본원에 참조로 도입된 미국 특허 출원 공개 2018/0057839에 보다 완전히 기재된 투여 형태를 포함한다.

[0107] **실시예**

[0108] 하기 실시예는 개시된 주제에 따른 조성물, 방법 및 결과를 예시하기 위해 아래에 제시된다. 이들 예는 본원에 개시된 주제의 모든 측면을 포함하는 것으로 의도되지 않고, 오히려 대표적인 방법 및 결과를 예시하기 위한 것이다. 이들 실시예는 당업자에게 자명한 본 발명의 균등물 및 변형을 배제하도록 의도되지 않는다.

[0109] **실시예 1: 방법 및 재료**

[0110] **인간 샘플 및 간세포 분리.** 비식별화된 정상 인간 간 조직 및/또는 세포는 NIH 계약 번호 HSN276201200017C 하에 지원된 피츠버그 대학의 인간 연구 검토 위원회(Human Research Review Committee)에서 승인한 프로토콜에 의해 서면 동의를 얻은 후 간 조직 세포 분배 시스템(Liver Tissue Cell Distribution System)(Pittsburgh, PA)을 통해 얻었다. 성인 인간 간 조직 및/또는 세포는 또한 피츠버그 대학의 인간 연구 검토 위원회 및 기관 검토 위원회(Institutional Review Board)(식별자가 제거된 면제; IRB#: PRO12090466)에서 승인한 프로토콜에 따라 UPMC 아동 병원의 Ira J Fox 연구소에서 얻었다 (표 1). 간세포는 이전에 설명한 대로 3단계 콜라게나제 소화 기술을 사용하여 분리되었다(Gramignoli R et al., 2012). 트리판 블루 배제를 사용하여 이전에 설명한 대로 분리 후 세포 생존율을 평가했으며 생존율이 >80%인 세포 제제만 분석에 사용했다.

표 1

표 1. 본 연구에서 환자의 임상 매개변수

| 병인/코드(code), n | 나이 (년) | 성별 (M,F) | 총 빌리루빈 (mg/dL) | 알부민 (g/dL) | MELD | 차일드-피 |
|----------------|---------|----------|----------------|------------|------------|------------|
| 건강한 간, 4 | 64 ± 13 | 2,2 | — | — | — | — |
| NASH, 5 | 60 ± 10 | 2,3 | 1.76 ± 0.6 | 2.9 ± 0.1 | 12.4 ± 6.1 | 9.4 ± 0.9 |
| NASH 1 | 71 | 0,1 | 1.7 | 3.0 | 8 | 9 (B) |
| NASH 2 | 68 | 0,1 | 1.7 | 2.9 | 8 | 9 (B) |
| NASH 3 | 53 | 1,0 | 1.1 | 2.8 | 8 | 9 (B) |
| NASH 4 | 59 | 1,0 | 1.6 | 2.8 | 21 | 11 (C) |
| NASH 5 | 48 | 0,1 | 2.7 | 3.0 | 17 | 9 (B) |
| Alcohol, 2 | 55 ± 5 | 3,0 | 2.03 ± 1.0 | 2.53 ± 0.6 | 14.0 ± 2.6 | 10.0 ± 1.0 |
| Alcohol 1 | 52 | 1,0 | 1.5 | 1.9 | 16 | 11 (C) |
| Alcohol 2 | 61 | 1,0 | 1.4 | 3.2 | 11 | 9 (B) |

[0111]

[0112]

인 실리코 HNF4a 번역 후 변형(PTM) 분석. HNF4a 세포 국소화를 조절하는 PTM을 식별하기 위해 데이터베이스 및 간행물에서 컴퓨터 검색을 통해 인 실리코 분석을 수행했다(도 5). 프로세스는 식별(identification), 선별(screening) 및 선택(selection)의 세 단계로 나뉜다. 처음에는 51개의 PTM이 식별되었다. 다음으로 선별 단계에서 두 가지 제거 기준을 적용하여 23개의 PTM을 선택했다(도 5). 2개의 인산화 및 1개의 아세틸화 변형이 평가할 수 있는 HNF4a 국소화와 관련된 가장 그럴듯한 PTM으로 선택 단계에서 확인되었다.

[0113]

기체 크로마토그래피-질량 분석법을 사용한 안정 동위원소 분석. 100만개의 인간 간세포를 13-C6-표지 글루코스 및 글루타민 동위원소 추적자의 존재 하에 DMEM(Dulbecco's Modified Eagle's Medium) F12에서 96시간 동안 배양하였다(Thermo Fisher Scientific, San Jose, CA). 배지를 제거하고 세포를 얼음처럼 차가운 인산완충생리식염수로 세척하였다. 다음으로, 세포를 400 µL 메탄올 및 1 µL 노르발린을 함유하는 400 µL 물에 켄칭하고, 급어내고, 800 µL 얼음처럼 차가운 클로로포름으로 세척하고, 4°C에서 30분간 불텍싱하고, 4°C에서 7,300 rpm으로 10분간 원심분리하였다. 대사 산물(metabolite) 분석을 위해 상부 수상(aqueous phase)을 수집했다. 대사 산물 추출물을 14,000g에서 10분간 원심분리하여 극성상, 단백질 간상, 클로로포름상을 분리하였다. 물/메탄올 상을 포함하는 극성 대사 산물을 새로운 미세원심분리 튜브로 옮기고 SpeedVac에서 건조시키고 기체 크로마토그래피-질량 분석(GC-MS) 분석까지 -80°C에서 보관하였다. 그런 다음, 30 µL의 메톡시아민 하이드로클로라이드(Thermo Scientific)를 건조된 샘플에 첨가하고 간헐적으로 불텍싱하면서 30°C에서 2시간 동안 인큐베이션하였다. 샘플에 총 45 µL의 MMBTSTFA + 1% 텀트-부틸디메틸클로로실란을 첨가하고 55°C에서 1시간 동안 인큐베이션하였다. 유도체화된 샘플을 유리 삽입물이 있는 가스 크로마토그래피(GC) 바이알로 옮기고 GC-MS 자동샘플러에 추가했다. GC-MS 분석은 Agilent 5977B 질량 분석기에 연결된 30-m HP-5MSUI 모세관 컬럼이 장착된 Agilent 7890 GC(Santa Clara, CA)를 사용하여 수행되었다. 극성 대사 산물의 경우 GC 오븐에 대해 다음 가열 사이클이 사용되었다: 3분 동안 100°C, 이어서 300°C까지 5°C/분의 램프 및 48분의 총 실행 시간 동안 300°C에서 유지하였다. 스캔 모드에서 데이터를 수집했다. 대사 산물의 상대적 풍부함은 각 대사 산물 단편에 대한 잠재적으로 표시된 모든 이온의 통합 신호로부터 계산되었다. 질량 동위원소 분포는 모델로 분석하기 전에 IsoCorrector를 사용하여 본래 풍부도에 대해 수정되었다. 대사 산물 수준은 내부 표준 노르발린의 신호로 정규화되었다. 부분 농축 계산은 기질에서 중간 대사 산물에 대한 13C의 부분 기여도를 나타낸다. 이는 다음과 같이 계산된다:

[0114]

$$MPE = \left(\sum_{i=0}^{N_C} i * x_i \right) / N_C$$

[0115]

여기서 NC는 13C로 표시할 수 있는 탄소의 수이고, xi는 (M + i)번째 동위 원소의 비율이다.

[0116]

면역조직화학 및 HNF4a 정량화. 파라핀이 포매된 간 조직을 자일렌으로 탈파라핀화하고 에탄올로 탈수하였다. 항원 언마스킹(unmasking)은 pH 6.0의 시트르산 완충액에서 끓임으로써 수행되었다. 그런 다음 슬라이드를 3% 과산화수소에서 인큐베이션하고 정상 동물 혈청으로 차단한 다음, 1차 항체와 함께 4°C에서 밤새 인큐베이션했다. 사용된 1차 항체는 표 2에 나열되어 있다. 그런 다음 조직 섹션을 1차 항체의 동물 종에 해당하는 비오틴화된 2차 항체와 함께 인큐베이션하고(BA-1000; Vector Laboratories, Burlingame, CA), 3,3'-디아미노벤지딘에 노출시켜서(SK-4105; Vector Laboratories) 퍼옥시다제 활성을 시각화했다. 대조염색은 Richard-Allan Scientific Signature Series Hematoxylin(Thermo Scientific, Waltham, MA)으로 수행되었다. 정량화를 위해 2

명의 간 병리학자가 핵 및 세포질 HNF4 α의 면역 반응성을 독립적으로 등급을 매겼으며, 샘플당 3개의 고-전력 필드에서 1,000개의 간세포를 계산했다. 정상 간(n = 2), 차일드-퍼 B(n = 4) 및 차일드-퍼 C(n = 2)가 이러한 분석에 포함되었다. 차일드-퍼 B 및 C는 간경변 인간 간으로 그룹화되고 결과는 계산된 총 세포 수에 대한 백분율로 표시된다.

표 2

| 항체/표적 | 종 | 면역조직화학 희석배율 | 웨스턴 블롯 희석배율 | 회사/카탈로그 번호 |
|---------------------|----|----------------|----------------|-----------------------|
| HNF4 α | 생쥐 | 1:200 | 1:1000 | Abcam/Ab41898 |
| 아세틸-HNF4 α (Lys106) | 토끼 | - | 1:500 | Cusabio/CSB-PA727840 |
| cMET | 토끼 | 1:200 | 1:1000 | Cell Signaling/8198T |
| 총 AKT | 토끼 | - | 1:1000 | Cell Signaling/ 4691S |
| 포스포-AKT(Ser473) | 토끼 | 1:50 | 1:1000 | Cell Signaling/4060S |
| 포스포-AKT(Thr308) | 토끼 | 1:100 | 1:1000 | Cell Signaling/9275S |
| 총 AMPK α | 토끼 | - | 1:1000 | Cell Signaling/5831S |
| 포스포-AMPK α (Thr172) | 토끼 | 1:100 | 1:1000 | Cell Signaling/2535S |
| 총 EGFR | 토끼 | - | 1:1000 | Cell Signaling/ 4267S |
| 포스포-EGFR(Y1086) | 토끼 | - | 1:1000 | Cell Signaling/2220S |
| 포스포-H3(Ser10) | 토끼 | 1:200 | 1:1000 | Cell Signaling/9701S |
| 카스파제 3 활성 | 토끼 | 1:100 | 1:500 | Abcam/Ab32042 |
| HDAC1 | 토끼 | - | 1:100 | Cell Signaling/2062 |
| 히스톤 H3 | 생쥐 | - | 1:100 | Santa Cruz/SC517576 |
| B-액틴 | 토끼 | - | 1:2000 | Cell Signaling/4970S |

[0117]

[0118]

단백질 추출 및 웨스턴 블로팅. 단백질 발현 분석을 수행하기 위해, 분리된 간세포를 두 부분으로 나누었다; 한 분획은 이전에 설명한 표준 절차(Bell AW et al., 2006)에 따라 총 단백질 추출에 사용되었고, 다른 분획은 핵 단백질 분리에 사용되었다. 핵 단백질 분리를 위해, 환자당 1×10^7 에서 5×10^7 사이의 분리된 간세포를 세척하고 40mmol/L Tris(pH 7.6), 14mmol/L NaCl 및 1mmol/L EDTA에서 수거한 다음 원심분리(5분, 100g)하였다. 세포 펠렛을 2mL의 저장성 완충액[10mmol/L HEPES (pH 7.9), 10mmol/L NaH₂PO₄, 1.5mmol/L MgCl₂, 1mmol/L DTT, 0.5mmol/L 스퍼미딘, 및 프로테아제 및 포스파타제 억제제 카테일을 포함하는 1mol/L NaF (Sigma, St. Louis, MO)]에 현탁시켰다. 얼음 위에서 10분간 인큐베이션 후, 샘플을 Dounce 균질화기에서 균질화한 다음 원심분리했다(5분, 800g). 트리판 블루 염색으로 세포 용해를 모니터링했다. 상층액은 세포질 추출물로 저장되었다. 핵 펠렛을 동일한 완충액에서 추가로 2회 세척하였다.

[0119]

핵 단백질은 연속 교반과 함께 4°C에서 45분 동안 고장성 완충액[30mmol/L HEPES (pH 7.9), 25% 글리세롤, 450mmol/L NaCl, 12mmol/L MgCl₂, 1mmol/L DTT, 및 프로테아제 및 포스파타제 억제제 카테일을 포함하는 0.1mmol/L EDTA (Sigma, St. Louis, MO)] 50-100 μL에서 추출되었다. 추출물을 30,000g에서 원심분리하고, 상층액을 수집하고 150mmol/L NaCl을 함유하지만 동일한 용액에 대해 2시간 동안 투석하였다. 단백질 농도는 Bicinchoninic Acid assay(Sigma, St. Louis, MO)에 의해 결정되었다.

[0120]

표준 절차에 따라 웨스턴 블롯 분석을 수행했다(Natarajan A et al., 2007). 각 단백질 밴드의 강도는 National Institutes of Health Image J 소프트웨어를 사용하여 정량화되었다. 1차 항체 및 그 희석액은 상기 표 2에 나와 있다.

[0121]

RNA 시퀀싱 및 분석. 전체 게놈 가닥 특이적 RNA-seq를 사용하여 인간 분리된 1차 간세포로부터 RNA 발현 수준을 프로파일링했다. RNA-Seq 라이브러리는 이전(Hainer SJ et al., Genes Dev. 2015) 및 문헌(Kumar R et al., 2012)에 설명된 대로 준비되었다. TRIzol을 사용하여 장 세포에서 RNA를 추출한 후 제조업체의 지침에 따라 컬럼 정제(Zymo RNA clean and concentrator column)를 수행했다. 총 RNA는 풀링된 안티센스 올리고 혼성화를 사용하여 rRNA를 고갈시켰고, 이전에 설명한 대로 RNaseH 소화를 통해 고갈되었다 (Morlan JD et al., 2012; Adiconis X et al., 2013). Zymo RNA clean and concentrator column을 통한 정제 후, 첫 번째 가닥 cDNA가 합성되었다. 이어서, 두 번째 가닥 cDNA를 합성, 정제 및 단편화하였다. Illumina 기술을 사용하여 RNA-seq 라이브러리를 준비했다. 간단히 말해서, 말단 수선, A-테일링(A-tailing) 및 바코드 어댑터 결합 후 PCR 증폭 및 크기 선택을 수행했다. 라이브러리의 완전성은 quBit 정량화, 단편 분석기 크기 분포 평가 및 각 라이브러리에서

~10개 단편의 생어(Sanger) 시퀀싱으로 확인되었다. 라이브러리는 페어드엔드(paired-end) Illumina 시퀀싱을 사용하여 시퀀싱되었다.

- [0122] 페어드 엔드 리드(Paired-end read)는 QIAGEN의 CLC Genomics 워크벤치를 사용하여 hg38에 정렬되었고 TPM(transcript per million)으로 평가되었다. 데이터를 정렬하기 위해, Cluster 3.0(De Hoon MJ et al., 2004)을 사용하여 K-평균 클러스터링을 수행하고 Java TreeView(Saldanha AJ, 2004)를 사용하여 히트맵을 생성했다. 미스매치(Murphy SL et al., 2015)에 대한 기본 설정으로 삽입 비용(Goldman L et al., 2016)과 결실 비용(Goldman L et al., 2016)이 사용되었다. IPA (Ingenuity Pathway Analysis)는 차등적으로 발현된 유전자를 식별하고, 다운스트림 효과를 예측하고, 표적을 식별하는 데 사용되었다(QIAGEN Bioinformatics; www.qiagen.com/ingenuity). IPA 내의 조절 효과 분석은 상류 조절자와 생물학적 기능 간의 관계를 식별하는 데 사용되었다. 기본 설정이 분석에 사용되었다(즉, 상류 조절자는 유전자, RNA 및 단백질로 제한됨). RNA-seq 데이터는 Gene Expression Omnibus(등록 번호 www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE134422)에서 사용할 수 있다.
- [0123] 정상 인간 간세포에서 AKT 억제. 정상 인간 간세포(1백만 세포/웰)를 콜라겐 코팅된 웰에서 배양하였다. 세포를 성장 인자 또는 혈청 없이 6시간 동안 배양하였다. 그런 다음 세포를 AKT 억제제인 5 μM의 MK-2206(Cayman Chemical, Ann Arbor, Michigan)으로 24시간 동안 처리하였다. 총, 세포질 및 핵 단백질은 앞에서 설명한 대로 웨스턴 블로팅을 위해 추출되었다.
- [0124] 통계 분석. 데이터는 평균 ± SD로 표현되었다. 2개의 통계 그룹에 대한 웨스턴 블롯 결과는 Mann-Whitney 비모수 검정에 의해 평가되었고 3개의 통계 그룹에 대해서는 Kruskal Wallis 비모수 검정에 의해 평가되었다. 그룹 간의 비교는 Dunn 다중 비교 테스트에 의해 수행되었다. 분석된 단백질 간의 연관성은 스피어만 순위 상관관계수 테스트를 사용하여 평가되었다. 차일드-피 점수와 MELD 점수로 측정된 임상 상태와 단백질 발현 사이의 관계를 설명하기 위해 선형 회귀가 사용되었다. 통계는 Prism 4.0(GraphPad Software Inc., San Diego, California, USA)을 사용하여 수행되었다. 차이는 P < 0.05일 때 유의한 것으로 간주되었다.
- [0125] 웨스턴 블롯으로 분석한 단백질 간의 직접적인 의존성을 확인하기 위해, 경로분석(구조방정식모델)을 이용하였다. 경로 분석 모델은 종속 변수와 다중 독립 변수 간의 관찰된 상관관계 사이의 가능한 인과 관계를 설명하는 것을 목표로 한다. 경로 모델은 연구 프레임워크와 회귀 가중치 및 모델 적합도 결과를 기반으로 경로를 추가 및 제거하여 테스트 및 수정되었다. 결과는 연구 시스템에 대한 변수의 직접 및 간접 효과를 보여주는 다이어그램으로 표시되었다. 변수 간의 상관관계와 선형 관계는 P < 0.05와 중요도를 나타내는 임의의 계수(숫자가 클수록 큰 관계를 나타냄)에 의해 결정된다. 경로 분석은 InfoStat 버전 2013(Grupo InfoStat, FCA, Universidad Nacional de Cordoba, Cordoba, Argentina)을 사용하여 수행되었다.
- [0126] 샘플의 임상 상태를 구별하는 단백질 그룹을 밝히기 위해 자율(Unsupervised) 다변량 주성분 분석(PCA)이 웨스턴 블롯 데이터에 적용되었다. 대부분의 분산을 설명하는 주성분(PC)의 산점도(Scatter plot)가 그려졌다. PCA 분석에는 통계 소프트웨어 JMP 버전 14(SAS Institute, Cary, NC, USA)가 사용되었다.
- [0127] **실시예 2. 말기 간부전이 있는 인간 간에서 HNF4 α 핵 국소화는 감소하지만 세포질 국소화는 증가된다.**
- [0128] HNF4 α는 전사 인자로 기능하며 활성을 위해서는 핵 국소화가 필요하다 (Babeu JP et al., 2014; Chellappa K et al., 2012; Guo H, 2014; Hong YH et al., 2003; Lu H et al., 2016, Song Y et al., 2015; Soutoglou E et al., 2000; Sun K et al., 2007; Yokoyama A et al., 2011; Zhou W et al., 2012; Bell AW et al., 2006; Kritis AA et al., 1996; Tanaka T et al., 2006; Walesky C et al., 2015). 따라서, HNF4 α의 위치를 결정하고 발현을 간 대상부전(decompensation)과 연관시키기 위해 병든 간 표본의 간세포에 대한 면역조직화학(immunohistochemistry) 및 웨스턴 블롯을 수행했다. 말기 간부전이 있는 간의 간세포의 약 78%는 HNF4 α의 세포질 발현 또는 세포질 및 약한 핵 발현을 모두 나타내는 반면, 정상 인간 간은 75%의 간세포가 HNF4 α의 강한 핵 국소화를 나타냈다. 웨스턴 블롯에 의해 평가된 분리된 간세포로부터의 총 HNF4 α 단백질 발현은 말기 간과 정상 대조군 사이에 통계적 차이를 나타내지 않았다(P = 0.166; 도 1A). 퇴행성 질환이 있는 환자와 기능적 대상부전 정도에 따른 대조군의 간에서 HNF4 α 발현의 차이를 식별하는 능력은 대규모 환자 코호트에 대한 연구가 필요했기 때문에 이 결과는 놀라운 일이 아니다(Guzman-Lepe J et al., 2018). 그러나 본 연구에서는, HNF4 α 위치에 따라 통계적으로 유의한 차이가 관찰되었다. HNF4 α는 정상 대조군에서 분리된 간세포와 비교할 때, 기능적으로 보상되지 않은 간에서 분리된 간세포의 세포질에서 높은 수준(P = 0.023; 도 1B) 및 핵에서 낮은 수준(P = 0.023; 도 1C)으로 검출되었다.

- [0129] HNF4 α 의 기능과 안정성은 많은 번역-후 변형자에 의해 조절되기 때문에(Chellappa K et al., 2012; Guo H, 2014; Hong YH et al., 2003; Lu H et al., 2016, Song Y et al., 2015; Soutoglou E et al., 2000; Sun K et al., 2007; Yokoyama A et al., 2011; Zhou W et al., 2012), 그리고 핵 국소화가 활성화에 중요하기 때문에, HNF4 α 국소화를 조절하는 변형자를 평가하기 위한 인 실리코 분석을 수행했다. AMPK α 활성화가 HNF4 α 전사를 조절하는 것으로 밝혀졌다(Hong YH et al., 2003). 또한, AKT 경로에 의해 매개될 수 있는 HNF4 α 의 아세틸화는 분자를 안정화시켜 핵에서의 유지를 촉진한다(Soutoglou E et al., 2000)(도 5).
- [0130] **실시예 3: HNF4 α 는 진행성 간 질환에서 인간 간세포 기능의 주요 조절자이다.**
- [0131] 다음 평가는 정상 대조군(n=4)의 인간 간세포와 간경변 및 말기 간부전 환자로부터 회복된 간세포(n=4)(차일드-퍼 C) 간의 유전자 발현 차이를 비교하기 위해 수행되었으며, 연구를 NASH 및 알코올 매개 Laennec's 간경변 환자로 제한했다. RNA-seq 데이터의 계층적 클러스터링은 K-평균 클러스터링의 히트맵에서 계산된 간경변 및 간 기능 장애와 관련된 세 가지 주요 동적 패턴을 나타냈다(대조군에 대한 \log_2 FC; K=3). 클러스터 I(3478개 유전자) 및 III(1669개 유전자)는 대조군 인간 간세포에 비해 말기 간부전 환자의 간세포에서 중간 내지 고도로 상향 조절된 유전자를 나타낸다. 이 클러스터의 대부분의 유전자는 자가포식(authphagy) 및 세포사살 신호 전달과 관련이 있다(데이터는 표시되지 않음).
- [0132] 그러나 클러스터 II는, 말기 간세포에서 유의하게 하향 조절된 1669개의 유전자로 구성되었으며, 세린-트레오닌 단백질 키나아제(AKT1), 사이토크롬 P450(cytochrome P450 [CYP]c8, CYP2c9, CYP2e1, CYP3A4) 및 간세포 핵 인자(HNF4 α , forkhead box a1[FOXA1])을 암호화하는 유전자를 포함했다. 이 클러스터에 표시된 상위 경로에는 파르네소이드 X 수용체/레티노이드 X 수용체(RXR) 및 간 X 수용체/RXR 활성화, 미토콘드리아 기능 장애, 산화적 인산화 및 RXR 기능 억제가 포함된다. 이러한 하향 조절된 유전자에 대한 경로 분석은 HNF4 α 가 중심 상류 조절자임을 보여주었으며, 히트맵은 NASH 및 알코올 매개 Laennec's 간경변 환자의 간세포의 유전자 발현 프로파일에서 많은 유사성을 확인했다(데이터는 표시되지 않음). 이러한 결과는 말기 부전이 있는 간경변성 간에서 회수된 쥐 간세포의 유전자 발현 프로파일과 거의 동일했다(Liu L et al., 2012).
- [0133] **실시예 4: cMET 및 AKT 인산화는 말기 간부전 환자의 인간 간세포에서 HNF4 α 핵 국소화와 상관관계가 있다.**
- [0134] EGFR 및 cMET은 인 실리코 및 RNAseq 분석에서 확인된 AMPK 및 AKT 경로(Komposch K et al., 2015; Paranjpe S et al., 2016; Tsagianni A et al., 2018)를 조절하는 것으로 나타났기 때문에, 간 표본에서 이러한 분자에 대한 HNF4 α 항체-기반 분석의 중요한 조절자로 수행되었다. 비대상 간 표본에서 cMET 발현은 면역조직화학 및 웨스턴 블롯(P = 0.023; 도 2A 및 2B) 모두에 의해 측정될 때 분리된 대조군 인간 간세포와 비교할 때 현저히 감소되었다. 정상 및 병든 간 표본 간에 EGFR 발현에는 유의한 차이가 없었다(도 2A 및 2B). 그러나 EGFR의 활성 형태인 포스포-EGFR(Y1086)은 정상 인간 간세포와 비교할 때 비대상 질환 환자의 병든 간에서 유래한 간세포에서 높게 발현되었다(도 6A 및 6B). 또한, 세포 사멸과 간세포 복제의 연속적인 주기가 간경변의 특징이므로(Tsochatzis EA et al., 2014), 환자의 간세포에서 이러한 관찰은 면역조직화학 및 웨스턴 블롯을 사용하여 말기 간세포에서 복제 [포스포-H3(Ser10)] 및 세포 사멸(활성 카스파제 3) 마커의 발현을 입증하는 말기 간부전으로 입증되었다.
- [0135] cMET는 AMPK α 및 AKT를 제어할 수 있고 cMET는 기능적으로 비대상된 간의 간세포에서 유의하게 하향조절되므로, AMPK α 및 AKT의 활성화 과정을 분석했다. 총 AMPK α , 활성화된 AMPK α (Thr172) 및 그 기능 비대상 간세포와 정상 인간 간세포 간에 통계적으로 차이가 없었다(도 2A 및 2B). 그러나, 총 AKT, 활성화된 AKT(Thr308) 및 그 비율은 간 표본 및 간 기능 비대상 환자의 분리된 간세포에서 유의하게 감소했다(도 2A 및 2B). 또 다른 AKT 인산화 부위(Ser473)는 간 표본이나 정상 대조군 또는 비대상 표본에서 분리된 간세포에서 변하지 않았다(도 2A 및 B).
- [0136] HNF4 α 의 핵 국소화와 번역 후 변형 간의 관계를 추가로 분석하기 위해 스피어만 순위 상관 테스트를 수행했다. cMET 발현은 총 HNF4 α ($r = 0.76$; $P = 0.021$; 도 2C) 및 핵 HNF4 α ($r = 0.71$; $P = 0.037$; 도 2C)와 통계적으로 양의 유의한 상관관계를 보였다. 활성화된 AKT(Thr308)는 또한 총 HNF4 α ($r = 0.73$; $P = 0.031$; 도 2C) 및 핵 HNF4 α ($r = 0.82$; $P = 0.011$; 도 2C)와 통계적으로 양의 유의한 상관관계를 보인 반면, 세포질 HNF4 α 는 cMET($r = -0.80$; $P = 0.014$; 도 2C) 및 활성화된 AKT(Thr308)($r = -0.77$; $P = 0.021$; 도 2C)와 음의 상관관계를 보였다. 또한, 포스포-AKT(Thr308)/총 AKT의 비율은 cMET($r = 0.80$; $P = 0.014$; 도 6A) 및 총 AKT($r = 0.71$; $P = 0.037$; 도 6A)와 양의 상관관계가 있다. 따라서, 감소된 cMET은 AKT 경로의 감소된 활성화, 핵에서 감소된 HNF4 α 및 세포질에서 HNF4 α 의 더 많은 발현과 관련이 있었다.

[0137] **실시예 5: HNF4 α의 핵 국소화는 cMET/AKT 축에 의해 영향을 받고 간 기능 장애의 정도와 상관관계가 있다.**

[0138] 도 3A에서 볼 수 있듯이, 경로 분석은 cMET 발현과 핵 HNF4 α 발현(0.56; P = 0.004) 사이의 유의한 인과 관계와 핵 HNF4 α 수준과 활성화된 AKT(Thr308)/총 AKT 비율(0.05; P = 0.006) 사이의 직접적인 관계를 보였다. 모델링은 또한 총 HNF4 α 발현 수준이 핵 국소화에 기여한다는 것을 보여주었다(0.60; P = 0.042). 그러나, cMET 발현은 총 HNF4 α 발현(-0.37; P = 0.024)(도 3A)과 음의 관계를 가져서, cMET 발현이 총 HNF4 α 발현에 직접적인 영향을 미치지 않고 핵 국소화에만 영향을 미친다는 것을 나타낸다. 핵 발현 수준이 간 기능 장애의 정도(차일드-피 점수)와 상관관계가 있는지 여부를 평가하기 위해, 선형 회귀 분석을 수행했는데, 핵 HNF4 α 발현 수준이 차일드-피 점수($R^2 = 0.80$; P = 0.007)(도 3B)와 유의한 역 관계가 있음을 보여주었다. 함께, 이러한 경로 및 단백질 발현의 선형 회귀 통계 분석은 HNF4 α 국소화가 간 질환 진행과 연관되고, cMET 발현 및 AKT 인산화가 간세포 HNF4 α 핵 국소화 및 기능을 유지하는 데 중심적 역할을 한다는 것을 보여준다.

[0139] 다음으로, 주성분 분석(PCA)을 수행하여 HNF4 α 번역 후 변형자 관련 분자를 평가하고, 분자 패턴이 말기 간세포에서 간 기능과 상관관계가 있는지(차일드-피 점수) 묘사했다(도 3C 및 3D). 도 3C에 도시된 바와 같이, PC1(69.9%) 및 PC2(30.1%)는 연구된 분리된 간세포에서 간 기능 수준(차일드-피 점수)의 가변성을 100% 구별한다. 정상 인간 간세포를 특징으로 하는 벡터는 cMET 및 활성화된 AKT(Thr308)/총 AKT 비율, 총 HNF4 α 및 핵 HNF4 α인 반면, 인간 간세포 부전을 특징으로 하는 음성 특성은 세포질 HNF4 α 및 활성 카스파제 3 발현이다(도 3C 및 3D). 이와 함께, 이 통계 분석은 말기 간부전이 있는 인간 간세포의 분자 프로파일링을 확증하고 cMET, 활성화된 AKT(Thr308), 및 전체 및 핵 HNF4 α의 발현 사이의 인과 관계를 확립한다.

[0140] **실시예 6: 아세틸화 감소를 통해 말기 간부전 환자에서 핵에서의 HNF4 α 보유가 감소한다.**

[0141] AKT의 표적 중 하나는 CREB 결합 단백질의 활성화이다(Dekker FJ et al., 2009). CREB 결합 단백질은 뉴클레오솜 DNA에 대한 전사 인자의 접근을 증가시켜 핵 내 전사 인자의 전사 및 보유를 활성화시키는 뉴클레오솜 히스톤에 대한 고유 아세틸화 활성을 갖는 것으로 잘 알려져 있다. 따라서 이 축은 HNF4 α 핵 보유와 관련될 수 있다(Soutoglou E et al., 2000). 게놈 전체 전사체 분석 및 인 실리코 분석에서 cMET/AKT 키나아제 축 경로가 CREB 결합 단백질의 활성화를 통해 HNF4 α의 국소화 및 안정성을 제어할 수 있음을 나타내므로(Kumar R et al., 2012), 아세틸화된 HNF4 α의 핵 발현을 측정했고(도 4A-4C), 정상 대조군과 비교할 때 말기 간부전 환자의 간세포에서 핵의 HNF4 α 아세틸화가 유의하게 감소되었음을 발견했다(P = 0.024; 도 4A). 그리고 HNF4 α의 아세틸화가 간 기능 장애의 정도(차일드-피 점수)와 관련이 있음을 확인하기 위해, 선형 회귀 분석을 수행했으며, 아세틸화된 HNF4 α의 감소가 간 기능 장애와 직접적으로 유의한 상관관계가 있음을 보였다($R^2 = 0.71$; P = 0.004; 도 4C). 원리의 증명으로, HNF4 α 핵 국소화에서 활성화된 AKT(Thr308)의 역할을 확증하기 위해 AKT 신호 전달을 억제하여 갓 분리된 정상 인간 간세포에서 예비 실험을 수행했다. 갓 분리된 정상 인간 간세포를 강력한 알로스테릭 범-AKT 억제제(allosteric pan-AKT inhibitor)인 MK-2206으로 처리했다. AKT 억제 처리 24시간 후, 비처리 대조군과 비교하여 활성화된 AKT(Thr308)의 80%, 핵 HNF4 α의 25% 및 아세틸화된 핵 HNF4 α 발현의 14%가 감소되었다.

[0142] **실시예 7: 핵에서 HNF4 α의 보유는 여러 신호 분자에 의해 조절되며, 이는 말기 간부전과 유의하게 음의 연관성을 나타낸다.**

[0143] HNF4 α는 간 기능의 마스터 조절자이다(Babeu JP et al., 2014; Chellappa K et al., 2012; Guo H et al., 2014; Lu H et al., 2016, Song Y et al., 2015; Soutoglou E et al., 2000; Sun K et al., 2007; Xu Z et al., 2007; Zhou W et al., 2012; Bell AW et al., 2006). HNF4 α 발현의 변화는 암, B형 및 C형 간염, 알코올 매개 간경변, NASH와 같은 여러 병인을 가진 간 질환과 관련이 있다(Babeu JP et al., 2014; Chellappa K et al., 2012; Guo H et al., 2014; Lu H et al., 2016, Song Y et al., 2015; Soutoglou E et al., 2000; Sun K et al., 2007; Xu Z et al., 2007; Zhou W et al., 2012; Bell AW et al., 2006). TLF가 있는 동물 모델에서 HNF4 α 발현의 강력한 감소가 확인되었으며, 유전자 요법을 사용하여 HNF4 α 생산을 복원함으로써 간 세포를 정상 기능으로 재부팅한다(Nishikawa T et al., 2014). 이러한 관찰이 인간에게 적용되는지 여부를 평가하기 위해, 간 기능이 비대상된 대규모 환자 코호트의 간에서 간-강화 전사 인자 발현에 대한 연구가 수행되었다(Guzman-Lepe J et al., 2018). HNF4 α mRNA 수준이 하향 조절되고 차일드-피 분류에 기초한 간 기능 장애의 정도와 상관관계가 있는 것으로 밝혀졌다. 이러한 연구에서 HNF4 α의 핵 국소화는 균일하지 않았다(Guzman-Lepe J et al., 2018).

[0144] 인간 간세포는 NASH 및 알코올 매개 Laennec's 간경변으로 인한 간경변 및 말기(차일드-피 B, C) 간부전 환자의

의식된 간에서 분리되었다. 이러한 분리된 간세포에서 대조군 정상 간세포에서 볼 수 있는 것과 비교하여 세포질에 위치한 HNF4α의 증가와 핵의 HNF4α의 감소가 있었다. 또한, 인간 부전 간경변성 간세포에서 세포질 또는 핵으로의 HNF4α의 국소화는 간세포 기능 장애의 정도와 상관관계가 있었다. 이러한 데이터는 HNF4α 핵 수송 또는 보유를 조절하는 경로가 말기 간부전 치료의 표적이 될 수 있음을 나타낸다.

[0145] AMPK 및 AKT 키나아제는 HNF4α 국소화를 조절할 수 있는 주요 구성요소이다(Hong YH et al., 2003; Song Y et al., 2015; Soutoglou E et al., 2000). AMPK는 에너지 항상성을 유지하고, ATP(아데노신 삼인산) 생산 경로를 촉진하며, ATP 소비를 줄이는 데 중심 역할을 하는 반면(Woods A et al., 2017), AKT 활성화는 세포 증식, 생존 및 성장을 촉진한다(Manning BD et al., 2017, Morales-Ruiz M et al., 2017). AKT 활성화는 트레오닌 308 및/또는 세린 473에서의 인산화에 의해 매개된다(Praveen P et al., 2016; Inoue J et al., 2017). 본원에 개시된 것은 활성화된 AKT(Thr308) 수준이 말기 인간 간세포에서 유의하게 감소되었기 때문에 활성화된 AKT(Thr308)와 HNF4α 국소화 사이의 유의한 상관관계를 보여준다. 이러한 발견은 Thr308에서의 AKT 인산화가 간 질환의 말기 단계에서 간세포 부전에 중요한 역할을 할 수 있음을 나타낸다.

[0146] 분석은 AKT와 AMPK의 상류 조절자이며 간 기능 및 재생과 관련된 두 가지 중심 수용체인 cMET 및 EGFR에 대해 수행되었다(Natarajan A et al., 2007; Komposch K et al., 2015; Paranjpe S et al., 2016; Tsagianni A et al., 2018; Alam A et al., 2017). 생쥐(mouse)에서 cMET와 EGFR의 결합된 파괴(disruption)는 간 항상성을 변경하고 말기 간부전으로 이어진다(Tsagianni A et al., 2018). 인간 부전 간경변 간세포에서, cMET의 발현 감소가 있었고 발현 감소는 HNF4α 국소화와 직접적으로 상관관계가 있었다. 대조적으로, 인간 부전 간경변 간세포 또는 대조군 인간 간세포에서 총 EGFR 발현에는 차이가 없었다. 따라서, 인간 간에서 cMET는 AKT 경로 활성화와 HNF4α 국소화 및 기능을 조절하는 데 중요한 역할을 할 수 있다. 또한, 다양한 통계 분석(스피어만 순위 상관 테스트, 경로 분석, 선형 회귀 분석 및 주성분 분석)을 사용하여 인간 부전 간경변 및 정상 간세포에서 분석된 번역 후 변형자 간의 연관성을 설정할 수 있다. 이러한 통계적 분석은 cMET과 활성화된 AKT(Thr308)가 핵 HNF4α의 발현 수준과 직접적인 관련이 있음을 보여주었다. cMET 단백질 발현과 총 HNF4α 사이에서 발견된 음의 연관성은 cMET 발현이 총 HNF4α 발현에 직접적인 영향을 미치지 않고 핵 국소화에만 영향을 미친다는 것을 나타낸다.

[0147] HNF4α의 아세틸화는 인간 부전 간경변 간세포에서 활성화된 AKT(Thr308)의 낮은 발현 수준과, 뉴클레오솜 DNA에 대한 전사 인자 결합을 증가시키는 고유 아세틸화 활성을 갖는 분자인 CREB 결합 단백질에 대한 AKT의 직접적인 효과에 기초하여, 간경변 부전 간세포에서 영향을 받을 수 있다. 실제로, 핵 HNF4α 아세틸화는 간경변성 간의 인간 간세포에서 현저하게 감소되었으며 그 수준은 간 기능 장애의 정도와 관련이 있었다. 이러한 관찰은 트레오닌 308에서의 AKT 인산화가 CREB 결합 단백질을 조절함으로써 아세틸화를 통해 HNF4α 핵 보유를 매개한다는 것을 나타낸다(Soutoglou E et al., 2000).

[0148] 요약하면, 세포질에서 HNF4α의 국소화는 간 질환의 진행 단계 동안 핵에서 HNF4α를 유지하는 분자 경로의 변경으로 인해 발생한다. cMET 및 활성화된 AKT(Thr308)는 하향 조절되고 HNF4α의 아세틸화 및 핵 보유에 영향을 미친다. 이러한 데이터는 HNF4α를 핵에 국소화하는 것을 통해 만성 간 질환에서 간세포 기능의 회복을 보여준다.

[0149] **실시예 8: 전사 인자 렌티바이러스(LV) 구축물(construct)을 사용한 1차 인간 간세포의 형질도입.**

[0150] 전사 인자 LV를 사용한 1차 인간 간세포의 형질도입. 간세포는 이중 콜라겐(두꺼운 층) 시스템에서 배양되어 간세포의 역분화를 방지한다. 콜라겐 샌드위치 프로토콜은 이후에 사용된다. 다음을 준비한다:

[0151] WARM: dPBS, HMM(기본+SingleQuotes), HCM(HBM 기본 + HCM SingleQuotes); ON ICE: 녹색 형광 단백질(GFP) LV*, 전사 인자(TF) LV*, Max Enhancer, TransDux; 기타: 1.5mL 튜브, 50mL 튜브, 팁, 피펫.

[0152] 1. 두꺼운 층의 콜라겐에 웰당 5e5 간세포를 플레이팅하고 4시간 동안 세포를 부착시킨다.

[0153] 2. 따뜻한 dPBS 2X로 웰을 세척하여 죽은 세포를 제거하고 배지를 500uL HMM(FBS 없음)으로 교체한다.

[0154] 3. 5개의 튜브를 준비하고 다음과 같이 라벨을 붙인다: a. GFPLV-2; b. GFPLV-10; c. TFLV-0; d. TFLV-2; e. TFLV-10.

[0155] 4. 50mL 튜브에 HMM/Max Enhancer/TransDux(HMT) 용액을 준비한다: 12.323mL HMM + 3.100mL Max Enhancer + 77.5uL TransDux.

[0156] 5. 미리 라벨이 붙은 튜브에 LV 용액을 준비한다: a. GFPLV-2: 618.3uL HMM + 1.7uL GFP LV; b. GFPLV-10:

611.6uL HMM + 8.4uL GFP LV; c. TFLV-0: 620uL HMM; d. TFLV-2: 618.6uL HMM + 1.44uL TF4 LV; e. TFLV-10: 612.82uL HMM + 7.18uL TF4 LV.

- [0157] 6. 웰의 HMM 배지를 500 μL HMT 용액으로 교체한다.
- [0158] 7. 각 LV 용액 100uL를 웰에 분배한다. 혼합할 접시를 흔들어준다.
- [0159] 8. 다음날 따뜻한 dPBS 2X로 웰을 세척하여 죽은 세포와 남아 있는 LV 용액을 제거한다.
- [0160] 9. 두꺼운 콜라겐으로 세포를 덮고 콜라겐이 겔화되도록 2시간 동안 둔다.
- [0161] 10. 500uL HCM을 웰에 추가하고 매일 신선한 HCM으로 교체한다.
- [0162] 11. 형질도입 후 72시간 및 96시간에 샘플을 얻는다(72시간 프로토콜 참조).
- [0163] 12. 96시간에 얻은 샘플의 경우 따뜻한 dPBS로 세포를 세척하고 배지를 500uL HCM(FBS 없음)으로 교체한다.
- [0164] 참고: *37°C의 수조에서 LV를 빠르게 해동한다. 후드로 옮겨 회전, 반전 또는 부드럽게 볼텍싱하여 혼합하고 얼음 위에 보관한다. 사용하지 않은 LV는 분취하여 -80°C에서 재냉동할 수 있으며 각 재냉동 시 바이러스 활성이 10-20% 손실된다.
- [0165] 전사 인자(TF) LV 구축물.
- [0166] TF LV 구축물은 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, POM121C 또는 HNF4 α를 암호화하는 폴리뉴클레오타이드 또는 DNAJB1/HSP40, ATF6, ATF4 또는 PERK에 해당하는 RNAi를 함유하는 렌티바이러스 벡터(Systems Bioscience, Cat#CS970S-1)이다.
- [0167] ELISA(72 및 96시간)에 대해 조정 배지(conditioned medium) 수집. 각 그룹에 대해 2개의 웰에서 1200uL의 조정 배지를 수집하고 1.5mL 튜브로 옮긴다. 수집된 배지를 따뜻한 HMM으로 교체한다. 조정 배지를 2분 동안 20,000xg에서 원심분리 한다. 상층액을 새 튜브에 옮기고 -20°C에서 보관한다.
- [0168] GFP 발현 및 브라이트필드(Brightfield)의 사진을 찍어 형질감염 효율을 결정한다(72 및 96시간). 따뜻한 dPBS로 부유(floating) 세포를 세척한다. 따뜻한 HCM으로 교체한다. GFP 발현에는 필터 2(녹색 여기(excitation))를 사용하고 브라이트필드에는 필터 6을 사용하여 사진을 찍는다.
- [0169] RNA 추출을 위해 Qiazol에서 세포 용해물을 수집한다(72 및 96시간). 따뜻한 dPBS 2X로 세포를 세척한다. 600uL Qiazol로 웰을 코팅하고 1분 동안 인큐베이션한다. P1000을 사용하여 플레이트에서 세포를 긁어내고 1.5mL 튜브로 옮긴다. RNA가 분리될 때까지 -20°C에서 보관한다.
- [0170] IF로 대한 웰을 고정한다(72시간). 따뜻한 dPBS 2X로 세포를 세척한다. 750uL 4% PFA 용액으로 웰을 코팅하고 40분 동안 인큐베이션한다. 1mL dPBS 3X로, 세척당 10분간 웰을 세척한다. dPBS 1mL를 추가한다. HNF4A가 염색될 때까지 4C에서 보관한다.
- [0171] 웨스턴 블롯에 대한 세포 용해물을 수집한다(72 및 96시간). 다음을 함유하는 얼음처럼 차가운 용해 용액을 준비한다:

표 3

| [0172] | 1mL | 2mL | 3mL |
|---------------------------------|-------|--------|--------|
| a. NP40 세포 용해 완충액 - 90% | 900uL | 1800uL | 2700uL |
| b. Complete (1 tab/1.5mL) - 6% | 60uL | 120uL | 180uL |
| c. Halt protease inhibitor - 2% | 20uL | 40uL | 60uL |
| d. 0.1M PMSF in ethanol - 2% | 20uL | 40uL | 60uL |

- [0173] 따뜻한 dPBS 2X로 세포를 세척한다. 200uL의 얼음처럼 차가운 용해 용액으로 웰을 코팅하고 워크인 냉장에서 흔들면서 30분 동안 인큐베이션한다. 고무 청소기(rubber policeman)를 사용하여 세포를 분리하고 침전물 및 고체를 포함하는 용해물을 미리 라벨이 붙은 튜브에 수집하고 옮긴다. 4°C에서 10분 동안 20,000xg에서 회전시킨다. 상층액을 미리 라벨이 붙은 깨끗한 1.5mL 튜브에 옮긴다. -80°C 냉동고에 용해물을 보관한다.
- [0174] TF 면역형광 공동 염색(두꺼운 콜라겐 샌드위치 층이 있는 12-웰 플레이트용). 세포의 분리를 방지하기 위해 부드럽게 샘플을 흡인한다. 샘플을 건조시키지 않는다. * 최대 속도에서 5분 동안 회전을 줄인다. * 2차 항체는

실험에 따라 적절한 다른 항체로 대체될 수 있다.

- [0175] 고정(세포가 이미 고정된 경우 차단 및 투과 단계로 이동). 실온에서 40분 동안 PBS pH 7.4에서 4% 파라포름알데히드로 샘플을 고정한다. 샘플을 얼음처럼 차가운 PBS로 각 세척당 10분간 3X 세척한다. 염색을 진행하거나 염색이 될 때까지 4°C에서 보관한다(최대 2주).
- [0176] 차단 및 투과. 샘플을 1mL PBS로 2X 세척한다. 샘플을 1mL 세척 완충액(PBS, 0.1% BSA 및 0.1% Tween)으로 각 세척당 10분간 3X 세척한다. 1mL 차단 완충액(PBS, 10% 정상 당나귀 혈청, 1% BSA, 0.1% Tween 및 0.1% Triton X-100)에서 샘플을 2시간 동안 인큐베이션하여 차단하고 투과시킵니다.
- [0177] 항체 인큐베이션: *마우스 항-TF stock 1° Ab를 볼텍싱 및 스프린다운. 차단 완충액에서 1:500 마우스 항-TF 1° Ab 희석을 준비한다. 600uL의 희석된 1° Ab로 샘플을 코팅하고 가슴 챔버에서 실온에서 6시간 동안 또는 4°C에서 밤새 인큐베이션한다. 1° Ab 용액을 흡인하고 세척 완충액으로 세포를 각 세척당 10분간 3X 세척한다. *당나귀 항-마우스IgG-AF594(Invitrogen A21203)를 볼텍싱 및 스프린다운한다. 차단 완충액에 1:250으로 희석된 2° Ab를 준비한다. 600 μL의 희석된 2° Ab로 샘플을 코팅하고 실온에서 2시간 동안 가슴 챔버에서 인큐베이션한다. 2° Ab 용액을 흡인하고 세척 완충액으로 세포를 각 세척당 10분간 3X 세척한다.
- [0178] 카운터염색 및 마운팅: PBS 3X로 샘플을 세척한다. 암실의 PBS에서 1ug/ml Hoechst 33342의 1 mL에서 2분 동안 세포를 인큐베이션한다. PBS로 샘플을 3X 세척한다. 샘플은 4°C의 암실에서 보관될 수 있다. RED 채널을 사용하여 TF를 테스트한다.
- [0179] **실시예 9: 전사 인자(TF) mRNA(50, 100, 500ng)를 사용한 1차 인간 간세포의 형질감염.**
- [0180] 1차 인간 간세포의 형질감염. 다음을 준비한다:
- [0181] WARM: DPBS, HMM(HMM 기본 배지 + HMM SingleQuotes), Opti-MEM → RT; ON ICE: Lipofectamine Messenger Max → RT, mRNA → RT; 기타: 1.5mL 튜브, 50mL 튜브, 팁, 피펫.
- [0182] 1. 따뜻한 DPBS로 세포를 세척하고 배지를 500uL HMM(FBS 없음)으로 교체한다.
- [0183] 2. 다음과 같이 라벨이 붙은 두 세트의 튜브를 준비한다: a. GFP-50. b. GFP-100. c. GFP-500. d. TF-0. e. TF-50. f. TF-100. g. TF-500. TF는 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300, POM121C 및 HNF4 a에서 선택된다.
- [0184] 3. Opti-MEM에서 두 번 희석된 mRNA를 준비한다: a. DilGFP: 32.5uL OptiMEM + 3.61uL GFP mRNA. b. DilHNF: 40.3uL OptiMEM + 4.47uL TF mRNA.
- [0185] 4. 미리 라벨이 붙은 튜브의 첫 번째 세트에서 mRNA-OptiMEM 혼합물을 준비한다: a. GFP-50: 312.2uL Opti-MEM + 2.8uL DilGFP. b. GFP-100: 309.4uL Opti-MEM + 5.6uL DilGFP. c. GFP-500: 287.3uL Opti-MEM + 27.7uL DilGFP. d. HNF-0: 315uL Opti-MEM. e. TF-50: 311.5uL Opti-MEM + 3.5uL DilTF. f. TF-100: 308.1uL Opti-MEM + 6.9uL DilTF. g. TF-500: 280.7uL Opti-MEM + 34.3uL DilTF.
- [0186] 5. 희석된 Lipo 혼합물을 준비한다: a. 2,220.75uL Opti-MEM + 141.75uL Lipo.
- [0187] 6. 2-3초 동안 희석된 Lipo 혼합물을 볼텍싱하고 미리 라벨이 붙은 두 번째 세트의 튜브에 310uL를 분배한다. 실온에서 10분 동안 인큐베이션한다.
- [0188] 7. 각 mRNA-OptiMEM 혼합물의 310uL를 희석된 Lipo 혼합물이 함유된 튜브에 옮긴다. 실온에서 5분 동안 인큐베이션한다.
- [0189] 8. 각 웰에 각 혼합물 100uL를 분배한다.
- [0190] 9. 인큐베이션하고 다음날 아침, 따뜻한 dPBS로 세척하고 500uL HMM(FBS 없음)으로 교체한다.
- [0191] 10. 형질감염 후 24* 및 48*시간에 샘플을 얻는다(언제 간세포가 플레이팅 되는지에 따라 다름).
- [0192] ELISA(24 및 48시간)에 대해 조정 배지(conditioned medium) 수집. 각 그룹에 대해 2개의 웰에서 1200uL의 조정 배지를 수집하고 1.5mL 튜브로 옮긴다. 수집된 배지를 따뜻한 HMM으로 교체한다. 조정 배지를 2분 동안 20,000xg에서 원심분리 한다. 상층액을 새 튜브에 옮기고 -20°C에서 보관한다.
- [0193] GFP 발현 및 브라이트필드(Brightfield)의 사진을 찍어 형질감염 효율을 결정한다(24 및 48시간). 따뜻한 dPBS로 부유 세포를 세척한다. 따뜻한 HMM으로 교체한다. GFP 발현에는 필터 2(녹색 여기)를 사용하고 브라이트필드

에는 필터 6을 사용하여 사진을 찍는다.

[0194] RNA 추출을 위해 Qiazol에서 세포 용해물을 수집한다(24 및 48시간). 따뜻한 dPBS 2X로 세포를 세척한다. 600uL Qiazol로 웰을 코팅하고 1분 동안 인큐베이션한다. P1000을 사용하여 플레이트에서 세포를 긁어내고 1.5mL 튜브로 옮긴다. RNA가 분리될 때까지 -20°C에서 보관한다.

[0195] IF로 웰을 고정한다(24시간). 따뜻한 dPBS 2X로 세포를 세척한다. 750uL 4% PFA 용액으로 웰을 코팅하고 20분 동안 인큐베이션한다. 1mL dPBS 3X로, 세척당 5분간 웰을 세척한다. dPBS 1mL를 추가한다. TF가 염색될 때까지 4°C에서 보관한다.

[0196] 웨스턴 블롯에 대한 세포 용해물을 수집한다(24 및 48시간). 다음을 함유하는 얼음처럼 차가운 용해 용액을 준비한다:

표 4

| | 1mL | 2mL | 3mL |
|---------------------------------|-------|--------|--------|
| a. NP40 세포 용해 완충액 - 90% | 900uL | 1800uL | 2700uL |
| b. Complete (1 tab/1.5mL) - 6% | 60uL | 120uL | 180uL |
| c. Halt protease inhibitor - 2% | 20uL | 40uL | 60uL |
| d. 0.1M PMSF in ethanol - 2% | 20uL | 40uL | 60uL |

[0198] 따뜻한 dPBS 2X로 세포를 세척한다. 200uL의 얼음처럼 차가운 용해 용액으로 웰을 코팅하고 워크인 냉장에서 흔들면서 30분 동안 인큐베이션한다. 고무 청소기(rubber policeman)를 사용하여 세포를 분리하고 침전물 및 고체를 포함하는 용해물을 미리 라벨이 붙은 튜브에 수집하고 옮긴다. 4°C에서 10분 동안 20,000xg에서 회전시킨다. 상층액을 미리 라벨이 붙은 깨끗한 1.5mL 튜브에 옮긴다. -80°C 냉동고에 용해물을 보관한다.

[0199] 면역형광 동시 염색(염색, 고정, 차단 및 투과, 항체 인큐베이션, 카운터염색 및 마운팅)은 상기 실시예 9에서 언급된 바와 동일하다.

[0200] 실시예 10. 전사 인자 및 조절자 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300 및 POM121C는 말기 간부전이 있는 간경변성 간세포의 HNF4α 핵 발현을 개선한다.

[0201] 간-강화 전사 인자는 말기 간경변이 있는 쥐의 간세포에서 안정적으로 하향 조절되며, 그 중 하나인 간세포 핵 인자 4 알파(HNF4α)의 강제 재발현이 기능을 회복하기 위해 기능 장애 간세포를 배양과 생체내 모두에서 재프로그래밍하는 것으로 밝혀졌다. 진행성 간 질환 환자의 대규모 코호트에서 병든 간에서 HNF4α mRNA 발현 수준이 간 기능 장애의 정도(차일드-피 분류)와 상관관계가 있으며 그 발현이 쥐 연구의 경우처럼 핵에 국한되지 않는 것으로 나타났다. 진행성 간경변 환자의 간에서는, 간 기능이 저하됨에 따라 HNF4α RNA 발현량이 감소하고, 세포질에서 단백질 발현이 발견된다. 이러한 발견은 퇴행성 간 질환 환자의 간 기능 손상을 설명할 수 있다. 더욱이, RNA-seq 분석은 HNF4α 및 핵 단백질 전위에 관여하는 기타 전사 인자/조절자-관련 경로가 말기 부전 환자의 간경변 세포에서 하향 조절되는 것으로 나타났으며, 여기서 HNF4α의 핵 수준은 유의하게 감소되었으며, HNF4α의 세포질 발현이 증가된 것으로 밝혀졌다. 또한, 소포체(ER) 스트레스의 4가지 주요 전사 조절자가 유의하게 상향 조절되는 것으로 밝혀졌다. 이 연구는 HNF4α 및 번역 후 변형과 관련된 경로의 조작이 말기 간부전 환자에서 간세포 기능을 회복시킬 수 있음을 나타낸다.

[0202] PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300 및 POM121C의 단백질 발현은 말기 간부전 환자의 간 기능 장애 정도와 상관관계가 있다. HNF4α가 제대로 기능하려면 핵에서 발현되어야 한다; 따라서 HNF4α의 핵 국소화와 관련된 신호 전달 경로는 기능이 비대상된 외식된 인간 간에서 분리된 간세포에서 분석되었다. 전사 인자 및 조절자 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300 및 POM121C가 RNA-seq 분석에서 HNF4α의 중요한 조절자로 확인되었기 때문에, NASH 또는 알코올 유발 간경변(차일드 피 "B" 및 "C")을 위해 간 이식을 겪은 환자의 간으로부터 분리한 1차 인간 간세포 또는 대조군 정상 간세포에서 이러한 분자에 대한 항체 기반 분석이 수행되었다. HNF4α 발현은 분리된 대조군 인간 간세포와 비교하면 웨스턴 블롯(도 8A)으로 측정할 때 비대상 간 표본에서 현저히 감소했다. 간부전이 진행됨에 따라 MTF1 발현에도 유의한 차이가 있었다. 또한, 단순 선형 회귀를 사용하여 차일드-피로 채점된 인간 간세포는 HNF4α 및 MTF1의 단백질 발현과 상관관계가 있었으며, HNF4α 및 MTF1 모두 간부전 정도와 유의하게 상관관계가 있음을 발견했다(p=0.007)(도 8B-8F). 또한, NROB2(도 9A-D), NR5A2(도 10A-D), PROX1(도 11A-D)의 단백질 발현은 차일드 피 C 간세포에서 유의하게 낮았으며 이들의 발현은 간세

포 기능 장애의 정도와 상관관계가 있었다.

- [0203] HNF4 α 핵 발현 및 위치에 대한 이러한 확인된 전사 인자 및 조절자의 역할을 더 이해하기 위해, 인간 간세포주를 CRISPR/Cas9를 사용하여 유전자 편집하여 PROX1 또는 NR5A2 또는 NROB2 또는 MTF1 또는 SREBP1 또는 EP300 및 POM121C의 발현을 녹아웃(KO)시켰다(도 15A-B). PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300 또는 POM121C의 KO에 의해 HNF4 α 핵 발현이 유의하게 감소하는 것으로 나타났다(도 15A). 특히 HNF4 α 의 높은 비-핵 발현은 PROX1 또는 SREBP1이 KO일 때 관찰되었다. 세포질에서 HNF4 α 의 발현은 말기 간부전이 있는 인간 간세포에 대한 이전 연구에서 유사하게 확인되었다. 또한 HNF4 α 단독 또는 PROX1 또는 NR5A2 또는 NROB2 또는 MTF1 또는 SREBP1 또는 POM121C와 조합하여 HNF4 α 의 핵 발현을 유도하는 효과를 테스트하기 위해, 간 이식을 겪는 NASH로 인한 말기 간부전 환자의 외식된 간에서 분리된 인간 간세포에서 치료를 수행했다(도 16). HNF4 α -AAV 단독으로 처리한 지 96시간 후에 대조군(GFP-AAV)과 비교하여 HNF4 α 의 핵 발현이 약 1배 증가됨을 발견했다. 그러나 HNF4 α -AAV 치료가 PROX1-AAV 또는 NR5A2-AAV 또는 NROB2-AAV 또는 MTF1-AAV 또는 SREBP1-AAV 또는 POM121C-AAV와 조합되었을 때, 모든 조합은 HNF4 α 의 유의한 핵 발현을 유도했으며(도 16), 특히 조합이 HNF4 α + PROX1 또는 SREBP1과 관련된 경우 유의한 핵 발현을 유도했다.
- [0204] 따라서, 이 연구는 전사 인자 및 조절자 PROX1, NR5A2, NROB2, MTF1, SREBP1, EP300 및 POM121C가 말기 간부전이 있는 간경변 간세포에서 HNF4 α 의 핵 발현을 개선함을 입증한다. 더욱이, 이 결과는 HNF4 α 와 하나 이상의 전사 인자 및 조절자(PROX1 또는 NR5A2 또는 NROB2 또는 MTF1 또는 SREBP1 또는 EP300 및 POM121C)를 포함하는 모든 조합이 HNF4 α 의 핵 발현 및 말기 간 부전을 치료하는 HNF4 α 의 재프로그래밍 능력을 향상시킨다는 것을 보여준다.
- [0205] **참고문헌:**
- [0206] Tsochatzis EA, Bosch J, Burroughs AK. Liver cirrhosis. *Lancet* 2014; 383:1749-1761.
- [0207] Murphy SL, Xu J, Kochanek KD, Curtin SC, Arias E. Deaths: Final Data for 2015. In: *Prevention CfDCA*, editor.: National Vital Statistics Reports; 2017.
- [0208] Goldman L, Schafer AI. *Goldman-Cecil medicine*. 25th edition. ed. Philadelphia, PA: Elsevier/Saunders, 2016: 2 volumes (x1, 2722, 12108 pages).
- [0209] Lopez PM, Martin P. Update on liver transplantation: indications, organ allocation, and long-term care. *Mt Sinai J Med* 2006;73:1056-1066.
- [0210] Archambeaud I, Auble H, Nahon P, Planche L, Fallot G, Faroux R, Gournay J, et al. Risk factors for hepatocellular carcinoma in Caucasian patients with non-viral cirrhosis: the importance of prior obesity. *Liver Int* 2015;35:1872-1876.
- [0211] Donato F, Gelatti U, Limina RM, Fattovich G. Southern Europe as an example of interaction between various environmental factors: a systematic review of the epidemiologic evidence. *Oncogene* 2006;25:3756-3770.
- [0212] Gelatti U, Covolo L, Talamini R, Tagger A, Barbone F, Martelli C, Cremaschini F, et al. N-Acetyltransferase-2, glutathione S-transferase M1 and T1 genetic polymorphisms, cigarette smoking and hepatocellular carcinoma: a case-control study. *Int J Cancer* 2005;115:301-306.
- [0213] Kuper H, Tzonou A, Kaklamani E, Hsieh CC, Laggiou P, Adami HO, Trichopoulos D, et al. Tobacco smoking, alcohol consumption and their interaction in the causation of hepatocellular carcinoma. *Int J Cancer* 2000;85:498-502.
- [0214] Guzman-Lepe J, Cervantes-Alvarez E, Collin de l'Hortet A, Wang Y, Mars WM, Oda Y, Bekki Y, et al. Liver-enriched transcription factor expression relates to chronic hepatic failure in humans. *Hepatol Commun* 2018;2:582-594.
- [0215] Hernaez R, Sola E, Moreau R, Gines P. Acute-on-chronic liver failure: an update. *Gut* 2017;66:541-553.
- [0216] Lee YA, Wallace MC, Friedman SL. Pathobiology of liver fibrosis: a translational success story. *Gut* 2015;64:830-841.

- [0217] Pessayre D, Lebrec D, Descatoire V, Peignoux M, Benhamou JP. Mechanism for reduced drug clearance in patients with cirrhosis. *Gastroenterology* 1978;74:566-571.
- [0218] Cichoż-Lach H, Michalak A. Oxidative stress as a crucial factor in liver diseases. *World J Gastroenterol* 2014;20:8082-8091.
- [0219] Simoes ICM, Fontes A, Pinton P, Zischka H, Wieckowski MR. Mitochondria in non-alcoholic fatty liver disease. *Int J Biochem Cell Biol* 2018;95:93-99.
- [0220] Malhi H, Kaufman RJ. Endoplasmic reticulum stress in liver disease. *J Hepatol* 2011;54:795-809.
- [0221] Zhang XQ, Xu CF, Yu CH, Chen WX, Li YM. Role of endoplasmic reticulum stress in the pathogenesis of nonalcoholic fatty liver disease. *World J Gastroenterol* 2014;20:1768-1776.
- [0222] Wang K. Molecular mechanisms of hepatic apoptosis. *Cell Death Dis* 2014;5:e996.
- [0223] Seki E, Schwabe RF. Hepatic inflammation and fibrosis: functional links and key pathways. *Hepatology* 2015;61:1066-1079.
- [0224] Zhang BH, Weltman M, Farrell GC. Does steatohepatitis impair liver regeneration? A study in a dietary model of non-alcoholic steatohepatitis in rats. *J Gastroenterol Hepatol* 1999;14:133-137.
- [0225] Michalopoulos GK, Khan Z. Liver Stem Cells: Experimental Findings and Implications for Human Liver Disease. *Gastroenterology* 2015;149:876-882.
- [0226] Dubuquoy L, Louvet A, Lassailly G, Truant S, Boleslawski E, Artru F, Maggiorotto F, et al. Progenitor cell expansion and impaired hepatocyte regeneration in explanted livers from alcoholic hepatitis. *Gut* 2015;64:1949-1960.
- [0227] Nishikawa T, Bell A, Brooks JM, Setoyama K, Melis M, Han B, Fukumitsu K, et al. Resetting the transcription factor network reverses terminal chronic hepatic failure. *J Clin Invest* 2015;125:1533-1544.
- [0228] Babeu JP, Boudreau F. Hepatocyte nuclear factor 4-alpha involvement in liver and intestinal inflammatory networks. *World J Gastroenterol* 2014;20:22-30.
- [0229] Chellappa K, Jankova L, Schnabl JM, Pan S, Brelivet Y, Fung CL, Chan C, et al. Src tyrosine kinase phosphorylation of nuclear receptor HNF4alpha correlates with isoform-specific loss of HNF4alpha in human colon cancer. *Proc Natl Acad Sci U S A* 2012;109:2302-2307.
- [0230] Guo H, Gao C, Mi Z, Wai PY, Kuo PC. Phosphorylation of Ser158 regulates inflammatory redox-dependent hepatocyte nuclear factor-4a transcriptional activity. *Biochem J* 2014;461:347.
- [0231] Hong YH, Varanasi US, Yang W, Leff T. AMP-activated protein kinase regulates HNF4alpha transcriptional activity by inhibiting dimer formation and decreasing protein stability. *J Biol Chem* 2003;278:27495-27501.
- [0232] Lu H. Crosstalk of HNF4alpha with extracellular and intracellular signaling pathways in the regulation of hepatic metabolism of drugs and lipids. *Acta Pharm Sin B* 2016;6:393-408.
- [0233] Song Y, Zheng D, Zhao M, Qin Y, Wang T, Xing W, Gao L, et al. Thyroid-Stimulating Hormone Increases HNF-4alpha Phosphorylation via cAMP/PKA Pathway in the Liver. *Sci Rep* 2015;5:13409.
- [0234] Soutoglou E, Katrakili N, Talianidis I. Acetylation regulates transcription factor activity at multiple levels. *Mol Cell* 2000;5:745-751.
- [0235] Sun K, Montana V, Chellappa K, Brelivet Y, Moras D, Maeda Y, Parpura V, et al. Phosphorylation of a conserved serine in the deoxyribonucleic acid binding domain of nuclear receptors alters intracellular localization. *Mol Endocrinol* 2007;21:1297-1311.
- [0236] Xu Z, Tavares-Sanchez OL, Li Q, Fernando J, Rodriguez CM, Studer EJ, Pandak WM, et al. Activation of bile acid biosynthesis by the p38 mitogen-activated protein kinase (MAPK): hepatocyte nuclear factor-

4alpha phosphorylation by the p38 MAPK is required for cholesterol 7alpha-hydroxylase expression. *J Biol Chem* 2007;282:24607-24614.

- [0237] Yokoyama A, Katsura S, Ito R, Hashiba W, Sekine H, Fujiki R, Kato S. Multiple post-translational modifications in hepatocyte nuclear factor 4alpha. *Biochem Biophys Res Commun* 2011;410:749-753.
- [0238] Zhou W, Hannoun Z, Jaffray E, Medine CN, Black JR, Greenhough S, Zhu L, et al. SUMOylation of HNF4alpha regulates protein stability and hepatocyte function. *J Cell Sci* 2012;125:3630-3635.
- [0239] Gramignoli R, Green ML, Tahan V, Dorko K, Skvorak KJ, Marongiu F, Zao W, et al. Development and application of purified tissue dissociation enzyme mixtures for human hepatocyte isolation. *Cell Transplant* 2012;21:1245-1260.
- [0240] Bell AW, Michalopoulos GK. Phenobarbital regulates nuclear expression of HNF-4alpha in mouse and rat hepatocytes independent of CAR and PXR. *Hepatology* 2006;44:186-194.
- [0241] Natarajan A, Wagner B, Sibilia M. The EGF receptor is required for efficient liver regeneration. *Proc Natl Acad Sci U S A* 2007;104:17081-17086.
- [0242] Rasband WS. ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA,. In.
- [0243] Hainer SJ, Gu W, Carone BR, Landry BD, Rando OJ, Mello CC, Fazzio TG. Suppression of pervasive noncoding transcription in embryonic stem cells by esBAF. *Genes Dev* 2015;29:362-378.
- [0244] Kumar R, Ichihashi Y, Kimura S, Chitwood DH, Headland LR, Peng J, Maloof JN, et al. A High-Throughput Method for Illumina RNA-Seq Library Preparation. *Front Plant Sci* 2012;3:202.
- [0245] Morlan JD, Qu K, Sinicropi DV. Selective depletion of rRNA enables whole transcriptome profiling of archival fixed tissue. *PLoS One* 2012;7:e42882.
- [0246] Adiconis X, Borges-Rivera D, Satija R, DeLuca DS, Busby MA, Berlin AM, Sivachenko A, et al. Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nat Methods* 2013;10:623-629.
- [0247] de Hoon MJ, Imoto S, Nolan J, Miyano S. Open source clustering software. *Bioinformatics* 2004;20:1453-1454.
- [0248] Saldanha AJ. Java Treeview--extensible visualization of microarray data. *Bioinformatics* 2004;20:3246-3248.
- [0249] Kritis AA, Argyrokastritis A, Moschonas NK, Power S, Katrakili N, Zannis VI, Cereghini S, et al. Isolation and characterization of a third isoform of human hepatocyte nuclear factor4. *Gene* 1996;173:275-280.
- [0250] Tanaka T, Jiang S, Hotta H, Takano K, Iwanari H, Sumi K, Daigo K, et al. Dysregulated expression of P1 and P2 promoter-driven hepatocyte nuclear factor-4alpha in the pathogenesis of human cancer. *J Pathol* 2006;208:662-672.
- [0251] Walesky C, Apte U. Role of hepatocyte nuclear factor 4alpha (HNF4alpha) in cell proliferation and cancer. *Gene Expr* 2015;16:101-108.
- [0252] Liu L, Yannam GR, Nishikawa T, Yamamoto T, Basma H, Ito R, Nagaya M, et al. The microenvironment in hepatocyte regeneration and function in rats with advanced cirrhosis. *Hepatology* 2012;55:1529-1539.
- [0253] Komposch K, Sibilia M. EGFR Signaling in Liver Diseases. *Int J Mol Sci* 2015;17.
- [0254] Paranjpe S, Bowen WC, Mars WM, Orr A, Haynes MM, DeFrances MC, Liu S, et al. Combined systemic elimination of MET and epidermal growth factor receptor signaling completely abolishes liver regeneration and leads to liver decompensation. *Hepatology* 2016;64:1711-1724.
- [0255] Tsagianni A, Mars WM, Bhushan B, Bowen WC, Orr A, Stoops J, Paranjpe S, et al. Combined Systemic Disruption of MET and Epidermal Growth Factor Receptor Signaling Causes Liver Failure in Normal Mice.

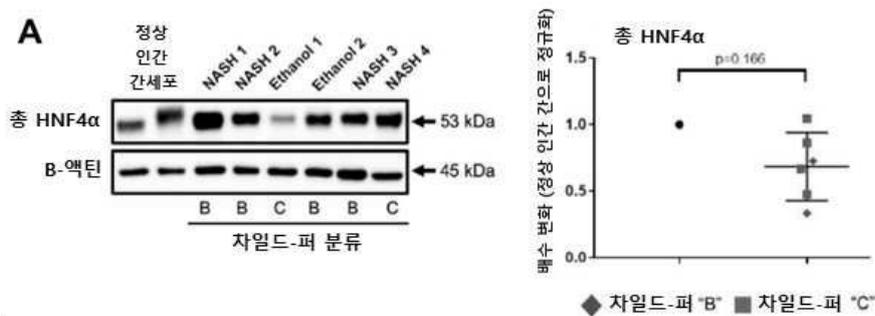
Am J Pathol 2018;188:2223-2235.

- [0256] Dekker FJ, Haisma HJ. Histone acetyl transferases as emerging drug targets. *Drug Discov Today* 2009;14:942-948.
- [0257] Argemi J, Latasa MU, Atkinson SR, Blokhin IO, Massey V, Gue JP, Cabezas J, et al. Defective HNF4alpha-dependent gene expression as a driver of hepatocellular failure in alcoholic hepatitis. *Nat Commun* 2019;10:3126.
- [0258] Woods A, Williams JR, Muckett PJ, Mayer FV, Liljevald M, Bohlooly YM, Carling D. Liver-Specific Activation of AMPK Prevents Steatosis on a High-Fructose Diet. *Cell Rep* 2017;18:3043-3051.
- [0259] Manning BD, Toker A. AKT/PKB Signaling: Navigating the Network. *Cell* 2017;169:381-405.
- [0260] Morales-Ruiz M, Santel A, Ribera J, Jimenez W. The Role of Akt in Chronic Liver Disease and Liver Regeneration. *Semin Liver Dis* 2017;37:11-16.
- [0261] Alam A, Chun Suen K, Ma D. Acute-on-chronic liver failure: recent update. *J Biomed Res* 2017;31:283-300.
- [0262] B. Sosa-Pineda, J.T. Wigle, G. Oliver Hepatocyte migration during liver development requires Prox1 *Nat. Genet.*, 25 (2000), pp. 254-255.
- [0263] Song KH, Li T and Chiang JY (2006) A prospero-related homeodomain protein is a novel co-regulator of hepatocyte nuclear factor 4alpha that regulates the cholesterol 7alpha-hydroxylase gene. *J Biol Chem* 281, 10081-10088.
- [0264] Xu SZ. Prox1 Facilitates Transfected CHO Cell Proliferation through Activation of the AKT Signaling Pathway. *Int J Biomed Sci.* 2010 Mar;6(1):49-59.
- [0265] Rausa FM, Galarneau L, Belanger L, Costa RH. 1999. The nuclear receptor fetoprotein transcription factor is coexpressed with its target HNF-3b in the developing murine liver intestine and pancreas. *Mech Dev* 89: 185-188.
- [0266] Labelle-Dumais C, Jacob-Wagner M, Pare JF, Belanger L, Dufort D. Nuclear receptor NR5A2 is required for proper primitive streak morphogenesis. *Dev Dyn.* 2006 Dec;235(12):3359-69.
- [0267] Kim DH, Kwon S, Byun S, Xiao Z, Park S, Wu SY, Chiang CM, Kemper B, Kemper JK. Critical role of RanBP2-mediated SUMOylation of Small Heterodimer Partner in maintaining bile acid homeostasis. *Nat Commun.* 2016 Jul 14;7:12179.
- [0268] Lindert U, Cramer M, Meuli M, Georgiev O, Schaffner W. Metal-responsive transcription factor 1 (MTF-1) activity is regulated by a nonconventional nuclear localization signal and a metal-responsive transactivation domain. *Mol Cell Biol.* 2009 Dec;29(23):6283-93. doi: 10.1128/MCB.00847-09.
- [0269] Rutherford JC, Bird AJ. Metal-responsive transcription factors that regulate iron, zinc, and copper homeostasis in eukaryotic cells. *Eukaryot Cell.* 2004 Feb;3(1):1-13.
- [0270] Shi Q, Hoffman B, Liu Q, PI3K-Akt signaling pathway upregulates hepatitis C virus RNA translation through the activation of SREBPs. *Virology.* 2016 Mar; 490():99-108.
- [0271] Porstmann T, Santos CR, Griffiths B, Cully M, Wu M, Leever S, Griffiths JR, Chung YL, Schulze A. SREBP activity is regulated by mTORC1 and contributes to Akt-dependent cell growth. *Cell Metab.* 2008 Sep;8(3):224-36.
- [0272] Breaux M, Lewis K, Valanejad L, Iakova P, Chen F, Mo Q, Medrano E, Timchenko L, Timchenko N. p300 Regulates Liver Functions by Controlling p53 and C/EBP Family Proteins through Multiple Signaling Pathways. *Mol Cell Biol.* 2015 Sep 1;35(17):3005-16.
- [0273] He L, Cao J, Meng S, Ma A, Radovick S, Wondisford FE, Activation of basal gluconeogenesis by coactivator p300 maintains hepatic glycogen storage. *Mol Endocrinol.* 2013 Aug; 27(8):1322-32.

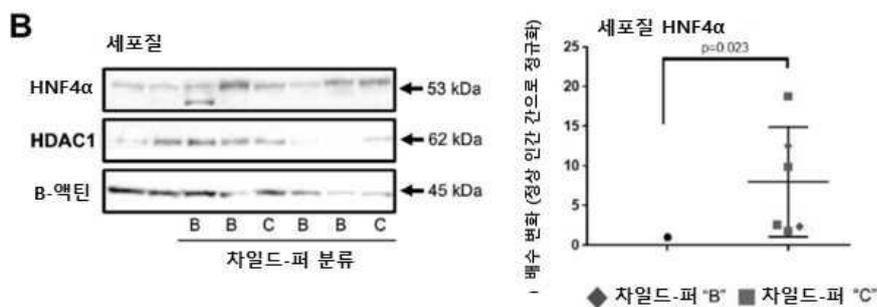
- [0274] Stavru F, Nautrup-Pedersen G, Cordes VC, Gørlich D. Nuclear pore complex assembly and maintenance in POM121- and gp210-deficient cells. *J Cell Biol.* 2006 May 22;173(4):477-83.
- [0275] Fusakio ME, Willy JA, Wang Y, Mirek ET, Al Baghdadi RJ, Adams CM, Anthony TG, Wek RC. Transcription factor ATF4 directs basal and stress-induced gene expression in the unfolded protein response and cholesterol metabolism in the liver. *Mol Biol Cell.* 2016 May 1;27(9):1536-51.
- [0276] Malhi H, Kaufman RJ. Endoplasmic reticulum stress in liver disease. *J Hepatol.* 2011 Apr;54(4):795-809.
- [0277] Florentino RM, Fraunhoffer NA, Morita K, Takeishi K, Ostrowska A, Achreja A, Animasahun O, Haep N, Arazov S, Agarwal N, Collin de l'Hortet A, Guzman-Lepe J, Tafaleng EN, Mukherjee A, Troy K, Banerjee S, Paranjpe S, Michalopoulos GK, Bell A, Nagrath D, Hainer SJ, Fox IJ, Soto-Gutierrez A. Cellular Location of HNF4 α is Linked With Terminal Liver Failure in Humans. *Hepatol Commun.* 2020 Apr 21;4(6):859-875.
- [0278] Guzman-Lepe J, Cervantes-Alvarez E, Collin de l'Hortet A, Wang Y, Mars WM, Oda Y, Bekki Y, Shimokawa M, Wang H, Yoshizumi T, Maehara Y, Bell A, Fox IJ, Takeishi K, Soto-Gutierrez A. Hepatol Commun. Liver-enriched transcription factor expression relates to chronic hepatic failure in humans. 2018 Mar 23;2(5):582-594.
- [0279] Nishikawa T, Bellance N, Damm A, Bing H, Zhu Z, Handa K, Yovchev MI, Sehgal V, Moss TJ, Oertel M, Ram PT, Pipinos II, Soto-Gutierrez A, Fox IJ, Nagrath D. A switch in the source of ATP production and a loss in capacity to perform glycolysis are hallmarks of hepatocyte failure in advance liver disease. *J Hepatol.* 2014 Jun;60(6):1203-11.
- [0280] Nishikawa T, Bell A, Brooks JM, Setoyama K, Melis M, Han B, Fukumitsu K, Handa K, Tian J, Kaestner KH, Vodovotz Y, Locker J, Soto-Gutierrez A, Fox IJ. Resetting the transcription factor network reverses terminal chronic hepatic failure. *J Clin Invest.* 2015 Apr;125(4):1533-44.

도면

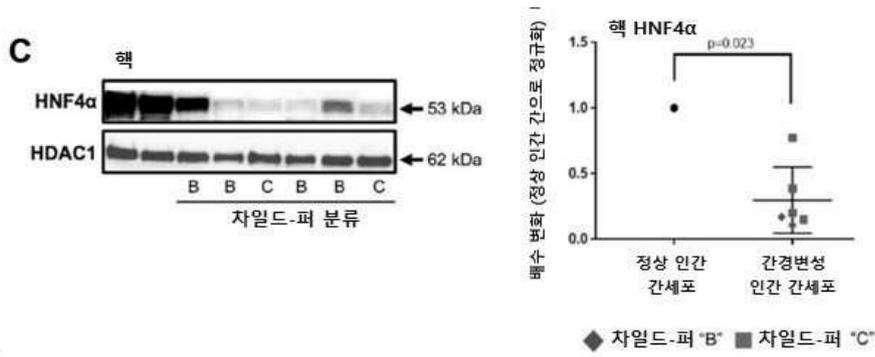
도면1a



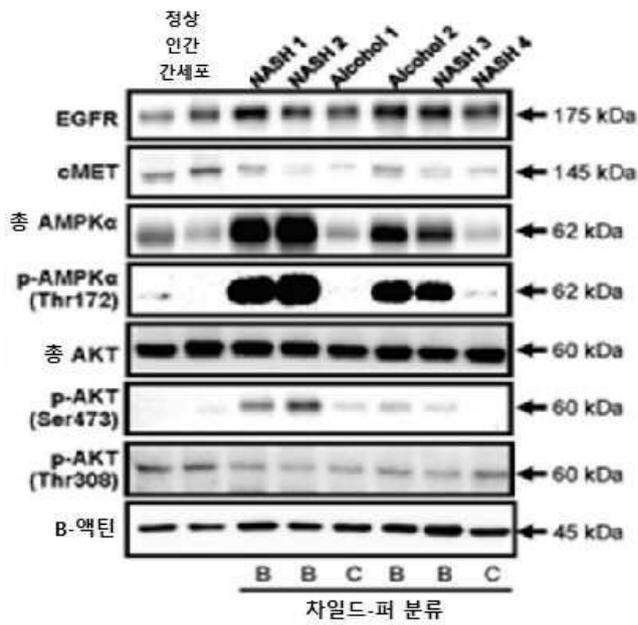
도면1b



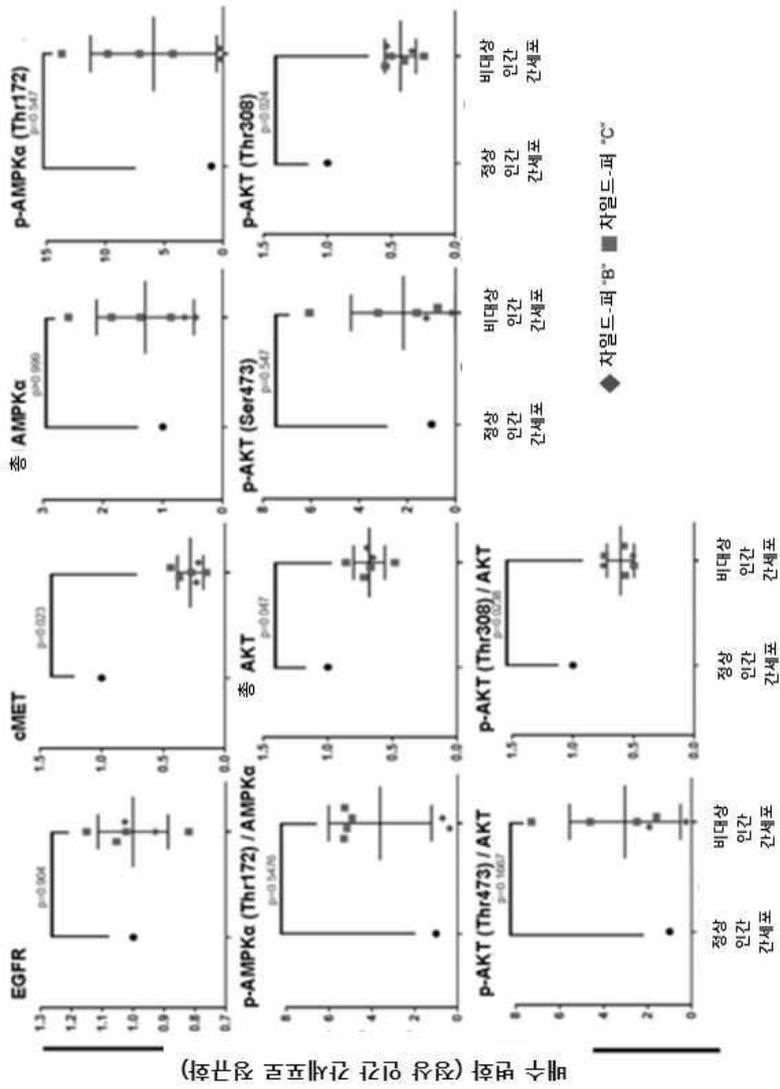
도면1c



도면2a



도면2b

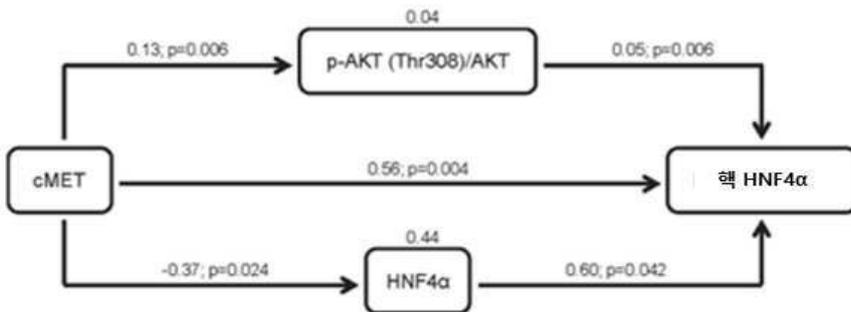


도면2c

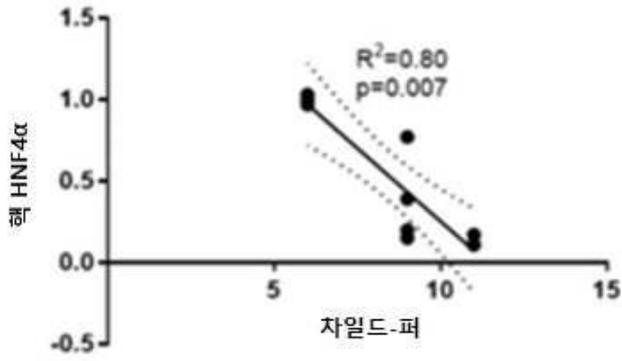
HNF4α에 대한 스피어만 상관관계

| 유전자 | HNF4α | | 세포질 HNF4α | | 핵 HNF4α | |
|-----------------------|-------------|--------------|--------------|--------------|-------------|--------------|
| | r | p-값 | r | p-값 | r | p-값 |
| HNF4α | | | | | | |
| Cytoplasmatic HNF4α | -0.67 | 0.059 | -0.67 | 0.059 | 0.67 | 0.059 |
| Nuclear HNF4α | 0.67 | 0.059 | -0.62 | 0.086 | -0.62 | 0.086 |
| EGFR | 0.43 | 0.250 | -0.12 | 0.776 | -0.27 | 0.493 |
| cMET | 0.76 | 0.021 | -0.80 | 0.014 | 0.71 | 0.037 |
| 총 AMPKα | 0.45 | 0.230 | -0.18 | 0.644 | 0.25 | 0.521 |
| p-AMPKα (Thr172) | 0.25 | 0.521 | 0.22 | 0.581 | 0.22 | 0.744 |
| p-AMPKα (Thr172)/AMPK | 0.25 | 0.521 | 0.22 | 0.581 | 0.13 | 0.744 |
| 총 AKT | 0.88 | 0.003 | -0.73 | 0.031 | 0.65 | 0.067 |
| p-AKT (Ser473) | 0.35 | 0.359 | -0.13 | 0.744 | -0.20 | 0.613 |
| p-AKT (Ser473)/AKT | 0.02 | 0.982 | 0.18 | 0.194 | -0.38 | 0.644 |
| p-AKT (Thr308) | 0.73 | 0.031 | -0.77 | 0.021 | 0.82 | 0.011 |
| p-AKT (Thr308)/AKT | 0.62 | 0.086 | -0.72 | 0.037 | 0.73 | 0.031 |

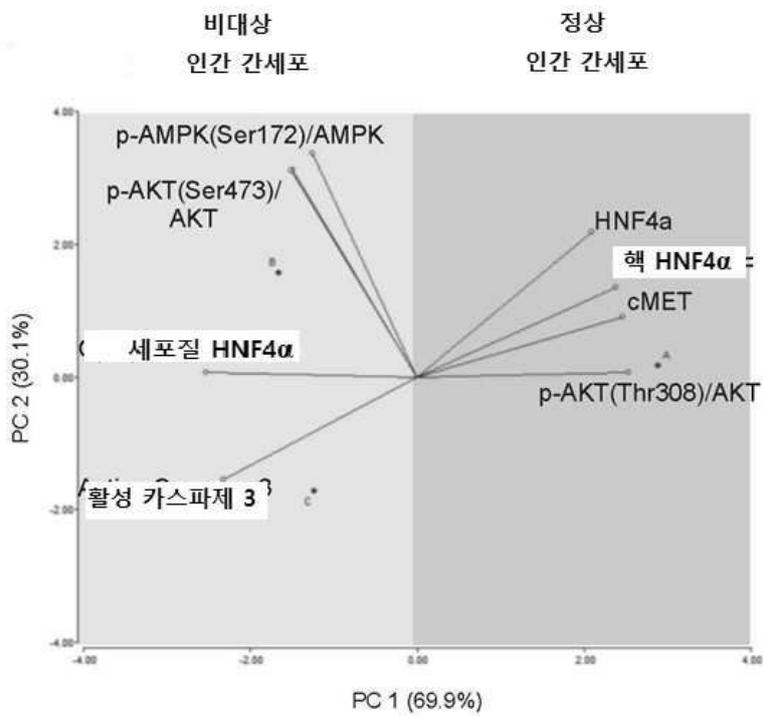
도면3a



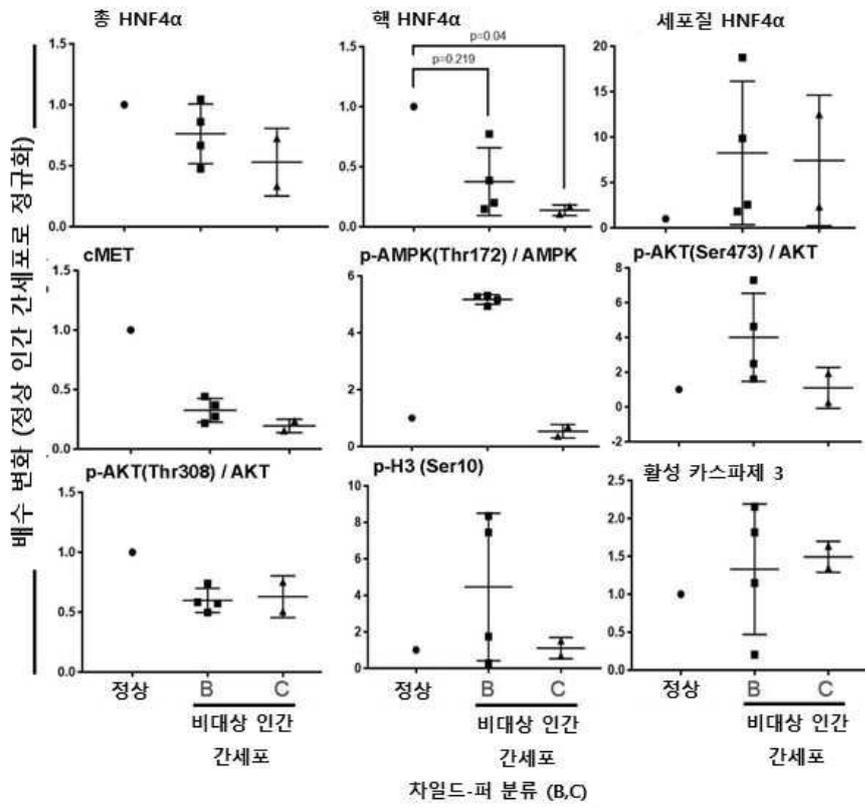
도면3b



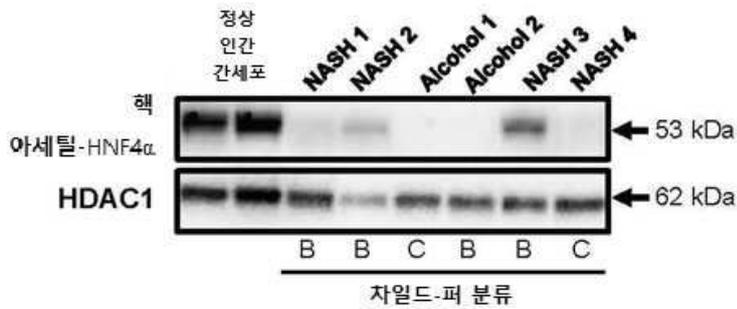
도면3c



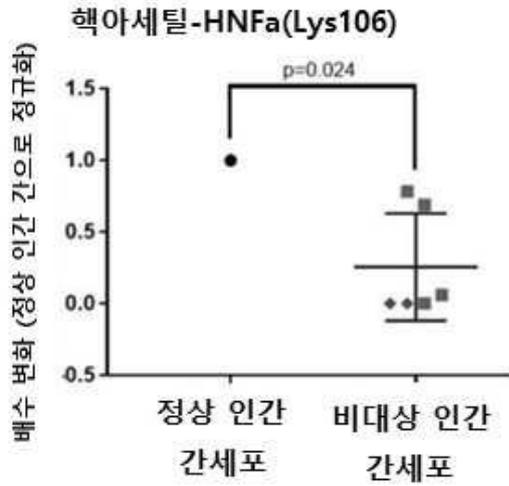
도면3d



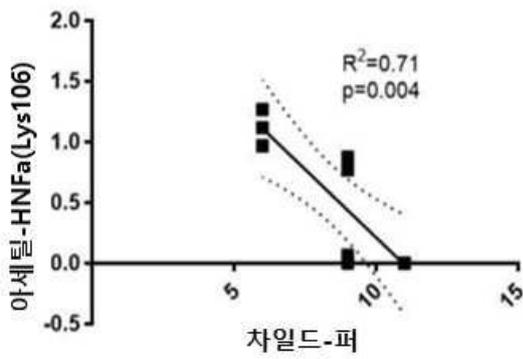
도면4a



도면4b



도면4c



도면5

| PTM | 국소화 (Localization) | 기능 | 참조 |
|---------------------------|-----------------------|----------------------|----------|
| 인산화 (phosphorylation) | Y14 | 감소된 전이활성 | 22308320 |
| | S78 | 감소된 전이활성 | 17389749 |
| | S87 | DNA 결합 감소 | 17389749 |
| | S133 | DNA 결합 감소 | 26302721 |
| | S134 | DNA 결합 감소 | 28196117 |
| | S142 | 전사 억제 | 28196117 |
| | S143 | 전사 억제 | 28196117 |
| | S158 | DNA 결합 및 전이활성 안정화 | 16351573 |
| | T166 | 전사 억제 | 21708125 |
| | S167 | 전사 억제 | 17603092 |
| | Y277 | DNA 결합 감소 | 22308320 |
| | Y279 | DNA 결합 감소 | 22308320 |
| | Y286 | 전사 억제 | 27709008 |
| | T288 | 전사 억제 | 22308320 |
| | S304 | 감소된 결합 및 이량체화 | 12740371 |
| S313 | 전사 억제 | 12740371 | |
| 아세틸화 (acetylation) | K106 | DNA 결합 안정화 | 10882110 |
| | K458 | 전사 억제 | 21708125 |
| 수모일화 (sumoylation) | K365 | 단백질 안정성 | 22505616 |
| | D367 | 단백질 안정성 | 22505616 |
| 유비퀴틴화 (ubiquitination) | K234 | 단백질 분해 | 25700366 |
| | K307 | 단백질 분해 | 25700366 |
| | K309 | 단백질 분해 | 25700366 |

도면6a

P-AKT(Thr308)/AKT에 대한 스피어만 상관관계

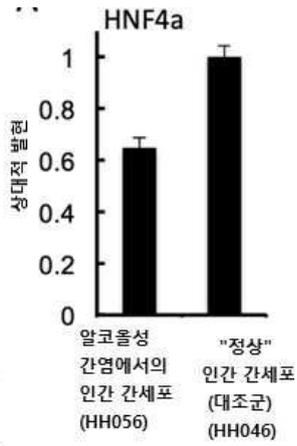
| Genes | p-AKT (Thr308)/AKT | |
|-----------------------|--------------------|--------------|
| | r | p-값 |
| EGFR | 0.02 | 0.982 |
| cMET | 0.80 | 0.014 |
| 총 AMPKα | 0.08 | 0.843 |
| p-AMPKα (Thr172) | -0.20 | 0.613 |
| p-AMPKα (Thr172)/AMPK | -0.20 | 0.613 |
| 총 AKT | 0.71 | 0.037 |
| p-AKT (Ser473) | -0.32 | 0.410 |
| p-AKT (Ser473)/AKT | -0.41 | 0.086 |

도면6b

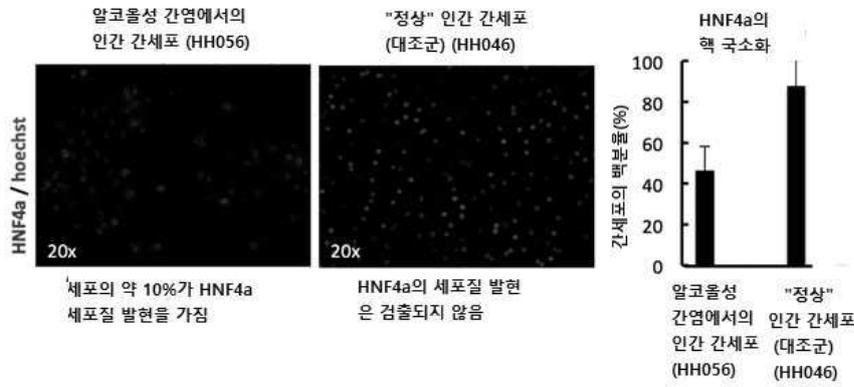
P-AKT(Ser473)/AKT에 대한 스피어만 상관관계

| Genes | p-AKT (Ser473)/AKT | |
|-----------------------|--------------------|-------------------|
| | r | p-값 |
| EGFR | 0.23 | 0.56 |
| cMET | -0.36 | 0.35 |
| 총 AMPKα | 0.93 | <0.0001 |
| p-AMPKα (Thr172) | 0.94 | <0.0001 |
| p-AMPKα (Thr172)/AMPK | 0.71 | 0.032 |
| 총 AKT | -0.04 | 0.919 |
| p-AKT (Thr308) | -0.32 | 0.394 |
| p-AKT (Thr308)/AKT | -0.41 | 0.268 |

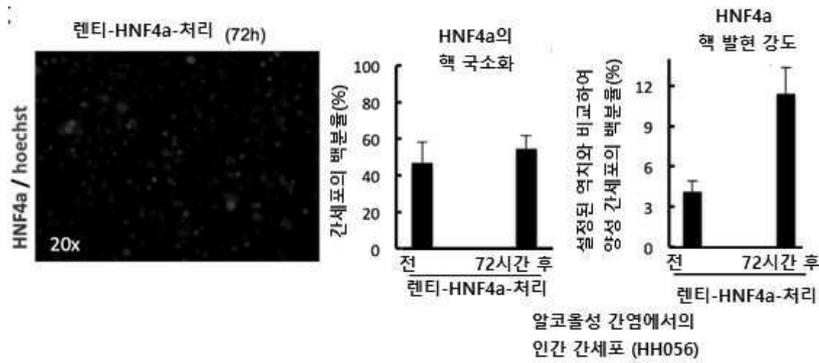
도면7a



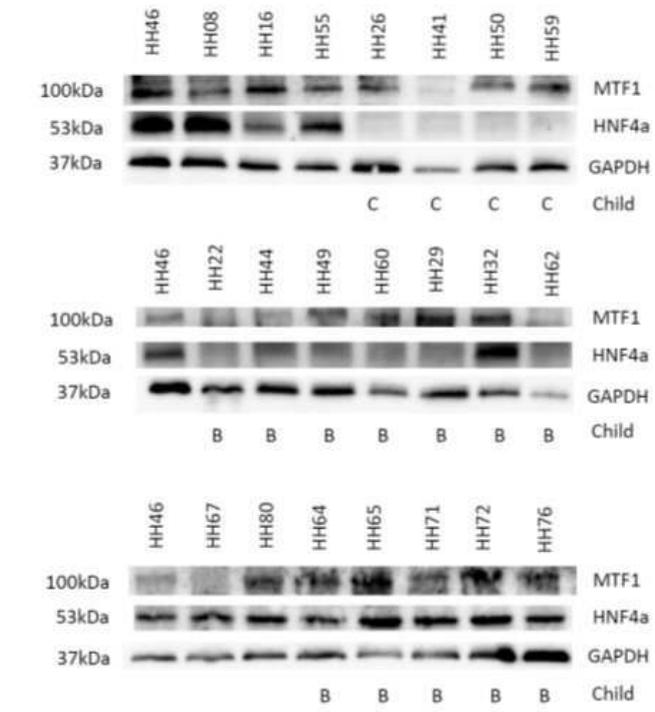
도면7b



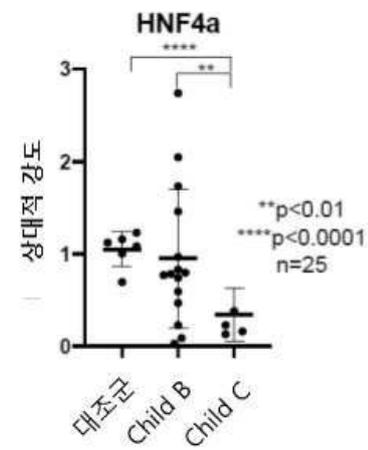
도면7c



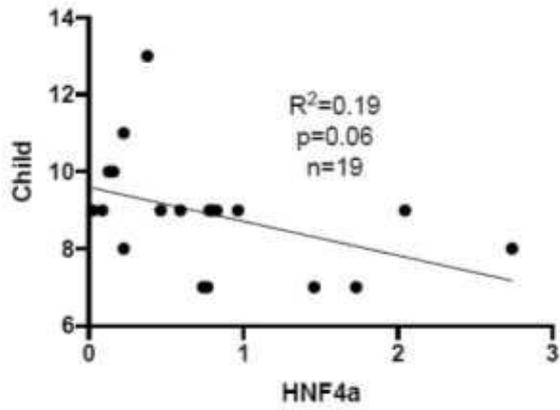
도면8a



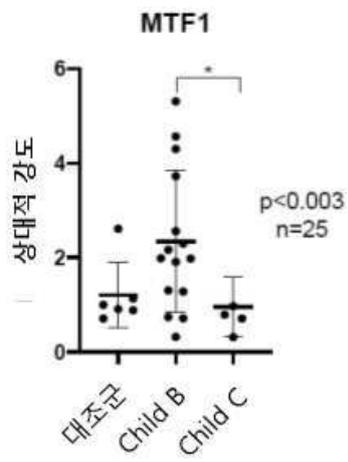
도면8b



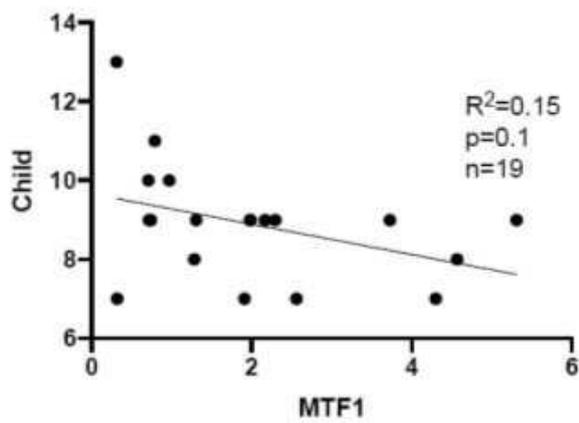
도면8c



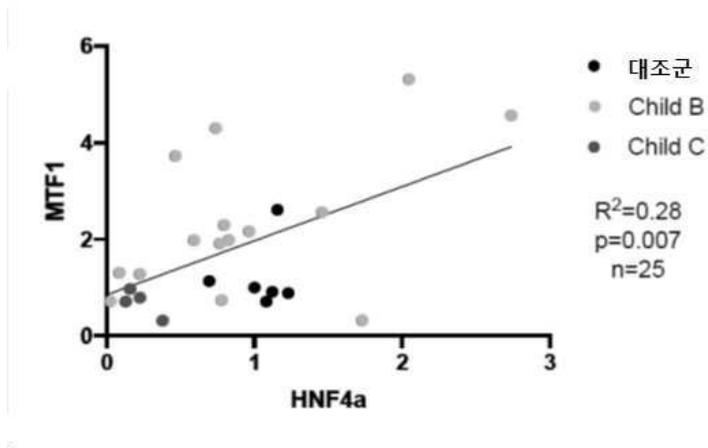
도면8d



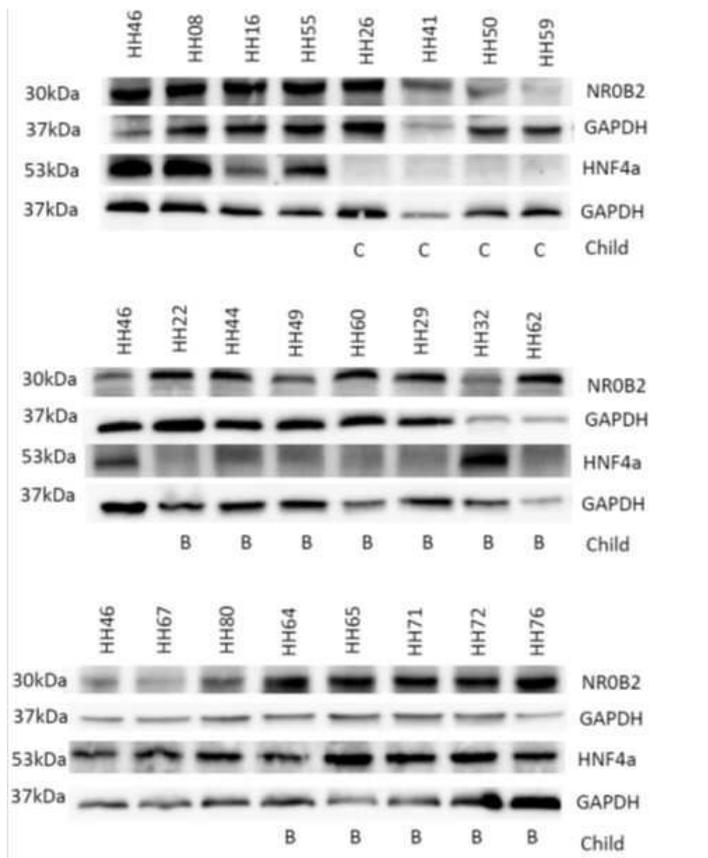
도면8e



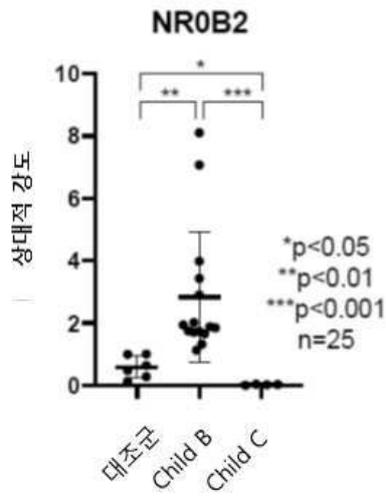
도면8f



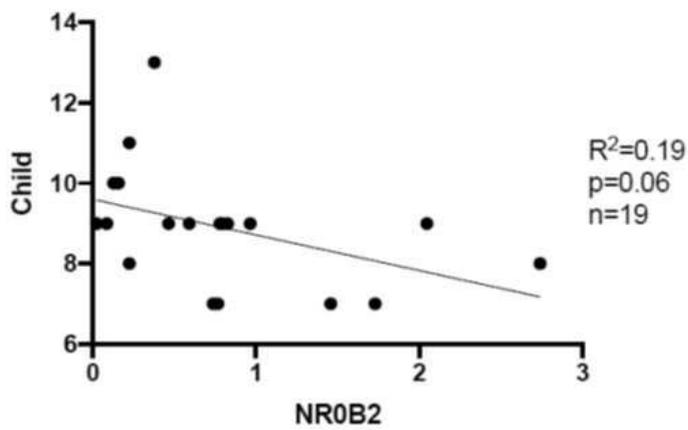
도면9a



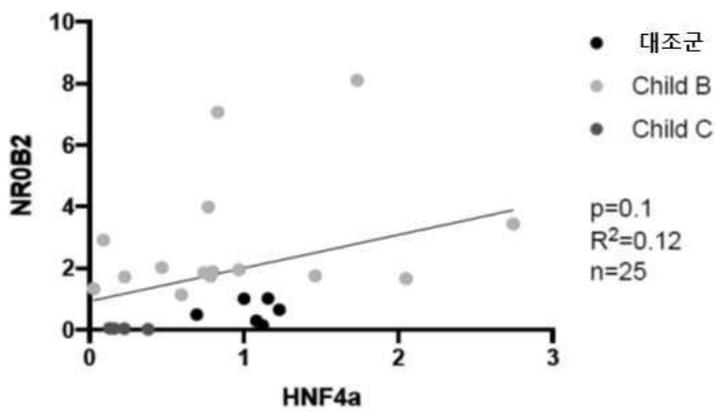
도면9b



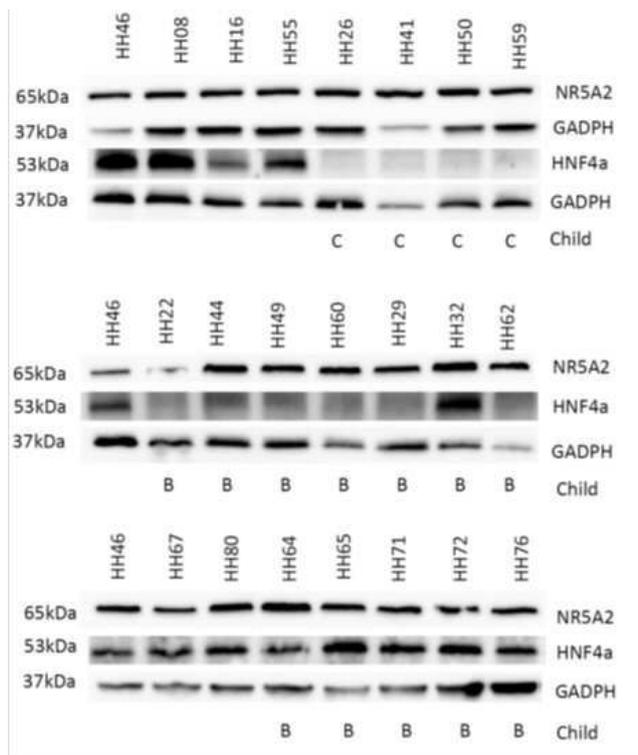
도면9c



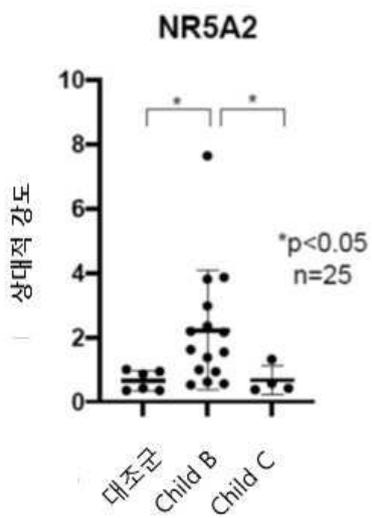
도면9d



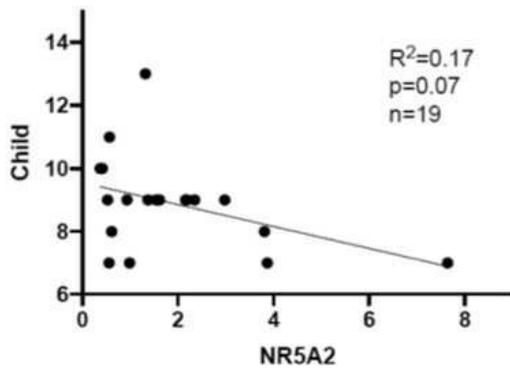
도면10a



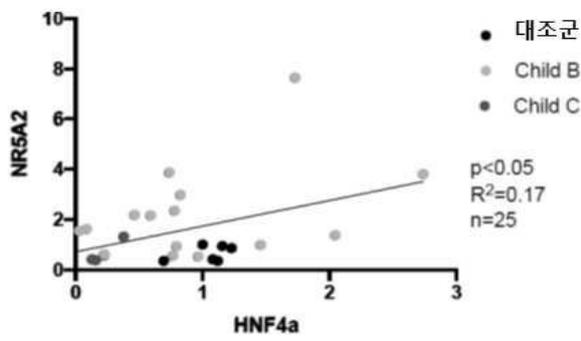
도면10b



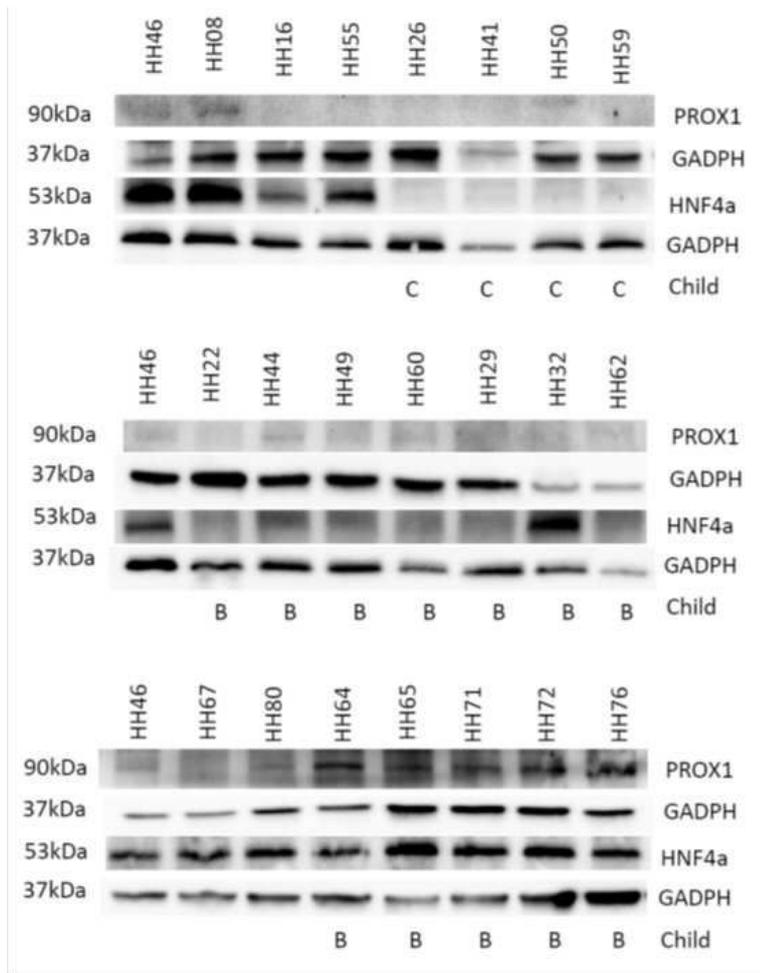
도면10c



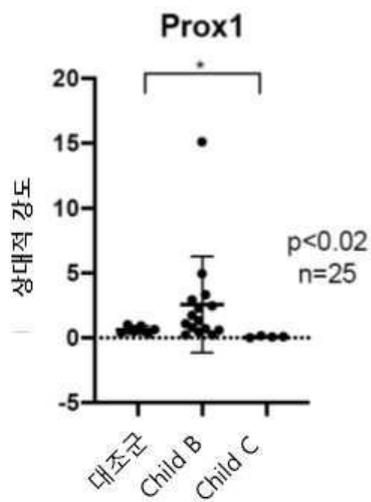
도면10d



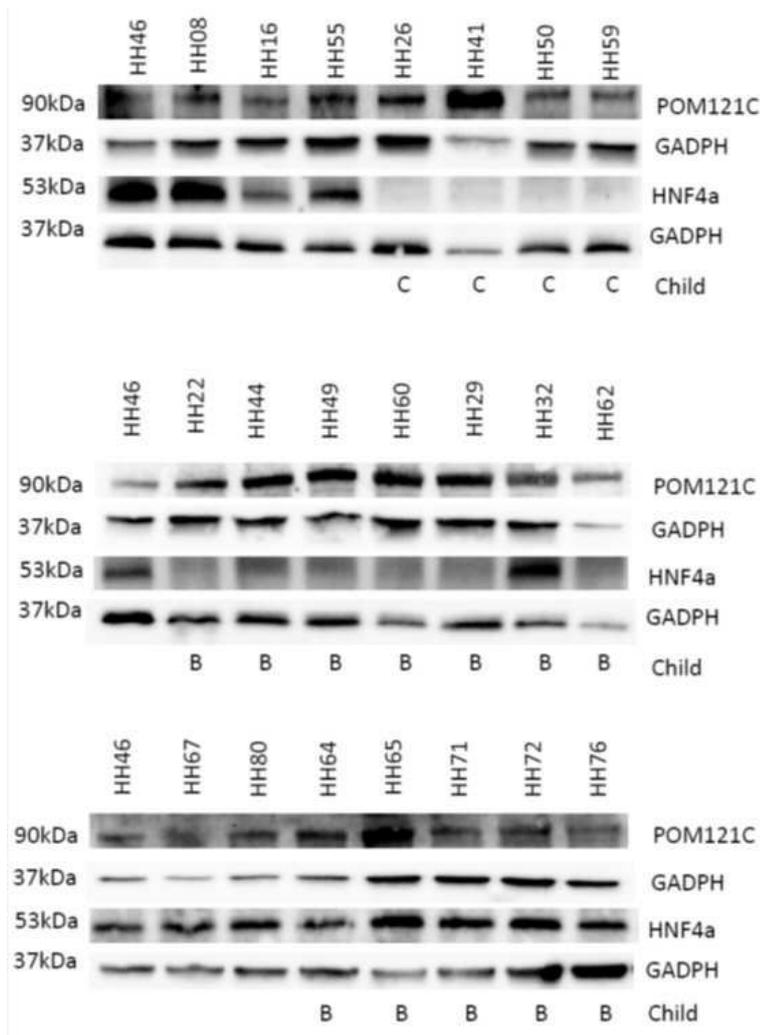
도면11a



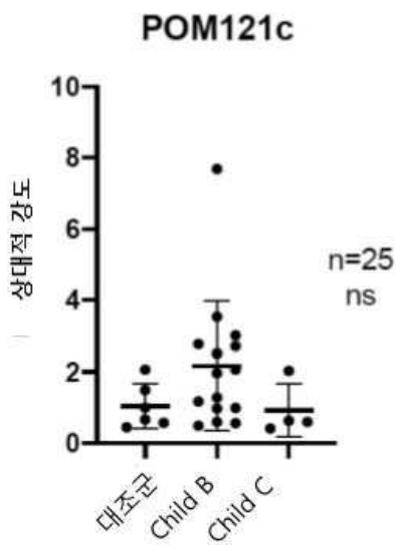
도면11b



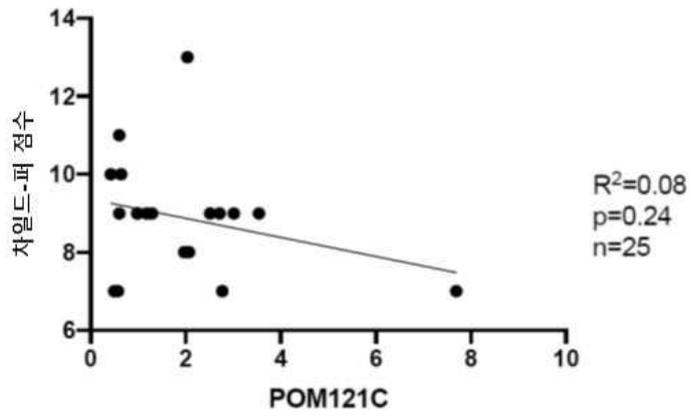
도면12a



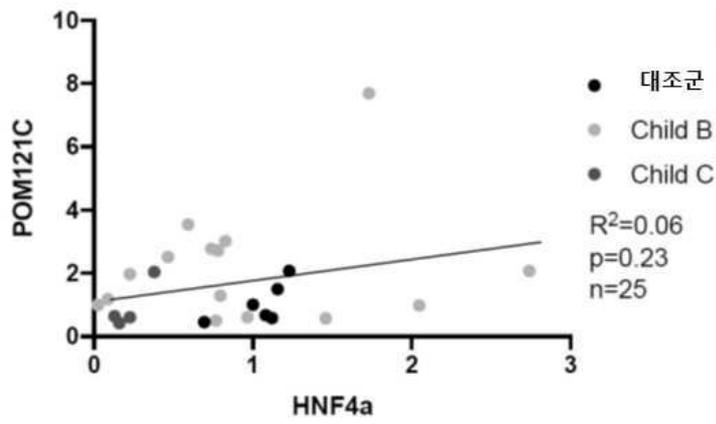
도면12b



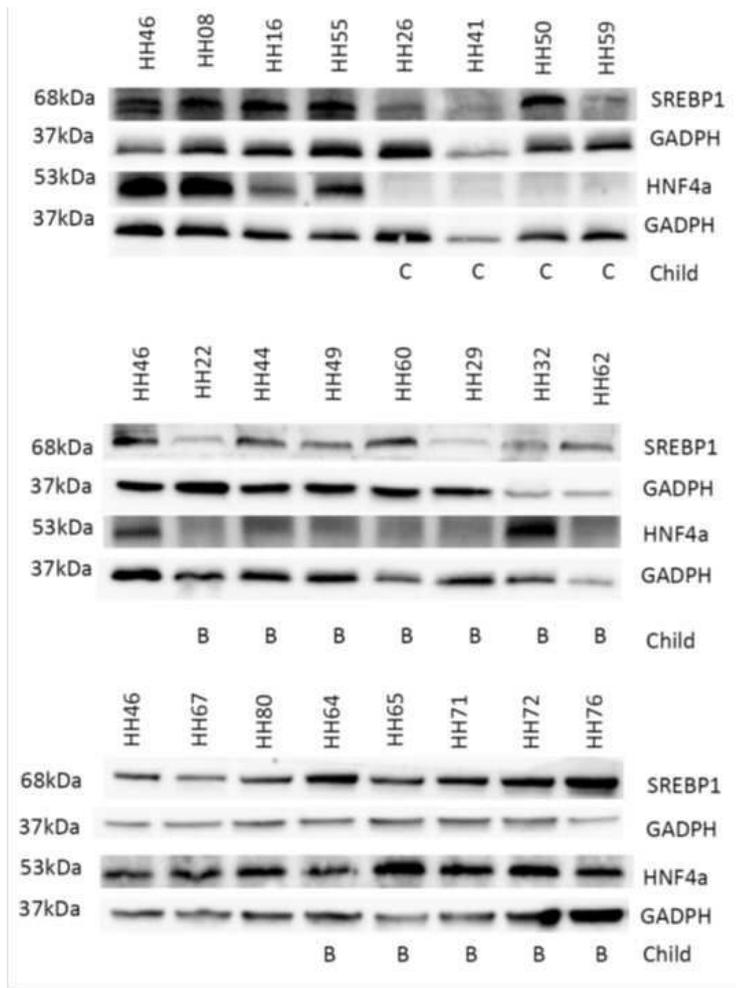
도면12c



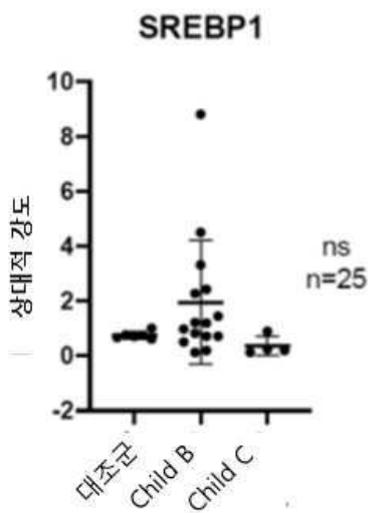
도면12d



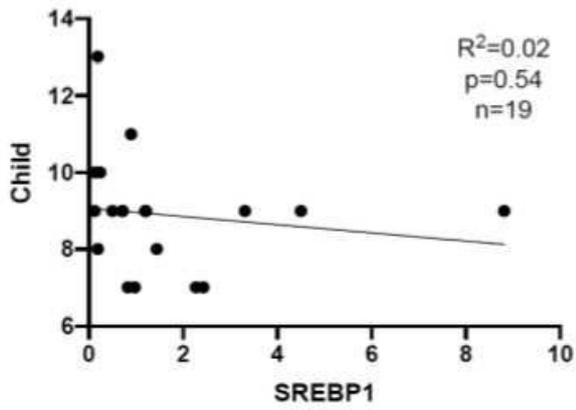
도면13a



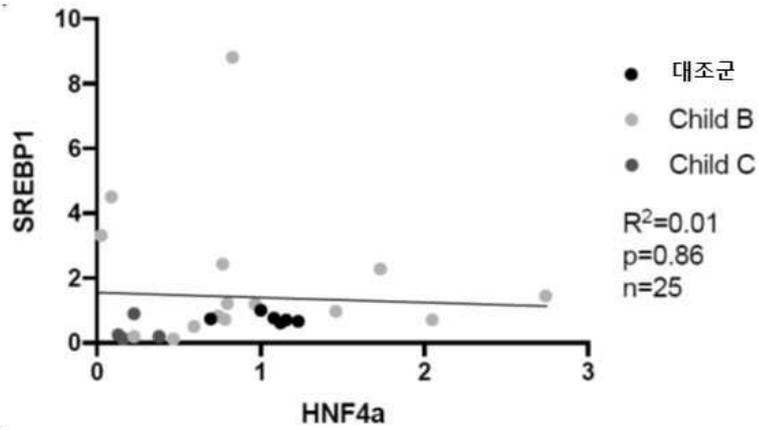
도면13b



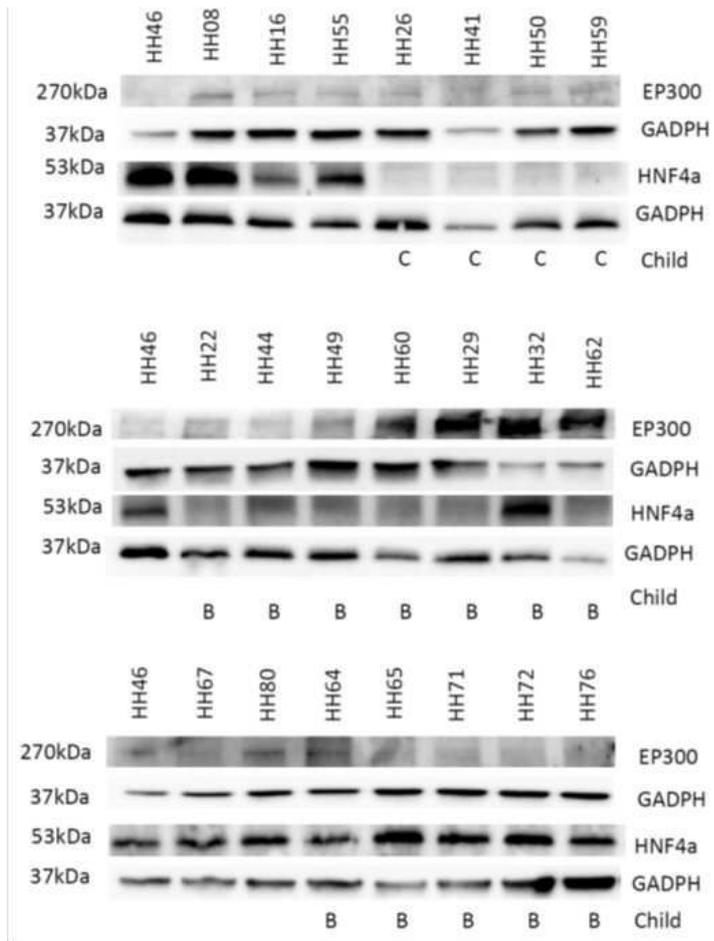
도면13c



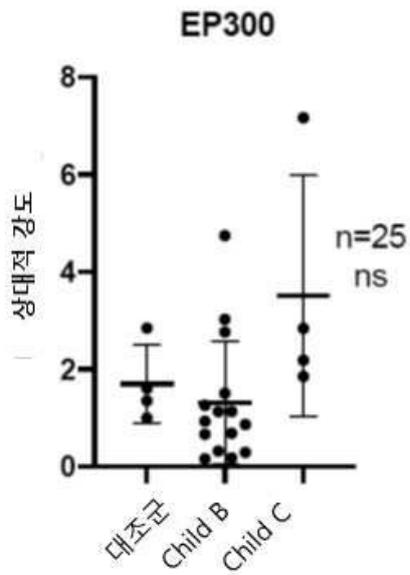
도면13d



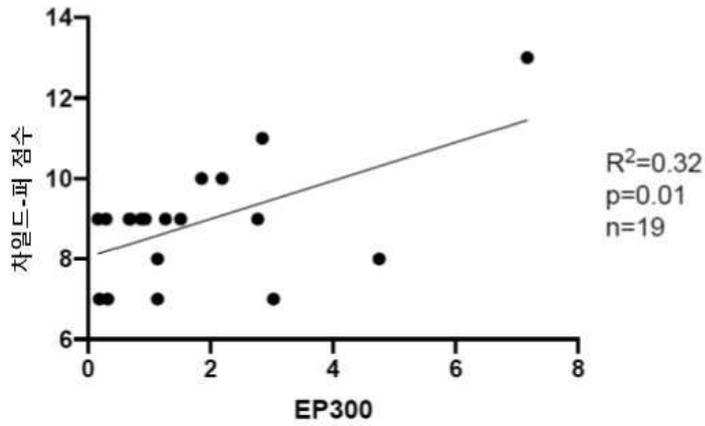
도면14a



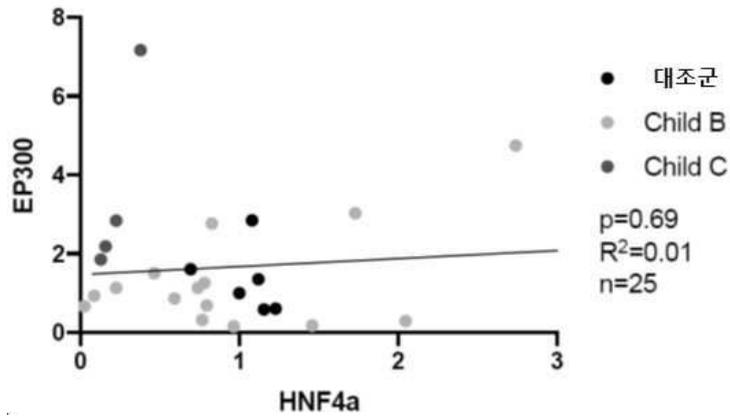
도면14b



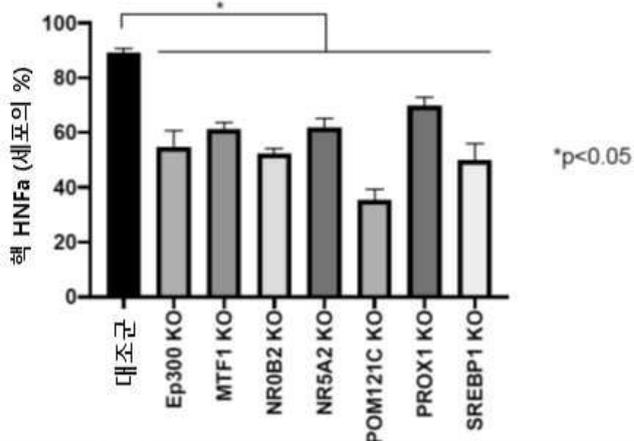
도면14c



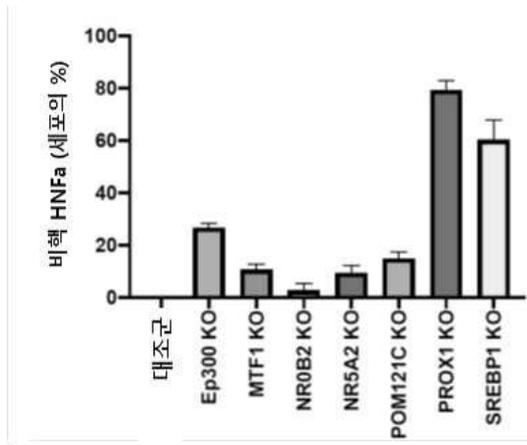
도면14d



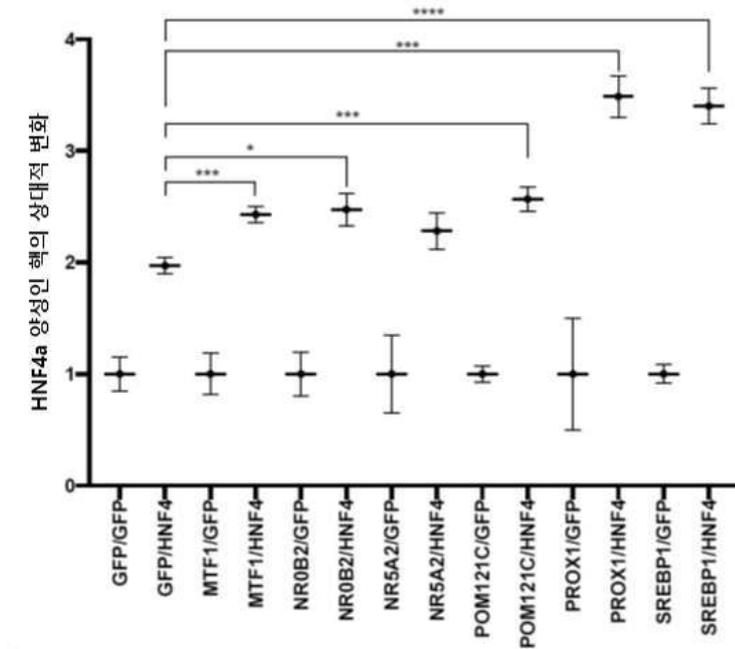
도면15a



도면15b



도면16



서열 목록

SEQUENCE LISTING

<110> UNIVERSITY OF PITTSBURGH-OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION

<120> COMPOSITIONS AND METHODS FOR TREATING LIVER DISEASE

<130> 10504-041W01

<160> 34

<170> PatentIn version 3.5

<210> 1

<211> 474

<212> PRT

<213> Homo sapiens

<400> 1

Met Arg Leu Ser Lys Thr Leu Val Asp Met Asp Met Ala Asp Tyr Ser

1 5 10 15

Ala Ala Leu Asp Pro Ala Tyr Thr Thr Leu Glu Phe Glu Asn Val Gln

 20 25 30

Val Leu Thr Met Gly Asn Asp Thr Ser Pro Ser Glu Gly Thr Asn Leu

 35 40 45

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

 50 55 60

Asp Arg Ala Thr Gly Lys His Tyr Gly Ala Ser Ser Cys Asp Gly Cys

65 70 75 80

Lys Gly Phe Phe Arg Arg Ser Val Arg Lys Asn His Met Tyr Ser Cys

 85 90 95

Arg Phe Ser Arg Gln Cys Val Val Asp Lys Asp Lys Arg Asn Gln Cys

 100 105 110

Arg Tyr Cys Arg Leu Lys Lys Cys Phe Arg Ala Gly Met Lys Lys Glu

 115 120 125

Ala Val Gln Asn Glu Arg Asp Arg Ile Ser Thr Arg Arg Ser Ser Tyr

 130 135 140

Glu Asp Ser Ser Leu Pro Ser Ile Asn Ala Leu Leu Gln Ala Glu Val

145 150 155 160

Leu Ser Arg Gln Ile Thr Ser Pro Val Ser Gly Ile Asn Gly Asp Ile

 165 170 175

Arg Ala Lys Lys Ile Ala Ser Ile Ala Asp Val Cys Glu Ser Met Lys

 180 185 190

Glu Gln Leu Leu Val Leu Val Glu Trp Ala Lys Tyr Ile Pro Ala Phe

 195 200 205

Cys Glu Leu Pro Leu Asp Asp Gln Val Ala Leu Leu Arg Ala His Ala

 210 215 220

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

 Leu Lys Ala Ile Ile Phe Phe Asp Pro Asp Ala Lys Gly Leu Ser Asp
 290 295 300
 Pro Gly Lys Ile Lys Arg Leu Arg Ser Gln Val Gln Val Ser Leu Glu
 305 310 315 320
 Asp Tyr Ile Asn Asp Arg Gln Tyr Asp Ser Arg Gly Arg Phe Gly Glu
 325 330 335
 Leu Leu Leu Leu Leu Pro Thr Leu Gln Ser Ile Thr Trp Gln Met Ile
 340 345 350

 Glu Gln Ile Gln Phe Ile Lys Leu Phe Gly Met Ala Lys Ile Asp Asn
 355 360 365
 Leu Leu Gln Glu Met Leu Leu Gly Gly Ser Pro Ser Asp Ala Pro His
 370 375 380
 Ala His His Pro Leu His Pro His Leu Met Gln Glu His Met Gly Thr
 385 390 395 400
 Asn Val Ile Val Ala Asn Thr Met Pro Thr His Leu Ser Asn Gly Gln
 405 410 415

 Met Cys Glu Trp Pro Arg Pro Arg Gly Gln Ala Ala Thr Pro Glu Thr
 420 425 430
 Pro Gln Pro Ser Pro Pro Gly Gly Ser Gly Ser Glu Pro Tyr Lys Leu
 435 440 445
 Leu Pro Gly Ala Val Ala Thr Ile Val Lys Pro Leu Ser Ala Ile Pro
 450 455 460
 Gln Pro Thr Ile Thr Lys Gln Glu Val Ile
 465 470

<210> 2

<211> 737

<212>

PRT

<213> Homo sapiens

<400> 2

Met Pro Asp His Asp Ser Thr Ala Leu Leu Ser Arg Gln Thr Lys Arg

1 5 10 15

Arg Arg Val Asp Ile Gly Val Lys Arg Thr Val Gly Thr Ala Ser Ala

 20 25 30

Phe Phe Ala Lys Ala Arg Ala Thr Phe Phe Ser Ala Met Asn Pro Gln

 35 40 45

Gly Ser Glu Gln Asp Val Glu Tyr Ser Val Val Gln His Ala Asp Gly

 50 55 60

Glu Lys Ser Asn Val Leu Arg Lys Leu Leu Lys Arg Ala Asn Ser Tyr

65 70 75 80

Glu Asp Ala Met Met Pro Phe Pro Gly Ala Thr Ile Ile Ser Gln Leu

 85 90 95

Leu Lys Asn Asn Met Asn Lys Asn Gly Gly Thr Glu Pro Ser Phe Gln

 100 105 110

Ala Ser Gly Leu Ser Ser Thr Gly Ser Glu Val His Gln Glu Asp Ile

 115 120 125

Cys Ser Asn Ser Ser Arg Asp Ser Pro Pro Glu Cys Leu Ser Pro Phe

 130 135 140

Gly Arg Pro Thr Met Ser Gln Phe Asp Met Asp Arg Leu Cys Asp Glu

145 150 155 160

His Leu Arg Ala Lys Arg Ala Arg Val Glu Asn Ile Ile Arg Gly Met

 165 170 175

Ser His Ser Pro Ser Val Ala Leu Arg Gly Asn Glu Asn Glu Arg Glu

 180 185 190

Met Ala Pro Gln Ser Val Ser Pro Arg Glu Ser Tyr Arg Glu Asn Lys

 195 200 205

Arg Lys Gln Lys Leu Pro Gln Gln Gln Gln Gln Ser Phe Gln Gln Leu
 210 215 220
 Val Ser Ala Arg Lys Glu Gln Lys Arg Glu Glu Arg Arg Gln Leu Lys
 225 230 235 240
 Gln Gln Leu Glu Asp Met Gln Lys Gln Leu Arg Gln Leu Gln Glu Lys
 245 250 255

 Phe Tyr Gln Ile Tyr Asp Ser Thr Asp Ser Glu Asn Asp Glu Asp Gly
 260 265 270
 Asn Leu Ser Glu Asp Ser Met Arg Ser Glu Ile Leu Asp Ala Arg Ala
 275 280 285
 Gln Asp Ser Val Gly Arg Ser Asp Asn Glu Met Cys Glu Leu Asp Pro
 290 295 300
 Gly Gln Phe Ile Asp Arg Ala Arg Ala Leu Ile Arg Glu Gln Glu Met
 305 310 315 320

 Ala Glu Asn Lys Pro Lys Arg Glu Gly Asn Asn Lys Glu Arg Asp His
 325 330 335
 Gly Pro Asn Ser Leu Gln Pro Glu Gly Lys His Leu Ala Glu Thr Leu
 340 345 350
 Lys Gln Glu Leu Asn Thr Ala Met Ser Gln Val Val Asp Thr Val Val
 355 360 365
 Lys Val Phe Ser Ala Lys Pro Ser Arg Gln Val Pro Gln Val Phe Pro
 370 375 380

 Pro Leu Gln Ile Pro Gln Ala Arg Phe Ala Val Asn Gly Glu Asn His
 385 390 395 400
 Asn Phe His Thr Ala Asn Gln Arg Leu Gln Cys Phe Gly Asp Val Ile
 405 410 415
 Ile Pro Asn Pro Leu Asp Thr Phe Gly Asn Val Gln Met Ala Ser Ser
 420 425 430
 Thr Asp Gln Thr Glu Ala Leu Pro Leu Val Val Arg Lys Asn Ser Ser
 435 440 445

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

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

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

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

 Gln Ala Ile Asn Asp Gly Val Thr Ser Thr Glu Glu Leu Ser Ile Thr
 645 650 655
 Arg Asp Cys Glu Leu Tyr Arg Ala Leu Asn Met His Tyr Asn Lys Ala
 660 665 670
 Asn Asp Phe Glu Val Pro Glu Arg Phe Leu Glu Val Ala Gln Ile Thr
 675 680 685
 Leu Arg Glu Phe Phe Asn Ala Ile Ile Ala Gly Lys Asp Val Asp Pro
 690 695 700

Ser Trp Lys Lys Ala Ile Tyr Lys Val Ile Cys Lys Leu Asp Ser Glu
 705 710 715 720
 Val Pro Glu Ile Phe Lys Ser Pro Asn Cys Leu Gln Glu Leu Leu His
 725 730 735
 Glu

<210> 3

<211> 541

<212> PRT

<213> Homo sapiens

<400> 3

Met Ser Ser Asn Ser Asp Thr Gly Asp Leu Gln Glu Ser Leu Lys His
 1 5 10 15
 Gly Leu Thr Pro Ile Gly Ala Gly Leu Pro Asp Arg His Gly Ser Pro
 20 25 30
 Ile Pro Ala Arg Gly Arg Leu Val Met Leu Pro Lys Val Glu Thr Glu
 35 40 45
 Ala Leu Gly Leu Ala Arg Ser His Gly Glu Gln Gly Gln Met Pro Glu
 50 55 60
 Asn Met Gln Val Ser Gln Phe Lys Met Val Asn Tyr Ser Tyr Asp Glu
 65 70 75 80
 Asp Leu Glu Glu Leu Cys Pro Val Cys Gly Asp Lys Val Ser Gly Tyr
 85 90 95
 His Tyr Gly Leu Leu Thr Cys Glu Ser Cys Lys Gly Phe Phe Lys Arg
 100 105 110
 Thr Val Gln Asn Asn Lys Arg Tyr Thr Cys Ile Glu Asn Gln Asn Cys
 115 120 125
 Gln Ile Asp Lys Thr Gln Arg Lys Arg Cys Pro Tyr Cys Arg Phe Gln
 130 135 140
 Lys Cys Leu Ser Val Gly Met Lys Leu Glu Ala Val Arg Ala Asp Arg
 145 150 155 160
 Met Arg Gly Gly Arg Asn Lys Phe Gly Pro Met Tyr Lys Arg Asp Arg

165 170 175
Ala Leu Lys Gln Gln Lys Lys Ala Leu Ile Arg Ala Asn Gly Leu Lys
180 185 190
Leu Glu Ala Met Ser Gln Val Ile Gln Ala Met Pro Ser Asp Leu Thr
195 200 205
Ile Ser Ser Ala Ile Gln Asn Ile His Ser Ala Ser Lys Gly Leu Pro
210 215 220
Leu Asn His Ala Ala Leu Pro Pro Thr Asp Tyr Asp Arg Ser Pro Phe
225 230 235 240
Val Thr Ser Pro Ile Ser Met Thr Met Pro Pro His Gly Ser Leu Gln
245 250 255
Gly Tyr Gln Thr Tyr Gly His Phe Pro Ser Arg Ala Ile Lys Ser Glu
260 265 270
Tyr Pro Asp Pro Tyr Thr Ser Ser Pro Glu Ser Ile Met Gly Tyr Ser
275 280 285
Tyr Met Asp Ser Tyr Gln Thr Ser Ser Pro Ala Ser Ile Pro His Leu
290 295 300
Ile Leu Glu Leu Leu Lys Cys Glu Pro Asp Glu Pro Gln Val Gln Ala
305 310 315 320
Lys Ile Met Ala Tyr Leu Gln Gln Glu Gln Ala Asn Arg Ser Lys His
325 330 335
Glu Lys Leu Ser Thr Phe Gly Leu Met Cys Lys Met Ala Asp Gln Thr
340 345 350
Leu Phe Ser Ile Val Glu Trp Ala Arg Ser Ser Ile Phe Phe Arg Glu
355 360 365
Leu Lys Val Asp Asp Gln Met Lys Leu Leu Gln Asn Cys Trp Ser Glu
370 375 380
Leu Leu Ile Leu Asp His Ile Tyr Arg Gln Val Val His Gly Lys Glu
385 390 395 400
Gly Ser Ile Phe Leu Val Thr Gly Gln Gln Val Asp Tyr Ser Ile Ile
405 410 415

Ala Ser Gln Ala Gly Ala Thr Leu Asn Asn Leu Met Ser His Ala Gln
 420 425 430

Glu Leu Val Ala Lys Leu Arg Ser Leu Gln Phe Asp Gln Arg Glu Phe
 435 440 445

Val Cys Leu Lys Phe Leu Val Leu Phe Ser Leu Asp Val Lys Asn Leu
 450 455 460

Glu Asn Phe Gln Leu Val Glu Gly Val Gln Glu Gln Val Asn Ala Ala
 465 470 475 480

Leu Leu Asp Tyr Thr Met Cys Asn Tyr Pro Gln Gln Thr Glu Lys Phe
 485 490 495

Gly Gln Leu Leu Leu Arg Leu Pro Glu Ile Arg Ala Ile Ser Met Gln
 500 505 510

Ala Glu Glu Tyr Leu Tyr Tyr Lys His Leu Asn Gly Asp Val Pro Tyr
 515 520 525

Asn Asn Leu Leu Ile Glu Met Leu His Ala Lys Arg Ala
 530 535 540

<210> 4
 <211> 257
 <212> PRT
 <213> Homo sapiens
 <400> 4

Met Ser Thr Ser Gln Pro Gly Ala Cys Pro Cys Gln Gly Ala Ala Ser
 1 5 10 15

Arg Pro Ala Ile Leu Tyr Ala Leu Leu Ser Ser Ser Leu Lys Ala Val
 20 25 30

Pro Arg Pro Arg Ser Arg Cys Leu Cys Arg Gln His Arg Pro Val Gln
 35 40 45

Leu Cys Ala Pro His Arg Thr Cys Arg Glu Ala Leu Asp Val Leu Ala
 50 55 60

Lys Thr Val Ala Phe Leu Arg Asn Leu Pro Ser Phe Trp Gln Leu Pro
 65 70 75 80

Pro Gln Asp Gln Arg Arg Leu Leu Gln Gly Cys Trp Gly Pro Leu Phe

85 90 95
 Leu Leu Gly Leu Ala Gln Asp Ala Val Thr Phe Glu Val Ala Glu Ala
 100 105 110

Pro Val Pro Ser Ile Leu Lys Lys Ile Leu Leu Glu Glu Pro Ser Ser
 115 120 125

Ser Gly Gly Ser Gly Gln Leu Pro Asp Arg Pro Gln Pro Ser Leu Ala
 130 135 140

Ala Val Gln Trp Leu Gln Cys Cys Leu Glu Ser Phe Trp Ser Leu Glu
 145 150 155 160

Leu Ser Pro Lys Glu Tyr Ala Cys Leu Lys Gly Thr Ile Leu Phe Asn
 165 170 175

Pro Asp Val Pro Gly Leu Gln Ala Ala Ser His Ile Gly His Leu Gln
 180 185 190

Gln Glu Ala His Trp Val Leu Cys Glu Val Leu Glu Pro Trp Cys Pro
 195 200 205

Ala Ala Gln Gly Arg Leu Thr Arg Val Leu Leu Thr Ala Ser Thr Leu
 210 215 220

Lys Ser Ile Pro Thr Ser Leu Leu Gly Asp Leu Phe Phe Arg Pro Ile
 225 230 235 240

Ile Gly Asp Val Asp Ile Ala Gly Leu Leu Gly Asp Met Leu Leu Leu
 245 250 255

Arg

<210> 5

<211> 753

<212> PRT

<213> Homo sapiens

<400> 5

Met Gly Glu His Ser Pro Asp Asn Asn Ile Ile Tyr Phe Glu Ala Glu
 1 5 10 15

Glu Asp Glu Leu Thr Pro Asp Asp Lys Met Leu Arg Phe Val Asp Lys
 20 25 30

Asn Gly Leu Val Pro Ser Ser Ser Gly Thr Val Tyr Asp Arg Thr Thr
 35 40 45
 Val Leu Ile Glu Gln Asp Pro Gly Thr Leu Glu Asp Glu Asp Asp Asp
 50 55 60
 Gly Gln Cys Gly Glu His Leu Pro Phe Leu Val Gly Gly Glu Glu Gly
 65 70 75 80
 Phe His Leu Ile Asp His Glu Ala Met Ser Gln Gly Tyr Val Gln His
 85 90 95
 Ile Ile Ser Pro Asp Gln Ile His Leu Thr Ile Asn Pro Gly Ser Thr
 100 105 110
 Pro Met Pro Arg Asn Ile Glu Gly Ala Thr Leu Thr Leu Gln Ser Glu
 115 120 125
 Cys Pro Glu Thr Lys Arg Lys Glu Val Lys Arg Tyr Gln Cys Thr Phe
 130 135 140
 Glu Gly Cys Pro Arg Thr Tyr Ser Thr Ala Gly Asn Leu Arg Thr His
 145 150 155 160
 Gln Lys Thr His Arg Gly Glu Tyr Thr Phe Val Cys Asn Gln Glu Gly
 165 170 175
 Cys Gly Lys Ala Phe Leu Thr Ser Tyr Ser Leu Arg Ile His Val Arg
 180 185 190
 Val His Thr Lys Glu Lys Pro Phe Glu Cys Asp Val Gln Gly Cys Glu
 195 200 205
 Lys Ala Phe Asn Thr Leu Tyr Arg Leu Lys Ala His Gln Arg Leu His
 210 215 220
 Thr Gly Lys Thr Phe Asn Cys Glu Ser Glu Gly Cys Ser Lys Tyr Phe
 225 230 235 240
 Thr Thr Leu Ser Asp Leu Arg Lys His Ile Arg Thr His Thr Gly Glu
 245 250 255
 Lys Pro Phe Arg Cys Asp His Asp Gly Cys Gly Lys Ala Phe Ala Ala
 260 265 270
 Ser His His Leu Lys Thr His Val Arg Thr His Thr Gly Glu Arg Pro

Leu Pro Ala Met Val Gln Thr Leu Pro Leu Gly Ala Asn Ser Val Leu
 530 535 540
 Thr Asn Asn Pro Thr Ile Thr Ile Thr Pro Thr Pro Asn Thr Ala Ile

 545 550 555 560
 Leu Gln Ser Ser Leu Val Met Gly Glu Gln Asn Leu Gln Trp Ile Leu
 565 570 575
 Asn Gly Ala Thr Ser Ser Pro Gln Asn Gln Glu Gln Ile Gln Gln Ala
 580 585 590
 Ser Lys Val Glu Lys Val Phe Phe Thr Thr Ala Val Pro Val Ala Ser
 595 600 605
 Ser Pro Gly Ser Ser Val Gln Gln Ile Gly Leu Ser Val Pro Val Ile

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

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

 740 745 750
 Lys

<210> 6

<211> 1147

<212> PRT

<213> Homo sapiens

<400> 6

Met Asp Glu Pro Pro Phe Ser Glu Ala Ala Leu Glu Gln Ala Leu Gly

1 5 10 15

Glu Pro Cys Asp Leu Asp Ala Ala Leu Leu Thr Asp Ile Glu Asp Met

 20 25 30

Leu Gln Leu Ile Asn Asn Gln Asp Ser Asp Phe Pro Gly Leu Phe Asp

 35 40 45

Pro Pro Tyr Ala Gly Ser Gly Ala Gly Gly Thr Asp Pro Ala Ser Pro

 50 55 60

Asp Thr Ser Ser Pro Gly Ser Leu Ser Pro Pro Pro Ala Thr Leu Ser

65 70 75 80

Ser Ser Leu Glu Ala Phe Leu Ser Gly Pro Gln Ala Ala Pro Ser Pro

 85 90 95

Leu Ser Pro Pro Gln Pro Ala Pro Thr Pro Leu Lys Met Tyr Pro Ser

 100 105 110

Met Pro Ala Phe Ser Pro Gly Pro Gly Ile Lys Glu Glu Ser Val Pro

 115 120 125

Leu Ser Ile Leu Gln Thr Pro Thr Pro Gln Pro Leu Pro Gly Ala Leu

 130 135 140

Leu Pro Gln Ser Phe Pro Ala Pro Ala Pro Pro Gln Phe Ser Ser Thr

145 150 155 160

Pro Val Leu Gly Tyr Pro Ser Pro Pro Gly Gly Phe Ser Thr Gly Ser

 165 170 175

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

 180 185 190

Pro Pro Gly Val Pro Pro Val Ser Leu His Thr Gln Val Gln Ser Val

 195 200 205

Val Pro Gln Gln Leu Leu Thr Val Thr Ala Ala Pro Thr Ala Ala Pro

 210 215 220

Val Thr Thr Thr Val Thr Ser Gln Ile Gln Gln Val Pro Val Leu Leu
 225 230 235 240

Gln Pro His Phe Ile Lys Ala Asp Ser Leu Leu Leu Thr Ala Met Lys
 245 250 255

Thr Asp Gly Ala Thr Val Lys Ala Ala Gly Leu Ser Pro Leu Val Ser
 260 265 270

Gly Thr Thr Val Gln Thr Gly Pro Leu Pro Thr Leu Val Ser Gly Gly
 275 280 285

Thr Ile Leu Ala Thr Val Pro Leu Val Val Asp Ala Glu Lys Leu Pro
 290 295 300

Ile Asn Arg Leu Ala Ala Gly Ser Lys Ala Pro Ala Ser Ala Gln Ser
 305 310 315 320

Arg Gly Glu Lys Arg Thr Ala His Asn Ala Ile Glu Lys Arg Tyr Arg
 325 330 335

Ser Ser Ile Asn Asp Lys Ile Ile Glu Leu Lys Asp Leu Val Val Gly
 340 345 350

Thr Glu Ala Lys Leu Asn Lys Ser Ala Val Leu Arg Lys Ala Ile Asp
 355 360 365

Tyr Ile Arg Phe Leu Gln His Ser Asn Gln Lys Leu Lys Gln Glu Asn
 370 375 380

Leu Ser Leu Arg Thr Ala Val His Lys Ser Lys Ser Leu Lys Asp Leu
 385 390 395 400

Val Ser Ala Cys Gly Ser Gly Gly Asn Thr Asp Val Leu Met Glu Gly
 405 410 415

Val Lys Thr Glu Val Glu Asp Thr Leu Thr Pro Pro Pro Ser Asp Ala
 420 425 430

Gly Ser Pro Phe Gln Ser Ser Pro Leu Ser Leu Gly Ser Arg Gly Ser
 435 440 445

Gly Ser Gly Gly Ser Gly Ser Asp Ser Glu Pro Asp Ser Pro Val Phe
 450 455 460

Glu Asp Ser Lys Ala Lys Pro Glu Gln Arg Pro Ser Leu His Ser Arg

465 470 475 480
 Gly Met Leu Asp Arg Ser Arg Leu Ala Leu Cys Thr Leu Val Phe Leu
 485 490 495

 Cys Leu Ser Cys Asn Pro Leu Ala Ser Leu Leu Gly Ala Arg Gly Leu
 500 505 510
 Pro Ser Pro Ser Asp Thr Thr Ser Val Tyr His Ser Pro Gly Arg Asn
 515 520 525
 Val Leu Gly Thr Glu Ser Arg Asp Gly Pro Gly Trp Ala Gln Trp Leu
 530 535 540
 Leu Pro Pro Val Val Trp Leu Leu Asn Gly Leu Leu Val Leu Val Ser
 545 550 555 560

 Leu Val Leu Leu Phe Val Tyr Gly Glu Pro Val Thr Arg Pro His Ser
 565 570 575
 Gly Pro Ala Val Tyr Phe Trp Arg His Arg Lys Gln Ala Asp Leu Asp
 580 585 590
 Leu Ala Arg Gly Asp Phe Ala Gln Ala Ala Gln Gln Leu Trp Leu Ala
 595 600 605
 Leu Arg Ala Leu Gly Arg Pro Leu Pro Thr Ser His Leu Asp Leu Ala
 610 615 620

 Cys Ser Leu Leu Trp Asn Leu Ile Arg His Leu Leu Gln Arg Leu Trp
 625 630 635 640
 Val Gly Arg Trp Leu Ala Gly Arg Ala Gly Gly Leu Gln Gln Asp Cys
 645 650 655
 Ala Leu Arg Val Asp Ala Ser Ala Ser Ala Arg Asp Ala Ala Leu Val
 660 665 670
 Tyr His Lys Leu His Gln Leu His Thr Met Gly Lys His Thr Gly Gly
 675 680 685

 His Leu Thr Ala Thr Asn Leu Ala Leu Ser Ala Leu Asn Leu Ala Glu
 690 695 700
 Cys Ala Gly Asp Ala Val Ser Val Ala Thr Leu Ala Glu Ile Tyr Val
 705 710 715 720

Ala Ala Ala Leu Arg Val Lys Thr Ser Leu Pro Arg Ala Leu His Phe
 725 730 735

Leu Thr Arg Phe Phe Leu Ser Ser Ala Arg Gln Ala Cys Leu Ala Gln
 740 745 750

Ser Gly Ser Val Pro Pro Ala Met Gln Trp Leu Cys His Pro Val Gly
 755 760 765

His Arg Phe Phe Val Asp Gly Asp Trp Ser Val Leu Ser Thr Pro Trp
 770 775 780

Glu Ser Leu Tyr Ser Leu Ala Gly Asn Pro Val Asp Pro Leu Ala Gln
 785 790 795 800

Val Thr Gln Leu Phe Arg Glu His Leu Leu Glu Arg Ala Leu Asn Cys
 805 810 815

Val Thr Gln Pro Asn Pro Ser Pro Gly Ser Ala Asp Gly Asp Lys Glu
 820 825 830

Phe Ser Asp Ala Leu Gly Tyr Leu Gln Leu Leu Asn Ser Cys Ser Asp
 835 840 845

Ala Ala Gly Ala Pro Ala Tyr Ser Phe Ser Ile Ser Ser Ser Met Ala
 850 855 860

Thr Thr Thr Gly Val Asp Pro Val Ala Lys Trp Trp Ala Ser Leu Thr
 865 870 875 880

Ala Val Val Ile His Trp Leu Arg Arg Asp Glu Glu Ala Ala Glu Arg
 885 890 895

Leu Cys Pro Leu Val Glu His Leu Pro Arg Val Leu Gln Glu Ser Glu
 900 905 910

Arg Pro Leu Pro Arg Ala Ala Leu His Ser Phe Lys Ala Ala Arg Ala
 915 920 925

Leu Leu Gly Cys Ala Lys Ala Glu Ser Gly Pro Ala Ser Leu Thr Ile
 930 935 940

Cys Glu Lys Ala Ser Gly Tyr Leu Gln Asp Ser Leu Ala Thr Thr Pro
 945 950 955 960

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

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

260 265 270
 Gln Ile Gln Thr Lys Thr Val Leu Ser Asn Asn Leu Ser Pro Phe Ala
 275 280 285
 Met Asp Lys Lys Ala Val Pro Gly Gly Gly Met Pro Asn Met Gly Gln

 290 295 300
 Gln Pro Ala Pro Gln Val Gln Gln Pro Gly Leu Val Thr Pro Val Ala
 305 310 315 320
 Gln Gly Met Gly Ser Gly Ala His Thr Ala Asp Pro Glu Lys Arg Lys
 325 330 335
 Leu Ile Gln Gln Gln Leu Val Leu Leu Leu His Ala His Lys Cys Gln
 340 345 350
 Arg Arg Glu Gln Ala Asn Gly Glu Val Arg Gln Cys Asn Leu Pro His

 355 360 365
 Cys Arg Thr Met Lys Asn Val Leu Asn His Met Thr His Cys Gln Ser
 370 375 380
 Gly Lys Ser Cys Gln Val Ala His Cys Ala Ser Ser Arg Gln Ile Ile
 385 390 395 400
 Ser His Trp Lys Asn Cys Thr Arg His Asp Cys Pro Val Cys Leu Pro
 405 410 415
 Leu Lys Asn Ala Gly Asp Lys Arg Asn Gln Gln Pro Ile Leu Thr Gly

 420 425 430
 Ala Pro Val Gly Leu Gly Asn Pro Ser Ser Leu Gly Val Gly Gln Gln
 435 440 445
 Ser Ala Pro Asn Leu Ser Thr Val Ser Gln Ile Asp Pro Ser Ser Ile
 450 455 460
 Glu Arg Ala Tyr Ala Ala Leu Gly Leu Pro Tyr Gln Val Asn Gln Met
 465 470 475 480
 Pro Thr Gln Pro Gln Val Gln Ala Lys Asn Gln Gln Asn Gln Gln Pro

 485 490 495
 Gly Gln Ser Pro Gln Gly Met Arg Pro Met Ser Asn Met Ser Ala Ser
 500 505 510

Pro Met Gly Val Asn Gly Gly Val Gly Val Gln Thr Pro Ser Leu Leu
 515 520 525
 Ser Asp Ser Met Leu His Ser Ala Ile Asn Ser Gln Asn Pro Met Met
 530 535 540
 Ser Glu Asn Ala Ser Val Pro Ser Leu Gly Pro Met Pro Thr Ala Ala
 545 550 555 560
 Gln Pro Ser Thr Thr Gly Ile Arg Lys Gln Trp His Glu Asp Ile Thr
 565 570 575
 Gln Asp Leu Arg Asn His Leu Val His Lys Leu Val Gln Ala Ile Phe
 580 585 590
 Pro Thr Pro Asp Pro Ala Ala Leu Lys Asp Arg Arg Met Glu Asn Leu
 595 600 605
 Val Ala Tyr Ala Arg Lys Val Glu Gly Asp Met Tyr Glu Ser Ala Asn
 610 615 620
 Asn Arg Ala Glu Tyr Tyr His Leu Leu Ala Glu Lys Ile Tyr Lys Ile
 625 630 635 640
 Gln Lys Glu Leu Glu Glu Lys Arg Arg Thr Arg Leu Gln Lys Gln Asn
 645 650 655
 Met Leu Pro Asn Ala Ala Gly Met Val Pro Val Ser Met Asn Pro Gly
 660 665 670
 Pro Asn Met Gly Gln Pro Gln Pro Gly Met Thr Ser Asn Gly Pro Leu
 675 680 685
 Pro Asp Pro Ser Met Ile Arg Gly Ser Val Pro Asn Gln Met Met Pro
 690 695 700
 Arg Ile Thr Pro Gln Ser Gly Leu Asn Gln Phe Gly Gln Met Ser Met
 705 710 715 720
 Ala Gln Pro Pro Ile Val Pro Arg Gln Thr Pro Pro Leu Gln His His
 725 730 735
 Gly Gln Leu Ala Gln Pro Gly Ala Leu Asn Pro Pro Met Gly Tyr Gly
 740 745 750
 Pro Arg Met Gln Gln Pro Ser Asn Gln Gly Gln Phe Leu Pro Gln Thr

Ser Thr Glu Thr Glu Glu Arg Ser Thr Glu Leu Lys Thr Glu Ile
 1010 1015 1020
 Lys Glu Glu Glu Asp Gln Pro Ser Thr Ser Ala Thr Gln Ser Ser
 1025 1030 1035
 Pro Ala Pro Gly Gln Ser Lys Lys Lys Ile Phe Lys Pro Glu Glu
 1040 1045 1050
 Leu Arg Gln Ala Leu Met Pro Thr Leu Glu Ala Leu Tyr Arg Gln
 1055 1060 1065
 Asp Pro Glu Ser Leu Pro Phe Arg Gln Pro Val Asp Pro Gln Leu
 1070 1075 1080
 Leu Gly Ile Pro Asp Tyr Phe Asp Ile Val Lys Ser Pro Met Asp
 1085 1090 1095
 Leu Ser Thr Ile Lys Arg Lys Leu Asp Thr Gly Gln Tyr Gln Glu
 1100 1105 1110
 Pro Trp Gln Tyr Val Asp Asp Ile Trp Leu Met Phe Asn Asn Ala
 1115 1120 1125
 Trp Leu Tyr Asn Arg Lys Thr Ser Arg Val Tyr Lys Tyr Cys Ser
 1130 1135 1140
 Lys Leu Ser Glu Val Phe Glu Gln Glu Ile Asp Pro Val Met Gln
 1145 1150 1155
 Ser Leu Gly Tyr Cys Cys Gly Arg Lys Leu Glu Phe Ser Pro Gln
 1160 1165 1170
 Thr Leu Cys Cys Tyr Gly Lys Gln Leu Cys Thr Ile Pro Arg Asp
 1175 1180 1185
 Ala Thr Tyr Tyr Ser Tyr Gln Asn Arg Tyr His Phe Cys Glu Lys
 1190 1195 1200
 Cys Phe Asn Glu Ile Gln Gly Glu Ser Val Ser Leu Gly Asp Asp
 1205 1210 1215
 Pro Ser Gln Pro Gln Thr Thr Ile Asn Lys Glu Gln Phe Ser Lys
 1220 1225 1230
 Arg Lys Asn Asp Thr Leu Asp Pro Glu Leu Phe Val Glu Cys Thr

| | | |
|---|------|------|
| 1235 | 1240 | 1245 |
| Glu Cys Gly Arg Lys Met His Gln Ile Cys Val Leu His His Glu | | |
| 1250 | 1255 | 1260 |
| Ile Ile Trp Pro Ala Gly Phe Val Cys Asp Gly Cys Leu Lys Lys | | |
| 1265 | 1270 | 1275 |
| Ser Ala Arg Thr Arg Lys Glu Asn Lys Phe Ser Ala Lys Arg Leu | | |
| 1280 | 1285 | 1290 |
| Pro Ser Thr Arg Leu Gly Thr Phe Leu Glu Asn Arg Val Asn Asp | | |
| 1295 | 1300 | 1305 |
| Phe Leu Arg Arg Gln Asn His Pro Glu Ser Gly Glu Val Thr Val | | |
| 1310 | 1315 | 1320 |
| Arg Val Val His Ala Ser Asp Lys Thr Val Glu Val Lys Pro Gly | | |
| 1325 | 1330 | 1335 |
| Met Lys Ala Arg Phe Val Asp Ser Gly Glu Met Ala Glu Ser Phe | | |
| 1340 | 1345 | 1350 |
| Pro Tyr Arg Thr Lys Ala Leu Phe Ala Phe Glu Glu Ile Asp Gly | | |
| 1355 | 1360 | 1365 |
| Val Asp Leu Cys Phe Phe Gly Met His Val Gln Glu Tyr Gly Ser | | |
| 1370 | 1375 | 1380 |
| Asp Cys Pro Pro Pro Asn Gln Arg Arg Val Tyr Ile Ser Tyr Leu | | |
| 1385 | 1390 | 1395 |
| Asp Ser Val His Phe Phe Arg Pro Lys Cys Leu Arg Thr Ala Val | | |
| 1400 | 1405 | 1410 |
| Tyr His Glu Ile Leu Ile Gly Tyr Leu Glu Tyr Val Lys Lys Leu | | |
| 1415 | 1420 | 1425 |
| Gly Tyr Thr Thr Gly His Ile Trp Ala Cys Pro Pro Ser Glu Gly | | |
| 1430 | 1435 | 1440 |
| Asp Asp Tyr Ile Phe His Cys His Pro Pro Asp Gln Lys Ile Pro | | |
| 1445 | 1450 | 1455 |
| Lys Pro Lys Arg Leu Gln Glu Trp Tyr Lys Lys Met Leu Asp Lys | | |
| 1460 | 1465 | 1470 |

Ala Val Ser Glu Arg Ile Val His Asp Tyr Lys Asp Ile Phe Lys
 1475 1480 1485
 Gln Ala Thr Glu Asp Arg Leu Thr Ser Ala Lys Glu Leu Pro Tyr
 1490 1495 1500
 Phe Glu Gly Asp Phe Trp Pro Asn Val Leu Glu Glu Ser Ile Lys
 1505 1510 1515
 Glu Leu Glu Gln Glu Glu Glu Glu Arg Lys Arg Glu Glu Asn Thr
 1520 1525 1530
 Ser Asn Glu Ser Thr Asp Val Thr Lys Gly Asp Ser Lys Asn Ala
 1535 1540 1545
 Lys Lys Lys Asn Asn Lys Lys Thr Ser Lys Asn Lys Ser Ser Leu
 1550 1555 1560
 Ser Arg Gly Asn Lys Lys Lys Pro Gly Met Pro Asn Val Ser Asn
 1565 1570 1575
 Asp Leu Ser Gln Lys Leu Tyr Ala Thr Met Glu Lys His Lys Glu
 1580 1585 1590
 Val Phe Phe Val Ile Arg Leu Ile Ala Gly Pro Ala Ala Asn Ser
 1595 1600 1605
 Leu Pro Pro Ile Val Asp Pro Asp Pro Leu Ile Pro Cys Asp Leu
 1610 1615 1620
 Met Asp Gly Arg Asp Ala Phe Leu Thr Leu Ala Arg Asp Lys His
 1625 1630 1635
 Leu Glu Phe Ser Ser Leu Arg Arg Ala Gln Trp Ser Thr Met Cys
 1640 1645 1650
 Met Leu Val Glu Leu His Thr Gln Ser Gln Asp Arg Phe Val Tyr
 1655 1660 1665
 Thr Cys Asn Glu Cys Lys His His Val Glu Thr Arg Trp His Cys
 1670 1675 1680
 Thr Val Cys Glu Asp Tyr Asp Leu Cys Ile Thr Cys Tyr Asn Thr
 1685 1690 1695
 Lys Asn His Asp His Lys Met Glu Lys Leu Gly Leu Gly Leu Asp

1700 1705 1710
 Asp Glu Ser Asn Asn Gln Gln Ala Ala Ala Thr Gln Ser Pro Gly

 1715 1720 1725
 Asp Ser Arg Arg Leu Ser Ile Gln Arg Cys Ile Gln Ser Leu Val
 1730 1735 1740
 His Ala Cys Gln Cys Arg Asn Ala Asn Cys Ser Leu Pro Ser Cys
 1745 1750 1755
 Gln Lys Met Lys Arg Val Val Gln His Thr Lys Gly Cys Lys Arg
 1760 1765 1770
 Lys Thr Asn Gly Gly Cys Pro Ile Cys Lys Gln Leu Ile Ala Leu

 1775 1780 1785
 Cys Cys Tyr His Ala Lys His Cys Gln Glu Asn Lys Cys Pro Val
 1790 1795 1800
 Pro Phe Cys Leu Asn Ile Lys Gln Lys Leu Arg Gln Gln Gln Leu
 1805 1810 1815
 Gln His Arg Leu Gln Gln Ala Gln Met Leu Arg Arg Arg Met Ala
 1820 1825 1830
 Ser Met Gln Arg Thr Gly Val Val Gly Gln Gln Gln Gly Leu Pro

 1835 1840 1845
 Ser Pro Thr Pro Ala Thr Pro Thr Thr Pro Thr Gly Gln Gln Pro
 1850 1855 1860
 Thr Thr Pro Gln Thr Pro Gln Pro Thr Ser Gln Pro Gln Pro Thr
 1865 1870 1875
 Pro Pro Asn Ser Met Pro Pro Tyr Leu Pro Arg Thr Gln Ala Ala
 1880 1885 1890
 Gly Pro Val Ser Gln Gly Lys Ala Ala Gly Gln Val Thr Pro Pro

 1895 1900 1905
 Thr Pro Pro Gln Thr Ala Gln Pro Pro Leu Pro Gly Pro Pro Pro
 1910 1915 1920
 Ala Ala Val Glu Met Ala Met Gln Ile Gln Arg Ala Ala Glu Thr
 1925 1930 1935

Gln Arg Gln Met Ala His Val Gln Ile Phe Gln Arg Pro Ile Gln
 1940 1945 1950

His Gln Met Pro Pro Met Thr Pro Met Ala Pro Met Gly Met Asn
 1955 1960 1965

Pro Pro Pro Met Thr Arg Gly Pro Ser Gly His Leu Glu Pro Gly
 1970 1975 1980

Met Gly Pro Thr Gly Met Gln Gln Gln Pro Pro Trp Ser Gln Gly
 1985 1990 1995

Gly Leu Pro Gln Pro Gln Gln Leu Gln Ser Gly Met Pro Arg Pro
 2000 2005 2010

Ala Met Met Ser Val Ala Gln His Gly Gln Pro Leu Asn Met Ala
 2015 2020 2025

Pro Gln Pro Gly Leu Gly Gln Val Gly Ile Ser Pro Leu Lys Pro
 2030 2035 2040

Gly Thr Val Ser Gln Gln Ala Leu Gln Asn Leu Leu Arg Thr Leu
 2045 2050 2055

Arg Ser Pro Ser Ser Pro Leu Gln Gln Gln Gln Val Leu Ser Ile
 2060 2065 2070

Leu His Ala Asn Pro Gln Leu Leu Ala Ala Phe Ile Lys Gln Arg
 2075 2080 2085

Ala Ala Lys Tyr Ala Asn Ser Asn Pro Gln Pro Ile Pro Gly Gln
 2090 2095 2100

Pro Gly Met Pro Gln Gly Gln Pro Gly Leu Gln Pro Pro Thr Met
 2105 2110 2115

Pro Gly Gln Gln Gly Val His Ser Asn Pro Ala Met Gln Asn Met
 2120 2125 2130

Asn Pro Met Gln Ala Gly Val Gln Arg Ala Gly Leu Pro Gln Gln
 2135 2140 2145

Gln Pro Gln Gln Gln Leu Gln Pro Pro Met Gly Gly Met Ser Pro
 2150 2155 2160

Gln Ala Gln Gln Met Asn Met Asn His Asn Thr Met Pro Ser Gln

| | | |
|-----------------------------|---------------------|-------------|
| 2165 | 2170 | 2175 |
| Phe Arg Asp Ile Leu Arg Arg | Gln Gln Met Met Gln | Gln Gln Gln |
| 2180 | 2185 | 2190 |
| Gln Gln Gly Ala Gly Pro Gly | Ile Gly Pro Gly Met | Ala Asn His |
| 2195 | 2200 | 2205 |
| Asn Gln Phe Gln Gln Pro Gln | Gly Val Gly Tyr Pro | Pro Gln Gln |
| 2210 | 2215 | 2220 |
| Gln Gln Arg Met Gln His His | Met Gln Gln Met Gln | Gln Gly Asn |
| 2225 | 2230 | 2235 |
| Met Gly Gln Ile Gly Gln Leu | Pro Gln Ala Leu Gly | Ala Glu Ala |
| 2240 | 2245 | 2250 |
| Gly Ala Ser Leu Gln Ala Tyr | Gln Gln Arg Leu Leu | Gln Gln Gln |
| 2255 | 2260 | 2265 |
| Met Gly Ser Pro Val Gln Pro | Asn Pro Met Ser Pro | Gln Gln His |
| 2270 | 2275 | 2280 |
| Met Leu Pro Asn Gln Ala Gln | Ser Pro His Leu Gln | Gly Gln Gln |
| 2285 | 2290 | 2295 |
| Ile Pro Asn Ser Leu Ser Asn | Gln Val Arg Ser Pro | Gln Pro Val |
| 2300 | 2305 | 2310 |
| Pro Ser Pro Arg Pro Gln Ser | Gln Pro Pro His Ser | Ser Pro Ser |
| 2315 | 2320 | 2325 |
| Pro Arg Met Gln Pro Gln Pro | Ser Pro His His Val | Ser Pro Gln |
| 2330 | 2335 | 2340 |
| Thr Ser Ser Pro His Pro Gly | Leu Val Ala Ala Gln | Ala Asn Pro |
| 2345 | 2350 | 2355 |
| Met Glu Gln Gly His Phe Ala | Ser Pro Asp Gln Asn | Ser Met Leu |
| 2360 | 2365 | 2370 |
| Ser Gln Leu Ala Ser Asn Pro | Gly Met Ala Asn Leu | His Gly Ala |
| 2375 | 2380 | 2385 |
| Ser Ala Thr Asp Leu Gly Leu | Ser Thr Asp Asn Ser | Asp Leu Asn |
| 2390 | 2395 | 2400 |

Ser Asn Leu Ser Gln Ser Thr Leu Asp Ile His
 2405 2410

<210> 8
 <211> 1229
 <212> PRT
 <213> Homo sapiens
 <400> 8

Met Ser Pro Ala Ala Ala Ala Ala Gly Ala Gly Glu Arg Arg Arg Pro
 1 5 10 15
 Ile Ala Ser Val Arg Asp Gly Arg Gly Arg Gly Cys Gly Gly Pro Ala
 20 25 30
 Gly Ala Ala Leu Leu Gly Leu Ser Leu Val Gly Leu Leu Leu Tyr Leu
 35 40 45
 Val Pro Ala Ala Ala Ala Leu Ala Trp Leu Ala Val Gly Thr Thr Ala
 50 55 60
 Ala Trp Trp Gly Leu Ser Arg Glu Pro Arg Gly Ser Arg Pro Leu Ser
 65 70 75 80
 Ser Phe Val Gln Lys Ala Arg His Arg Arg Thr Leu Phe Ala Ser Pro
 85 90 95
 Pro Ala Lys Ser Thr Ala Asn Gly Asn Leu Leu Glu Pro Arg Thr Leu
 100 105 110
 Leu Glu Gly Pro Asp Pro Ala Glu Leu Leu Leu Met Gly Ser Tyr Leu
 115 120 125
 Gly Lys Pro Gly Pro Pro Gln Pro Ala Pro Ala Pro Glu Gly Gln Asp
 130 135 140
 Leu Arg Asn Arg Pro Gly Arg Arg Pro Pro Ala Arg Pro Ala Pro Arg
 145 150 155 160
 Ser Thr Pro Pro Ser Gln Pro Thr His Arg Val His His Phe Tyr Pro
 165 170 175
 Ser Leu Pro Thr Pro Leu Leu Arg Pro Ser Gly Arg Pro Ser Pro Arg
 180 185 190
 Asp Arg Gly Thr Leu Pro Asp Arg Phe Val Ile Thr Pro Arg Arg Arg

Ala Asp Lys Glu Ser Gln Gly Glu Lys Ala Ala Asp Thr Thr Pro Arg
 450 455 460

Lys Lys Gln Asn Ser Asn Ser Gln Ser Thr Pro Gly Ser Ser Gly Gln

465 470 475 480

Arg Lys Arg Lys Val Gln Leu Leu Pro Ser Arg Arg Gly Glu Gln Leu
 485 490 495

Thr Leu Pro Pro Pro Gln Leu Gly Tyr Ser Ile Thr Ala Glu Asp
 500 505 510

Leu Asp Leu Glu Lys Lys Ala Ser Leu Gln Trp Phe Asn Gln Ala Leu
 515 520 525

Glu Asp Lys Ser Asp Ala Ala Ser Asn Ser Val Thr Glu Thr Pro Pro

530 535 540

Thr Thr Gln Pro Ser Phe Thr Phe Thr Leu Pro Ala Ala Ala Thr Ala
 545 550 555 560

Ser Pro Pro Thr Ser Leu Leu Ala Pro Ser Thr Asn Pro Leu Leu Glu
 565 570 575

Ser Leu Lys Lys Met Gln Thr Pro Pro Ser Leu Pro Pro Cys Pro Glu
 580 585 590

Ser Ala Gly Ala Ala Thr Thr Glu Ala Leu Ser Pro Pro Lys Thr Pro

595 600 605

Ser Leu Leu Pro Pro Leu Gly Leu Ser Gln Ser Gly Pro Pro Gly Leu
 610 615 620

Leu Pro Ser Pro Ser Phe Asp Ser Lys Pro Pro Thr Thr Leu Leu Gly
 625 630 635 640

Leu Ile Pro Ala Pro Ser Met Val Pro Ala Thr Asp Thr Lys Ala Pro
 645 650 655

Pro Thr Leu Gln Ala Glu Thr Ala Thr Lys Pro Gln Ala Thr Ser Ala

660 665 670

Pro Ser Pro Ala Pro Lys Gln Ser Phe Leu Phe Gly Thr Gln Asn Thr
 675 680 685

Ser Pro Ser Ser Pro Ala Ala Pro Ala Ala Ser Ser Ala Ser Pro Met

Pro Gln Pro Ala Phe Gly Ala Ala Glu Gly Gln Pro Pro Gly Ala Ala
 945 950 955 960
 Lys Pro Ala Leu Thr Pro Ser Phe Gly Ser Ser Phe Thr Phe Gly Asn
 965 970 975
 Ser Ala Ala Pro Ala Pro Ala Thr Ala Pro Thr Pro Ala Pro Ala Ser
 980 985 990
 Thr Ile Lys Ile Val Pro Ala His Val Pro Thr Pro Ile Gln Pro Thr
 995 1000 1005
 Phe Gly Gly Ala Thr His Ser Ala Phe Gly Leu Lys Ala Thr Ala
 1010 1015 1020
 Ser Ala Phe Gly Ala Pro Ala Ser Ser Gln Pro Ala Phe Gly Gly
 1025 1030 1035
 Ser Thr Ala Val Phe Ser Phe Gly Ala Ala Thr Ser Ser Gly Phe
 1040 1045 1050
 Gly Ala Thr Thr Gln Thr Ala Ser Ser Gly Ser Ser Ser Ser Val
 1055 1060 1065
 Phe Gly Ser Thr Thr Pro Ser Pro Phe Thr Phe Gly Gly Ser Ala
 1070 1075 1080
 Ala Pro Ala Gly Ser Gly Ser Phe Gly Ile Asn Val Ala Thr Pro
 1085 1090 1095
 Gly Ser Ser Ala Thr Thr Gly Ala Phe Ser Phe Gly Ala Gly Gln
 1100 1105 1110
 Ser Gly Ser Thr Ala Thr Ser Thr Pro Phe Thr Gly Gly Leu Gly
 1115 1120 1125
 Gln Asn Ala Leu Gly Thr Thr Gly Gln Ser Thr Pro Phe Ala Phe
 1130 1135 1140
 Asn Val Gly Ser Thr Thr Glu Ser Lys Pro Val Phe Gly Gly Thr
 1145 1150 1155
 Ala Thr Pro Thr Phe Gly Gln Asn Thr Pro Ala Pro Gly Val Gly
 1160 1165 1170
 Thr Ser Gly Ser Ser Leu Ser Phe Gly Ala Ser Ser Ala Pro Ala

1175 1180 1185
 Gln Gly Phe Val Gly Val Gly Pro Phe Gly Ser Ala Ala Pro Ser
 1190 1195 1200
 Phe Ser Ile Gly Ala Gly Ser Lys Thr Pro Gly Ala Arg Gln Arg
 1205 1210 1215
 Leu Gln Ala Arg Arg Gln His Thr Arg Lys Lys
 1220 1225

<210> 9

<211> 339

<212> PRT

<213> Homo sapiens

<400> 9

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

 50 55 60
 Arg Tyr Gly Glu Glu Gly Leu Lys Gly Ser Gly Pro Ser Gly Gly Ser
 65 70 75 80
 Gly Gly Gly Ala Asn Gly Thr Ser Phe Ser Tyr Thr Phe His Gly Asp
 85 90 95
 Pro His Ala Met Phe Ala Glu Phe Phe Gly Gly Arg Asn Pro Phe Asp
 100 105 110
 Thr Phe Phe Gly Gln Arg Asn Gly Glu Glu Gly Met Asp Ile Asp Asp

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

Pro Pro Val Thr His Asp Leu Arg Val Ser Leu Glu Glu Ile Tyr Ser
 165 170 175

Gly Cys Thr Lys Lys Met Lys Ile Ser His Lys Arg Leu Asn Pro Asp
 180 185 190

Gly Lys Ser Ile Arg Asn Glu Asp Lys Ile Leu Thr Ile Glu Val Lys
 195 200 205

Lys Gly Trp Lys Glu Gly Thr Lys Ile Thr Phe Pro Lys Glu Gly Asp
 210 215 220

Gln Thr Ser Asn Asn Ile Pro Ala Asp Ile Val Phe Val Leu Lys Asp
 225 230 235 240

Lys Pro His Asn Ile Phe Lys Arg Asp Gly Ser Asp Val Ile Tyr Pro
 245 250 255

Ala Arg Ile Ser Leu Arg Glu Ala Leu Cys Gly Cys Thr Val Asn Val
 260 265 270

Pro Thr Leu Asp Gly Arg Thr Ile Pro Val Val Phe Lys Asp Val Ile
 275 280 285

Arg Pro Gly Met Arg Arg Lys Val Pro Gly Glu Gly Leu Pro Leu Pro
 290 295 300

Lys Thr Pro Glu Lys Arg Gly Asp Leu Ile Ile Glu Phe Glu Val Ile
 305 310 315 320

Phe Pro Glu Arg Ile Pro Gln Thr Ser Arg Thr Val Leu Glu Gln Val
 325 330 335

Leu Pro Ile

<210> 10
 <211> 670
 <212> PRT
 <213> Homo sapiens
 <400> 10

Met Gly Glu Pro Ala Gly Val Ala Gly Thr Met Glu Ser Pro Phe Ser
 1 5 10 15

Pro Gly Leu Phe His Arg Leu Asp Glu Asp Trp Asp Ser Ala Leu Phe

Val Pro Ala Pro Ser Ala Asn Ser Pro Val Asn Gly Lys Leu Ser Val
 275 280 285

Thr Lys Pro Val Leu Gln Ser Thr Met Arg Asn Val Gly Ser Asp Ile
 290 295 300

Ala Val Leu Arg Arg Gln Gln Arg Met Ile Lys Asn Arg Glu Ser Ala
 305 310 315 320

Cys Gln Ser Arg Lys Lys Lys Lys Glu Tyr Met Leu Gly Leu Glu Ala
 325 330 335

Arg Leu Lys Ala Ala Leu Ser Glu Asn Glu Gln Leu Lys Lys Glu Asn
 340 345 350

Gly Thr Leu Lys Arg Gln Leu Asp Glu Val Val Ser Glu Asn Gln Arg
 355 360 365

Leu Lys Val Pro Ser Pro Lys Arg Arg Val Val Cys Val Met Ile Val
 370 375 380

Leu Ala Phe Ile Ile Leu Asn Tyr Gly Pro Met Ser Met Leu Glu Gln
 385 390 395 400

Asp Ser Arg Arg Met Asn Pro Ser Val Ser Pro Ala Asn Gln Arg Arg
 405 410 415

His Leu Leu Gly Phe Ser Ala Lys Glu Ala Gln Asp Thr Ser Asp Gly
 420 425 430

Ile Ile Gln Lys Asn Ser Tyr Arg Tyr Asp His Ser Val Ser Asn Asp
 435 440 445

Lys Ala Leu Met Val Leu Thr Glu Glu Pro Leu Leu Tyr Ile Pro Pro
 450 455 460

Pro Pro Cys Gln Pro Leu Ile Asn Thr Thr Glu Ser Leu Arg Leu Asn
 465 470 475 480

His Glu Leu Arg Gly Trp Val His Arg His Glu Val Glu Arg Thr Lys
 485 490 495

Ser Arg Arg Met Thr Asn Asn Gln Gln Lys Thr Arg Ile Leu Gln Gly
 500 505 510

Ala Leu Glu Gln Gly Ser Asn Ser Gln Leu Met Ala Val Gln Tyr Thr

50 55 60
 Asp Gly Leu Val Ser Pro Ser Asn Asn Ser Lys Glu Asp Ala Phe Ser
 65 70 75 80
 Gly Thr Asp Trp Met Leu Glu Lys Met Asp Leu Lys Glu Phe Asp Leu
 85 90 95
 Asp Ala Leu Leu Gly Ile Asp Asp Leu Glu Thr Met Pro Asp Asp Leu
 100 105 110
 Leu Thr Thr Leu Asp Asp Thr Cys Asp Leu Phe Ala Pro Leu Val Gln

 115 120 125
 Glu Thr Asn Lys Gln Pro Pro Gln Thr Val Asn Pro Ile Gly His Leu
 130 135 140
 Pro Glu Ser Leu Thr Lys Pro Asp Gln Val Ala Pro Phe Thr Phe Leu
 145 150 155 160
 Gln Pro Leu Pro Leu Ser Pro Gly Val Leu Ser Ser Thr Pro Asp His
 165 170 175
 Ser Phe Ser Leu Glu Leu Gly Ser Glu Val Asp Ile Thr Glu Gly Asp

 180 185 190
 Arg Lys Pro Asp Tyr Thr Ala Tyr Val Ala Met Ile Pro Gln Cys Ile
 195 200 205
 Lys Glu Glu Asp Thr Pro Ser Asp Asn Asp Ser Gly Ile Cys Met Ser
 210 215 220
 Pro Glu Ser Tyr Leu Gly Ser Pro Gln His Ser Pro Ser Thr Arg Gly
 225 230 235 240
 Ser Pro Asn Arg Ser Leu Pro Ser Pro Gly Val Leu Cys Gly Ser Ala

 245 250 255
 Arg Pro Lys Pro Tyr Asp Pro Pro Gly Glu Lys Met Val Ala Ala Lys
 260 265 270
 Val Lys Gly Glu Lys Leu Asp Lys Lys Leu Lys Lys Met Glu Gln Asn
 275 280 285
 Lys Thr Ala Ala Thr Arg Tyr Arg Gln Lys Lys Arg Ala Glu Gln Glu
 290 295 300

Ala Leu Thr Gly Glu Cys Lys Glu Leu Glu Lys Lys Asn Glu Ala Leu

305 310 315 320

Lys Glu Arg Ala Asp Ser Leu Ala Lys Glu Ile Gln Tyr Leu Lys Asp

325 330 335

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

340 345 350

<210> 12

<211> 1116

<212> PRT

<213> Homo sapiens

<400> 12

Met Glu Arg Ala Ile Ser Pro Gly Leu Leu Val Arg Ala Leu Leu Leu

1 5 10 15

Leu Leu Leu Leu Leu Gly Leu Ala Ala Arg Thr Val Ala Ala Gly Arg

20 25 30

Ala Arg Gly Leu Pro Ala Pro Thr Ala Glu Ala Ala Phe Gly Leu Gly

35 40 45

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

50 55 60

Val Ala Ala Ala Glu Val Thr Val Glu Asp Ala Glu Ala Leu Pro Ala

65 70 75 80

Ala Ala Gly Glu Gln Glu Pro Arg Gly Pro Glu Pro Asp Asp Glu Thr

85 90 95

Glu Leu Arg Pro Arg Gly Arg Ser Leu Val Ile Ile Ser Thr Leu Asp

100 105 110

Gly Arg Ile Ala Ala Leu Asp Pro Glu Asn His Gly Lys Lys Gln Trp

115 120 125

Asp Leu Asp Val Gly Ser Gly Ser Leu Val Ser Ser Ser Leu Ser Lys

130 135 140

Pro Glu Val Phe Gly Asn Lys Met Ile Ile Pro Ser Leu Asp Gly Ala

145 150 155 160

Leu Phe Gln Trp Asp Gln Asp Arg Glu Ser Met Glu Thr Val Pro Phe
 165 170 175
 Thr Val Glu Ser Leu Leu Glu Ser Ser Tyr Lys Phe Gly Asp Asp Val
 180 185 190
 Val Leu Val Gly Gly Lys Ser Leu Thr Thr Tyr Gly Leu Ser Ala Tyr
 195 200 205

 Ser Gly Lys Val Arg Tyr Ile Cys Ser Ala Leu Gly Cys Arg Gln Trp
 210 215 220
 Asp Ser Asp Glu Met Glu Gln Glu Glu Asp Ile Leu Leu Leu Gln Arg
 225 230 235 240
 Thr Gln Lys Thr Val Arg Ala Val Gly Pro Arg Ser Gly Asn Glu Lys
 245 250 255
 Trp Asn Phe Ser Val Gly His Phe Glu Leu Arg Tyr Ile Pro Asp Met
 260 265 270

 Glu Thr Arg Ala Gly Phe Ile Glu Ser Thr Phe Lys Pro Asn Glu Asn
 275 280 285
 Thr Glu Glu Ser Lys Ile Ile Ser Asp Val Glu Glu Gln Glu Ala Ala
 290 295 300
 Ile Met Asp Ile Val Ile Lys Val Ser Val Ala Asp Trp Lys Val Met
 305 310 315 320
 Ala Phe Ser Lys Lys Gly Gly His Leu Glu Trp Glu Tyr Gln Phe Cys
 325 330 335

 Thr Pro Ile Ala Ser Ala Trp Leu Leu Lys Asp Gly Lys Val Ile Pro
 340 345 350
 Ile Ser Leu Phe Asp Asp Thr Ser Tyr Thr Ser Asn Asp Asp Val Leu
 355 360 365
 Glu Asp Glu Glu Asp Ile Val Glu Ala Ala Arg Gly Ala Thr Glu Asn
 370 375 380
 Ser Val Tyr Leu Gly Met Tyr Arg Gly Gln Leu Tyr Leu Gln Ser Ser
 385 390 395 400

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

Ala Trp Leu Glu Ala Pro Pro Glu Lys Trp Gln Glu Lys Met Asp Glu
 660 665 670

Ile Trp Leu Lys Asp Glu Ser Thr Asp Trp Pro Leu Ser Ser Pro Ser
 675 680 685

Pro Met Asp Ala Pro Ser Val Lys Ile Arg Arg Met Asp Pro Phe Ala
 690 695 700

Thr Lys Glu His Ile Glu Ile Ile Ala Pro Ser Pro Gln Arg Ser Arg
 705 710 715 720

Ser Phe Ser Val Gly Ile Ser Cys Asp Gln Thr Ser Ser Ser Glu Ser
 725 730 735

Gln Phe Ser Pro Leu Glu Phe Ser Gly Met Asp His Glu Asp Ile Ser
 740 745 750

Glu Ser Val Asp Ala Ala Tyr Asn Leu Gln Asp Ser Cys Leu Thr Asp
 755 760 765

Cys Asp Val Glu Asp Gly Thr Met Asp Gly Asn Asp Glu Gly His Ser
 770 775 780

Phe Glu Leu Cys Pro Ser Glu Ala Ser Pro Tyr Val Arg Ser Arg Glu
 785 790 795 800

Arg Thr Ser Ser Ser Ile Val Phe Glu Asp Ser Gly Cys Asp Asn Ala
 805 810 815

Ser Ser Lys Glu Glu Pro Lys Thr Asn Arg Leu His Ile Gly Asn His
 820 825 830

Cys Ala Asn Lys Leu Thr Ala Phe Lys Pro Thr Ser Ser Lys Ser Ser
 835 840 845

Ser Glu Ala Thr Leu Ser Ile Ser Pro Pro Arg Pro Thr Thr Leu Ser
 850 855 860

Leu Asp Leu Thr Lys Asn Thr Thr Glu Lys Leu Gln Pro Ser Ser Pro
 865 870 875 880

Lys Val Tyr Leu Tyr Ile Gln Met Gln Leu Cys Arg Lys Glu Asn Leu
 885 890 895

Lys Asp Trp Met Asn Gly Arg Cys Thr Ile Glu Glu Arg Glu Arg Ser

<212> DNA

<213> Homo sapiens

<400> 13

```

agctgaggga gcgctctgaa ataatacacc attgcagccg gggaaagcag agcggcgcaa      60
aagagctctc gccgggtccg cctgctccct ctccgcttcg ctctcttct ctcttttacc      120
cttctctctt ctctctctct gctgctctct ctctctctcc cgctcttctc tctctctctc      180
tctgctctc tctctctccc ttagctctc ttcttttctt ctctcttct tccctctct      240

cgctctccc ctgctctctt tctctctctt cccctcccct cccgctctc tctcccctct      300
ccctctccca ctgccccgc tcgctcgctc gctgtcgcac agactcaccg tcccttgtec      360
aattatcata ttcatcacc gcaagatac accgtgtgtg cactcgcgtg ttttctctc      420
tctgccgggg gaaaaaaaa agagagagag agatagagag agagagagag agagagagag      480
agaggctcgg tccactgct ccctgcaccg cgggccggg attcttgagc tgtgcccagc      540
tgacgagctt ttgaagatgg cacaataacc gtccagtgat gcctgacat gacagcacag      600
ccctcttaag ccggcaaacc aagaggagaa gattgacat tggagtgaaa aggacggtag      660

ggacagcatc tgcattttt gctaaggcaa gagcaacgtt ttttagtgcc atgaatccc      720
aaggttctga gcaggatgtt gagtattcag tggcagca tgcagatggg gaaaagtcaa      780
atgtactccg caagctctg aagaggcgca actcgatga agatgccatg atgcctttc      840
caggagcaac cataatttc cagctgttga aaaataacat gaacaaaaat ggtggcacgg      900
agcccagttt ccaagccagc ggtctctcta gtacaggctc cgaagtacat caggaggata      960
tatgcagcaa ctctcaaga gacagcccc cagagtgtct tcccccttt ggacgccta      1020
ctatgagcca gttgatatg gatcgcttat gtgatgagca cctgagagca aagcgcgcc      1080

gggttgagaa tataattcgg ggtatgagcc attccccag tgtggcatta aggggcaatg      1140
aaaatgaaag agagatggcc ccgcagtctg tgagtcccc agaaagtac agagaaaaca      1200
aacgcaagca aaagcttccc cagcagcagc aacagagttt ccagcagctg gtttcagccc      1260
gaaaagaaca gaagcgagag gagcggccgac agctgaaaca gcagctggag gacatgcaga      1320
aacagctgcg ccagctgcag gaaaagtctt accaaatcta tgacagcact gattcggaaa      1380
atgatgaaga tggtaacctg tctgaagaca gcatgcgctc ggagatcctg gatgccaggg      1440
cccaggactc tgcggaagg tcagataatg agatgtgca gctagacca ggacagtta      1500

ttgaccgagc tcgagccctg atcagagagc aggaaatggc tgaaaacaag ccgaagcgag      1560
aaggcaaca caaagaaaga gaccatgggc caactcctt acaaccggaa ggcaaacatt      1620
tggtgagac ctgaaacag gaactgaaca ctccatgctc gcaagttgtg gacactgtgg      1680

```

tcaaagcttt ttcggccaag ccctcccgcc aggttctca ggtcttccca cctctccaga 1740
 tccccagcg cagatttga gtcaatgggg aaaaccacaa tttccacacc gccaccagc 1800
 gcctgcagtg ctttggcgac gtcatcattc cgaaccccct ggacaccttt ggcaatgtgc 1860
 agatggccag ttccactgac cagacagaag cactgccctt ggttgtccgc aaaaactcct 1920

ctgaccagtc tgcctccggc cctgccgctg gcggccacca ccagcccctg caccagtgcg 1980
 ctctctctgc caccacgggc ttaccacagt ccaccttccg ccacccttc ccccttccct 2040
 tgatggccta tccatttcag agcccattag gtgctccctc cggctccttc tctggaaaag 2100
 acagagcctc tctgaatcc ttagacttaa ctagggatac cacgagtctg aggaccaaga 2160
 tgtcatctca ccacctgagc caccacctt gttcaccagc acaccgccc agcaccgccc 2220
 aagggtctc cttgtcgtc ataaagtccg agtgcggcga tcttcaagat atgtctgaaa 2280
 tatcacctta ttcgggaagt gcaatgcagg aaggattgtc acccaatcac ttgaaaaaag 2340

caaagctcat gttttttat acccgttatc ccagctcaa tatgtgaag acctactct 2400
 ccgacgtaaa gttcaacaga tgcattacct ctgagctcat caagtggttt agcaatttcc 2460
 gtgagtttta ctacattcag atggagaagt acgcacgtca agccatcaac gatggggtea 2520
 ccagtactga agagctgtct ataaccagag actgtgagct gtacagggct ctgaacatgc 2580
 actacaataa agcaaatgac tttgaggttc cagagagatt cctggaagtt gctcagatca 2640
 cattacggga gtttttcaat gccattatcg caggcaaaga tgttgatcct tectggaaga 2700
 aggccatata caaggtcatc tgcaagctgg atagtgaagt ccctgagatt ttcaaatccc 2760

cgaactgcct acaagagctg cttcatgagt agaaatttca acaactcttt ttgaatgtat 2820
 gaagagtagc agtccccttt ggatgtccaa gttatatgtg tctagatttt gatttcatat 2880
 atatgtgtat gggaggcatg gatatgtat gaaatcagct ggtaattcct cctcatcacg 2940
 tttctctcat tttcttttgt tttccattgc aaggggatgg ttgttttctt tetgccttta 3000
 gtttgctttt gcccaaggcc cttaacattt ggacacttaa aatagggtta attttcaggg 3060
 aaaaagaatg ttggcgtgtg taaagtctct attagcaatg aagggaattt gttaacgatg 3120
 catccacttg attgatgact tattgcaaat ggcggttggc tgaggaaaac ccatgacaca 3180

gcacaactct acagacagtg atgtgtctct tgttttact gctaagaagg tctgaaaatt 3240
 taatgaaacc acttcataca ttttaagtatt ttgtttgggt tgaactcaat cagtagcttt 3300
 tccttacatg tttaaaaata attccaatga cagatgagca gctcactttt ccaaagtacc 3360
 ccaaaggcc aaattaaaa agaaaaataa tactctcaa gccttgtcta agaaaagagg 3420
 caaactctga aagtcgtacc agtttcttct ggaggcaaag caattttgca caaaaccagc 3480
 tctctcaaga tgagactaga aattcatacc tggctttgta gccacctctc taaactgaa 3540

aataggttct tcttcataag tgagcttaca tcattcttca taaagaaaa tcctataact 3600

tgttatcatt tttgcttcag atactaaaag gcactaagtt tccaatttac gctgctcaac 3660
tttgtttata tgcctaaaag gattctgttt acttaacaat tttttccct aaaatactat 3720
tttctgaata ctccctcca gtaaggaata aaggaaagcc caacttgcc ataaaattct 3780
tgctacact agaagtttgt tgacagccat tagctgactt gatcgtcatc tcctaagagg 3840
aacacatata tttcacaag caattccaca ctatcctgat gggtatgcaa agtggtgaca 3900
gtctaactca gtgtttcttc attttaggta taacatttta aagcaattga taatgcctct 3960
tccaattcag aagctagtat tgaccaaaat gtgagaagag tgtatagcat aggaaaattt 4020

ggggttaacc caaaagacac aattccagca cacataagaa agctagctgc tattttatgc 4080
tttcttccat ggttctctc tttttccct tttattttc cctgttttc aatgatgtac 4140
agtgttccct acttgcatg aaaaaactcg tatggcattc aactttttt tcttaggtgg 4200
gtttttgtgt ccagatgcag taagaattca ttgttcatcc taaaactggt tccagaccc 4260
ttcttcccc ttaggtaatt tgatatacac ctctaaaat gacacagtaa caaatctggt 4320
atttagaaca tatagaacat aaatgccatt ttttaattca actttaataa gaattacatt 4380
tgactttgga gaatacaggt cttgacccat gtgactgact agctgaccg atcgctgtaa 4440

tttaacgtca ttataaatt ctgctgatgg acaggaatgt atgaactcaa ttattgtcag 4500
caciaagcct taaaacctgc tgactttaa ttaaatggtg cagtcctatg atgccctgca 4560
ccatccaggg gactaacagg gcctcgcagt gtagacagag ggtgcagcca cacgggcggg 4620
ggcaccagcc acctcaactc gcaaccgagg cctcacacat ctcccagctc aactctact 4680
aatgcacaga gtcattagat ccaatttgtt atttttctca cttgctttaa aaaaaagcag 4740
tttgataat catgacattg gaataaagtg ggaaggaaaa attccatcag cacaaaatag 4800
ggaagtaate ccaacttga gtcacagttt tctgactggc tttgttttaa aagaggatgg 4860

cagtccttgt tcgtgtcagt gtgccactgg gtttttctg ttccgtgtaa ttcatatcaa 4920
ctttgtgtt ccatttgcaa ggtaaaagc aaagctgtag tgtattcacc tatgtagaca 4980
gattgctaga taicttttg atctggggcg agttcaatat tgattccaga cttatttgg 5040
tttttttagt attattttc cctcccttc taatttaaat agacaaatta agcaaaagtg 5100
tgtgttca accaaatgtt gatgccctta tctactgata atatcctctc aatgttcaact 5160
gaggcataga aattattca gagtagaaat tgcagcatga ggataaactc acctcttgt 5220
tctgaaaata gaactttatc actatgcttt ccggtggttt tcccttttac aatcgaatc 5280

ttgtgcctcc caagtgcatt ggaaaatgac aaaagcctgt ctctccaaat tcctatntaa 5340
 cagtttgatt ttttttttt aatcaccatc tttcaaatct tagctcaact ctaccaagt 5400
 gaaaattggc tacttgggag aaagttaact ttctatggtg ggatggtgaa ggatgagggga 5460
 cagtttaccat aggaaaagaa aaaaaaaagt ctaaagtcca tgttgaaaaa ccacactacc 5520
 acttatnttc tgctaacct aaattatntt tgcgtatag cttgaggtta tagtctgtgc 5580
 ctagacctaa aatgcaccag cgggggggat tttaaaaat cttcaaat accagttntt 5640
 tccaacaag tacaattgtt cttgtgcctt ctgtggcttt cgatttcac tttttgactt 5700

 tatttccaat tactacagct gcaataaaca ctagatnttt tttctggctg ttgacataa 5760
 cgttgatagc tatgcatatt ttgtgtcttt ttaaaacaaa gcgggagaat acgtntttga 5820
 agaagagaat tttagaaca gtttgatacc gcaattatt ttttctcaa ttgtttgagc 5880
 agcattcgag tttgaaaa tctttagaaa gccaatnttt tgtaactgtg gtgcaaatct 5940
 tgtgtnttct tagcctaag aaaagtagta tagaagcaat attcatacc atgtgtata 6000
 tatgtgtgcg cagatgtgtg aacataaaat cacatacaca catatacaca catgtaaaaa 6060
 tatacatata tatatatgcg tgtgaagtgg aaagcttacc ttttctatc tagatntaag 6120

 aacctatntt agacatttgt tatgtnttgt gaaaagaatg ttctatnttc aacaaaacat 6180
 ttaattctta ctgtatctct ggctgtntta ttaggacgtt tcacattaaa tggtaaaaca 6240
 catggaagat gttagaatgt agtaattatt taagtaaagc ttcaccaca tnttctgaa 6300
 gtttgctntt tgctccgag tattatntta ttaaagaagt gttntatgtt tgcagaatct 6360
 ttgtcactgt actagggatg tgggtgaata tcattntaaa aaattntaaa caacaaaaaa 6420
 aaagcaaac agaaacacta aagcaagagg ggaactntta taaagcaatg taaatntta 6480
 acctcatggc tgcattatg taagacatga gattnttaata aataactaca ttctcacgac 6540

 atctgtttaa tttactagga aactacagct gactgtatag acagttgaaa gcattcttga 6600
 aaatcctgct ctctcctntt aaaagttaac aatctctntt atcagatgtc aagggaagg 6660
 gtaatgcagt tctgtntaat ttatgaaatt tctntntcta tgtacatgaa gacattntgt 6720
 aagtaacacc ccccttccc atgcgcacat gtgcgcatac acacacacac acacacacac 6780
 acacacacaa acacacacac tgcataaag ctaatgattt ggggactnta aaaaatagga 6840
 tgtcctccag gaacaatcat aaattntatga aagaaagagt agttntacaga ctcccctgaa 6900
 agaagcagtg tatatgtgaa gacagtgcaa aaatctctnt gccatgtata ttatagcgtta 6960

 ttcatntgtg tgaatagtac aaatgnttcc tctggtaca aactctgtgt ttgcaaatnt 7020
 acaagaagca ttgtntntca aaagctcccc ttaaaaaatg taactgnttt atatgagtaa 7080
 gcagntaccg tnttgcactt aaatgntatg ttgaaggaaa tgcagntntt tnttctgtag 7140

atctgttggg tgtaaaccat ctataaaact aaagctaaaa tgctcatatt cagagctggg 7200
 atcaaaactg gtatttaacc ttgcatctt cttataatta tccttctaag aatataacag 7260
 aatgtggaag tgtctggact ttgagtcttt tcaactgagc cttctctcaa atctgacacc 7320
 ccctcagaat gcacaaacat aagcagaaaa ggcaacaag cttaccttct tttgtgaaaa 7380

 cgtattcatt ctgtatTTTT ttaaataattc aattccccta aaaatgggga gaaaatattt 7440
 taaaattgta tattacgact tcaaatttag aactaagaaa aaaatgtatt tgggattggg 7500
 ctacgcgcta cctagaagaa tcaaagggtca tggcttcct caatattgtc ccagccattt 7560
 ctcatatgta tatagtataa accgtgacaa aacctgcct ttatattatt tagcaatag 7620
 ttgtaaatag cattattaag ctcttttttg taataaagac cctttgattt gaatatagta 7680
 caataactga actgataaag tcaatTTTT atTTTTgtt gTTTTTTta gctagaggca 7740
 atttcaattg tgaatTTTT ttgttgtcta ttgttctgaa gactttgcat aattttattgg 7800

 ttaattttat cctaatttat ttgatgaagg tgtacaattt tgtattacca aggatgtact 7860
 gtaatattaa ttgatatgat aaacacaatg agactccctg tccatattaa aaagaaaata 7920
 aaaagggtca gtagacaatt gatTTTTaaag gaaaagttaa aaaaattagt ttggcagcta 7980
 ctaaatttta aaacaggaaa aaaaaagtT gtTgtgggga gggTgggaaa ggggttttac 8040
 tttgtgtgtt ttaagctttt gtatactctc caaactttta ccttttgctt tgtaccactt 8100
 aaaggataca gtagtccaat tgccttgtgt gccttccate tccctttaa ctgaatgtat 8160
 gtgcagtata tatgcaagct tgtgcaaaat aaaatataca ttacaagctc agtgccgttt 8220

 gatTTTctta aagaaagagt gactTTtaat tttTggacct gtatccaatt gtaggacagt 8280
 aggctagtTg tgccagtaat gtcaagtatg gagattttct ttcactaaa ttcttcattc 8340
 tgttagccta acgtgcagct cctagaaca acctTTTTa ctttagatgc ttggaataat 8400
 tgctTggatt tctctctctg aaacatcttt caggcttaac tttatttagc cctgaaactt 8460
 aaaaaaaa 8468
 <210> 14
 <211> 4969
 <212> DNA
 <213> Homo sapiens
 <400> 14
 gctTTTTctt aactttcact aagggttact gtagtctgat gtgtccttcc caaggccacg 60

 aaatttgaca agctgcactt ttcttttget caatgatttc tgctttaagc caaagaactg 120
 cctataattt cactaagaat gtcttctaata tcagatactg gggatttaca agagtcttta 180

aagcacggac ttacacctat tgggtgctggg cttccggacc gacacggatc ccccatcccc 240
gccccgggtc gccttgtcat gctgcccaaa gtggagacgg aagccctggg actggctcga 300
tcgcatgggg aacagggcca gatgccgaa aacatgcaag tgtctcaatt taaaatggtg 360
aattactcct atgatgaaga tctggaagag ctttgtcccg tgtgtggaga taaagtgtct 420
gggtaccatt atgggctcct cacctgtgaa agctgcaagg gatTTTTTaa gcgaacagtc 480

caaaataata aaaggtacac atgtatagaa aaccagaact gccaaattga caaaacacag 540
agaaagcgtt gtccttactg tcgttttcaa aaatgtctaa gtgttggaat gaagctagaa 600
gctgtaaggg ccgaccgaat gcgtggagga aggaataagt ttgggccaat gtacaagaga 660
gacagggccc tgaagcaaca gaaaaagcc ctcacccgag ccaatggact taagctagaa 720
gccatgtctc aggtgatcca agctatgcc tctgacctga ccatttcctc tgcaattcaa 780
aacatccact ctgcctcaa aggcctacct ctgaacctg ctgccttgcc tectacagac 840
tatgacagaa gtcctttgt aacatcccc attagcatga caatgcccc tcacggcagc 900

ctgcaaggtt accaaacata tggccacttt ctagccggg ccatcaagtc tgagtacca 960
gacctata ccagctcacc cgagtcata atgggctatt catatagga tagttaccag 1020
acgagctctc cagcaagcat cccacatctg atactggaac tttgaagtg tgagccagat 1080
gagcctcaag tccaggctaa aatcatggc tatttcagc aagagcagc taaccgaagc 1140
aagcacgaaa agctgagcac ctttgggctt atgtgcaaaa tggcagatca aactctctc 1200
tccattgtcg agtgggccag gagtagtatc ttcttcagag aacttaaggt tgatgacaa 1260
atgaagctgc ttcagaactg ctggagtgag ctcttaatcc tcgaccacat ttaccgaaa 1320

gtggtacatg gaaaggaagg atccatcttc ctggttactg ggcaacaagt ggactattcc 1380
ataatagcat cacaagccgg agccaccctc aacaacctca tgagtcagc acaggagtta 1440
gtggcaaac ttcgttctct ccagttgat caacgagagt tcgtatgtct gaaattcttg 1500
tgctcttta gtttagatgt caaaaacctt gaaaacttc agctggtaga aggtgtccag 1560
gaacaagtca atgccgcct gctggactac acaatgtgta actaccgca gcagacagag 1620
aaatttgac agctacttct tcgactacc gaaatccggg ccatcagat gcaggetgaa 1680
gaatacctct actacaagca cctgaacggg gatgtgcct ataataacct tctcattgaa 1740

atgttcatg caaaagagc ataagttaca acccctagga gctctgcttt caaaacaaaa 1800
agagattggg ggagtggga ggggaagaa gaacaggaag aaaaaagta ctctgaactg 1860
ctccaagtaa cgtaattaa aaacttgctt taaagatatt gaatttaaa aggcataata 1920
atcaaatact taatagcaaa taaatgatgt atcagggtat ttgtattgca aactgtgaat 1980
caaaggcttc acagccccag aggattccat ataaaagaca ttgtaatgga gtggattgaa 2040

ctcacagatg gataccaaca cggtcagaag aaaaacggac agaacggttc ttgtatattt 2100
 aaactgatct cactatgaa gaaatttagg aactaatctt attaattagg cttatacagc 2160

gggggatattg agcttacagg attcctccat ggtaaagctg aactgaaaca attctcaaga 2220
 atgcatcagc tgiacctaca atagcccctc cctcttcctt tgaaggcccc agcacctctg 2280
 ccctgtggtc accgaatctg tactaaggac ctgtgttcag ccacaccag tggtagctcc 2340
 accaaatcat gaacagccta attttgagtg tctgtgtctt agacctgcaa acagctaata 2400
 ggaaattcta ttaatattgt agcttgccat tttaaatag ttctgagggt tgttttctct 2460
 cgtgttcagc atgttaagaa aatgcaggca gtatccctca tcttatgtaa gtgttaatta 2520
 atattaaggg aatgactac aaactttcaa agcaaatgct ccatagctaa agcaacttag 2580

accttatttc tgctactgtt gctgaaatgt ggctttggca ttgttggatt tcataaaaaa 2640
 tttctggcag gaagtcttgt tagtatacat cagtcttttt catcatccaa gttttagtct 2700
 catttaaaaa tacaacatta aacacatttt gctaggatgt caaatagtca cagttctaag 2760
 tagttggaaa caaaattgac gcatgttaat ctatgcaaag agaaaggaaa ggatgaggtg 2820
 atgtattgac tcaaggttca ttcttgctgc aattgaacat cctcaagagt tgggatggaa 2880
 atggtgattt ttacatgtgt cctggaaaga tattaaagta attcaaatct tccccaaagg 2940
 ggaaaggaag agagtgatac tgaccttttt aagtcataga ccaaagtctg ctgtagaaca 3000

aatatgggag gacaaagaat cgcaaattct tcaaatgact attatcagta ttattaacat 3060
 gcgatgccac aggtatgaaa gtcttgccct atttcacaat tttaaaaggt agctgtgcag 3120
 atgtggatca acatttgttt aaaataaagt attaatactt taaagtcaa taagatatag 3180
 tgtttacatt ctttaggtcc tgaggggcag ggggatctgt gatataacaa aatagcaaaa 3240
 gcggtaatth ccttaattgt atttttctga ttggtaatta ttttaacag tacttaatta 3300
 ttctatgtcg tgagacacta aaatcaaaaa cggaatctc atttagactt taattttttt 3360
 gagattatcg gcggcacaat cactttgtag aaactgtaa aaataaaagt atctcctagt 3420

cccttaattt tttcataaat atttctggct tttgagtagt gtatttatat tgtatatcat 3480
 actttcaact gtagacaatt atgatgctaa tttattgttt ctgggtttca cctttgtata 3540
 agatatagcc aagactgaag aaaccaaata tafgtgttta ctgtagcatg tcttcaaat 3600
 agtggaaact agttcagga catagaagag tcttaatgaa ttaaaatcat tcaactgatt 3660
 aaatgtctgt aaactttcat cattcctact gtagtttatt taatatctat tgtaaattat 3720
 gigacttgta gcttcctctg gttttcaagt aaactcaaca aggtggagtc ttacctggtt 3780

ttcctttcca agcattgtaa attgtatacc aaagatatta gttattactt ctgtgtgtac 3840

 aaagaggatt attttattat gtttattaat cacctctaact actcatccac atgaagggta 3900
 cacattaggt aagctgggcg ttgactcatg cgcagtetca gtcacccgtg ttatcttctg 3960
 ggctcaaagg acaatgcaaa atcgccgatc agagctcata cccaaagcat tacagagaac 4020
 agcagcatca ttgccctccc cagctgaaaa acaagtggc tagaagatac atggagagga 4080
 atggtgtggt caacagttaa tgaaacgggt ctatcatgca tgtgtaatgt ggatggagac 4140
 aattataaga ttgactata actatttggg gggctcttaa cattgccaaa aaaacaata 4200
 tgttgatfff tttttatff tttttttat ttttaagaggc gggatcttga tctcacatgt 4260

 tgcccaggct ggccttgaac tcctgggctc aagcattcct cctgcctcag cctccccat 4320
 agctgggact aggggtgcat gccagcatac ctggctacgt tgactcttaa aatctatggt 4380
 ctcttatttt aaagatacag tgctcccccac tgaaaattaa acctaaaaaa tgtcacatat 4440
 tggatgttg ttaacctggt agattaaatc atgagaatga ttagaaagac gggcaacaca 4500
 gcgggttaca tccacactgc tgatcacacc aacgacagga gctgataagc aagaaagcgt 4560
 cacagccage gtctgttacc ccaaggttga caagtgaagt ttctctaag ttgattgtta 4620
 gccgatttgt aacctggcat ttacttagca actgccttat caattacagg atttgccggt 4680

 aaaagcagac tcaaatataa aggtttttgg ctttaacttgg tttattatag ttgctctatg 4740
 tttgtaaaca gacaatctct aatgctgat ttttgtatc acagatctgc agctgccttg 4800
 gacttgaate catgcaatgt ttagagtgtg aagtcagtta cttgttgatg ttttcttact 4860
 gtatcaatga aatacatatt gtcatgtcag ttcttgccag gaacttctca acaaaatgga 4920
 atttttttt tcagtatttc aataaatatt gatatgccca gcctgataa 4969

 <210> 15
 <211> 4831
 <212> DNA
 <213> Homo sapiens
 <400> 15

 gctttttctt aactttcact aagggttact gtagtctgat gtgtccttcc caaggccacg 60

 aaatttgaca agctgcactt ttcttttgc tcaatgattc tgctttaagc caagaactg 120
 cctataatff cactaagaat gtcttctaact tcagatactg gggatttaca agagtcttta 180
 aagcacggac ttacacctat tgtgtctcaa tttaaaatgg tgaattactc ctatgatgaa 240
 gatctggaag agctttgtcc cgtgtgtgga gataaagtgt ctgggtacca ttatgggctc 300
 ctcacctgtg aaagctgcaa gggatfffft aagcgaacag tccaaaataa taaaaggtag 360

acatgtatag aaaaccagaa ctgccaaatt gacaaaacac agagaaagcg ttgtccttac 420
 tgtcgttttc aaaaatgtct aagtgttga atgaagctag aagctgtaag ggccgaccga 480

 atgctggag gaaggaataa gtttgggcca atgtacaaga gagacagggc cctgaagcaa 540
 cagaaaaag ccctcatccg agccaatgga cttagctag aagccatgtc tcaggatgatc 600
 caagctatgc cctctgacct gaccatttcc tctgcaattc aaaacatcca ctctgcctcc 660
 aaaggcctac ctctgaacca tgctgccttg cctcctacag actatgacag aagtcctttt 720
 gtaacatccc ccattagcat gacaatgccc cctcacggca gcctgcaagg ttaccaaaaca 780
 tatggccact ttcttagccg ggccatcaag tctgagtacc cagacccta taccagctca 840
 cccgagtcca taatgggcta ttcatatatg gatagttacc agacgagctc tccagcaagc 900

 atcccacatc tgatactgga acttttgaag tgtgagccag atgagcctca agtccaggct 960
 aaaaatcatgg cctatttga gcaagagcag gctaacccgaa gcaagcacga aaagctgagc 1020
 acctttgggc ttatgtgcaa aatggcagat caaactctct tctccattgt cgagtgggcc 1080
 aggagtagta tcttcttcag agaacttaag gttgatgacc aatgaagct gcttcagaac 1140
 tgctggagtg agctcttaat cctcgaccac atttaccgac aagtggtaaca tggaaaggaa 1200
 ggatccatct tcctggttac tgggcaacaa gtggactatt ccataatagc atcacaagcc 1260
 ggagccacce tcaacaacct catgagtcac gcacaggagt tagtggcaaa acttcgttct 1320

 ctccagtttg atcaacgaga gttcgtatgt ctgaaattct tggtgctctt tagtttagat 1380
 gtcaaaaacc ttgaaaactt ccagctggta gaaggtgtcc aggaacaagt caatgccgcc 1440
 ctgctggact acacaatgtg taactaccgg cagcagacag agaaatttgg acagctactt 1500
 cttcgactac ccgaaatccg ggccatcagt atgcaggctg aagaatacct ctactacaag 1560
 cacctgaacg gggatgtgcc ctataataac cttctcattg aaatgttga tgccaaaaga 1620
 gcataagtta caaccctag gagctctgct ttcaaaaaca aaagagattg ggggagtggg 1680
 gagggggaag aagaacagga agaaaaaag tactctgaac tgctccaagt aacgctaatt 1740

 aaaaacttgc tttaaagata ttgaatttaa aaaggcataa taatcaaata cttaatagca 1800
 aataaatgat gtatcagggt atttgtattg caaactgtga atcaaaggct tcacagcccc 1860
 agaggattcc atataaaaga cattgtaatg gactggattg aactcacaga tggataccaa 1920
 cacggtcaga agaaaaacgg acagaacggt tcttgtatat ttaaactgat ctccactatg 1980
 aagaaattta ggaactaatc ttattaatga ggcttataca gcgggggatt tgagcttaca 2040
 ggattcctcc atggtaaagc tgaactgaaa caattctcaa gaatgatca gctgtaccta 2100

caatagcccc tcctcttcc tttgaaggcc ccagcacctc tgcctgtgg tcaccgaatc 2160

tgtactaagg acctgtgttc agccacaccc agtggtagct ccaccaaatc atgaacagcc 2220

taatlttgag tgtctgttc ttagacctgc aaacagctaa taggaaatc tattaatag 2280

ttagcttggc attttaata tgttctgagg gttgtttgt ctcgtgttca tgatgtaag 2340

aaaatgcagg cagtatccct catcttatgt aagtgttaat taatattaag ggaaatgact 2400

acaaacttcc aaagcaaatg ctccatagct aaagcaactt agaccttatt tetgctactg 2460

ttgctgaaat gtggctttgg cattgttggc ttccataaaa aatttctggc aggaagtctt 2520

gttagtatac atcagtcttt ttcacatcc aagttttag ttcatttaa aatacaacat 2580

taaacacatt ttgctaggat gtcaaatagt cacagtctca agtagttgga aacaaaattg 2640

acgcatgtta atctatgcaa agagaaagga aaggatgagg tgatgtattg actcaaggtt 2700

cattcttgc gcaattgaac atcctcaaga gttgggatgg aaatggtgat ttttacctg 2760

gtcctggaaa gatattaag taattcaaat ctccccaaa ggggaaagga agagagtgat 2820

actgacctt ttaagtcata gaccaaagtc tgctgtagaa caaatatggg aggacaaaga 2880

atcgcaaat ctcaaatga ctattatcag tattattaac atgcgatgcc acaggtatga 2940

aagtcttggc ttatttcaca attttaaaag gtagctgtgc agatgtggat caacattgt 3000

ttaaaataaa gtattaatac tttaaagtca aataagatat agtgtttaca ttcttttagt 3060

cctgaggggc aggggatct gtgatatac aaaatagcaa aagcggaat ttccttaatg 3120

ttattttct gattggtaat tatttttaac agtacttaat tattctatgt cgtgagacac 3180

taaaatcaaa aacgggaatc tcatttagac tttaatttt ttgagattat cggcggcaca 3240

atcactttgt agaaactgta aaaaataaaa gtatctccta gtccttaat tttttcataa 3300

atatttctgg cttttgagta gtgtatttat attgtatc atactttcaa ctgtagacaa 3360

ttatgatgct aatttattgt ttcttggttt cacctttgta taagatatag ccaagactga 3420

agaaaccaa tatatgtgt tactgtagca tgccttcaa ttagtggaac ttagttcagg 3480

gacatagaag agtcttaatg aattaaaatc attcactga ttaaagtct gtaaacttc 3540

atcattccta ctgtagttta ttaatatct attgtaaat attgacttg tagcttctc 3600

tggttttcaa gtaaaccaa caagtgag tcttacctgg ttttcttc caagcattgt 3660

aaattgtata ccaagatat tagttattac ttctgtgtgt acaaagagga ttattttatt 3720

atgtttatta atcacctca atactcatcc acatgaaggg tacacattag gtaagctggg 3780

cgttgactca tgcgcagtct cagtcaccg tgttatctc gtggctcaa ggacaatgca 3840

aaatcgccga tcagagctca tacccaaagc attacagaga acagcagcat cattgcctc 3900
cccagctgaa aaacaagttg gctagaagat acatggagag gaatggtgtg gtcaacagtt 3960
aatgaaacgg ttctatcatg catgtgtaat gtggatggag acaattataa gatttgacta 4020
taactatttg gagggctctt aacattgcc aaaaaacaaa tatgttgatt tttattttat 4080
tttatttttt attttaagag gcgggatctt gatctcacat gttgccagg ctggccttga 4140
actcctgggc tcaagcattc ctctgcctc agcctceccc atagctggga ctagggtgc 4200
atgccagcat acctggctac gttgactctt aaaatctatg ttctcttatt ttaaagatac 4260

agtgctcccc actgaaaatt aaacctaaaa aatgtcacat attggtatgt tgtaacctg 4320
gtagattaaa tcatgagaat gattagaaag acgggcaaca cagcgggtta catccacact 4380
gctgatcaca ccaacgacag gagctgataa gcaagaaagc gtcacagcca gcgtctgttc 4440
accaaggtt gacaagtgaa gtttctctaa tgttgattgt tagccgattt gtaacctggc 4500
atttacttag caactgcctt atcaattaca ggatttgccg gtaaaagcag actcaaatat 4560
aaaggttttt ggcttaactt ggtttattat agttgctcta tgtttgtaa cagacaatct 4620
ctaatgtctg attatttgta tcacagatct gcagctgcct tggacttgaa tccatgcaat 4680

gtttagagtg tgaagtcagt tacttgttga tgttttctta ctgtatcaat gaaatacata 4740
ttgtcatgtc agttcttgcc aggaacttct caacaaaatg gaattttttt tttcagtatt 4800
tcaataaata ttgatatgcc cagcctgata a 4831

<210> 16

<211> 4787

<212> DNA

<213> Homo sapiens

<400> 16

atagtttgtc atctttttag ctgcgaagac gccttcgtgg gtctgcgtcc ggagcaaggc 60
ggttaccga tcccgcagc gagccgcgcc gcgcgtcgtg gcggcctgat ttctgtttaa 120
ctttaatagg gcgatctcga gtgtctcaat ttaaaatggt gaattactcc tatgatgaag 180

atctggaaga gctttgtccc gtgtgtggag ataaagtgtc tgggtacat tatgggctcc 240
tcacctgtga aagctgcaag ggatttttta agcgaacagt ccaaataat aaaaggtaca 300
catgtataga aaaccagaac tgccaaattg acaaacaca gagaaagcgt tgtccttact 360
gtcgttttca aaaatgtcta agtgttgaa tgaagctaga agctgtaagg gccgaccgaa 420
tgcgtggagg aaggaataag tttgggcaa tgtacaagag agacagggcc ctgaagcaac 480
agaaaaaagc cctcatccga gccaatggac ttaagctaga agccatgtct caggtgatcc 540

aagctatgcc ctctgacctg accatttctt ctgcaattca aaacatccac tctgcctcca 600

aaggcctacc tctgaacat gctgccttgc ctctacaga ctatgacaga agtccctttg 660
taacatcccc cattagcatg acaatgcccc ctcacggcag cctgcaaggt taccaaacat 720
atggccactt tcctagccgg gccatcaagt ctgagtacc agaccctat accagctcac 780
ccgagtccat aatgggctat tcatatatgg atagttacca gacgagctct ccagcaagca 840
tcccacatct gatactggaa cttttgaagt gtgagccaga tgagcctcaa gtccaggcta 900
aatcatggc ctatttgtag caagagcagg ctaaccgaag caagcacgaa aagctgagca 960
cctttgggct tatgtgcaaa atggcagatc aaactctctt ctccattgtc gagtgggcca 1020

ggagtagtat ctcttcaga gaacttaagg ttgatgacca aatgaagctg cttcagaact 1080
gctggagtga gctcttaatc ctgaccaca ttaccgaca agtggtagat ggaaaggaag 1140
gatccatctt cctggttact gggcaacaag tggactattc cataatagca tcacaagccg 1200
gagccacct caacaacctc atgagtcag cacaggagtt agtggcaaaa cttcgttctc 1260
tccagtttga tcaacgagag ttcgtatgtc tgaattctt ggtgctctt agtttagatg 1320
tcaaaaaact tgaaaacttc cagctggtag aagggttcca ggaacaagtc aatgccgccc 1380
tgctggacta cacaatgtgt aactaccgc agcagacaga gaaatttga cagctacttc 1440

ttcgactacc cgaatccgg gccatcagta tgcaggctga agaatactc tactacaagc 1500
acctgaacgg ggatgtgccc tataataacc ttctcattga aatgttgcag gccaaaagag 1560
cataagttac aacccttagg agctctgctt tcaaaaacaaa aagagattgg gggagtgggg 1620
agggggaaga agaacaggaa gaaaaaagt actctgaact gctccaagta acgctaatta 1680
aaaacttgct ttaaagatat tgaatttaa aaggcataat aatcaaac ttaatagcaa 1740
ataaatgatg tatcagggta tttgtattgc aaactgtgaa tcaaaggctt cacagcccca 1800
gaggattcca tataaaagac attgtaatgg agtggattga actcacagat ggataccaac 1860

acggtcagaa gaaaaacgga cagaacggtt cttgtatatt taaactgac tccactatga 1920
agaaatttag gaactaatct tattaattag gcttatacag cgggggattt gagcttacag 1980
gattcctcca tggtaaagct gaactgaaac aattctcaag aatgcatcag ctgtacctac 2040
aatagcccc cctcttctt ttgaaggccc cagcacctct gccctgtggt caccgaatct 2100
gtactaagga cctgtgttca gccacacca gtggtagctc caccaaatca tgaacagcct 2160
aattttagt gtctgtgtct tagacctgca aacagctaat aggaaattct attaatatgt 2220
tagcttgcca ttttaaatat gttctgaggg ttgtttgtc tcgtgttcat gatgtaaga 2280

aaatgcaggc agiatccctc atcttatgta agtgtaatt aatattaagg gaaatgacta 2340
 caaaccttca aagcaaatgc tccatagcta aagcaactta gaccttattt ctgctactgt 2400
 tgctgaaatg tggctttggc attgttggat ttcataaaaa atttctggca ggaagtcttg 2460
 ttagtataca tcagtctttt tcatcatcca agttttagt tcatttaaaa atacaacatt 2520
 aaacacattt tgctaggatg tcaaatagtc acagttctaa gtagtggaa acaaaattga 2580
 cgcatgtaa tctatgcaaa gagaaaggaa aggatgaggt gatgtattga ctcaaggttc 2640
 attcttgctg caattgaaca tctcaagag ttgggatgga aatggtgatt tttacatgtg 2700

 tcctggaaag atattaaagt aattcaaatc ttcccaaag gggaaaggaa gagagtata 2760
 ctgacctttt taagtcatag accaaagtct gctgtagaac aaataggga ggacaaagaa 2820
 tcgcaaatc ttcaaatgac tattatcagt attattaaca tgcgatcca caggatgaa 2880
 agtcttgctt tatttcaca ttttaaaagg tagctgtgca gatgtggatc aacatttgtt 2940
 taaaataaag tattaact ttaaagtcaa ataagatata gtgtttacat tctttaggtc 3000
 ctgaggggca ggggatctg tgatataaca aaatagcaaa agcggtaatt tccttaatgt 3060
 ttttttctg attgtaatt attttaaca gtacttaatt attctatgtc gtgagacact 3120

 aaaatcaaaa acgggaatct catttagact ttaatTTTT tgagattatc ggcggcacia 3180
 tcactttgta gaaactgtaa aaaataaaag tatctcctag tcccttaatt ttttcataaa 3240
 ttttctggc ttttagtag tgtatttata ttgtatatca tactttcaac tgtagacaat 3300
 tatgatgcta atttattgtt tcttggtttc acctttgtat aagatatagc caagactgaa 3360
 gaaaccaa atatgtgtt actgtagcat gtcttcaaat tagtggact tagttcaggg 3420
 acatagaaga gtcttaatga attaaaatca ttcacttgat taaatgtctg taaatctca 3480
 tcattctac ttagtttat ttaatatcta ttgtaatta tgtgacttgt agcttctct 3540

 ggttttcaag taaactcaac aagggtgagt cttacctggt tttcctttcc aagcattgta 3600
 aattgtatac caagatatt agttattact tctgtgtgta caaagaggat tttttatta 3660
 tgtttattaa tcaccttaa tactcatcca catgaagggt acacattagg taagctgggc 3720
 gttgactcat gcgcagtctc agtcaccegt gttatcttcg tggetcaaag gacaatgcaa 3780
 aatcgccgat cagagctcat acccaaagca ttacagagaa cagcagcatc attgccctcc 3840
 ccagctgaaa aacaagtgg ctagaagata catggagagg aatggtgtgg tcaacagtta 3900
 atgaaacggt tctatcatgc atgtgtaatg tggatggaga caattataag atttgactat 3960

 aactatttgg aggtcttta acattgcaa aaaaacaaat atgttgattt ttattttatt 4020
 ttatttttta ttttaagagg cgggatcttg atctcacatg ttgccaggc tggccttgaa 4080
 ctctgggct caagcattcc tctgcctca gcctcccca tagctgggac taggggtgca 4140

tgccagcata cctggctacg ttgactctta aaatctatgt tctcttattt taaagataca 4200
 gtgctcccca ctgaaaatta aacctaaaaa atgtcacata ttggatggtt gttaacctgg 4260
 tagattaaat catgagaatg attagaaaga cgggcaacac agcgggttac atccacactg 4320
 ctgatcacac caacgacagg agctgataag caagaaagcg tcacagccag cgtctgttca 4380

cccaaggttg acaagtgaag tttctctaataa gttgattggt agccgatttg taacctggca 4440
 tttactttagc aactgcctta tcaattacag gatttgccgg taaaagcaga ctcaaatata 4500
 aaggtttttg gcttaacttg gtttattata gttgctctat gtttgtaaac agacaatctc 4560
 taatgtctga ttatttgtat cacagatctg cagctgcctt ggacttgaat ccatgcaatg 4620
 tttagagtgt gaagtcagtt acttgttgat gttttcttac tgtatcaatg aaatacatat 4680
 tgtcatgtca gttcttgcca ggaacttctc aacaaaatgg aatttttttt ttcagtattt 4740
 caataaatat tgatatgcc agcctgataa tttttaaaaa aaaaaaa 4787

<210> 17

<211> 1165

<212> DNA

<213> Homo sapiens

<400> 17

gccagagagc tggaaagtgag agcagatccc taacctgag caccagccaa ccaggggcct 60
 gcccatgcca gggagctgca agccgccccg ccattctcta cgcacttctg agctccagcc 120
 tcaaggctgt cccccagcc cgtagccgct gcctatgtag gcagcaccgg cccgtccagc 180
 tatgtgcacc tcatcgacc tgccgggagg ccttggatgt tctggccaag acagtggcct 240
 tcctcaggaa cctgccatcc ttctggcagc tgcttccca ggaccagcgg cggctgctgc 300
 agggttgctg gggccccctc ttctgcttg ggttggecca agatgctgtg acctttgagg 360

tggtgaggc cccggtgcc agcatactca agaagattct gctggaggag ccagcagca 420
 gtggaggcag tggccaactg ccagacagac ccagccctc cctggctgcg gtgcagtggc 480
 ttcaatgctg tctggagtcc ttctggagcc tggagcttag cccaaggaa tatgectgcc 540
 tgaaaggac catcctctc aacccgatg tgccagcct ccaagccgc tcccacattg 600
 ggacactgca gcaggaggct cactgggtgc tgtgtgaagt cctggaacce tgggcccag 660
 cageccaagg ccgectgacc cgtgtcctc tcacggctc caccetcaag tccattccga 720
 ccagcctgct tggggacctc ttctttgcc ctatcattgg agatgttgac atcgctggcc 780

ttcttgggga catgcttttg ctcaggtgac ctgttccagc ccaggcagag atcaggtggg 840
 cagaggctgg cagtgtgat tcagcctggc catccccaga ggtgacccaa tgctcctgga 900

gggggcaagc ctgtatagac agcacttggc tccttaggaa cagctcttca ctgagccaca 960
 cccacattg gacttccttg gtttggacac agtgttccag ctgcctggga ggcttttgg 1020
 ggtccccaca gcctctgggc caagactcct gtccttctt gggatgagaa tgaaagctta 1080
 ggctgcttat tggaccagaa gtcctatcga ctttatacag aactgaatta agttattgat 1140
 ttttgaata aaaggtatga aacac 1165

<210> 18

<211> 7937

<212> DNA

<213> Homo sapiens

<400> 18

aatcatggtg ccgctgggga ggggagaagc tgcctgctgcc gccgttgccg ggagcccgcg 60
 agacaagtca ttacgttttc atttctcaca actgggctga gcacaactga accatggggg 120
 aacacagtcc agacaacaac atcatctact ttgaggcaga ggaagatgag ctgacccccg 180
 atgataaaat gctcaggttt ttggataaaa acggactggt gccttctca tctggaactg 240
 tttatgatag gaccactgtt cttattgagc aggaccctgg cactttggag gatgaagatg 300
 acgacggaca gtgcggagaa cacttgccct ttctagtagg gggatgaagag ggctttcacc 360

tgatagatca tgaagcaatg tcccagggtt atgtgcagca cattatctca ccagatcaga 420
 ttcatttgac aataaacctt ggttccacac ccatgccaag aaatattgaa ggtgcaacct 480
 tcactctgca gtcggaatgt ccggaacaa aacgtaaaga agtaaagcgg taccaatgta 540
 cctttgaggg ctgtccccgc acctacagca cagcaggcaa cctgcgaacc caccagaaga 600
 ctccaggagg agagtacacc tttgtctgta atcaggaggg ctgtggcaaa gccttctta 660
 cctcttacag cctcaggatc cacgtgcgag tgcacacgaa ggagaagcca tttgagtgtg 720
 acgtgcaggg ctgtgagaag gcattcaaca cactgtacag gctgaaagca catcagaggc 780

ttcacacagg gaaaacgttt aactgtgaat ctgaaggctg cagcaatac ttcaccacac 840
 tcagtgatct gaggaagcac attcgaactc atacagggga aaagccattt cgggtcgatc 900
 acgatggctg tggaaaagca tttgcagca gccaccacct taaaactcac gttcgtacac 960
 atactggtga aagacccttc ttctgcccc gtaatggctg tgagaaaaca ttcagcactc 1020
 aatacagtct caaaagtcac atgaaaggtc atgataacaa aggacactca tacaatgcac 1080
 ttccacaaca caatggatca gaggatacaa atcactcact ttgtctaagt gacttgagcc 1140
 ttctgtccac agattctgaa ttgcgagaaa attccagtac gaccagggc caggacctca 1200

gcacaatttc accagcaatc atctttgaat caatgttcca gaattcagat gatacggcaa 1260

ttcaggaaga tcctcaacag acagcttctc tgactgaaag ttttaatggt gatgcagagt 1320
 cagtcagtga tgttccgcca tccacaggaa attcagcadc tttatctctt ccaactgtac 1380
 tgcaacctgg cctctccgag ccaccccagc ctctaactacc tgcctcagct cegtctgctc 1440
 ctcccctgc tcctcccta ggacctggct cccagcaagc tgcatttggc aacccccctg 1500
 ctctcttaca acctccagaa gtgcctgttc cccacagcac acagtttgct gctaatac 1560
 aagagtttct tccgcacccc caggcacccg ageccattgt accaggactt tctgttgttg 1620

 ctggggcttc tgcatcagca gcggcagtgg catcagctgt ggccagacca gccccaccac 1680
 aaagtactac tgagcccctg ccagccatgg tccagactct gccctgggt gccaaactctg 1740
 tcctaactaa taatcccaca ataaccatca ccccaactcc caacacagct atcctgcagt 1800
 ccagcctagt catgggagaa cagaacttac aatggatatt aaatggtgcc accagttctc 1860
 cacaaaacca agaacaaatt cagcaagcat ctaaagtga gaaggtgttt ttaccactg 1920
 cagtaccagt agccagtgc ccaggagct ctgtccagca gattggcctc agtgctcctg 1980
 tgatcatcat caaacaagaa gaggcatgtc agtgtcagtg tgcatgccgg gactctgcaa 2040

 aggagcgggc atccagcagg agaaagggtc gctcctcccc acccctcca gagccgagcc 2100
 cccaggctcc tgatgggccc agcctgcagc tcccagcga gactttctct tcagcccctg 2160
 ttcccgggtc atcactctct acctgcccct cctcctgtga gcaaagccga caagcagaga 2220
 ctcttcaga cctcagaca gaaacattaa gtgccatgga tgtgtcagag ttctatccc 2280
 tccagagcct ggacaccccg tccaatctga ttccattga agcactactg cagggggagg 2340
 aggagatggg cctcaccagc agcttctcca agtgaagggc ccatgtgtgc tcacctctgg 2400
 gaaaagcggg tgagcaggag gcatgaggta caatgcctgc catcatgggt cagaaatttg 2460

 aaggatgaag aaatctactg tttgaaatcc tcaccttca gacgtatfff cttattcac 2520
 atcccaggag catccatfff aaggaactat tctttgaaa aaaacaaaaa acaaaaaaaaa 2580
 caacaaaaaa agctaagtta taagtgaact gtttgctgc actgtatgtc acttttgctt 2640
 gttgtcatgt gaacttgaa actaaggta ctctgtgca taaaaattct aatgaaagg 2700
 gtgtgttct catcaatctg atgtgccc caatctgca tgggtcttt gtggatcggg 2760
 caggagtttt cagtgtgtg ggtgtgtc cttctatgt gcttttgaa tctgaggctg 2820
 acatttgctt ggaaggccag accctgtc catcagagag ggcagtggca aaggccagtg 2880

 aggcagctgt gaggtgaca gggttcaggt gagatggtgt tgcatttgt gcttagtgtt 2940
 ggtggtgctc aggttgata acacgggtcg ttctgcagcc cgcttcagca caaataggca 3000
 gcttaaggcc tggtcacag gctgtgggt tgatctggct ctgcagagcc cctagccagc 3060
 ttgttgactg ctgtctgtg atgactgtg tgcaaacag gctctagcaa catgatcact 3120

gtccttgcc tctctgttct tctctcgggt tggttgccag ggcttgcaga tgcagtgaa 3180
 ttttccttgg ggaacatcgc tgttttgtcc tagagtgaac ttgtggctta tggccagtgc 3240
 tgtttgggtg tctgccttct ttttaatggt attttcttcc tcagagcaga agggctgcat 3300

 tttgcttata agaagaaggt gcagatttaa gggaaattcat atgaggtggc atgtaattgg 3360
 caggccaggt gtcttggttc caggttccag ccaggctttg ggttgccccc tccatctctg 3420
 cccccctctg gatthttgat acagcctcat acagtgcaaa caaggatgtg acttgctcag 3480
 cttagtcatg tgatttattt aaaaaaaaa aaaaaagaaa cacaaaacga tgatcttcta 3540
 ctccaggtat agcaaaacaa aaaaattccc ttccaccaaa aaagcctgaa atgttgcaat 3600
 aagttatctc atttggatg ttccattaag ttgtgttata ggaaaaaatt gtgtgtgtgt 3660
 gttatagaat tatatccata tgtctgcctt tggtccaag tcattgcctc ttaaaataaa 3720

 agatacaate catactagca tgaaaagttt cctcaacag gctatattaa catagtcatg 3780
 agtgctgacc aaactcaccg agctcagagg ccaggcatgg cctgaggtgc agaataggcc 3840
 tctgcctccc aagagccctt tcttgcctt gagcaaggag tgggtttcca caaacaaggc 3900
 tgctcttcta agccaacagt gtcaggcagg aagcagccat aattttgcct tgcattttca 3960
 ttccctaatg taaagggatc tgcatgtgtc actctcctgt tctctgagcc attgctcagg 4020
 gccagccaag atattattga gaacagataa ttaccttgg agccagagcc cctccctgcc 4080
 tttagcaagg atgttcaggg acagacaaaag agggcagtggt tggatgatgt tgttactgcc 4140

 atgaggagaa atggcagtaa gaaatcttaa ctacaagcag ccaatttctc attccaggac 4200
 cctagccaga ataattgact tctttttttt tttagacag agttttgctt ttgttgcctt 4260
 ggctggagtg cagtggcgca atcttggctc accgcaacct ccacttccca ggttcaagca 4320
 attctgcctc agcctcccga gtagctggga ttacagccat gcgccaccac gcctggctaa 4380
 ttttgtattt ttagtagaga cggggtttct ccatgttgggt caggtctgtc tcgaactccc 4440
 aacgtcaggt gatcctccc cctcagctc ccaaagtct gggattacag gtgtgagcca 4500
 ccatgcctgg tctagggat tgaactctt ttgaccttct gcactcctt ccccaaaagg 4560

 attgtggctt ctgtgacca ttgacctcag cagaagtga taaggcagga ggtttccagt 4620
 cctcctggaa aaccaactgc tgagcatgag ttgtccttgg ctgtctctgg gcctcgcacc 4680
 catgggaagt tttgagtag ggccttctt gtttacaact tctgagagag tagtgatggg 4740
 accccaaagt aagcttctat cagagggcag taatgacctt cccatctccc catcctgtgg 4800
 tcacctggga tttgggttcc ctggggcgga agtgggaaat aggagcccag ggaaggacat 4860
 ctgaagcacc cacagtttaa atagcgactc ttcttgggtt gccagggctg tgctcatgtt 4920

gggcctacc caggtcagtg gccttttcac tctcaaagat cgggggtgac gaagcagcat 4980

cttaaacagt gtttagctg acagatTTTT ccagttgaag aagctgaagt atctgccttt 5040

tgagggtgac tctagattac aagagagact atcagatcca atcagctttt gaaaaatcag 5100

gaatgtgctt aaaatgccag tgagtgggtg tgaagatcag gtttatttcc caggataggc 5160

agtctttctt tccctctttt tccagtctc ttttctacc tgtccttcca taagcctggt 5220

ctctagtgcc gaggacctg gagaagagag ggcccctagc taagctggag ccagaaagaa 5280

ccttcatggt gagtgtggtt tcccaactt gggaaatgaa cctagggtga ggttagggag 5340

gattttaga attctagcag gtaaaaatca agcctttttc ccctccata gtaacttaa 5400

ccattttcac ccgtacgtt atcttctagc tctactaac atctccaatt agacaaccgc 5460

ttcaaaggtt gcttagcacc agggattggg gttcatggac ttaggaggtg gtaaggaaac 5520

tggctggcac caccaaactg ccttagtgaa ctggccctt cccaggcagg taggttgggc 5580

attgagggaa ggtggcccag tctttgctgg cagggtctgg ccagtctcat gggggcacca 5640

ccagtttggg gacagaagtg gtgtcattta ttgaattcca cctccgttta gaaggagatc 5700

atggtacaaa gccaggagg ggccttaaga tgtgaagatc tctctgagta aaaggcaagg 5760

ggcttctctt cttcacctct gggatagttg gtagatcgga gagttttatt ttcagggtca 5820

aggctgtgga ctgatgggga tattggaggg tgggtgggtt ttctgagag actttgtata 5880

atgctgaatg tgcagagg gacaagttt gacaacctca tattggtata ttaaagaaat 5940

aataaataa aaaagcactt taggttattt tatctttaac ccgattgctg caatttcttt 6000

tgtgtgtata tatacatata tatactttcc acaaagtitt atttttgct cagaataaaa 6060

agttaaattg aggtgtgaaa agaaaagcac ttaccttggg gcaatatgtg tagcttgatg 6120

gtcgtgtgcc catgtggccc tggcctggca gcgtttttcc gctcaatcag ccctgtgctg 6180

tgagactgtc catagggaaa cactattatg cattctcagc aaccgctcaa tetatgcaag 6240

ccttccctgt gtgccccagg gcgccccctc aggctctctg aagaactgct gtgggtcctg 6300

ttttctgctg actgttgagg ccttttttca tcacttcttg gtctctcgcc atcttttccc 6360

tcttccat tacaaaatga tgcctgaaag gaaggaacag attgttctta gtagaaacc 6420

tggcaccttc tagactttta tatttgaat cacatccatt gtccttaaag acttttccag 6480

agtgattgaa accattgatt tgtggaactg caacaatatt tctcaagagt ttacaattgt 6540

cttacaccac caatcagaaa tatgtttggg gaggcatggg tggcaggggg caacatgagc 6600

catttccata ctgctccca tactacctgt gttggttctt taagcaaaaa ggcctccagc 6660

cttcattaaa tctataaac aacattaca tcccatgaat cattatcaac tttcttcctt 6720
 tactccacct cccccatfff atcttcaatt ctcaggctat agaatagaca agttttagaa 6780
 tatacttcag ccaaagcaga aggatfffca tagatccaat atgcaaatag ttctgctgtg 6840
 acatagatca gaaattgtct gttccttaat atttgggggt ggaggcaaca aaactgaagc 6900
 atatttctga ttggttagcc attttggcct gtttcctctt gcatgttttt tagggagaga 6960
 gggaaacagt cctgtatttc ttctgatgcc ctttggggaa gccgacgacc actgggcatt 7020
 tctcactggt actcctgttc aagagagggc ttctcagtct gcaactgaaaa atgcaaat 7080

aactggatct ttatgtcaat gtgtacatag tacaagcttt tttactggaa ttgaggttta 7140
 aaaccacaca ctgccctfff ggtggtgtgc ctgttgggcc aaaaattggg tgataatgta 7200
 gtgtcacttt ctcagctcaa tgcagtttct actttttctt atgggaaaat tttcataaa 7260
 acctttttgc accaaaacc aggggtgttt ttgcaatat ccttgtttatc ctctagtagt 7320
 gccaaagttag aggcctttctc ttgccctfff cctgctgtgt tctcaggcct cccaagggt 7380
 atttgactca acagtctaca tcttctgttg tgttttggag aatgtggggg tgggggtcag 7440
 agttcaaggt gtctgttccc ttttctgtg aactctttct agtccctatt tggggagggt 7500

ggctggaaac agatfffctg tgaatfffctg gctcagatct tctagccagg aaaggcaaga 7560
 gcccccaaga gccctfffctt ttgacataca ctaatcattg gccgggtctc ttgtgacaac 7620
 ttttaaaate ccaaatagtt ttatttggat tatgtaaaag taagtgtgaa acatgggaac 7680
 aacggacttc cactgagcga tgtgaaaacg ttacaggttc agtacttcca aaggaagaaa 7740
 cctccaaacc caaaaaagaa taaatatgaa ttgtatfff tgaagaatgt gaaataatgg 7800
 tgtttgctta attgctcatt ttgtataaac ttaatattgt actttaaaat atctgctaaa 7860
 aagtgaaaat ttaactfff ggaattgaaa aagcaatatt aaatactaat gaaatcctaa 7920

ttaaatgctt atttaaa 7937

<210> 19
 <211> 4991
 <212> DNA
 <213> Homo sapiens
 <400> 19

agtttccgag gaactfffctg ccggcgcggg gccgcctctg aggccagggc aggacacgaa 60
 cgcgcgagc ggcggcgagg actgagagcc gggggcggcg cggcgctccc taggaagggc 120
 cgtacgagc ggcgggcccg cggggcctcc cggaggagc ggctgcgcca tggacgagcc 180
 acccttcagc gaggcggctt tggagcagc gctgggcgag ccgtgcgctc tggacgcggc 240

gctgctgacc gacatcgaag gtgaagtcgg cgcggggagg ggtagggcca acggcctgga 300

cgccccaagg gcgggcgacg atcgcggagc catggattgc actttcgaag acatgcttca 360

gcttatcaac aaccaagaca gtgacttecc tggcctatit gaccaccct atgctgggag 420

tggggcaggg ggcacagacc ctgccagccc cgataccagc tccccaggca gcttgtctcc 480

acctcctgcc acattgagct cctctcttga agccttctg agcggggccgc aggcagcgcc 540

ctcaccctg tcccctccc agcctgcacc cactccattg aagatgtacc cgtccatgcc 600

cgctttctcc cctgggcctg gtatcaagga agagtcatg cactgagca tctgcagac 660

ccccaccca cagccccgc caggggcct cctgccacag agcttcccag cccagcccc 720

accgcagttc agctccacc ctgtgttagg ctaccccagc cctccgggag gcttctctac 780

aggaagccct cccgggaaca cccagcagcc gctgcctggc ctgccactgg ctccccgcc 840

aggggtcccc cccgtctct tgcacacca ggtccagagt gtggtcccc agcagctact 900

gacagtcaca gctgccccca cggcagcccc tgtaacgacc actgtgacct cgcagatcca 960

gcaggtcccg gtctgtctgc agccccactt catcaaggca gactcgtgc ttctgacagc 1020

catgaagaca gacggagcca ctgtgaagc ggcaggtctc agtccccctgg tctctggcac 1080

cactgtgcag acagggcctt tgccgaccct ggtgagtggc ggaaccatct tggcaacagt 1140

cccactggtc gtagatgagg agaagctgcc tatcaaccgg ctgcagctg gcagcaaggc 1200

cccggcctct gccagagcc gtggagagaa gcgcacagcc cacaacgcca ttgagaagcg 1260

ctaccctcc tccatcaatg acaaaatcat tgagctcaag gatctggtgg tgggactga 1320

ggcaaagctg aataaatctg ctgtcttgcg caaggccatc gactacattc gctttctgca 1380

acacagcaac cagaaactca agcaggagaa cctaagtctg cgcactgctg tccacaaaag 1440

caaatctctg aaggatctgg tgtcggcctg tggcagtgga gggaacacag acgtgctcat 1500

ggagggcgtg aagactgagg tggaggacac actgacceca cccccctcg atgctggctc 1560

acctttccag agcagcccct tgtcccttgg cagcaggggc agtggcagcg gtggcagtgg 1620

cagtgactcg gagcctgaca gccagctctt ttaggacagc aaggcaaagc cagagcagcg 1680

gccgtctctg cacagccggg gcatgctgga ccgctcccgc ctggccctgt gcacgctcgt 1740

cttctctgce ctgtctgca accccttggc ctcttctgctg ggggcccggg ggcttcccag 1800

cccctcagat accaccagcg tctaccatag cctggggcgc aacgtgctgg gcaccgagag 1860

cagagatggc cctggctggg cccagtggct gctgccccca gtggtctggc tgctcaatgg 1920

gctgttgggt ctcgtctct tgggtcttct ctttctctac ggtgagccag tcacacggcc 1980

ccaactcaggc cccgccgtgt acttctggag gcatcgcaag caggctgacc tggacctggc 2040
 ccggggagac ttgcccagg ctgcccagca gctgtggctg gccctgcggg cactgggccc 2100
 gccctgccc acctcccacc tggacctggc ttgtagcctc ctctggaacc tcatccgtca 2160
 cctgctgcag cgtctctggg tgggccgctg gctggcaggc cgggcagggg gcctgcagca 2220
 ggactgtgct ctgcgagtgg atgctagcgc cagcggccga gacgcagccc tggcttacca 2280
 taagctgcac cagctgcaca ccatggggaa gcacacaggc gggcacctca ctgccaccaa 2340
 cctggcgctg agtgcctga acctggcaga gtgtgcaggg gatgccgtgt ctgtggcgac 2400

 gctggccgag atctatgtgg cggtgcatt gagagtgaag accagtctcc cacgggcctt 2460
 gcattttctg acacgcttct tcttgagcag tgcccgcag gccctgcctgg cacagagtgg 2520
 ctcagtgcct cctgccatgc agtggctctg ccacccctg gccaccgtt tcttcgtgga 2580
 tggggactgg tccgtgctca gtaccctatg ggagagcctg tacagcttgg ccgggaacct 2640
 agtggacccc ctggcccagg tgactcagct attccgggaa catctcttag agcgagcact 2700
 gaactgtgtg acccagccca accccagccc tgggtcagct gatggggaca aggaattctc 2760
 ggatgccctc gggtaacctgc agctgctgaa cagctgttct gatgctgcgg gggctcctgc 2820

 ctacagcttc tccatcagtt ccagcatggc caccaccacc ggctagacc cggtgccaa 2880
 gtggtgggccc tctctgacag ctgtggtgat ccaactggctg cggcgggatg aggaggcggc 2940
 tgagcggctg tgcccgtgg tggagcacct gccccgggtg ctgcaggagt ctgagagacc 3000
 cctgccagg gcagctctgc actcctcaa ggctgcccgg gccctgctgg gctgtgccaa 3060
 ggagagctct ggtccagcca gcctgacat ctgtgagaag gccagtgggt acctgcagga 3120
 cagcctggct accacaccag ccagcagctc cattgacaag gccgtgcagc tgttctctgt 3180
 tgacctgctt ctgtggtgc gcaccagcct gtggcggcag cagcagcccc cggccccggc 3240

 cccagcagcc cagggcacca gcagcaggcc ccaggcttcc gcccttgagc tgcgtggctt 3300
 ccaacgggac ctgagcagcc tgaggcggct ggacacagagc ttccggccc ccatgcggag 3360
 ggtgttctca catgaggcca cggcccggct gatggcgggg gccagccca cacggacaca 3420
 ccagctctc gaccgcagtc tgaggcggcg gccaggeccc ggtggcaaag gaggcgcggt 3480
 ggccggagctg gagccgcggc ccacgcggcg ggagcacgcg gaggccttgc tgctggcctc 3540
 ctgctacctg cccccggct tctgtcggc gccccggcag cgcgtgggca tgctggctga 3600
 ggccggcgcgc aactcagaga agcttggcga tcgccggctg ctgcacgact gtcagcagat 3660

 gctcatgcgc ctgggcggtg ggaccactgt cacttcagc tagaccccgt gtccccggcc 3720
 tcagcacccc tgtctctagc cactttggtc ccgtgcagct tctgtcctgc gtcgaagctt 3780
 tgaaggccga aggcagtgca agagactctg gcctccacag ttcgacctgc ggctgctgtg 3840

tgccttcgcg gtggaaggcc cgagggggc gatcttgacc ctaagaccgg cggccatgat 3900
 ggtgctgacc tctggtggcc gatcggggca ctgcaggggc cgagccattt tggggggccc 3960
 ccctccttgc tctgcaggca ccttagtggc ttttttcctc ctgtgtacag ggaagagagg 4020
 ggtacatttc cctgtgctga cggaagccaa cttggctttc ccggactgca agcagggctc 4080

 tgccccagag gcctctctct cgtcgtggg agagagacgt gtacatagt taggtcagcg 4140
 tgcttagcct cctgacctga ggctcctgtg ctactttgcc ttttgcaaac tttattttca 4200
 tagattgaga agttttgtac agagaattaa aaatgaaatt atttataatc tgggttttgt 4260
 gtcttcagct gatggatgtg ctgactagt agagtgttg ggccctcccc cagcacctag 4320
 ggaaaggctt ccctccccc tccggccaca aggtacacaa cttttaactt agctcttccc 4380
 gatgtttgtt tgttagtggg aggagtgggg agggctggct gtatggcctc cagcctacct 4440
 gttccccctg ctcccagggc acatggttgg gctgtgtcaa cccttagggc ctccatgggg 4500

 tcagttgtcc ctctcacct cccagctctg tccccatcag gtccttgggt ggcacgggag 4560
 gatggactga cttccaggac ctgttgtgtg acaggagcta cagcttgggt ctccctgcaa 4620
 gaagtctgga acgtctcacc tcccccatcc cggccctgg tcatctcaca gcaaagaagc 4680
 ctctccctc ccgacctgcc gccacactgg agagggggca cagggcgagg ggaggtttcc 4740
 tgttctgtga aaggccgact cctgactcc attcatgccc cccccccag cccctccctt 4800
 cattccatt cccaaccta aagcctggcc cggctcccag ctgaatctgg tcggaatcca 4860
 cgggctgcag attttcaaa acaatcgttg tatctttatt gactttttt ttttttttt 4920

 tctgaatgca atgactgttt tttactctta aggaaaataa acatctttta gaaacagctc 4980
 gatacacaca a 4991
 <210> 20
 <211> 4923
 <212> DNA
 <213> Homo sapiens
 <400> 20
 agcagagctg cggccggggg aaccagttt ccgaggaact tttcggcgc gccgggcccg 60
 ctctgaggcc agggcaggac acgaacgcgc ggagcggcgg cggcgactga gagccggggc 120
 cgccggcggc ctccctagga agggccgtac gaggcggcgg gcccggggg cctcccggag 180
 gagcgggctg cgccatggac gagccacct tcagcgaggc ggctttggag caggcgtgg 240

 gcgagccgtg cgatctggac gcggcgtgc tgaccgacat cgaagacatg cttcagctta 300
 tcaacaacca agacagtac ttcctggcc tatttgacc accctatgct gggagtgggg 360

cagggggcac agaccctgcc agccccgata ccagctcccc aggcagcttg tctccacctc 420
 ctgccacatt gagctcctct cttgaagcct tctgagcgg gccgcaggca gcgccctcac 480
 ccctgtcccc tccccagcct gcaccacctc cattgaagat gtaccctgcc atgcccgctt 540
 tctccccctgg gcctggatc aaggaagagt cagtgccact gagcatcctg cagaccccca 600
 ccccacagcc cctgccaggg gccctcctgc cacagagctt cccagcccca gccccaccgc 660

 agttcagctc caccctgtg ttaggctacc ccagccctcc gggaggcttc tetacaggaa 720
 gccctcccgg gaacaccag cagccgctgc ctggcctgcc actggcttcc ccgccagggg 780
 tccccccgt ctcttgac acccaggtcc agagtgtgt cccccagcag ctactgacag 840
 tcacagctgc ccccacggca gccctgtaa cgaccactgt gacctcgcag atccagcagg 900
 tccccgtcct gctgcagccc cacttcatca aggcagactc gctgcttctg acagccatga 960
 agacagacgg agccactgtg aaggcggcag gtctcagtc cctggtctct ggcaccactg 1020
 tgcagacagg gcctttgccg accctgggtga gtggcggaac catcttgca acagtccac 1080

 tggctgtaga tgcggagaag ctgcctatca accggctcgc agctggcagc aaggccccgg 1140
 cctctgcccc gagccgtgga gagaagcgc cagcccacaa cgccattgag aagcgtacc 1200
 gctcctccat caatgacaaa atcattgagc tcaaggatct ggtggtgggc actgaggcaa 1260
 agctgaataa atctgctgtc ttgcgaagg ccatcgacta cattcgcttt ctgcaacaca 1320
 gcaaccagaa actcaagcag gagaacctaa gtctgcgcac tgctgtccac aaaagcaaat 1380
 ctctgaagga tctggtgtcg gcctgtggca gtggagggaa cacagacgtg ctcatggagg 1440
 gcgtgaagac tgaggtggag gacacactga ccccccccc ctcggatgct ggctcacctt 1500

 tcagagcag cccttgtcc cttggcagca ggggcagtgg cagcgggtgc agtggcagtg 1560
 actcggagcc tgacagccca gtctttgagg acagcaaggc aaagccagag cagcggccgt 1620
 ctctgcacag cggggcatg ctggaccgt cccgcctggc cctgtgcag ctctcttcc 1680
 tctgcctgtc ctgcaacccc ttggcctcct tgctgggggc cgggggctt cccagccct 1740
 cagataccac cagcgtctac catagccctg ggcgcaactg gctgggcacc gagagcagag 1800
 atggccctgg ctgggccag ttgctgtctc cccagtggt ctggctgctc aatgggctgt 1860
 tgggtctcgt ctcttggtg cttctctttg tctacggtga gccagtcaca cggccccact 1920

 caggccccgc cgtgtacttc tggaggcatc gcaagcaggc tgacctggac ctggccccgg 1980
 gagactttgc ccaggctgcc cagcagctgt ggctggccct gcgggcactg ggccggcccc 2040
 tgcccacctc ccacctggac ctggcttgta gcctcctctg gaacctcatc cgtcacctgc 2100
 tgcagcgtct ctgggtgggc cgctggctgg caggccgggc agggggcctg cagcaggact 2160
 gtgctctgcg agtggatgct agcggcagcg cccgagacgc agccctggtc taccataagc 2220

tgcaccagct gcacacatg gggaagcaca caggcgggca cctcactgcc accaacctgg 2280
 cgctgagtgc cctgaacctg gcagagtgtg caggggatgc cgtgtctgtg gcgacgctgg 2340

 ccgagatcta tgtggcgct gcattgagag tgaagaccag tctcccacgg gccttgcat 2400
 ttctgacacg ctcttctctg agcagtgcc gccaggcctg cctggcacag agtggctcag 2460
 tgctctctgc catgcagtgg ctctgccacc ccgtgggcca ccgtttcttc gtggatgggg 2520
 actggctcgt gctcagtacc ccatgggaga gcctgtacag cttggccggg aaccagtg 2580
 acccctggc ccaggtgact cagctattcc gggaacatct cttagagcga gcaactgaact 2640
 gtgtgacca gcccacccc agccctgggt cagctgatgg ggacaaggaa ttctcgatg 2700
 ccctcgggta cctgcagctg ctgaacagct gttctgatgc tgcggggct cctgcctaca 2760

 gcttctccat cagttccagc atggccacca ccaccgctg agaccggtg gccaaagtgt 2820
 gggcctctct gacagctgtg gtgatccact ggctgcggcg ggatgaggag gcggctgagc 2880
 ggctgtgccc gctgggtgag cacctgcccc gggtgctgca ggagtctgag agaccctgc 2940
 ccagggcagc tctgactcc ttcaaggtg cccgggcccct gctgggctgt gccaaaggcag 3000
 agtctggtec agccagcctg accatctgtg agaaggccag tgggtacctg caggacagcc 3060
 tggctaccac accagccagc agctccattg acaaggcctg gcagctgttc ctgtgtgacc 3120
 tgcttctgt ggtgcgcacc agcctgtggc ggcagcagca gccccggcc cggccccag 3180

 cagcccaggg caccagcagc aggccccag cttccgcct tgagctgctg ggcttccaac 3240
 gggacctgag cagcctgagg cggctggcac agagcttccg gcccctatg cggagggtgt 3300
 tctacatga ggccacggcc cggctgatgg cgggggcccag ccccacacgg acacaccagc 3360
 tctcgaccg cagtctgagg cggcgggag gccccgtgg caaaggagc gcggtggcgg 3420
 agctggagcc gggcccacg cggcgggagc acgcggagc cttgtctgtg gcctctgct 3480
 acctgcccc cggcttctg tggcgcccg ggcgagcgc tgggcatgct ggctgagcgc 3540
 gcgcgcacac tcgagaagct tggcgatgc cggctgctg acgactgtca gcagatgctc 3600

 atgcgectgg gcggtgggac cactgtcact tccagctaga cccctgtcc cggcctcag 3660
 caccctgtc tctagccact ttggctccgt gcagcttctg tctgctgctg aagctttgaa 3720
 ggccgaaggc agtgcaagag actctggcct ccacagttcg acctgcggct gctgtgtgcc 3780
 ttgcggtg aaggcccag gggcgcatc ttgacctaa gaccggcgcc catgatggtg 3840
 ctgacctgt gtggccgatc ggggactgc aggggcccag ccattttggg gggccccct 3900
 cttgtctgt caggcacctt agtggctttt ttctctgt gtacaggaa gagagggta 3960

catttccctg tgctgacgga agccaacttg gctttcccgg actgcaagca gggctctgcc 4020

ccagaggcct ctctctccgt cgtgggagag agacgtgtac atagttagg tcagcgtgct 4080

tagcctcctg acctgaggct cctgtgctac tttgcctttt gcaaacctta ttttcataga 4140

ttgagaagtt ttgtacagag aattaaat gaaattatit ataactggg ttttgtgtct 4200

tcagctgatg gatgtgctga ctagtgagag tgcctgggcc ctccccagc acctagggaa 4260

aggcttcccc tccccctcg gccacaaggt acacaacttt taacttagct cttcccgatg 4320

ttgtttgtt agtgggagga gtggggaggg ctggctgtat ggcctccagc ctacctgttc 4380

ccctgtctc cagggcacat ggttgggctg tgtcaacct tagggcctcc atggggtcag 4440

ttgtcccttc tcacctcca gctctgtccc catcaggtcc ctgggtggca cgggaggatg 4500

gactgacttc caggacctgt tgtgtgacag gagctacagc ttgggtctcc ctgcaagaag 4560

tctggcacgt ctcaactccc ccatccggc ccttggctcat ctcaacagaa agaagcctcc 4620

tccctcccga ctgcccga cactggagag ggggcacagg ggcgggggag gtttctgtt 4680

ctgtgaaagg ccgactcct gactccattc atgccccccc cccagcccc tcccttcatt 4740

cccattcccc aaactaaagc ctggccccgc tccagctga atctggctcg aatccacggg 4800

ctgcagattt tcaaaaaca tcgttgtatc tttattgact ttttttttt ttttttctg 4860

aatgcaatga ctgtttttta ctcttaagga aaataaacat ctttagaaa caaaaaaaaa 4920

aaa 4923

<210> 21

<211> 4754

<212> DNA

<213> Homo sapiens

<400> 21

aaaaatccgc cgcgccttga caggtgaagt cggcgcgggg aggggtaggg ccaacggcct 60

ggacgcccc agggcggcg cagatcgcgg agccatggat tgcactttcg aagacatgct 120

tcagcttate aacaaccaag acagtgactt ccttggccta tttgaccac cetatgctgg 180

gagtggggca gggggcacag accctgccag ccccgatacc agctccccag gcagcttgtc 240

tccacctct gccacattga gctctctct tgaagccttc ctgagcgggc cgcaggcagc 300

gccctcacc ctgtcccctc cccagcctgc acccactcca ttgaagatgt acccgccat 360

gcccgtttc tcccctgggc ctggtatcaa ggaagagtca gtgccactga gcacctgca 420

gacccccacc ccacagcccc tgccaggggc cctcctgcca cagagcttcc cagccccagc 480

cccaccgag ttcagctcca cccctgtgtt aggctacccc agccctccgg gaggttctc 540

tacaggaagc cctccccgga acaccagca gccgctgcct ggccctgccac tggcttcccc 600
gccaggggtc ccgcccgtct ccttgccacac ccaggtccag agtgtggtcc cccagcagct 660

actgacagtc acagctgccc ccacggcagc cctgtaacg accactgtga cctcgcagat 720
ccagcaggtc ccggtcctgc tgcagcccca cttcatcaag gcagactcgc tgcttctgac 780
agccatgaag acagacggag ccaactgtgaa ggcggcaggt ctcagtcccc tggctcttgg 840
caccactgtg cagacagggc ctttcccagc cctgggtgagt ggcggaacca tcttgccaac 900
agtcccactg gtcgtagatg cggagaagct gcctatcaac cggctcgcag ctggcagcaa 960
ggccccggcc tctgcccaga gccgtggaga gaagcgcaca gcccacaac ccattgagaa 1020
gcgctaccgc tctccatca atgacaaaat cattgagctc aaggatctgg tggtagggcac 1080

tgaggcaaag ctgaataaat ctgctgtctt gcgcaaggcc atcgactaca ttcgctttct 1140
gcaacacagc aaccagaaac tcaagcagga gaacctaatg ctgacgactg ctgtccacaa 1200
aagcaaatct ctgaaggatc tgggttcggc ctgtggcagt ggaggggaaca cagacgtgct 1260
catggagggc gtgaagactg aggtggagga cacactgacc ccaccccct cggatgctgg 1320
ctcaccttcc cagagcagcc ccttgtccct tggcagcagg ggcagtggca gcggtggcag 1380
tggcagtgac tcggagcctg acagcccagt ctttgaggac agcaaggcaa agccagagca 1440
gcggccgtct ctgcacagcc ggggatgct ggaccgctcc cgcctggccc tgtgcacgct 1500

cgcttctctc tgctgtctt gcaaccctt ggctccttg ctgggggccc gggggcttcc 1560
cagcccctca gataccacca gcgtctacca tagccctggg cgcaactgac tgggcaccga 1620
gagcagagat ggccttggct gggcccagtg gctgctgccc ccagtggctt ggctgctcaa 1680
tgggctgttg gtgctcgtct ccttgggtgct tctctttgtc tacggtgagc cagtacacg 1740
gccccactca gggcccgcg tgtacttctg gaggcacgc aagcaggctg acctggacct 1800
ggccccggga gactttgcc aggctgceca gcagctgtgg ctggccctgc gggcactggg 1860
ccggcccctg cccacctccc acctggacct ggctttagc ctctctgga acctcatccg 1920

tcacctgctg cagcgtctct gggtagggcg ctggctggca ggccgggcag ggggcctgca 1980
gcaggactgt gctctgcgag tggatgctag cgccagcgcc cgagacgcag ccctggtcta 2040
ccataagctg caccagctgc acaccatggg gaagcacaca ggcgggcacc tcaactgccac 2100
caacctggcg ctgagtgecc tgaacctggc agagtgtgca ggggatgccc tgtctgtggc 2160
gacgctggcc gagatctatg tggcggctgc attgagagtg aagaccagtc tcccacgggc 2220
cttgcatttt ctgacacgct tcttctgag cagtgcccgc caggcctgcc tggcacagag 2280

tggctcagtg cctcctgcca tgcagtggct ctgccacccc gtgggccacc gtttcttctg 2340

 ggatggggac tggtcctgfc tcagtacccc atgggagagc ctgtacagct tggccgggaa 2400
 cccagtggac cccttgccc aggtgactca gctattccgg gaacatctct tagagcgagc 2460
 actgaactgt gtgaccacgc ccaaccccag ccctgggtca gctgatgggg acaaggaatt 2520
 ctcgatgcc ctcggttacc tgcagctgct gaacagctgt tctgatgctg cgggggctcc 2580
 tgcttacage ttctccatca gttccageat ggccaccacc accggcgtag acccggtggc 2640
 caagtggtag gcctctctga cagctgtggt gatccactgg ctgcggcggg atgaggaggc 2700
 ggctgagcgg ctgtgcccgc tggtaggaca cctgccccgg gtgctgcagg agtctgagag 2760

 acccctgccc agggcagctc tgcactcctt caaggctgcc cgggccctgc tgggctgtgc 2820
 caaggcagag tctggtccag ccagcctgac catctgtgag aaggccagtg ggtacctgca 2880
 ggacagcctg gctaccacac cagccagcag ctccattgac aaggccgtgc agctgttctc 2940
 gtgtgacctg ctctttgtgg tgcgcaccag cctgtggcgg cagcagcagc ccccggcccc 3000
 ggccccagca gccagggca ccagcagcag gccccaggct tccccttg agctgctggtg 3060
 ctccaacgg gacctgagca gcctgaggcg gctggcacag agcttccggc ccgceatgcg 3120
 gagggtgttc ctacatgagg ccacggcccc gctgatggcg ggggccagcc ccacacggac 3180

 acaccagctc ctgaccgca gtctgaggcg gcgggcaggc cccggtggca aaggaggcgc 3240
 ggtggcggag ctggagccgc ggcccacgcg gcgggagcac gcggaggcct tgctgctggc 3300
 ctctgctac ctccccccg gcttctgtc ggccccggg cagcgcgtgg gcatgctggc 3360
 tgaggcggcg cgcacaactc agaagcttgg cgatcgccgg ctgctgcacg actgtcagca 3420
 gatgctcatg cgctgggcg gtgggaccac tgcacttcc agctagacc cgtgtccccg 3480
 gcctcagca cctgtctct agccactttg gtcccgtca gcttctgtcc tgcgtcgaag 3540
 ctttgaagge cgaaggcagt gcaagagact ctggcctcca cagttcgacc tgcggtgct 3600

 gtgtgccttc gcggtggaag gcccagggg cgcgatcttg accctaagac cggcggccat 3660
 gatggtgctg accttggtg gccgatcggg gcactgcagg ggccagcca ttttgggggg 3720
 cccccctct tgctctgag gcaccttagt ggcttttttc ctctgtgta cagggaagag 3780
 aggggtacat ttccctgfc tgacggaagc caacttgct ttcccggact gcaagcaggg 3840
 ctctgcecca gaggcctctc tctccgtctg gggagagaga cgtgtacata gtgtaggtca 3900
 gcgtgcttag cctcctgacc tgaggctcct gtgctacttt gccttttgca aactttattt 3960
 tcatagattg agaagttttg tacagagaat taaaaatgaa attatttata atctgggttt 4020

tgtgtcttca gctgatggat gtgctgacta gtgagagtgc ttgggccctc ccccagcacc 4080
 tagggaaagg ctccccctcc ccctccggcc acaaggtaca caacttttaa cttagctctt 4140
 cccgatgttt gtttgttagt gggaggagtg gggagggtg gctgtatggc ctccagccta 4200
 cctgttcccc ctgctcccag ggcacatggt tgggctgtgt caacccttag ggcctccatg 4260
 gggtcagttg tcctttctca cctcccagct ctgtcccctat caggccctg ggtggcacgg 4320
 gaggatggac tgacttccag gacctgttgt gtgacaggag ctacagcttg ggtctccctg 4380
 caagaagtct ggcacgtctc acctccccca tcccggcccc tggcatctc acagcaaaga 4440

agcctcctcc ctcccagact gccgccacac tggagagggg gcacaggggc gggggaggtt 4500
 tcctgttctg tgaaggccg actcctgac tccattcatg cccccccc cagccctcc 4560
 ctctattccc attccccaac ctaaagcctg gcccggtcc cagctgaatc tggtcggaat 4620
 ccacgggctg cagattttcc aaaacaatcg ttgtatcttt attgactttt ttttttttt 4680
 ttttctgaat gcaatgactg ttttttactc ttaaggaaaa taaacatctt ttagaaacag 4740
 ctcgatacac acaa 4754

<210> 22

<211> 8779

<212> DNA

<213> Homo sapiens

<400> 22

gagaaggagg aggacagcgc cgaggaggaa gaggttgatg gcggcggcgg agctccgaga 60
 gacctcggtt gggcaggggc cggccgtggc gggccgggga ctgctcctct agagcccgca 120
 gttctcggga attcgcgca gcggacgcgc tcggcgaatt tgtgtctctt gcccctctc 180
 cgggcttggg cccaggcccg gccctctgca cttgccetta cttttctat cgagtcgca 240
 tcctctcca gccactgca cccggcgaag agaaaaagga acttccccca cccctcggg 300
 tgccgtcga gccccccagc ccaccctgg gtgcggcgcg gggaccccgg gccgaagaag 360
 agatttctg aggattctgg ttttctctgc ttgtatctcc gaaagaatta aaaatggccg 420

agaatgtggt ggaaccgggg ccgcttcag ccaagcggcc taaactctca tctccggccc 480
 tctcggcgtc cgcagcgtat ggcacagatt ttgctctct atttgacttg gagcacgact 540
 taccagatga attaatcaac tctacagaat tgggactaac caatggtggt gatattaatc 600
 agcttcagac aagtcttggc atggtacaag atgcagcttc taaacataaa cagctgtcag 660
 aattgtctgc atctggtagt tcccctaacc tcaatatggg agttggtggc ccaggtcaag 720
 tcatggccag ccaggcccaa cagagcagtc ctggattagg tttgataaat agcatggtca 780

aaagccaat gacacaggca ggcttgactt ctccaacat ggggatgggc actagtggac 840

 caaatcaggg tcctacgcag tcaacaggta tgatgaacag tccagtaaat cagcctgcca 900
 tgggaatgaa cacagggatg aatgcgggca tgaatcctgg aatgttgct gcaggcaatg 960
 gacaagggat aatgcctaata caagtcatac acggttcaat tggagcaggc cgaggcgac 1020
 agaataatgca gtacccaac ccaggcatgg gaagtgtctg caacttactg actgagcctc 1080
 ttcagcaggg ctctccccag atgggaggac aaacaggatt gagaggcccc cagcctctta 1140
 agatgggaat gatgaacaac cccaatcctt atggttcacc atatactcag aatcctggac 1200
 agcagattgg agccagtggc cttggtctcc agattcagac aaaaactgta ctatcaata 1260

 acttatctcc atttgctatg gacaaaaagg cagttcctgg tggaggaatg cccaacatgg 1320
 gtcaacagcc agccccgag gtccagcagc caggcctggt gactccagtt gccaaggga 1380
 tgggttctgg agcacatata gctgatccag agaagcgaac gctcatccag cagcagcttg 1440
 ttctcctttt gcatgctcac aagtgccagc gccgggaaca ggccaatggg gaagtgaggc 1500
 agtgcaacct tccccactgt cgcacaatga agaatgtcct aaaccacatg acacactgcc 1560
 agtcaggcaa gtcttgccaa gtggcacact gtgcatcttc tcgacaaatc atttcacact 1620
 ggaagaattg tacaagacat gattgtcctg tgtgtctccc cctcaaaaat gctggtgata 1680

 agagaaatca acagccaatt ttgactggag caccctgtgg acttggaat cctagctctc 1740
 taggggtggg tcaacagtct gcccccaacc taagcactgt tagtcagatt gatcccagct 1800
 ccatagaaag agcctatgca gctcttgac tacctatca agtaaatcag atgccgacac 1860
 aaccccagggt gcaagcaaaag aaccagcaga atcagcagcc tgggcagctc cccaaggca 1920
 tgcggcccat gagcaacatg agtgctagtc ctatgggagt aaatggaggt gtaggagttc 1980
 aaacgccgag tcttctttct gactcaatgt tgcattcagc cataaattct caaaacccaa 2040
 tgatgagtga aaatgccagt gtgcccctcc tgggtctctat gccaacagca getcaacat 2100

 ccactactgg aattcggaaa cagtggcacg aagatattac tcaggatctt cgaatcatc 2160
 ttgttcacaa actcgtccaa gccatatttc ctacgccgga tctgctgct ttaaagaca 2220
 gacggatgga aaacctagtt gcatatgctc gaaaagtga aggggacatg tatgaatctg 2280
 caacaatcg agcggaaatc taccaccttc tagctgagaa aatctataag atccagaaag 2340
 aactagaaga aaaaagcaagg accagactac agaagcagaa catgctacca aatgctgcag 2400
 gcatggttcc agtttccatg aatccagggc ctaacatggg acagccgcaa ccaggaatga 2460
 cttctaattg cctctacct gaccaagta tgatccgtgg cagtgtgcca aaccagatga 2520

tgcctcgaat aactccacaa tctggtttga atcaatttgg ccagatgagc atggcccagc 2580
 ccctattgt accccggcaa accctctctc ttcagcacca tggacagtgt gctcaacctg 2640
 gagctctcaa cccgcctatg ggctatgggc ctctgatgca acagccttcc aaccagggcc 2700
 agttccttcc tcagactcag ttcccatcac agggaatgaa tgtaacaaat atccctttgg 2760
 ctccgtccag cggtaagct ccagtgtctc aagcacaat gtctagtctc tctgcccgg 2820
 tgaactctcc tataatgctc ccagggtctc aggggagcca cattcactgt ccccagcttc 2880
 ctcaaccagc tcttcatcag aattcacctc cgctgtacc tagtctgacc cccaccctc 2940

 accatactcc cccaagcata ggggctcagc agccaccagc aacaacaatt ccagcccctg 3000
 ttctacacc tctgccatg ccacctgggc cacagtccca ggctctacat cccctccaa 3060
 ggacagacc tacaccacca acaacacaac ttcccaaca agtgcagcct tcaactctg 3120
 ctgcaccttc tctgaccag cccagcagc agcctcgtc acagcagagc acagcagct 3180
 ctgttctac cccaacagca ccgtgcttc ctccgagcc tgcaactcca ctttccagc 3240
 cagctgtaag cattgaagga caggtatcaa atcctccatc tactagtagc acagaagtga 3300
 attctcagc cattgctgag aagcagcctt cccaggaagt gaagatggag gccaaaatgg 3360

 aagtggatca accagaacca gcagatactc agccggagga tatttcagag tctaaagtgg 3420
 aagactgtaa aatggaatct accgaacag aagagagaag cactgagtta aaaactgaaa 3480
 taaaagagga ggaagaccag ccaagtactt cagctacca gtcactcctg gctccaggac 3540
 agtcaaagaa aaagattttc aaaccagaag aactacgaca ggactgatg ccaactttgg 3600
 aggcacttta ccgtcaggat ccagaatccc ttccctttcg tcaacctgtg gaccctcagc 3660
 ttttaggaat ccctgattac ttgatattg tgaagagccc catggatctt tctaccatta 3720
 agaggaagtt agacactgga cagtatcagg agccctggca gtatgtgat gatattggc 3780

 ttatgttcaa taatgcctgg ttatataacc ggaaaacatc acgggtatac aaatactgct 3840
 ccaagctctc tgaggtcttt gaacaagaaa ttgaccagt gatgcaaagc cttggatact 3900
 gttgtggcag aaagttggag ttctctccac agacactgtg ttgctacgac aacagttgt 3960
 gcacaatacc tctgatgcc acttattaca gttaccagaa caggtatcat ttctgtgaga 4020
 agtgtttcaa tgatccaa ggggagagcg ttctttggg ggatgaccct tcccagcctc 4080
 aaactacaat aaataaagaa caattttcca agagaaaaaa tgacacactg gatcctgaac 4140
 tgtttgttga atgtacagag tgcggaagaa agatgcatca gatctgtgtc cttcaccatg 4200

 agatcatctg gcctgctgga ttctctgtg atggctgttt aaagaaaagt gcacgaacta 4260
 ggaaagaaaa taagttttct gctaaaaggt tgccatctac cagacttggc acctttctag 4320
 agaatcgtgt gaatgacttt ctgaggcgac agaatcacc tgagttagga gaggtcactg 4380

ttagagtagt tcatgcttct gacaaaaccg tggaaagtaa accaggcatg aaagcaaggt 4440
 ttgtggacag tggagagatg gcagaatcct ttccataaccg aaccaaagcc ctctttgcct 4500
 ttgaagaaat tgatggtgtt gacctgtgct tctttggcat gcatgttcaa gagtatggct 4560
 ctgactgccc tccaccaaac cagaggagag tatacatatc ttacctgat agtgttcatt 4620

 tcttccgtcc taaatgcttg aggactgcag tctatcatga aatcctaatt ggatatttag 4680
 aatatgtcaa gaaattaggt tacacaacag ggcatatttg ggcatgtcca ccaagtgagg 4740
 gagatgatta tatcttccat tgccatcctc ctgaccagaa gataccaag cccaagcgac 4800
 tgcaggaatg gtacaaaaaa atgcttgaca aggctgtatc agagcgtatt gtccatgact 4860
 acaaggatat ttttaacaa gctactgaag atagattaac aagtgcaaag gaattgcctt 4920
 atttcgaggg tgatttctgg cccaatgttc tggaaagaaag cattaaggaa ctggaacagg 4980
 aggaagaaga gagaaaacga gaggaaaaca ccagcaatga aagcacagat gtgaccaagg 5040

 gagacagcaa aaatgctaaa aagaagaata ataagaaaac cagcaaaaat aagagcagcc 5100
 tgagtggggg caacaagaag aaaccggga tgcccaatgt atctaacgac ctctcacaga 5160
 aactatatgc caccatggag aagcataaag aggtcttctt tgtgatccgc ctcatgtctg 5220
 gccctgctgc caactcctg cctcccattg ttgatcctga tcctctcacc cctgctgac 5280
 tgatggatgg tgggatgcg tttctcacgc tggcaaggga caagcacctg gatttctctt 5340
 cactccgaag agcccattgg tccaccatgt gcatgctggt ggagctgcac acgcagagcc 5400
 aggaccgctt tgtctacacc tgcaatgaat gcaagcacca tgtggagaca cgctggcact 5460

 gtactgtctg tgaggattat gacttgtgta tcacctgcta taactactaa aacctgacc 5520
 acaaaatgga gaaactaggg cttggcttag atgatgagag caacaaccag caggctgcag 5580
 ccaccagag cccagcgat tctcggcgc tgagtatcca gcctgcatc cagtctctgg 5640
 tccatgcttg ccagtgtcgg aatgccaatt gctcactgcc atcctgccag aagatgaagc 5700
 gggttgtgca gcataccaag ggttgcaaac ggaaaaccaa tggcgggtgc cccatctgca 5760
 agcagctcat tgcctctgc tgctaccatg ccaagcactg ccaggagaac aaatgcccgg 5820
 tgccgttctg cctaaacatc aagcagaagc tccggcagca acagctgcag caccgactac 5880

 agcaggccca aatgcttcgc aggaggatgg ccagcatgca gcggactggt gtggttgggc 5940
 agcaacaggg cctccctcc cccactctg ccaactcaac gacaccaact ggccaacagc 6000
 caaccacccc gcagacgccc cagcccactt ctgagcctca gcctaccct cccaatagca 6060
 tgccacccta cttgccagg actcaagctg ctggccctgt gtcccagggt aaggcagcag 6120
 gccaggtgac cctccaacc cctcctcaga ctgctcagcc acccttcca gggccccac 6180
 ctgcagcagt ggaatggca atgcagattc agagagcagc ggagacgcag cgccagatgg 6240

cccacgtgca aatTTTTcaA agGccaatcc aacaccagat gcccccgatg actcccatgg 6300

 ccccatggg tatgaacca cctcccatga ccagaggtcc cagtgggcat ttggagccag 6360
 ggatgggacc gacagggatg cagcaacagc caccttgag ccaaggagga ttgcctcagc 6420
 cccagcaact acagtctggg atgccaaggc cagccatgat gtcagtggcc cagcatggtc 6480
 aacctttgaa catggctcca caaccagat tgggccaggt aggtatcagc cactcaaac 6540
 caggcactgt gtctcaaca gccttcaaaa accttttgcg gactctcagg tetcccagct 6600
 ctcccctgca gcagcaacag gtgcttagta tccttcacgc caacccccag ctgttgctg 6660
 cattcatcaa gcagcgggct gccaaagtatg ccaactctaa tccacaacce atccctgggc 6720

 agcctggcat gccccagggg cagccagggc tacagccacc taccatgcca ggtcagcagg 6780
 ggtccactc caatccagcc atgcagaaca tgaatccaat gcaggcgggc gttcagaggg 6840
 ctggcctgcc ccagcagcaa ccacagcagc aactccagcc acccatggga gggatgagcc 6900
 cccaggtca gcagatgaac atgaaccaca acacatgcc ttcacaattc cgagacatct 6960
 tgagacgaca gcaaatgatg caacagcagc agcaacaggg agcagggcca ggaataggcc 7020
 ctggaatgga caaccataac cagttccagc aacccaagg agttggctac ccaccacagc 7080
 agcagcagcg gatgcagcat cacatgcaac agatgcaaca aggaatatg ggacagatag 7140

 gccagcttc ccaggcctg ggagcagagg caggtgccag tctacaggec tatcagcagc 7200
 gactccttca gcaacagatg gggctcccctg ttcagcccaa cccatgagc ccccagcagc 7260
 atatgctccc aaatcagccc cagtccccac acctacaagg ccagcagatc cctaattctc 7320
 tctcaatea agtgcgctct ccccagcctg tcctttctcc acggccacag tcccagcccc 7380
 cccactccag tccttccca aggatgcagc ctcagccttc tccacaccac gtttccccac 7440
 agacaagttc cccacatct ggactggtag ctgccaggc caacccatg gaacaagggc 7500
 attttgccag cccggaccag aattcaatgc tttctcagct tgctagcaat ccaggcattg 7560

 caaacctcca tggtgcaagc gccacggacc tgggactcag caccgataac tcagacttga 7620
 attcaaacct ctacagagat aactagaca tactagag acaccttgta gtattttggg 7680
 agcaaaaaaa ttattttctc ttaacaagac tttttgact gaaaacaatt tttttgaatc 7740
 tttcgtagcc taaaagacaa ttttcttg aacacataag aactgtgcag tagccgtttg 7800
 tggtttaag caaacatgca agatgaacct gagggatgat agaatacaaa gaatatattt 7860
 ttgttatggc tggttaccac cagcctttct tccccttgt gtgtgtggtt caagtgtgca 7920
 ctgggaggag gctgaggcct gtgaagccaa acaatatgct cctgccttgc acctccaata 7980

ggttttatta ttttttttaa attaatgaac atatgtaata ttaatagtta ttatttactg 8040
 gtgcagatgg ttgacatttt tccctatttt cctcacttta tggaagagtt aaaacatttc 8100
 taaccagag gacaaaaggg gttaatgtta ctttaaaatt acattctata tatatataaa 8160
 tatatataaa tatatattaa aataccagtt ttttttctct gggtgcaaag atgttcattc 8220
 ttttaaaaaa tgtttaaaaa aaaaaaaaaa ctgcctttct tcccctcaag tcaacttttg 8280
 tgctccagaa aattttctat tctgtaagtc tgagcgtaaa acttcaagta ttaaaataat 8340
 ttgtacatgt agagagaaaa atgacttttt caaaaatata caggggcagc tgccaaattg 8400

atgtattata tattgtggtt tctgtttctt gaaagaattt ttttcgttat ttttacatct 8460
 aacaaagtaa aaaaattaaa aagagggtta gaaacgattc cgggggatg attttaacat 8520
 gcaaaatgtc cctgggggtt tcttctttgc ttgtttctt cctccttacc ctacccccca 8580
 ctcacacaca cacacacaca cacacacaca cacacacaca cacactttct ataaaacttg 8640
 aaaatagcaa aaacctcaa ctgttgtaaa tcatgcaatt aaagttgatt acttataaat 8700
 atgaactttg gatcactgta tagactgtta aatttgattt cttattacct attgttaaat 8760
 aaactgtgtg agacagaca 8779

<210> 23

<211> 8701

<212> DNA

<213> Homo sapiens

<400> 23

gagaaggagg aggacagcgc cgaggaggaa gaggttgatg gcggcggcgg agctccgaga 60
 gacctcggct gggcaggggc cggccgtggc gggccgggga ctgcgcctct agagccgca 120
 gttctcggga attcgcgca gcggacgcgc tcggcgaatt tgtgctcttg tgccctctc 180
 cgggcttggg cccaggcccc gccctcgc ca ttgccctta cttttctat cgagtccgca 240
 tccctctcca gccactgca cccggcgaag agaaaaagga acttccccca cccctcggg 300
 tgccgtcgga gccccccagc ccaccctgg gtgcggcgcg gggaccccgg gccgaagaag 360

agatttctg aggattctgg ttttctcgc ttgtatctcc gaaagaatta aaaatggccg 420
 agaatgtggt ggaaccgggg cgccttcag ccaagcggcc taaactctca tctccggccc 480
 tctcggcgtc cgcagcgat ggcacagatt ttggctctct atttgacttg gagcagcact 540
 taccagatga attaatcaac tctacagaat tgggactaac caatggtggt gatattaatc 600
 agcttcagac aagtcttggc atggtacaag atgcagcttc taaacataaa cagctgtcag 660
 aattgctcgc atctggtagt tcccctaacc tcaatatggg agttggtggc ccaggicaag 720

tcatggccag ccaggcccaa cagagcagtc ctggattagg tttgataaat agcatggtca 780

 aaagcccaat gacacaggca ggcttgactt ctccaacat ggggatgggc actagtggac 840
 caaatcaggg tectacgcag tcaacaggta tgatgaacag tccagtaaat cagcctgcca 900
 tgggaatgaa cacagggatg aatgcgggca tgaatcctgg aatgttggct gcaggcaatg 960
 gacaagggat aatgcctaata caagtcatac acggttcaat tggagcaggc cgaggcgcac 1020
 agaatatgca gtacccaaac ccaggcatgg gaagtgcctg caacttactg actgagcctc 1080
 ttcagcaggg ctctccccag atgggaggac aaacaggatt gagaggcccc cagcctctta 1140
 agatgggaat gatgaacaac cccaatcctt atggttcacc atatactcag aatcctggac 1200

 agcagattgg agccagtggc cttggtctcc agattcagac aaaaactgta ctatcaaata 1260
 acttatctcc atttgctatg gacaaaaagg cagttcctgg tggaggaatg cccaacatgg 1320
 gtcaacagcc agccccgcag gtccagcagc caggcctggt gactccagtt gcccaaggga 1380
 tgggttctgg agcacataca gctgatccag agaagcgaac gctcatccag cagcagcttg 1440
 ttctcctttt gcatgctcac aagtgccagc gccgggaaca ggccaatggg gaagtgagge 1500
 agtgcaacct tcccactgt cgcacaatga agaatgtcct aaaccacatg acacactgcc 1560
 agtcaggcaa gtcttgccaa gtggcacact gtgcatcttc tcgacaaatc atttcacact 1620

 ggaagaattg tacaagacat gattgtctg tgtgtctccc cctcaaaaat gctggtgata 1680
 agagaaatca acagccaatt ttgactggag cacccgttgg acttggaat cctagctctc 1740
 taggggtggg tcaacagtct gccccaacc taagcactgt tagtcagatt gatcccagct 1800
 ccatagaaag agcctatgca gctcttggac taccctatca agtaaatcag atgccgacac 1860
 aacccaggt gcaagcaaa aaccagcaga atcagcagcc tgggcagtct cccaaggca 1920
 tgcggcccat gagcaacatg agtgcctagc ctatgggagt aaatggaggt gtaggagttc 1980
 aaacgccgag tcttctttct gactcaatgt tgcattcagc cataaattct caaaacccaa 2040

 tgatgagtga aaatgccagt gtgccctccc tgggtcctat gccaacagca gctcaacat 2100
 ccactactgg aattcggaag cagtggcacg aagatattac tcaggatctt cgaaatcatc 2160
 ttgttcacaa actcgtccaa gccatattc ctacgccgga tcctgctgct ttaaaagaca 2220
 gacggatgga aaacctagt gcatatgctc gaaaagtga aggggacatg tatgaatctg 2280
 caaacaatcg agcggaaatc taccacctc tagctgagaa aatctataag atccagaaag 2340
 aactagaaga aaaacgaagg accagactac agaagcagaa catgctacca aatgctgcag 2400
 gcatggttcc agtttccatg aatccagggc ctaacatggg acagccgcaa ccaggaatga 2460

cttctagttt gaatcaattt ggccagatga gcatggccca gcccctatt gtaccccggc 2520
 aaaccctcc tcttcagcac catggacagt tggctcaacc tggagctctc aaccgccta 2580
 tgggctatgg gcctcgtatg caacagcctt ccaaccaggg ccagttcctt cctcagactc 2640
 agttcccatc acagggaaatg aatgtaacaa atatcccttt ggctccgtcc agcggtaaac 2700
 ctccagtgtc tcaagcaca atgtctagtt ctctctgccc ggtgaactct cctataatgc 2760
 ctccagggtc tcaggggagc cacattcaact gtccccagct tctcaacca gctcttcac 2820
 agaattcacc ctgcctgta cctagtcgia ccccccccc tcaccatact ccccaagca 2880

 taggggctca gcagccacca gcaacaacaa ttccagcccc tgttctaca cctcctgcca 2940
 tgccacctgg gccacagtcc caggtcttac atccccctcc aaggcagaca cctacaccac 3000
 caaacacaca acttcccaa caagtgcagc ctctacttcc tgctgcacct tctgtgacc 3060
 agccccagca gcagcctgc tcaacagcaga gcacagcagc gtctgttctt acccaacag 3120
 caccgtgct tctcccgag cctgcaactc cactttcca gccagctgta agcattgaag 3180
 gacaggtatc aaatctcca tctactagta gcacagaagt gaattctcag gccattgctg 3240
 agaagcagcc ttcccaggaa gtgaagatgg aggccaaaat ggaagtggat caaccagaac 3300

 cagcagatac tcagccggag gatatttcag agtctaaagt ggaagactgt aaaatggaat 3360
 ctaccgaaac agaagagaga agcactgagt taaaaactga aataaaagag gaggaagacc 3420
 agccaagtac ttcagctacc cagtcatctc cggctccagg acagtcaaag aaaagattt 3480
 tcaaaccaga agaactacga caggcactga tgccaacttt ggaggcactt taccgtcagg 3540
 atccagaatc ccttcccttt cgtcaacctg tggaccctca gcttttagga atccctgatt 3600
 actttgatat tgtaagagc cccatggatc tttctacat taagaggaag ttagacactg 3660
 gacagtatca ggagccctgg cagtatgtcg atgatattg gcttatgttc aataatgcct 3720

 ggttatataa ccggaaaaca tcacgggtat acaaaactg ctccaagctc tetgaggtct 3780
 ttgaacaaga aattgacca gtgatgcaa gccttgata ctgttgtggc agaaagtgg 3840
 agttctctcc acagacactg tgttctacg gcaaacagtt gtgcacaata cctcgtgatg 3900
 ccacttatta cagttaccag aacaggtatc atttctgtga gaagtgttc aatgagatec 3960
 aaggggagag cgtttctttg ggggatgacc ctcccagcc tcaactaca ataaataaag 4020
 aacaatttc caagagaaa aatgacacac tggatcctga actgtttgtt gaatgtacag 4080
 agtgcggaag aaagatgcat cagatctgtg tccttcacca tgagatcatc tggcctgctg 4140

 gattcgtctg tgatggctgt ttaaagaaaa gtgcacgaac taggaaagaa aataagtttt 4200
 ctgctaaaag gttgcatct accagacttg gcacctttct agagaatcgt gtgaatgact 4260
 ttctgaggcg acagaatcac cctgagtcat gagaggtcac tgttagagta gttcatgctt 4320

ctgacaaaac cgtggaagta aaaccaggca tgaagcaag gtttggac agtggagaga 4380
 tggcagaatc ctttccatac cgaaccaaag ccctctttgc ctttgaagaa attgatggtg 4440
 ttgacctgtg cttctttggc atgcatgttc aagagtatgg ctctgactgc cctccacca 4500
 accagaggag agtatacata tcttacctcg atagtgtca tttcttccgt cctaaatgct 4560

 tgaggactgc agtctatcat gaaatcctaa ttggatattt agaatatgtc aagaaattag 4620
 gttacacaac agggcatatt tgggcatgtc caccaagtga gggagatgat tatatcttcc 4680
 attgccatcc tctgaccag aagatacca agcccaagcg actgcaggaa tggtaaaaa 4740
 aatgcttga caaggctgta tcagagcgtg ttgtccatga ctacaaggat atttttaaac 4800
 aagctactga agatagatta acaagtgcaa aggaattgcc ttatttcgag ggtgatttct 4860
 ggcccaatgt tctggaagaa agcattaagg aactggaaca ggaggaagaa gagagaaaac 4920
 gagaggaaaa caccagcaat gaaagcacag atgtgaccaa gggagacagc aaaaatgcta 4980

 aaaagaagaa taataagaaa accagcaaaa ataagagcag cctgagtagg ggcaacaaga 5040
 agaaacccgg gatgccaat gtatctaacg acctctcaca gaaactatat gccaccatgg 5100
 agaagcataa agaggtcttc ttttgatcc gcctcattgc tggccctgct gccactccc 5160
 tgctcccat tgtgatcct gatcctctca tcccctgca tctgatggat ggtcgggatg 5220
 cgtttctcac gctggcaagg gacaagcacc tggagtctc ttcactccga agagcccagt 5280
 ggtcccat gtgcatctg gtggagctgc acacgcagag ccaggaccgc tttgtctaca 5340
 cctgcaatga atgcaagcac catgtggaga cacgctggca ctgtactgtc tgtgaggatt 5400

 atgacttgtg taccactgc tataacacta aaaacatga ccacaaaatg gagaaactag 5460
 gccttggctt agatgatgag agcaacaacc agcaggctgc agccaccag agcccaggcg 5520
 attctgccg cctgagtatc cagcgtgca tccagtctct ggtccatgct tgccagtgtc 5580
 ggaatgcaa ttgctcactg ccatctgcc agaagatgaa gcggttggc cagcatacca 5640
 aggttggca acggaacc aatggcgggt gcccatctg caagcagctc attgccctct 5700
 gctgtacca tgccaagcac tgccaggaga acaaatgccc ggtgccgttc tgctaaaca 5760
 tcaagcagaa gctccggcag caacagctgc agcaccgact acagcaggcc caaatgcttc 5820

 gcaggaggat ggccagcatg cagcggactg gtgtggttgg gcagcaacag ggcctcctt 5880
 cccccactcc tgccactcca acgacaccaa ctggccaaca gccaaccacc cgcgagacgc 5940
 ccagcccac ttctcagct cagcctacc ctccaatag catgccacc tacttgccca 6000
 ggactcaage tctggcct gtgtcccagg gtaaggcagc aggcagggtg accctcca 6060
 cccctctca gactgtcag ccccccttc cagggecccc acctgcagca gtggaaatgg 6120
 caatgcagat tcagagagca gcggagacgc agcgcagat ggcccacgtg caaattttc 6180

aaaggccaat ccaacaccag atgccccga tgactcccat ggccccatg ggtatgaacc 6240

 cacctcccat gaccagaggt cccagtgggc atttggagcc agggatggga cgcacaggga 6300
 tgcagcaaca gccaccctgg agccaaggag gattgcctca gccccagcaa ctacagtctg 6360
 ggatgccaag gccagccatg atgtcagtgg cccagcatgg tcaacctttg aacatggctc 6420
 cacaaccagg attgggccag gtaggtatca gccactcaa accaggcact gtgtctcaac 6480
 aagccttaca aaaccttttg cggactctca ggtctcccag ctctcccctg cagcagcaac 6540
 aggtgcttag tatccttcac gccaaccccc agctgttggc tgcatcctc aagcagcggg 6600
 ctgccaagta tgccaactct aatccacaac ccatccctgg gcagcctggc atgccccagg 6660

 ggagccagg gctacagcca cctaccatgc caggtcagca ggggtccac tccaatccag 6720
 ccatgcagaa catgaatcca atgcaggcgg gcgttcagag ggctggcctg ccccagcagc 6780
 aaccacagca gcaactccag ccaccatgg gagggatgag ccccagcct cagcagatga 6840
 acatgaacca caacaccatg ccttcacaat tccgagacat cttgagacga cagcaaatga 6900
 tgcaacagca gcagcaacag ggagcagggc caggaatagg ccctggaatg gccaaccata 6960
 accagttcca gcaaccccaa ggagttggct acccaccaca gcagcagcag cggatgcagc 7020
 atcacatgca acagatgcaa caaggaaata tgggacagat aggccagctt ccccaggcct 7080

 tgggagcaga ggcaggtgcc agtctacagg cctatcagca gcgactcctt cagcaacaga 7140
 tggggtcccc tgttcagccc aaccccatga gccccagca gcatatgctc ccaaatcagg 7200
 cccagtcccc acacctaca ggccagcaga tcctaattc tctctcaat caagtgcct 7260
 ctccccagcc tgtcccttct ccacggccac agtcccagcc ccccactcc agtccctccc 7320
 caaggatgca gcctcagcct tctccacacc acgtttcccc acagacaagt tcccacatc 7380
 ctggactggt agctgccag gccaaccca tggacaagg gcattttgcc agccccgacc 7440
 agaattcaat gctttctcag cttgctagca atccaggcat ggcaaacctc catggtgcaa 7500

 gcgccacgga cctgggactc agcaccgata actcagactt gaattcaaac ctctcacaga 7560
 gtacactaga catacactag agacacctg tagtatattg ggagcaaaaa aattattttc 7620
 tcttaacaag actttttgta ctgaaaacaa ttttttgaa tctttcgtag cctaaaagac 7680
 aattttcctt ggaacacata agaactgtgc agtagccgtt tgtggtttaa agcaaacatg 7740
 caagatgaac ctgagggatg atagaataca aagaatata tttgtttatg gctggttacc 7800
 accagccttt ctccccctt gtgtgtgtgg ttcaagtgtg cactgggagg aggctgaggc 7860
 ctgtgaagcc aaacaatag ctctgcctt gcacctcaa taggttttat tattttttt 7920

aaattaatga acatatgtaa tattaatagt tattatttac tggcgcagat ggttgacatt 7980
 tttccctatt ttctcactt tatggaagag ttaaacatt tctaaccag aggacaaaag 8040
 gggttaatgt tactttaaaa ttacattcta tatatatata aatatatata aatatatatt 8100
 aaaataccag tttttttct ctgggtgcaa agatgttcat tcttttaaaa aatgtttaaa 8160
 aaaaaaaaa aactgccttt cttccctca agtcaacttt tgtgctccag aaaatttct 8220
 attctgtaag tctgagcgta aaacttcaag tattaaaata atttgtacat gtagagagaa 8280
 aatgacttt ttcaaaaata tacaggggca gctgccaaat tgatgtatta tatattgtgg 8340

tttctgttc ttgaaagaat tttttcgtt atttttacat ctaacaaagt aaaaaatta 8400
 aaaagagggt aagaacgat tccggtggga tgattttaac atgcaaaatg tccctggggg 8460
 tttctcttt gcttgcttc ttctcctta cctaccccc cactcacaca cacacacaca 8520
 cacacacaca cacacacaca cacacactt ctataaaact tgaaaatagc aaaaaccctc 8580
 aactgttgta aatcatgcaa ttaaagtga ttacttataa atatgaactt tggatcactg 8640
 tatagactgt taaattgat ttcttattac ctattgttaa ataaactgtg tgagacagac 8700
 a 8701

<210> 24

<211> 5867

<212> DNA

<213> Homo sapiens

<400> 24

actgcctcgc ctctctgtgc ctcttcaggt catcgcttgc tctcgttccc aggctttggc 60
 ttctagtgga cgagaatcac cgagctcgc gggctggatg ctgaccgcc ggaccagcac 120
 ctaggcgggc gggagctgtg cggcccaggg ttcgcgggg ccgggtagag gctcgagctg 180
 ggacccccga gcgtgaacct cggagccggc ggcgctgggg ccagaggggc gggagcccc 240
 agggaggcgg atctgggcc cgagaaggac accgcctgg atttgcccc taggcccggc 300
 ccgggccct cgggagcaga acagccttg tgaggtggac aggaggggac ctcgcgagca 360

gacgcgcgcg ccagcgacag cagcccgcc cggcctctcg ggagccgtgg ggcagaggct 420
 gcggagcccc aggagggggc cagtgtcatt caaagatgtg gctgtggatt tcaccagga 480
 ggagtggcgg caactggacc ctgatgagaa gataacatac ggggatgtga tgttgagaa 540
 ctacagccat ctagtttct tggcttatga ggtggcaaca tcttgtactt cggagattct 600
 gaagccgagc aacttgccca agtccttctt cttttcccat taacaagata tgatatcacc 660
 aagccaaacg tcatcattaa gttggagcag ggagaggagc tgtggataac gggaggtgaa 720

ttccatgtc aacatagtc tgggatcgtg ggactttacc agatcggttt gtaataacac 780

 ctggaagacg ctatccgac catcagacc agtattcctg tccggggta cttcccacag 840
 tgtgctggaa tggttatcac aagaaggctg tctgtceccc tgcgaactcc aggatggtgt 900
 gtagcccagt gactgtgagg atcgcccctc ctgacagaag attttcacgt tctgcgatac 960
 cagagcagat aatcagctca acactgtcgt caccatcaag taatgcccc gacctatgtg 1020
 caaaggagac tgtactgagt gccctcaaag agaagaagaa gaaaaggaca gtggaggaag 1080
 aagaccaa atccttcat gggcaggaaa ataaaagaag gcgccatgat agcagtggca 1140
 gtggacattc agcattttag cccctggtgg ccagtggagt ccccgttct tttgtgctca 1200

 agcctgggtc tctgaagaga ggcctcaatt ctgagagctc agatgaccac ttgaataaga 1260
 gatcccgaag ctcttccatg agctccttga caggcgctta cacaagtggc atccctagct 1320
 ccagccgcaa tgccattacc agttcttaca gctccactcg aggcatctca cagctgtgga 1380
 agagaaatgg cccagttca tcaccttct ctagcccagc ctcatcccgc tcccagacac 1440
 cggagaggcc agcaaagaaa ataagagaag aagagctgtg tcatcattcc agttcttcaa 1500
 ctccattggc agcagacaag gagtcccagg gagaaaaggc tgcagataca accccaagga 1560
 agaaacaaaa ctgaattct cagtctacac ctggcagctc tgggcagcgt aagcggaaag 1620

 ttcagctgct gccttctcgg cgagggaac agctgacctt gcctccacct ccccagcttg 1680
 gctattcgat cactgccgag gacctagact tagagaagaa ggcttcatta cagtgttca 1740
 accaggcctt ggaggacaag agtgatgctg cctcgaactc tgtcactgag accccaccta 1800
 ccaactcagcc ttcatttacc tttacctgc ctgctgctgc aactgcctcc ccaccacct 1860
 cctcctggc cccaagcacc aacctactgt tagagagctt gaagaagatg cagactcccc 1920
 cgagcctgcc accctgcccc gaatctgctg gacgagcaac cactgaggcc ctctcacctc 1980
 caaagacacc cagcctccta ccccgtggtg gtttatcaca gtcagggcg ccagggtgc 2040

 tccccagccc ctctttgac tccaaacccc cgaccacttt gctggggctg atccctgctc 2100
 catccatggt accagccact gacaccaagg cacctccaac ccttcaggca gagacggcta 2160
 ccaaacccca agccacatct gcccgtccc ccgccccaa gcaaagcttc ctgtttggaa 2220
 cacagaacac ctacattcc agcctgccc ccctgctgc atcttcagca tctccatgt 2280
 tcaagccat tttcaggtt ccaccaaga gtgagaagga aggcctcaca ccgctggcc 2340
 cttagctctc agccacagcg cctccagct cctccctccc cagcaccacc agcaccacag 2400
 cccgacctt ccagcctgct tttagcagca tggggccacc tgcatctgtg cccttgctg 2460

ctcccttctt caagcagaca actactcccg ccaactgctcc caccacaact gccccgctct 2520
tcaactggcct ggcagcggc acctctgctg tggctcccat cacctctgcc agtccatcca 2580
cagactctgc ttcgaagcct gcgtttggct ttggcataaa cagtgtgagc agcagcagtg 2640
tgagtaccac gaccagcacc gccactgccg cctcacagcc tttcctcttc ggggcgcccc 2700
aggcctctgc tgccagcttc accccggcca tgggctccat attccagttt ggcaaacctc 2760
ctgccttgcc cacaaccacc acagtacca ccttcagcca gtccttgccc actgcccgtgc 2820
caacggccac cagcagcagc gctgccgact ttagtggttt tggcagcacc ctgccacct 2880

ccgccccggc caccagcagc cagcccactc tgacgttcag taacacgagc acccccactg 2940
tcaacattcc ctttggctca agcgcgaagt ccccgtccc atcatatccg ggagccaacc 3000
cccagcccgc atttggggcc gctgaggggc agccaccggg ggccgccaag ccagccctta 3060
ccccagctt tggcagctct ttcacttttg gaaactctgc agccccggcc cgggctactg 3120
caccacacc tgcacctgcg tccacgatca agatcgtgcc tgcgcacgtg cctacgcca 3180
tccagcctac ctttggcggg gccacgcact cggcgtttgg attgaaagcc acggttccg 3240
ccttcggcgc tcccgccagc tcacagcccg ctttggcgg ctccactgct gtcttctct 3300

tcggtgcagc caccagctcc ggctttggag ccaccacca gaccgccagc agcgggagca 3360
gcagctcggg gtttggcagc acaacacat cacccttac gtttgggggt tcggcagccc 3420
ccgctggcag tgggagcttt gggatcaacg tggccacccc aggtccagc gccaccaccg 3480
gagctttcag ctttggagca ggacagagtg ggagcacagc cacctccacc cccttcacag 3540
ggggcttagg tcagaacgcc ctgggcacca ccggccagag cacaccgttt gccttcaacg 3600
tgggcagcac aactgagagc aaacctgtgt ttggaggcac cgccacccc acctttggtc 3660
agaacacccc tgcgcctgga gtgggcacat cgggcagcag cctctccttt ggggcatctt 3720

cagcaccgc ccaaggettt gttggtgttg gaccgttcgg atcggcggcc ctttcatttt 3780
ccattggtgc gggatccaag accccagggg ctgcacagcg actgcaggcc cgaaggcagc 3840
acaccgcaa aaagtagcct ttgtcccctg tcctgttcc ccccacct tcctaaatc 3900
tggaccttg cactgctag aaagagcctt ggacccttcc agctgcgtaa agcaaaccta 3960
ccccgatct ctggcttcag ccgccagggg gcagtggcag ccctggggcc ctttccttc 4020
tggaggaagc acaagcctca gggaagggga agcaggatgc ggagggccaa agccccggac 4080
ctctacttga acagttccac tggggaggct ggagaactaa ggacctgtac atagtgtccg 4140

ctgcctgac tcccgttag cgcaccctta ggcagcgccc ccttccacct ttccccgaga 4200
gcccgtctgc ctggaggggg cagggtccag cccgcctgga tcggtggtgt gcacctgatg 4260
ggatttggga aatgggttat ccctaaagct ttatcttctg ttgcttagct gtgagaagtg 4320

gttctcttcc tctggctcct tctggggact ctgtttcccc atttcttgct gctgtgtccc 4380
 tcaccggttc cttgcaggat tcctccttt ttaaagccc ttgaatctag ctttgccctg 4440
 gagaccccag tgggtgctgc tctgcccgtt ttcttctgc caagcctgaa tcaatgtttc 4500
 atctccaacc ctctgccagt ttggcccctc agagcttggg ggctcaagac tgtagcctg 4560

gcagagccag ggggaaggga agaagctctt ggagcaggca ggatgccac tctgtcttca 4620
 gctgcctcct cgcccagcta ccctttggcc ccattgggccc ctctgtctgccc tetccaggat 4680
 tgtatgtttc aagccttctc ctgtgttctt ttgtctgacg ctctgtgtat tgctctttga 4740
 atcgagtttg gaggaagagt tgagtgtat gaggggcggc atgttggttag tgccggactt 4800
 cctgtttcaa gttttctggg gcctcgctaa ttgaatgtgg aaagtagcac cacttgacgg 4860
 ctacaagtgc cgactcctga attttccat ggtgttctga cttcaagggc tggcagccag 4920
 ggagaatggg ccaggggaa gcaaagacct ctccctctg cggtttctgt cccactaac 4980

tgacctcact ggaggctacg tcacccaaag tagatgttag aaaacctaaa ttaatgaacc 5040
 atatttttaa aatcctattt ttcccaaaca gggccctctg cagcccatcc tttccttccg 5100
 tccttctgaa accacatacc ccaggcccaa ggccttctg gtcacgcca acctctttgg 5160
 gagaagtatg aatgcgtgtg tctaaattaa aagaaaaaaaa tatttaaagc ttttttaaca 5220
 aaaaattatt ttgtattta agctaaattg ctttttaaat tccttcaagc ttggttcatt 5280
 gaggtgggta agtataaatg ctattaacta ggaattagct gtatagttaa gttatgcctg 5340
 tgcaagaag aggctcaaat gctgtccccg gcagctttcc tgggggacta gagctcctc 5400

tggccatggt atatgaaatg taattcttat ttataaata atgtgatgta gatgtaaccg 5460
 gtgccccctc cccgttctga ctgagggcga gtgttatacc tggctgcgtg tgcagctgag 5520
 gagggagtga acctcaagcc taaatacctg ttaggattgg agggctctggg tggcctggg 5580
 cctagcaate aagcttctac ctgtacctta tgaaggttag acctcctag tgtcagtacc 5640
 tgagcttgtt tacctcagtt ccgcaggcag gacagccggg ccgggaacc ttagtgagag 5700
 tgagtgtgga tgtgtacagt acacacactg gacggcagcg ggaggctggg actttccatt 5760
 acaaatagag acttcattcc tgttgagtct agttggaatt tttagtatga atgtgagatt 5820

tttctctgc ttgtgacatt aagaataaaa aactgtgatc tatcgta 5867

<210> 25

<211> 2242

<212> DNA

<213> Homo sapiens

<400> 25

ggagccgggg gacggcgaca gcgggtcggc gggccgcagg aggggggtcat gggtaaagac 60
 tactaccaga cgttgggctt gccccgcggc gcgtcggacg aggagatcaa gcgggcctac 120
 cgcccagg cgctgcgcta ccaccggac aagaacaagg agcccggcgc cgaggagaag 180
 ttcaaggaga tcgtgaggc ctacgacgtg ctcagcgacc cgcgcaagcg cgagatcttc 240
 gaccgctacg gggaggaagg cctaaagggg agtggcccca gtggcggtag cggcgggtgt 300

 gccaatggta cctctttcag ctacacattc catggagacc ctcatgcat gtttgcctgag 360
 ttcttcgggtg gcagaaatcc ctttgacacc tttttgggc agcggaacgg ggaggaaggc 420
 atggacattg atgaccatt ctctggcttc cctatgggca tgggtggctt caccaactg 480
 aactttggcc gctcccgtc tgcccaagag cccgccgaa agaagcaaga tccccagtc 540
 acccagacc ttgagtctc cttgaagag atctacagcg gctgtaccaa gaagatgaaa 600
 atctcccaca agcggctaaa ccccgacgga aagagcattc gaaacgaaga caaatattg 660
 accatcgaag tgaagaagg gtggaagaa ggaacaaaa tcactttccc caaggaagga 720

 gaccagacct ccaacaacat tccagctgat atcgtctttg ttttaaagga caagccccac 780
 aatatcttta agagagatgg ctctgatgtc atttactctg ccaggatcag cctccgggag 840
 gctctgtgtg gctgcacagt gaactcccc actctggacg gcaggacgat acccgtcgt 900
 ttcaaagatg ttatcaggcc tggcatgcgg cgaaaaagttc ctggagaagg cctccccctc 960
 cccaaaaac cgagaaaagc tggggacctc attattgagt ttgaagtgat cttccccgaa 1020
 aggatcccc agacatcaag aaccgtactt gacgagttc ttccaatata gctatctgag 1080
 ctccccagg actgaccagg gacctttcca gagctcaagg atttctggac ctttctacca 1140

 gttgtggacc atgagagggt gggagggcc agggagggt ttcgtactgc tgaatgttt 1200
 ccagagcata tattacaatc tttcaaagtc gcacactaga cttcagtgtt ttttcgact 1260
 ataggcacc aggtgggtgg aacagcagga aaaggcattc cagtctgccc cactgggtct 1320
 ggcagccctc ccgggatggg cccacatcca cctccagtcc ctggccaggg gtgagaggca 1380
 gaccagcaga tggacttgat cctctgtgt ctttttgcct ctggctggt gataatgtca 1440
 acctgcagtc ttgattccca gacctgtac actcctctt tctgtttgt tgatecagttt 1500
 gtgctttatt ctgtattgt ctccatgtc ttgctcttct cctggagaat tctgtcttct 1560

 ctttggccat ctcaaatga gaacctaac tattctgca gaactgcctg gttggcgtcc 1620
 acaagcaata cctctcgtc cagcaggacc aaggagcca gcctccagtg agtgactcca 1680
 gcaagtgcag ccacctctc cttgatggtc tgggagcctg gcctcagcaa gggccttcc 1740
 tgacctctgg ctccagtga gctgaatgtc ctcaattgt gggtcacact ctttacattt 1800
 ctgtaaggca atcttggc acgtggggct taccagtggc ccaggttaatt tttgtttca 1860

tggactatgg acicctttcaa agggatctga tccttttgaa ttttgcacag ccctagatac 1920
aatccctttt gataaaaggg tctttgcttc tgattacagg agcactgtgg aacgtctgta 1980

aatatgtttt tataattcca tgtatagttg gtgtacactc aaaacctgtc cccggcagcc 2040
agtgcctctc gtatagggcc ataatggaat tctgaagaaa tcttggggag ggaaggggag 2100
ttggaacaaa tgtctgttcc ctggaggcca gtccagtctc cagaccttta gactcattgt 2160
aagttgccac tgccaacatg agaccaaagt gtgtgactag tcaatgaagt gcgacagcat 2220
taaagactga tgctaaacct ca 2242

<210> 26
<211> 2270
<212> DNA
<213> Homo sapiens
<400> 26

gtgcatgggg gtgggggagc cgggggtgac gacggggacg gcgcggcgga gcccgtgcg 60

gacccgggct cacctgggct cggcccgccg ggtcccgggg gcggcgccctc cggctcagct 120
gcggggcgag gggttgtgaa tgcaggagcc gaccccgctc gtgggcttgg gggctgggtt 180
gggataatc cggggaagtg atgacctggc cccggcaggg cacgcaggag gcagcggccg 240
cccaggtccg gagagcgggg ccgctcggga gggcctaaag gggagtggcc ccagtggcgg 300
tagcggcggg ggtgccaatg gtacctctt cagctacaca ttccatggag acctcatgc 360
catgtttgct gatttctcg gtggcagaaa tccctttgac acctttttg ggcagcggaa 420
cggggaggaa ggcatggaca ttgatgacc attctctggc ttcctatgg gcatgggtgg 480

cttcaccaac gtgaactttg gccgtcccc ctctgcecaa gagcccgccc gaaagaagca 540
agatceccca gtcacceacg accttcgagt ctcccttgaa gagatctaca gcggctgtac 600
caagaagatg aaaatctccc acaagcggct aaaccccgac ggaaagagca ttcgaaacga 660
agacaaaata ttgacctcg aagtgaagaa ggggtggaag gaaggaacca aatcacttt 720
ccccaaggaa ggagaccaga cctccaacaa cattccagct gatacgtct ttgttttaa 780
ggacaagccc cacaatatct ttaagagaga tggtctctgat gtcatttacc ctgccaggat 840
cagcctccgg gaggtctgt gtggctgcac agtgaacgtc cccactctgg acggcaggac 900

gataccgctc gtattcaaag atgttatcag gcctggcatg cggcgaagag ttcttgagga 960
aggcctcccc ctccccaaa caccgagaa acgtggggac ctcatattg agtttgaagt 1020
gatcttcccc gaaaggattc cccagacatc aagaaccgta cttgagcagg ttcttccaat 1080
atagctatct gagctcccca aggactgacc agggaccttt ccagagctca aggatttctg 1140

gacctttcta ccagttgtgg accatgagag ggtgggaggg cccagggagg gctttcgtac 1200
tgctgaatgt tttccagagc atatattaca atctttcaaa gtcgcacact agacttcagt 1260
ggtttttcga gctatagggc atcaggtggt gggaacagca ggaaaaggca ttccagtctg 1320

ccccactggg tctggcagcc ctcccgggat gggcccat ccacctccag tccttgcca 1380
ggggtgagag gcagaccagc agatggactt gatccctctg tgtctttttg cttctggctg 1440
gtagataatg tcaacctgca gtcttgattc ccagaccctg tacactcctc cttttctgtt 1500
gtgtgatcag tttgtgcttt attctgtatt tgtctccat gtcttgctct tctcctggag 1560
aattctgtct tctctttggc catctcaaat tgagaaccta aactattcct gcagaactgc 1620
ctggttggcg tccacaagca atacctctcg ttccagcagg accaaggag ccagcctcca 1680
gtgagtgact ccagcaagtg cagccacctc tccttgatg gtctgggagc ctggcctcag 1740

caagggcct tectgacctc tggctccagt gaagctgaat gteectactt tgtgggtcac 1800
actctttaca tttctgtaag gcaatcttgg cacacgtggg gcttaccagt ggcccaggta 1860
atTTTTgtt tcatggacta tggactcttt caaagggatc tgatcctttt gaattttgca 1920
cagccctaga tacaatccct tttgataaaa ggtcttttgc tctgattac aggagcactg 1980
tggaaagtct gtaaatatgt ttttataatt ccatgtatag ttggtgtaca ctcaaacct 2040
gtccccgca gccagtgtc tctgtatagg gccataatgg aattctgaag aaatcttggg 2100
gaggggaagg gagttggaac aaatgtctgt tcctggagg ccagtcagct gctcagacct 2160

ttagactcat tgtaagtgc cactgccaac atgagaccaa agtgtgtgac tagtcaatga 2220
agtgcgacag cattaaagac tgatgctaaa cctcagggga aaaaaaaaaa 2270

<210> 27
<211> 7470
<212> DNA
<213> Homo sapiens
<400> 27

agatattaat cacggagttc cagggagaag gaacttgtga aatgggggag cggctgggg 60
ttgccggcac catggagtca ctttttagcc cgggactctt tcacaggctg gatgaagatt 120
gggattctgc tctctttgtt gaactcggtt atttcacaga cactgatgag ctgcaattgg 180
aagcagcaaa tgagacgtat gaaaacaatt ttgataatct tgattttgat ttggatttga 240

tgcttggga gtcagacatt tgggacatca acaaccaaat ctgtacagtt aaagatatta 300
aggcagaacc tcagccactt tctccagcct cctcaagtta ttcagtctcg tctcctcggt 360
cagtggaactc ttattcttca actcagcatg ttctgagga gttggatttg tcttctagtt 420

ctcagatgtc tcccctttcc ttatatggtg aaaactctaa tagtctctct tcagcggagc 480
 cactgaagga agataagcct gtcactggtc cttaggaacaa gactgaaaat ggactgactc 540
 caaagaaaaa aattcaggtg aattcaaaac cttcaattca gccaagcct ttattgcttc 600
 cagcagcacc caagactcaa acaaactcca gtgttccagc aaaaaccatc attattcaga 660

cagtaccaac gcttatgcc a ttggcaaagc agcaaccaat tatcagttta caacctgcac 720
 ccactaaagg ccagacggtt ttgctgtctc agcctactgt ggtacaactt caagcacctg 780
 gagttctgcc ctctgctcag ccagtccttg ctgttgctgg gggagtcaca cagctcccta 840
 atcacgtggt gaatgtgta ccagccctt cagcgaatag cccagtgaat ggaaaacttt 900
 ccgtgactaa acctgtccta caaagtacca tgagaaatgt cggttcagat attgctgtgc 960
 taaggagaca gcaacgtatg ataaaaatc gagaatccgc ttgtcagtct cgcaagaaga 1020
 agaaagaata tatgctaggg ttagaggcga gattaaaggc tgccctctca gaaaacgagc 1080

aactgaagaa agaaaatgga acaactgaagc ggcagctgga tgaagtgtg tcagagaacc 1140
 agaggcttaa agtccctagt ccaaagcga gagttgtctg tgtgatgata gtattggcat 1200
 ttataatact gaactatgga cctatgagca tgttgaaca ggattccagg agaatgaacc 1260
 ctagtgtgag ccctgcaaat caaaggaggc accttctagg attttctgct aaagaggcac 1320
 aggacacatc agatggtatt atccagaaaa acagctacag atatgatcat tctgtttcaa 1380
 atgacaaagc cctgatggtg ctaactgaag aaccattgct ttacattcct ccacctcctt 1440
 gtcagcccct aattaacaca acagagtctc tcaggttaaa tcatgaactt cgaggatggg 1500

ttcatagaca tgaagtagaa aggaccaagt caagaagaat gacaaataat caacagaaaa 1560
 cccgtattct tcagggtgct ctggaacagg gctcaaattc tcagctgatg gctgttcaat 1620
 acacagaaac cactagtagt atcagcagga actcagggag tgagctacaa gtgtattatg 1680
 cttcaccag aagttatcaa gacttttttg aagccatccg cagaagggga gacacatfff 1740
 atgttgtgtc atttcgaagg gatcacctgc tgttaccagc taccacccat aacaagacca 1800
 caagacaaa aatgtcaatt gtgttaccag caataaacat aatgagaat gtgatcaatg 1860
 ggcaggacta cgaagtgatg atgcagattg actgtcaggt gatggacacc aggatcctec 1920

atatcaaaag ttgctcagtt cctccttacc tccgatca gcagaggaat caaaccaaca 1980
 ccttctttgg ctcccctccc gcagccacag aggcaacca cgttgtcagc accatcctg 2040
 agtcattaca atagcacct gcagctatgc tggaaaactg agcgtgggac cctgccagac 2100
 tgaagagcag gtgagcaaaa tgctgctttc tgccttgggtg gcaggcagag aactgtctcg 2160
 tactagaatt caaggaggaa agaagaagaa ataaaagaag ctgctccatt tttcatcctc 2220
 taccatcta ttggaaagc actggaattc agatgcaaga gaacaatggt tcttcagtgg 2280

caaatgtagc ccigcatcct ccagtgttac ctgggtgtaga ttttttttc tgtacctttc 2340

 taaacctctc ttcctctgt gatggttttg tgtttaaaca gtcacttctt tttaaataat 2400
 atccacctct cctttttgcc atttactta ttgattcata aagtgaattt ttttaagc 2460
 tatgccacac atgcatgttc aaatggtttc cactgattcg atttttcatt catttaatgc 2520
 aaaccattc tggatattgt gcttatttga gaaaacacat ttcaaaacca gaaaagccaa 2580
 aaacctcca aaaacaagca aaacaatttg gagctttaga taaaaggaaa aacteccagt 2640
 tggtaaagt tttctttact taggatttgt ggctcacacc taacaaagg gggtcagga 2700
 gtgggtacaa atttgagaaa atagaagggt aagggaaggg ccagtggtgg ggtttggaga 2760

 gaggagatag ctccattaat acacatgttt aaaagatgga aagttcacgc ctgtaatccc 2820
 agcactttgg gaggccgagg cgggtggatc acgaggtcag gagatcaaga ccatcccggc 2880
 taaaacggtg aaaccccgct tctactaaaa atacaaaaaa ttagccgggc gtagtgacgg 2940
 gcgcctgtag tcccagctac ttgggaggct gaggcaggag aatggcgtga acccgggagg 3000
 cggagcttgc agtgagccga gatcccga ctgcactcca gcctgggcga cagagcgaga 3060
 ctccgtctca aaaaaaaaaa aaaaaaaaaa aaaagatgga aagttcgatg tgactgcagt 3120
 atgagattaa agccacaact attgtttatt ttggggactc taggccacca agtattagca 3180

 cacatactta tgttttctct actaatctgg tccaggfctt catggaccac aggacaaagc 3240
 tttcattttc attcattctt ctattgaaat tataccaaat tcagctgagg aatatggaag 3300
 taactttaga cttaacaag acaaaagttt tttcactgaa gaattgacaa gtatttgctc 3360
 cttaaaacaa cgcagattag tgaacgtgga ttcttctga gggagtgcac cccataatat 3420
 ggcaataatt ttcagtttct ccaacgaaaa gatagtgaag gaattaaatc ttttgcctc 3480
 ccatggttaa aaaaaaaaaa aaagctgtgt tcatttttac tgtactatgc ctctttttc 3540
 accatagtag acaattatgt ttcatattgat gaattcatag aactggatct catacagcga 3600

 tgcctctct aatgttctac ctttcagttt ctaaagtgag tcttctccc tctctacaa 3660
 aacttttcaa tttttgatg taactcatc acaaaactg tttcttacc cagttgactt 3720
 gcctttgtca gatttctct tgttcacac tatagcaatc aatttctctt ctctcttaca 3780
 agaaagggaa cgagaaattg tagcaacctc tcaaggatta tatgcagcta gttagtttc 3840
 tgctgtgaa attaggtctg gctctaaat aattttaag aaccatcagc acttctaact 3900
 ctctggacag gtgcctcttt gtccaagctc gttaaatgct ttccaaggaa atcagttcaa 3960
 cttttgtgag cggggaaaag cagggttta ttgttgtgtt acctgggagt ctggagtttg 4020

aaaagtgcta attaaccttc cctttttcc acattacaaa cctttttaag cagcgcagca 4080
 ctccccttag atttggctat cctgggtgat tttcagacaa gaaccatttt ctctggggac 4140
 cattcttctg ctgggtgcca aggaatataa ggcaaatgcc cagaagacct tcaggtgact 4200
 gggcagtctt atcatgggat atttcttctg gcctgcccc tcccattct gtaatgtgaa 4260
 ttagccacac cagaggctgt gaccatggct agtagacagt ggcaacatag tcatcccaa 4320
 gatgctaate ttctgctgga actgcatcac gttatcatgg tcaatgtaa cctggtttgt 4380
 gtgggtgat tataaataga gttccctcc tctctgtgac agaatcacag gagaaggacc 4440

 catctcgtgg ccttcttgtt cttagcgtt cacttttact tcatcctcg attcccagct 4500
 ttttctatca tcattttgcc aactcctcag atgcaagact ttggttatgt catactcacc 4560
 aacgttagtc cctctctcc aggtgaaaag gtgggtagcg gttgggaggg agtctccact 4620
 gaagagcagg aagggtgtag cagggccggc agctctgcca cagagctagg ggtgcctgta 4680
 aggtgccgcc tagagcagcc tgggagcttt gccttctttt gtctctcact agcccttcta 4740
 ctctttgtca ttgcctgttc ttgagtggat cttgaaatg aggggacagg attctcctaa 4800
 gggtagagtt tcaggaaatg agtgaaagc aatgacaaa tgcaaagaag tagtcacttt 4860

 ttaaattgct ggcaaagcta taattaatcc ctaggcacia ttgtagtttt tattttaatg 4920
 tttgtatgca caaggccctt taggaaatga gaagttgcca tgccagatta atttttttt 4980
 ttttttttgg tgggattgcc ttttgggggt tgcagccaga aattgtgggt aatgtgtgta 5040
 tttttttatt tattaatatt taaacaggat tgtgcaagct tatgagacia ttagataaac 5100
 tcatggagga ggcaggtcct cctgttatta gatgatattg tgctcttggg gctgacaata 5160
 atacactctt gggaagtgat ggtagagact gatgggaata gtctttctgc ctggttgcaa 5220
 gtcccaaatt ttaagggtt aatggaagta agtggatgtt tcctcatgtt aactactgaa 5280

 tcagatgta ggagcttgc cctttggggt tgacttatgc ccagcagtac agggacacag 5340
 cttcattaga gtgttagtgt aaactaactc caaagttagg agttaatgtg aaaggatcat 5400
 ccttgaaca aatctgctgt ttgccatgct ttagtacag aaacttcaca tggagttttg 5460
 ggtgggattt gtgtttcac aagtaaaaaa tcctcacga ttataaaact cagagcatca 5520
 tctaattttt tttttaatg actacaagt ccagcaciaa actggcattt ctttgccatt 5580
 tcttgccagt aagaattga cacggaggta tttgaaagca atgttatgtg agtcattctt 5640
 aagtgtcca agtaagtta gaaacagaaa aggaacttgg gattcaaatt gatttttcaa 5700

 atcattttta aagagacatc atcctgacta aatcttagcc tgaaccttcc tcccctgtgt 5760
 gtattccccg gtagtcaccg cagcgagatg ctggtgagac tgccgtggtg gcatttagca 5820
 tcgttaaac tggaaaactc tcaagctctt tgccacttcc ctactatttt ttgattcttg 5880

ccattttacc aagcttaggt tgtgaaactt gacagaaatg tattacagga aaaacttata 5940
 attgtatttg actttctaac acattgcaaa gtttcaaagt gactttcact ttcaacaaca 6000
 tattagaagt aaccactttt gctttcacag cctgaagagt tagagcctga tctgatgccc 6060
 cctttcactc tgaagtcatg ggaaattttc cagccatgaa agccctcttt ccaactgcata 6120

ctgatgggct gactcagctt ccttcagccg actgagatct tttcatacta ttggctattt 6180
 cataccaatt aacctcttaa ataagattgt gaattgccaa aattgataga cacttattac 6240
 cacctgtgga ctccatattc cttaccacaa atgttatttt catcagtcct gactcatttt 6300
 aacttacaga aattaggatt gttgctgcta atatgaatac caattataac ttttagaaac 6360
 aagaataaag cctaaaagag aatgaaatat aagaaatgtt cgttcccacc cctaataaca 6420
 tttggaagtg aatattccca ttttcttcca cccacagggg ttgggattga tttttaattt 6480
 cctaggaaac aatactagac tacccaaaaa gatgttgcca gaatccaaaa ggaactatgc 6540

tcgtaaaaga aatgcagttt tctctacct aaaaaaaga aagtaaagtg tgttctgttc 6600
 ttatcttttt aatgactaag ctttaaacag tttattttgg gtaagactag aactttcggc 6660
 catttgttct aatatgtgtg ttattgatg caatagaatt tatgaaaaga agaatgacaa 6720
 aggtatctga ttagaaaatt tgatcttacg catgaatcca tgtcatggcc agccactgtc 6780
 acatagtggg tgccattctc aacatattgg tttgctaact ttaagcatta gggatttagc 6840
 aactaaaat acttttaatt atattagggt tggtaactaa ggagtaata aatcataatt 6900
 tatcatttgc caaggccaac aaacaacact attgtgtgtg ttgctctcaa tgaagttgaa 6960

taaaccagga ggettggeat atccccttta tgttaatccc agctagagat tagtaggttg 7020
 actttcacag caattgtata ttgatccatt ttaactcatc cttgccataa tttccaggcc 7080
 agtcaccagg acagaggaga tgatggggaa acagagcttt agatgaaaac tactatgcac 7140
 tactagcctt agaggcactg gtttctgtt accactttgg caagtatgga tggcttaagt 7200
 ccagtagggc ttcatccatg gagccattag aactgagggg ggagtgttag agatgccatt 7260
 tcaccaggat ctttttgctc aggttgtacc catgccaatt gaagaactg ttaaagatga 7320
 ggaggagaga tgiaccattc tctccttaa taatgatgtt ggtttgcaaa acctaaagaa 7380

ataataaca cagactattt catactttca agcaagtctt tatactacct gttatttctc 7440
 taaaattcaa ataagaatt tttaaactta 7470

<210> 28

<211> 2041

<212> DNA

<213> Homo sapiens

<400> 28

agccatttct actttgcccg cccacagatg tagttttctc tgcgctgtg cgttttcct 60
 cctccccgcc ctcagggtcc acggccacca tggcgtatta ggggcagcag tgcctcggc 120
 agcattggcc ttigcagcgg cggcagcagc accaggctct gcagcggcaa cccccagcgg 180
 cttaagccat ggcgtgagta ccggggcggg tcgtccagct gtgctcctgg gcccggcgcg 240

 ggttttggat tggtaggggtg cggcctgggg ccagggcggt gccgccaagg ggaagcgat 300
 ttaacgagcg cccgggacgc gtggtctttg cttgggtgtc cccgagacgc tcgctgcct 360
 gggatcggga aagcgtagtc ggggtccccg actgcttccc caggagccct acagccctcg 420
 gaccccagc ccccaaggg tcccagggt cttggctgtt gccccacgaa acgtggcagg 480
 aaccaagatg gcggcggcag ggcggcggcg cgggcgtgag tcaaggcgg gcggtggcg 540
 gggcgcggcc gccctggcgc tatttggacg tggggacgga gcgctttcct cttggcggcc 600
 ggtggaagaa tcccctggtc tccgtgagcg tccattttgt ggaacctgag ttgcaagcag 660

 ggaggggcaa atacaactgc cctgttcccg attctctaga tggccgatct agagaagtcc 720
 cgctcataa gtggaaggat gaaattctca gaacagctaa cctctaatgg gagttggctt 780
 ctgattctca ttcaggcttc tcacggcatt cagcagcagc gttgctgtaa ccgacaaaga 840
 caccttcgaa ttaagcacat tctctgattc cagcaaagca ccgcaacatg accgaaatga 900
 gcttctgag cagcagagtg ttggtggggg acttgatgtc ccccttcgac cagtcgggtt 960
 tggggctga agaaagccta ggtctcttag atgattacct ggaggtggcc aagcacttca 1020
 aacctcatgg gttctccagc gacaaggcta aggcgggctc ctccgaatgg ctggctgtgg 1080

 atgggttggc cagtccctcc aacaacagca aggaggatgc cttctccggg acagattgga 1140
 tgttggagaa aatgatttg aaggagtctg acttggatgc cctgttgggt atagatgacc 1200
 tggaaacat gccagatgac cttctgacca cgttggatga cacttgtgat ctctttgcc 1260
 ccctagtcca ggagactaat aagcagcccc ccagacggg gaaccaatt ggccatctcc 1320
 cagaaagttt aacaaaacc gaccaggttg ccccttcac cttcttaca cctcttccc 1380
 ttccccagg ggtcctgtcc tccaectcag atcattcctt tagtttagag ctgggcagtg 1440
 aagtggatat cactgaagga gataggaagc cagactacac tgcttacgtt gccatgatcc 1500

 ctcagtgcac aaaggaggaa gacacccctt cagataatga tagtggcatc tgtatgagcc 1560
 cagagtccta tctgggtct cctcagcaca gccctctac cagggctct ccaaatagga 1620
 gcctccatc tccaggtgtt ctctgtgggt ctgccgtcc caaaccttac gatcctcctg 1680
 gagagaagat ggtagcagca aaagtaagg gtgagaaact ggataagaag ctgaaaaaaa 1740
 tggagcaaaa caagacagca gccactaggt accgccagaa gaagaggcg gagcaggagg 1800

ctcttactgg tgagtgcaaa gagctggaaa agaagaacga ggctctaaaa gagaggcg 1860
 attccctggc caaggagatc cagtacctga aagatttgat agaagaggtc cgcaaggcaa 1920

gggggaagaa aagggtcccc tagttgagga tagtcaggag cgtcaatgtg cttgtacata 1980
 gagtgctgta gctgtgtgtt ccaataaatt atttttagg gaaagtaaaa aaaaaaaaaa 2040
 a 2041

<210> 29
 <211> 4536
 <212> DNA
 <213> Homo sapiens
 <400> 29

aggctgtcac tcaggtggca gcggcagagg ccgggctgag acgtggccag gggaacacgg 60
 ctggctgtcc aggccgtcgg ggcggcagta gggtccctag cacgtccttg ctttcttggg 120
 agctccaagc ggcgggagag gcaggcgtca gtggctgcgc ctccatgcct gcgcgcgggg 180

cgggacgctg atggagcgcg ccatcagccc ggggctgctg gtacgggccc tgctgtctgt 240
 gctgctgctg ctggggctcg cggcaaggac ggtggccgcg gggcgcgccc gtggcctccc 300
 agcggcagc gcggaggcgg cgttcggcct cggggcggcc gctgctccca cctcagcagc 360
 gcgagtaccg gcggcgggag ccgtggctgc ggccgaggtg actgtggagg acgctgaggc 420
 gctgccggca gccgcgggag agcaggagcc tcggggtccg gaaccagacg atgagacaga 480
 gttgcgaccg cgcggcaggt cattagtaat tatcagcact ttagatggga gaattgctgc 540
 cttggatcct gaaaatcatg gtaaaaagca gtgggatttg gatgtgggat ccggttcctt 600

ggtgtcatcc agccttagca aaccagaggt atttgggaat aagatgatca ttccttcctt 660
 ggatggagcc ctcttcagc gggaccaaga ccgtgaaagc atggaacag ttcctttcac 720
 agttgaatca ctcttgaat cttcttataa atttggagat gatgttgttt tggttggagg 780
 aaaatctctg actacatatg gactcagtgc atatagtga aaggtgaggt atatctgttc 840
 agctctgggt tgtcgccaat gggatagtga cgaatggaa caagaggaag acatcctgct 900
 tctacagcgt acccaaaaaa ctgttagagc tgcggacct gcagtgga atgagaagtg 960
 gaatttcagt gttggccact ttgaacttcg gtatattcca gacatggaaa cgagagccgg 1020

atatttgaa agcaccttta agcccaatga gaacacagaa gagtctaaaa ttatttcaga 1080
 tgtggaagaa caggaagctg ccataatgga catagtata aaggtttcgg ttgctgactg 1140
 gaaagtatg gcattcagta agaaggagg acatctggaa tgggagtacc agttttgtac 1200
 tccaattgca tctgcctggt tacttaagga tgggaaagtc attccatca gtctttttga 1260

tgatacaagt tatacatcta atgatgatgt tttagaagat gaagaagaca ttgtagaagc 1320
 tgccagagga gccacagaaa acagtgttta cttgggaatg tatagaggcc agctgtatct 1380
 gcagtcatca gtcagaatth cagaaaagtt tccttcaagt cccaaggctt tggaatctgt 1440

 cactaatgaa aacgcaatta ttcctttacc aacaatcaaa tggaaccct taattcattc 1500
 tccttccaga actcctgtct tggtaggatac tgatgaatth gacaaatgct tcagtaatga 1560
 taagttttct catgaagaat atagtaatgg tgcactttca atcttgcagt atccatataga 1620
 taatggttat tatctacat actacaagag ggagaggaac aaacgaagca cacagattac 1680
 agtcagattc ctgcacaacc cacattacaa caagaatata cgcacaaaagg atcctgttct 1740
 tcttttacac tggtaggaag aatagttgc aacgattttg tttgtatca tagcaacaac 1800
 gtttattgtg cgcaggctth tccatcctca tcctcacagg caaaggaagg agtctgaaac 1860

 tcagtgtaaa actgaaaata aatagattc tgtaagtggg gaagccaatg acagtagctg 1920
 gaatgacata aaaaactctg gatataatc acgataatc actgattttg agccaattca 1980
 atgcctggga cgtgggtgct ttggagttgt tttgaaagct aaaaacaaag tagatgactg 2040
 caattatgct atcaagagga tccgtctccc caatagggaa ttggctcggg aaaagtaat 2100
 gcgagaagtt aaagccttag ccaagcttga acacccgggc attgttagat atttcaatgc 2160
 ctggctcga gaccaccag agaagtggca agaaaagatg gatgaaatth ggctgaaaga 2220
 tgaaagcaca gactggccac tcagctctcc tagcccaatg gatgcacat cagttaaaat 2280

 acgcagaatg gatcctttcg ctacaaaaga acatattgaa atcatagctc cttcaccaca 2340
 aagaagcagg tctttttcag tagggatttc ctgtgaccag acaagttcat ctgagagcca 2400
 gttctacca ctggaattct caggaatgga ccatgaggac atcagtgagt cagtggatgc 2460
 agcatacaac ctccaggaca gttgccttac agactgtgat tggaagatg ggactatgga 2520
 tggcaatgat gaggggcact cctttgaact ttgtccttct gaagcttctc cttatgtaag 2580
 gtcaaggag agaacctcct cttcaatagt attgaaagat tctggctgtg ataatgcttc 2640
 cagtaagaa gagccgaaaa ctaatcgatt gcatattggc aaccattgtg ctaataaact 2700

 aactgctttc aagcccacca gtagcaaatc ttcttctgaa gctacattgt ctatttctcc 2760
 tccaagacca accactttaa gtttagatct cactaaaaac accacagaaa aactccagcc 2820
 cagttacca aagggtgatac ttacattca aatgcagctg tgcagaaaag aaaacctcaa 2880
 agactggatg aatggacgat gtacataga ggagagagag aggagcgtgt gtctgacat 2940
 cttctgcag atcgcagagg cagtggagtt tcttcacagt aaaggactga tgcacagga 3000
 cctcaagcca tccaacat tctttacaat ggatgatgtg gcaaggttg gagactttgg 3060

gttagtgact gcaatggacc aggatgagga agagcagacg gttctgaccc caatgccagc 3120

 ttatgccaga cacacaggac aagtagggac caaactgtat atgagcccag agcagattca 3180
 tggaaacage tattctcata aagtggacat cttttcttta ggccctgattc tttttgaatt 3240
 gctgtatcca ttcagactc agatggagag agtcaggacc ttaactgatg taagaaatct 3300
 caaatttcca ccattattta ctcagaaata tccttgtgag tacgtgatgg ttcaagacat 3360
 gctctctcca tccccatgg aacgacctga agctataaac atcattgaaa atgctgtatt 3420
 tgaggacttg gactttccag gaaaaacagt gctcagacag aggtctcgct ccttgagttc 3480
 atcggaaca aaacattcaa gacagtccaa caactcccat agccctttgc caagcaatta 3540

 gccttaagtt gtgctagcaa ccctaatagg tgatgcagat aatagcctac ttcttagaat 3600
 atgcctgtcc aaaattgcag acttgaaaag tttgttcttc gctcaatfff tttgtggact 3660
 acttttttta tatcaaattt aagctggatt tgggggcata acctaatttg agccaactec 3720
 tgagttttgc taacttaag gaaagggcta tctttgttct ttgttagtct cttgaaactg 3780
 gctgctggcc aagctttata gccctacca tttgcctaag gaggtagcag caatccctaa 3840
 tatatatata tagtgagaac taaaatggat atatffffat aatgcagaag aaggaaagtc 3900
 ccctgtgtg gtaactgtat tgttctagaa atatgctttc tagagatatg atgattttga 3960

 aactgatttc tagaaaaagc tgactccatt tttgtccctg gccggtaaat taggaatctg 4020
 cactatfff gaggacaagt agcacaact gtataacggg ttatgtccgt agttttatag 4080
 tcctatffgt agcattcaat agctttattc cttagatggg tctaggggtg gtttacagct 4140
 ttttactt ttacctcaa taaagggaaa atgaagcttt ttatgtaaat tggttgaaag 4200
 gtctagtttt gggaggaaaa aagccgtagt aagaaatgga tcatatatat tacaactaac 4260
 ttcttcaact atggactfff taagcctaat gaaatcttaa gtgtcttata tgtaatcctg 4320
 taggttggtta cttccccaa actgattata ggtaacagtt taatcatctc acttgctaac 4380

 atgtttttat tttcactgt aaatatgttt atgttttatt tataaaaatt ctgaaatcaa 4440
 tccatttggg ttggtgggtg acagaacaca cttaaagtgtg ttaacttgtg acttctttca 4500
 agtctaaatg atttaataaa acttttttta aattaa 4536

 <210> 30
 <211> 4233
 <212> DNA
 <213> Homo sapiens
 <400> 30
 agttcagatg agaggtttag tgtaaagtgg tataggacag aggggacaga gttgagatat 60

ttgatgcaaa atctaggttc ctggcaagtg aatgcattct tcctttcttg gtacccttgt 120
 ccagagattg agactgcgtg gccagaagt gacaaaagag gtttaaggga gcaggggaaga 180

 aaaggtcatt agtaattatc agcacttttag atgggagaat tgctgccttg gatcctgaaa 240
 atcatggtaa aaagcagtgg gatttggatg tgggatccgg ttccttgggt tcatccagcc 300
 ttagcaaacc agaggtattt gggaataaga tgatcattcc ttccttggat ggagccctct 360
 tccagtggga ccaagaccgt gaaagcatgg aacagttcc tttcacagtt gaatcacttc 420
 ttgaatcttc ttataaattt ggagatgatg ttgttttggg tggaggaaaa tctctgacta 480
 catatggact cagtgcata atgtgaaagg tgaggatata ctgttcagct ctgggttgtc 540
 gccaatggga tagtgacgaa atggaacaag aggaagacat cctgcttcta cagcgtaccc 600

 aaaaaactgt tagagctgtc ggacctcgca gtggcaatga gaagtggaat ttcagtgttg 660
 gccactttga acttcggtat attccagaca tggaaacgag agccggattt attgaaagca 720
 cctttaagcc caatgagaac acagaagagt ctaaaattat ttcagatgtg gaagaacagg 780
 aagctgcat aatggacata gtgataaagg ttcggttgc tgactggaaa gttatggcat 840
 tcagtaagaa gggaggacat ctggaatggg agtaccagtt ttgtactcca attgcatctg 900
 cctggttact taaggatggg aaagtcattc ccatcagtct ttttgatgat acaagttata 960
 catctaataga tgatgtttta gaagatgaag aagacattgt agaagctgcc agaggagcca 1020

 cagaaaacag tgtttacttg ggaatgtata gaggccagct gtatctgcag tcatcagtca 1080
 gaatttcaga aaagtttctt tcaagtcca aggctttgga atctgtcact aatgaaaacg 1140
 caattattcc tttaccaaca atcaaatgga aacccttaat tcattctcct tccagaactc 1200
 ctgtcttggg aggatctgat gaatttgaca aatgtctcag taatgataag ttttctcatg 1260
 aagaatatag taatggtgca ctttcaatct tgcagtatcc atatgataat ggttattatc 1320
 taccatacta caagaggag aggaacaaac gaagcacaca gattacagtc agattcctcg 1380
 acaaccaca ttacaacaag aatatccgca aaaaggatcc tgttcttctt ttacactggt 1440

 ggaaagaaat agttgcaacg attttgtttt gtatcatagc aacaacgttt attgtgcgca 1500
 ggcttttcca tcctcatcct cacaggcaaa ggaaggagtc tgaactcag tgtcaactg 1560
 aaaataaata tgattctgta agtggatgaag ccaatgacag tagctggaat gacataaaaa 1620
 actctggata tatatcacga tatctaaactg attttgagcc aattcaatgc ctgggacgtg 1680
 gtggctttgg agttgttttt gaagctaaaa acaaagtaga tgactgcaat tatgctatca 1740
 agaggatccg tctcccaat aggggaattgg ctcgggaaaa ggtaatgcga gaagttaaag 1800

ccttagccaa gcttgaacac ccgggcattg ttagatattt caatgcctgg ctggaagcac 1860

 caccagagaa gtggcaagaa aagatggatg aaatttggct gaaagatgaa agcacagact 1920
 ggccactcag ctctcctagc ccaatggatg caccatcagt taaaatacgc agaatggatc 1980
 ctttcgctac aaaagaacat attgaaatca tagctccttc accacaaaga agcaggtctt 2040
 tttcagtagg gatttctgtg gaccagacaa gttcatctga gagccagttc tcaccactgg 2100
 aattctcagg aatggacat gaggacatca gtgagtcagt ggatgcagca tacaacctcc 2160
 aggacagtgg ccttacagac tgtgatgtgg aagatgggac tatggatggc aatgatgagg 2220
 ggcaactcctt tgaactttgt ctttctgaag cttctcctta tgtaaggtca agggagagaa 2280

 cctcctcttc aatagtattt gaagattctg gctgtgataa tgcttccagt aaagaagagc 2340
 cgaaaactaa tcgattgcat attggcaacc attgtgctaa taaactaact gctttcaagc 2400
 ccaccagtag caaatcttct tctgaagcta cattgtctat ttctcctcca agaccaacca 2460
 cttaagttt agatctcact aaaaacacca cagaaaaact ccagcccagt tcaccaaagg 2520
 tgtatcttta cattcaaatg cagctgtgca gaaaagaaaa cctcaaagac tggatgaatg 2580
 gagatgtac catagaggag agagagagga gcgtgtgtct gcacatcttc ctgcagatcg 2640
 cagaggcagt ggagtttctt cacagtaaag gactgatgca cagggacctc aagccatcca 2700

 acatattctt tacaatggat gatgtggatca aggttggaga ctttgggtta gtgactgcaa 2760
 tggaccagga tgaggaagag cagacgggtc tgacccaat gccagcttat gccagacaca 2820
 caggacaagt agggacaaa ctgtatatga gccagagca gattcatgga aacagctatt 2880
 ctcataaagt ggacatcttt tctttaggcc tgattctatt tgaattgctg tatccattca 2940
 gcactcagat ggagagagtc aggaccttaa ctgatgtaag aaatctcaa tttccacat 3000
 tatttactca gaaatatct tgtgagtacg tgatgggtca agacatgctc tctccatccc 3060
 ccatggaacg acctgaagct ataaacatca ttgaaaatgc tgtatttgag gacttggact 3120

 ttccaggaaa aacagtgtc agacagaggt ctgcctcctt gagttcatcg ggaacaaaac 3180
 attcaagaca gtccaacaac tcccatagcc ctttgccaag caattagcct taagttgtgc 3240
 tagcaacct aataggatg gcagataata gcctacttct tagaatatgc ctgtcaaaaa 3300
 ttgcagactt gaaaagtttg ttcttcgtc aatTTTTTtg tggactactt tttttatc 3360
 aaatttaagc tggatttggg ggcataacct aatttgagcc aactcctgag ttttctata 3420
 ctaaggaaa gggctatctt tgttcttgt tagtctcttg aaactggctg ctggccaagc 3480
 tttatagccc tcaccattg cctaaggagg tagcagcaat ccctaataata tatatatagt 3540

gagaactaaa atggatatat ttttataatg cagaagaagg aaagtcctcc tgtgtggtaa 3600
 ctgtattggt ctagaaatat gctttctaga gatatgatga ttttgaaact gatttctaga 3660
 aaaagctgac tccatttttg tccctggcgg gtaaattagg aatctgact attttggagg 3720
 acaagtagca caaactgtat aacggtttat gtccgtagtt ttatagtcct atttgtagca 3780
 ttcaatagct ttattcctta gatggttcta gggtaggttt acagcttttt gtacttttac 3840
 ctccaataaa gggaaaatga agctttttat gtaaattggt tgaaaggctct agttttggga 3900
 ggaaaaaagc cgtagtaaga aatggatcat atataffaca actaacttct tcaactatgg 3960

actttttaag cctaatgaaa tcttaagtgt cttatatgta atcctgtagg ttggtacttc 4020
 ccccaactg attataggta acagttaat catctcactt gctaacatgt ttttatTTTT 4080
 cactgtaaat atgtttatgt tttatttata aaaattctga aatcaatcca tttgggttgg 4140
 tgggtacag aacacactta agtgtgttaa ctgtgactt ctttcaagtc taaatgattt 4200
 aataaaactt tttttaaatt aagaaaaaaa aaa 4233

<210> 31

<211> 4737

<212> DNA

<213> Homo sapiens

<400> 31

ggtttgaaag gaaggcagag agggcactgg gaggaggcag tgggagggcg gagggcgggg 60

gccttcgggg tgggcgcccc gggtagggca ggtggccgcg gcgtggaggc agggagaatg 120
 cgactctcca aaacctctgt cgacatggac atggccgact acagtgtctc actggacca 180
 gcctacacca ccttgaatt tgagaatgtg caggtgtga cgatgggcaa tgacactcc 240
 ccatcagaag gcaccaact caacgcgcc aacagcctgg gtgtcagcgc cctgtgtgcc 300
 atctgcgggg accgggccac gggcaaacac tacggtgcct cgagctgtga cggctgcaag 360
 ggcttcttcc ggaggagcgt gcggaagaac cacatgtact cctgcagatt tagccggcag 420
 tgcgtggtgg acaaagacaa gaggaaccag tccgctact gcaggctcaa gaaatgcttc 480

cgggctggca tgaagaagga agccgtccag aatgagcggg accggatcag cactcgaagg 540
 tcaagctatg aggacagcag cctgcctcc atcaatgcgc tctgcagc ggaggtcctg 600
 tcccagacaga tcacctccc cgtctccggg atcaacggcg acattcgggc gaagaagatt 660
 gccagcatcg cagatgtgtg tgagtccatg aaggagcagc tgctggttct cgttgagtgg 720
 gccaaagta tccagcttt ctgcgactc ccctggagc accaggtggc cctgctcaga 780
 gccatgctg gcgagcact gctgctcgga gccaccaaga gatccatggt gttcaaggac 840

gtgctgctcc taggcaatga ctacattgtc cctcggcact gcccggagct ggccggagatg 900

agccgggtgt ccafacgcat ccttgacgag ctggtgctgc ccttccagga gctgcagatc 960

gatgacaatg agtatgccta cctcaaagcc atcatcttct ttgaccaga tgccaagggg 1020

ctgagcgatc caggaagat caagcggctg cgttcccagg tgcaggtgag cttggaggac 1080

tacatcaacg accgccagta tgactcgcgt ggccgctttg gagagctgct gctgctgctg 1140

cccaccttgc agagcatcac ctggcagatg atcgagcaga tccagtcat caagctcttc 1200

ggcatggcca agattgacaa cctgttgacg gagatgctgc tgggagggtc ccccagcgat 1260

gcaccccatg cccaccacc cctgcacct cacctgatgc aggaacatat gggaaccaac 1320

gtcatcgttg ccaacacaat gcccactcac ctgagcaacg gacagatgtg tgagtggccc 1380

cgaccagggt gacaggcagc caccctgag accccacagc cctcaccgcc aggtggtca 1440

gggtctgagc cctataagct cctgccggga gccgtcgcca caatcgtaa gccctctct 1500

gccatcccc agccgacct caccaagcag gaagttatct agcaagccgc tggggcttgg 1560

gggtccact ggtcccccc agccccctaa gagagcacct ggtgatcacg tggtcacggc 1620

aaaggaagac gtgatgccg gaccagtecc agagcaggaa tgggaaggat gaagggcccg 1680

agaacatggc ctaagggcca catcccactg ccaccctga cgccctgctc tggataacaa 1740

gactttgact tggggagacc tctactgct tggacaactt ttctcatgtt gaagccactg 1800

ccttcacctt caccttcate catgtccaac ccccgacttc atcccaaagg acagccgctt 1860

ggagatgact tgaggcetta cttaaacca gctcccttct tccttagcct ggtgcttctc 1920

ctctcctage cctgtcatg gtgtccagac agagccctgt gaggtgggt ccaattgtgg 1980

cacttggggc accttgctcc tcttctgct gctgccccca cctctgctgc ctccctctgc 2040

tgtcaccttg ctgaccatc ccgtcttctc caaccacc tctccagagg ccaaggagge 2100

cttggaaacg attccccag tcattctggg aacatgtgt aagcactgac tgggaccagg 2160

caccaggcag ggtctagaag gctgtggtga gggaagacgc ctttctctc caaccaacc 2220

tcatctctct tcttcagga cttgggtggg tacttgggtg aggatccctg aaggccttca 2280

accgagaaa acaaaccag gttggcgact gcaacaggaa cttggagtgg agaggaaaag 2340

catcagaaag aggcagacca tccaccaggc ctttgagaaa gggtagaatt ctggctggtg 2400

gagcaggtga gatgggacat tccaaagaac agcctgagcc aaggcctagt ggtagtaaga 2460

atctagcaag aattgaggaa gaatggtgtg ggagaggat gatgaagaga gagaggcct 2520

gctggagagc ataggtctg gaacaccagg ctgaggtcct gatcagcttc aaggagtatg 2580

caggagctg ggcttccaga aaatgaacac agcagttctg cagaggacgg gaggctggaa 2640
 gctgggaggt caggtggggt ggatgatata atgccgggtga gagtaatgag gcttggggct 2700
 ggagaggaca agatgggtaa accctcacat cagagtgaca tccaggagga ataagctccc 2760
 agggcctgtc tcaagctctt ccttactccc aggcactgtc ttaaggcatc tgacatgcat 2820
 catctcattt aatcctcctt tcctccctat taacctagag attgtttttg ttttttattc 2880
 tcctcctccc tccccgcctt caccgcctcc actcctcctt aacctagaga ttgttacaga 2940
 agctgaaatt gcgttctaag aggtgaagtg attttttttc tgaaactcac acaactagga 3000

 agtggctgag tcaggacttg aaccaggctc tccttgatc agaacaggag ctcttaacta 3060
 cagtggctga atagcttctc caaaggctcc ctgtgttctc accgtgatca agttgagggg 3120
 ctccggctc cttctacag cctcagaaac cagactcgtt cttctgggaa cctgcccac 3180
 tcccaggacc aagattggcc tgaggctgca ctaaaattca cttaggtctg agcatcctgt 3240
 ttgctgataa atattaagga gaattcatga ctcttgacag cttttctctc ttcactcccc 3300
 aagtcaaggg gaggggtggc aggggtctgt ttcttgaag tcaggctcat ctggcctgtt 3360
 ggcatggggg tgggacagtg tgcacagtgt gggggcaggg gagggctaag caggcctggg 3420

 tttgagggtt gctccggaga ccgtcactcc aggtgcattc tggaagcatt agaccccagg 3480
 atggagcgac cagcatgtca tccatgtgga atcttgggtg ctttgaggac atcttgaaaa 3540
 atgccactga ccagtgtgaa caaaagggat gtgttatggg gctggaggtg tgattagta 3600
 ggagggaaac tgttggaccg actcctgccc cctgctcaac actgaccctt ctgagtgggt 3660
 ggaggcagtg cccagtgcc cagaaatccc accattagtg attgtttttt atgagaaaga 3720
 ggcgtggaga agtattgggg caatgtgtca gggaggaatc accacatccc tacggcagtc 3780
 ccagccaagc ccccaatccc agcggagact gtccctgct cagagctccc aagccttccc 3840

 ccaccaccte actcaagtgc ccctgaaatc cctgccagac ggctcagcct ggtctcgggt 3900
 aaggcagga ggctggaacc atttctgggc attgtgttca ttccactgt gttcctccac 3960
 ctctcctc cagcgttgtt cagacctctg tcttgggaga aaggttgaga taagaatgct 4020
 ccatggagtg ccgtgggcaa cagtggcctt tcatgggaac aatctgttgg agcagggggt 4080
 cagttctctg ctgggaatct accccttctt ggaggagaaa cccattccac ctttaataact 4140
 ttattgtaat gtgagaaaca caaaacaaag tttacttttt tgactctaag ctgacatgat 4200
 attagaaaat ctctcgtctt cttttttttt tttttttttt tttttggcta cttgagtgtt 4260

 ggtcctaaaa cataaaatct gatgacaaa cagagggttg ctggggggac aagcgtgggc 4320
 acaatttccc caccaagaca ccctgatctt caggcgggtc tcaggagctt ctaaaaatec 4380
 gcatggctct cctgagagtg gacagaggag aggagagggt cagaaatgaa cgctcttcta 4440

tttcttgtca ttaccaagcc aattactttt gccaaatfff tctgtgatct gccctgatta 4500
 agatgaattg tgaatttac atcaagcaat tatcaaagcg ggctgggtcc catcagaacg 4560
 acccacatct ttctgtgggt gtgaatgtca ttaggtcttg cgctgacccc tgagccccca 4620
 tcactgccgc ctgatggggc aaagaacaaa aaaacatttc ttactcttct gtgttttaac 4680

 aaaagtttat aaaacaaaat aaatggcgca tatgttttct aaaaaaaaa aaaaaaa 4737
 <210> 32
 <211> 6460
 <212> DNA
 <213> Homo sapiens
 <400> 32
 agatcttccc agaggacggt ttgaaaggaa ggcagagagg gcactgggag gaggcagtgg 60
 gagggcggag ggcgggggcc ttccgggttg gcgccaggg tagggcaggt ggcccggcg 120
 tggaggcagg gagaatgcga ctctcaaaa ccctcgtcga catggacatg gccgactaca 180
 gtgctgcact ggaccacgc tacaccacc tggaaattga gaatgtgcag gtgttgacga 240
 tgggcaatga cagtcacca tcagaaggca ccaacctcaa cgcgcccaac agcctgggtg 300

 tcagcgcct gtgtgccatc tgcggggacc gggccacggg caaacactac ggtgcctcga 360
 gctgtgacgg ctgcaagggc ttcttccgga ggagcgtgcg gaagaaccac atgtactcct 420
 gcagatttag ccggcagtcg gtggtggaca aagacaagag gaaccagtgc cgctactgca 480
 ggctcaagaa atgcttccgg gctggcatga agaaggaagc cgtccagaat gagcgggacc 540
 ggatcagcac tcgaagtica agctatgagg acagcgcct gccctccatc aatgcgctcc 600
 tgcagcggga ggtcctgtcc gcacagatca cctccccgt ctccgggatc aacggcgaca 660
 ttccggcgaa gaagattgcc agcatcgcag atgtgtgtga gtccatgaag gagcagctgc 720

 tggttctcgt ttagtgggcc aagtacatcc cagctttctg cgagctccc ctggacgacc 780
 aggtggccct gctcagagcc catgctggcg agcacctgct gctcggagcc accaagagat 840
 ccatggtgtt caaggacgtg ctgctcctag gcaatgacta cattgtccct cggcactgcc 900
 cggagctggc ggagatgagc cgggtgtcca tacgcatcct tgacgagctg gtgctgcctt 960
 tccaggagct gcagatcgat gacaatgagt atgctacct caaagccatc atcttctttg 1020
 acccagatgc caaggggctg agcgatccag ggaagatcaa gcggctgcgt tcccaggtgc 1080
 aggtgagctt ggaggactac atcaacgacc gccagtatga ctgcctggc cgctttggag 1140

 agctgctgct gctgctgcc accttgaga gcatcacctg gcagatgatc gagcagatec 1200
 agttcatcaa gctcttcggc atggccaaga ttgacaacct gttgcaggag atgctgctgg 1260

gagggcccc cagcgatgca ccccatgccc accaccccct gcacccctcac ctgatgcagg 1320
 aacatatggg aaccaacgtc atcgttgcca acacaatgcc cactcacctc agcaacggac 1380
 agatgtccac ccctgagacc ccacagccct caccgccagg tggctcaggg tctgagccct 1440
 ataagctcct gccgggagcc gtcgccacaa tctcaagcc cctctctgcc atccccagc 1500
 cgaccatcac caagcaggaa gttatctagc aagccctgg ggcttggggg ctccactggc 1560

 tccccccage ccctaagag agcacctggt gatcacgtgg tcacggcaaa ggaagacgtg 1620
 atgccaggac cagtcccaga gcaggaatgg gaaggatgaa gggcccgaga acatggccta 1680
 agggccacat cccactgcca ccttgacgc cctgctctgg ataacaagac tttgacttgg 1740
 ggagacctct actgccttgg acaacttttc tcatgttgaa gccactgcct tcacctcac 1800
 ctatccat gtccaacccc cgacttcac ccaaaggaca gccgcctgga gatgacttga 1860
 ggcttactt aaaccagct ccttcttcc ctagcctggt gcttctctc tectagcccc 1920
 tgtcatggtg tccagacaga gccctgtgag gctgggtcca attgtggcac ttggggcacc 1980

 ttgctcctcc ttctgctgct gccccacct ctgctgctc cctctgctgt caccttgctc 2040
 agccatcccg tcttctccaa caccacctct ccagaggcca aggaggcctt ggaaacgatt 2100
 cccccagtca ttctgggaac atgttgtaag cactgactgg gaccaggcac caggcaggg 2160
 ctagaaggct gtggtgaggg aagacgcctt tctctccaa cccaacctca tctctctct 2220
 tcagggactt ggggtgggac ttgggtgagg atccctgaag gccttcaacc cgagaaaaca 2280
 aaccaggtt ggcgactgca acaggaactt ggagtggaga ggaaaagcat cagaaagagg 2340
 cagaccatcc accaggcctt tgagaaaggg tagaatctg gctggtagag caggtgagat 2400

 gggacattcc aaagaacagc ctgagccaag gcctagtgg agtaagaatc tagcaagaat 2460
 tgaggaagaa tggigtggga gaggatgat gaagagagag agggcctgct ggagagcata 2520
 gggcttgga caccagctg aggtctgat cagcttcaag gagtatgcag ggagctgggc 2580
 ttccagaaaa tgaacacagc agttctgcag aggacgggag gctggaagct gggaggtcag 2640
 gtgggtgga tgatataat cgggtgagag taatgaggct tggggctgga gaggacaaga 2700
 tgggtaaac ctacatcag agtgacatcc aggaggaata agctcccagg gectgtctca 2760
 agctcttct tactcccagg cactgtctta aggcactga catgcatcat ctcattaat 2820

 cctccttcc tcctattaa cctagagatt gttttgttt tttattctcc tectcctcc 2880
 ccgcctcac ccgcccact cctcctaac ctagagattg ttacagaagc tgaattgcg 2940
 ttctaagagg tgaagtgatt tttttctga aactcacaca actaggaagt ggctgagtca 3000
 ggacttgaac ccaggtctcc ctggatcaga acaggagctc ttaactacag tggctgaata 3060
 gcttctccaa aggtcctctg tgttctcacc gtgatcaagt tgaggggctt ccggctccct 3120

tctacagcct cagaaaccag actcgttctt ctgggaaccc tgcccactcc caggaccaag 3180
 attggcctga ggctgcacta aaattcactt agggctcgagc atcctgtttg ctgataaata 3240

ttaaggagaa ttcatgactc ttgacagctt ttctctette actccccaag tcaaggggag 3300
 ggggtggcagg ggctctgttc ctggaagtca ggctcatctg gcctgttggc atgggggtgg 3360
 gacagtgtgc acagtgtggg ggcaggggag ggctaagcag gcctgggttt gagggctgct 3420
 ccggagaccg tcactccagg tgcatctctg aagcattaga ccccaggatg gagcgaccag 3480
 catgtcatcc atgtggaatc ttgggtggctt tgaggacatt ctggaaaatg cactgacca 3540
 gtgtgaacaa aagggatgtg ttatggggct ggaggtgtga ttaggtagga gggaaactgt 3600
 tggaccgact cctgcccct gctcaacact gaccctctg agtggttga ggcagtgcc 3660

cagtgccag aaatcccacc attagtatt gtttttatg agaaagagc gtggagaagt 3720
 atggggcaa tgtgtcaggg aggaatcacc acatccctac ggcagtcca gccaaagccc 3780
 caatcccagc ggagactgtg ccctgctcag agctccaag ccttcccca ccacctact 3840
 caagtcccc tgaatccct gccagacggc tcagcctggt ctgcgtaag gcagggagc 3900
 tgaaccatt tctgggcat gtggctctc cactgtgtt cctccacctc ctccctccag 3960
 cgttgctcag acctctgtc tgggagaaag gttgagataa gaatgtcca tggagtccc 4020
 tgggcaacag tggccctca tgggaacaat ctgttgagc agggggtcag ttctctctg 4080

ggaatctacc ctttctgga ggagaaaccc attccacctt aataacttta ttgtaatgtg 4140
 agaaacaaa aacaaagttt actttttga ctctaagctg acatgatatt agaaaatctc 4200
 tcgctctctt ttttttttt ttttttttt ttggctactt gagtgtgtgt cctaaaacat 4260
 aaaatctgat ggacaaacag aggggtgtctg gggggacaag cgtgggcaca atttcccac 4320
 caagacacc tgatcttcag gcgggtctca ggagcttcta aaaatccgca tggctctct 4380
 gagagtggac agaggagagg agagggtcag aaatgaacgc tcttctattt ctgtcatta 4440
 ccaagccaat tacttttgcc aaattttct gtgatctgcc ctgattaaga tgaattgtga 4500

aatttacatc aagcaattat caaagcgggc tgggtcccat cagaacgacc cacatcttc 4560
 tgtgggtgtg aatgtcatta ggtcttgcgc tgaccctga gcccctca ctgccctg 4620
 atggggcaaa gaaacaaaa acatttctta ctctctgtg ttttaacaaa agtttataaa 4680
 acaaaataaa tggcgcata gttttctaag tccttgata agtatcttt ctttcaggta 4740
 tcagaaataa gactgaatct tctggttcta cttgggggtt aaaaaattt ttttaagga 4800
 agaatgagaa tagttttata gttctttgtg atgtgcagaa tgtttttgtg tccattataa 4860

tttttcagtc ttcacatcaa gaggtaaagca gttagacatg attactccca ctttccagat 4920

 gagggagactg aggcttgggg gaagtgactt ctcttgaag gcagaggtgg acatctaacc 4980
 ctggtctctt gattccaagt acttagtata tcgagagagt gaaagttgat cccccctctt 5040
 gaagagggga gtgatgaggg gagagtgcaa tggcaagatc tggaagaatg gcaagaggtt 5100
 ccaaggtctt gtcatctcc accaaggttc aagacagaac cttttgctgg gtcacctcaa 5160
 tctgccagca atggaagatg agtagctgtg gggacatttc ataaaagcaa gtggtttttt 5220
 tgttttgttt tgttttgttt tttgtttttt tttctagaac aaggctgtgc acagtggctc 5280
 acctctgtaa tcccagcact ttgggaggct gaggcgggag gatcacttga gctcaggagt 5340

 tcgagaccag ccagagcaat ataaggagac cccatctcta caaaaaattt aaaaattagc 5400
 caggtatagt ggtgtgtgcc tatagtacca gctactctga aggctgaggt gggaagattg 5460
 cttcagccca ggagttcag gatgcagtga gctatgaatg caacactgca ctacagcctg 5520
 gatgaaagaa caagactctg tctcaaaaca taataataa gtaaaaagaa taaaagcaag 5580
 agatgcactt gagaatctcc agccagatct gtagccactg ggcttctctc caaggctaaa 5640
 ctattacagg agggtggcct tgtgtctcgg tcaccacaga ccacagcgtt ccattcactc 5700
 ggggttgtgc tggagctggc ttgtgagaac tgactgttag cttctcttcc caactccatg 5760

 ttgcccagtg ccacactgat agcttgaat tggttattgc cggagtgttt acaccacaag 5820
 gactagcaaa ctctacaaat ccgggctttt gttcctggag agcccgttgt taacattcac 5880
 cagcacacca cagcattcgg caatggctgg accatgggat gcctacatat ggggacatcc 5940
 tccttgggga tgaggttaga gcagggcgat ctttccact cttccttaag ggaggggaca 6000
 aaagtcttgg tctgggaagc acacgttttg ctgatcagcg taaccttggg caggtcactc 6060
 caccactccg agcctcatct gtaaagtggg aatgatatct ccctccaggg cagatgtcag 6120
 gattcaatgg aatgagatca cagtaactgt gagagctccc gttacatgag gagtacaagt 6180

 gaactcttca tgcgccctt tttagcgaga agttaacat taaactctcc aggcttcaga 6240
 gcaccattc gctgtctacc tgatccctag ggcgctccc gccttccct gtgccttccc 6300
 tccactagtc agcaccagga aatgttttcg ataacgttgc aacggaggcc ttgttcatgc 6360
 tgccgccatc ggggacaagc gcgggggggg ggggtggag gccagaggag actatttcag 6420
 tcctaaattg tgcttaataa acccatatca aaaccataaa 6460

<210> 33

<211> 6581

<212> DNA

<213> Homo sapiens

<400> 33

agatcttccc agaggacggt ttgaaaggaa ggcagagagg gcactgggag gaggcagtgg 60

 gagggcggag ggcgggggcc ttcggggtgg gcccaggg tagggcaggt ggcccggcg 120
 tggaggcagg gagaatgcga ctctcaaaa cctcgtcga catggacatg gccgactaca 180
 gtgctgcact ggaccacgcc tacaccacc tggaaattga gaatgtgcag gtgttgacga 240
 tgggcaatga ttttgttgc gctgcgtctc gccagattga ggcatcccct cggacatcac 300
 tggagcatal ctggaggggt ggacagttct ccacaggag acacgtcccc atcagaaggc 360
 accaacctca acgcgccaa cagcctgggt gtcagcgcct tgtgtgcat ctgccccgac 420
 cgggccacgg gcaaacacta cggtcctcg agctgtgacg gctgcaaggg cttcttccgg 480

 aggagcgtgc ggaagaacca catgtactcc tgcagattta gccggcagtg cgtggtggac 540
 aaagacaaga ggaaccagtg ccgctactgc aggtcaaga aatgcttccg ggctggcatg 600
 aagaaggaag ccgtccagaa tgagcgggac cggatcagca ctgaaggtc aagctatgag 660
 gacagcagcc tgcctccat caatgcgtct ctgcagcgg aggtcctgtc ccgacagatc 720
 acctccccg tctccggat caacggcgac attcggcgca agaagattgc cagcatcgca 780
 gatgtgtgtg agtccatgaa ggagcagctg ctggttctcg ttgagtgggc caagtacatc 840
 ccagctttct gcgagctccc cctggacgac caggtggccc tgctcagagc ccatgctggc 900

 gagcacctgc tgctcggagc caccaagaga tccatggtgt tcaaggacgt gctgctccta 960
 ggcaatgact acattgtccc tcggcactgc ccggagctgg cggagatgag ccgggtgtcc 1020
 atacgcatcc ttgacgagct ggtgctgccc ttccaggagc tgcagatcga tgacaatgag 1080
 tatgcctacc tcaaagccat catcttcttt gaccagatg ccaaggggct gagcgatcca 1140
 gggaagatca agcggctgcg ttcccagtg caggtgagct tggaggacta catcaacgac 1200
 cgccagtatg actcgcgtgg ccgctttgga gagctgctgc tgctgctgcc caccttgagc 1260
 agcatcacct ggcagatgat cgagcagatc cagttcatca agctcttcgg catggccaag 1320

 attgacaacc tgttgacgga gatgctgctg ggagggtccc ccagcagatgc acccatgcc 1380
 caccacccc tgcaccctca cctgatgcag gaacatatgg gaaccaacgt catcgttgcc 1440
 aacacaatgc cactcacct cagcaacgga cagatgtgtg agtggccccg acccagggga 1500
 caggcagcca cccctgagac cccacagccc tcaccgccag gtggctcagg gctgagccc 1560
 tataagctcc tgccgggagc cgtgccaca atcgtcaagc ccctctctgc catccccag 1620
 ccgaccatca ccaagcagga agttatctag caagccgctg gggcttgggg gctccactgg 1680

ctccccccag cccccaaga gagcacctgg tgatcacgtg gtcacggcaa aggaagacgt 1740

gatgccagga ccagtcccag agcaggaatg ggaaggatga agggccccgag aacatggcct 1800

aagggccaca tcccactgcc acccttgacg ccttgctctg gataacaaga ctttgacttg 1860

gggagacctc tactgccttg gacaactttt ctcatgttga agccactgcc ttcacettca 1920

ccttcatcca tgtccaacce ccgacttcat cccaaaggac agccgcctgg agatgacttg 1980

aggccttact taaaccacgc tcccttcttc cctagcctgg tgccttctct ctcttagccc 2040

ctgtcatggt gtccagacag agccctgtga ggctgggtcc aattgtggca cttggggcac 2100

cttgctcctc cttctgctgc tgcecccacc tctgctgctt ccctctgctg tcaccttget 2160

cagccatccc gtcttctcca acaccacctc tccagaggcc aaggaggcct tggaaacgat 2220

tccccagtc attctgggaa catgttgtaa gcactgactg ggaccaggca ccaggcaggg 2280

tctagaagge tgtggtgagg gaagacgctt ttctctcca acccaacctc atcctccttc 2340

ttcagggact tgggtgggta cttgggtgag gatccctgaa ggccttcaac ccgagaaaac 2400

aaaccaggt tggcgactgc aacaggaact tggagtggag aggaaaagca tcagaaagag 2460

gcagaccate caccaggcct ttgagaaaag gtagaattct ggctggtaga gcaggtgaga 2520

tgggacattc caaagaacag cctgagccaa ggcttagtgg tagtaagaat ctagcaagaa 2580

ttgaggaaga atggtgtggg agagggatga tgaagagaga gagggcctgc tggagagcat 2640

agggtctgga acaccaggct gaggtcctga tcagcttcaa ggagtatgca gggagctggg 2700

cttcagaaa atgaacacag cagtctgca gaggacggga ggctggaagc tgggaggtca 2760

ggtggggtgg atgataaat gcgggtgaga gtaatgaggc ttggggctgg agaggacaag 2820

atgggtaaac cctcacatca gagtacatc caggaggaat aagctcccag ggcctgtctc 2880

aagctcttcc ttactcccag gcactgtctt aaggcatctg acatgatca tctcatttaa 2940

tctcctcttc ctcctatta acctagagat tgtttttgtt ttttattctc ctctcctc 3000

cccgcctca cccgccccac tccctcctaa cctagagatt gttacagaag ctgaaattgc 3060

gttctaagag gtgaagtgat tttttttctg aaactcacac aactaggaag tggctgagtc 3120

aggacttgaa cccaggcttc cctggatcag aacaggagct cttactaca gtggctgaat 3180

agcttctcca aaggctccct gtgttctcac cgtgatcaag ttgaggggct tccggctccc 3240

ttctacagcc tcagaaacca gactcgttct tctgggaacc ctgcccactc ccaggaccaa 3300

gattggcctg aggctgcact aaaattcact tagggtcgag catcctgttt gctgataaat 3360

attaaggaga attcatgact cttgacagct tttctctctt cactccccaa gtcaagggga 3420

ggggtggcag gggctctgtt cctggaagtc aggcctcatct ggcctgttgg catgggggtg 3480
 ggacagtgtg cacagtgtgg gggcagggga gggctaagca ggcctgggtt tgaggctgc 3540
 tccggagacc gtcactccag gtgcattctg gaagcattag accccaggat ggagcgacca 3600
 gcatgtcatc catgtggaat cttgggtggct ttgaggacat tctggaaaat gccactgacc 3660
 agtgtgaaca aaagggatgt gttatggggc tggaggtgtg attaggtagg agggaaactg 3720
 ttggaccgac tectgcccc tgctcaacac tgaccctct gagtggttgg aggcagtgcc 3780
 ccagtgccca gaaatccac cattagtgat tgtttttat gaaaagagg cgtggagaag 3840

 tattggggca atgtgtcagg gaggaatcac cacatcccta cggcagtccc agccaagccc 3900
 ccaatcccag cggagactgt gccctgctca gagctccaa gccttcccc accacctcac 3960
 tcaagtgcc ctgaaatccc tgccagacgg ctcagcctgg tctgcggtaa ggcagggagg 4020
 ctggaacat ttctgggat tgtgttcatt cccactgtgt tectccacct cctccctca 4080
 gcgttgctca gacctctgc ttgggagaaa ggttgagata agaatgtccc atggagtgcc 4140
 gtgggcaaca gtggccctc atgggaaca tctgttgag caggggtca gttctctgt 4200
 gggaatctac cctttctgg aggagaaacc cattccacct taataacttt attgtaatgt 4260

 gagaacaca aaacaagt tactttttg actctaagct gacatgatat tagaaaatct 4320
 ctgctctct ttttttttt ttttttttt ttggctact tgagttgtg tctaaaaca 4380
 taaaatctga tggacaaca gagggttgct ggggggacaa gcgtgggcac aatttccca 4440
 ccaagacacc ctgatctca ggccggtctc aggagctct aaaaatccgc atggctctc 4500
 tgagagtgga cagaggagag gagagggtca gaaatgaacg ctcttctatt tcttgcatt 4560
 accaagccaa ttacttttg caaattttc tgtgatctgc cctgattaag atgaattgtg 4620
 aaatttcat caagcaatta tcaaagcggg ctgggtcca tcagaacgac ccacatctt 4680

 ctgtgggtgt gaatgtcatt aggtcttgcg ctgaccectg agccccatc actgccgct 4740
 gatggggcaa agaacaanaa aacatttct actcttctgt gtttaacaa aagtttataa 4800
 acaaaaata atggcgcata tgttttctaa gtccttggat aagtatctt tcttcaggt 4860
 atcagaaata agactgaatc ttctggttct acttgggggt taaaaattt ttttaagg 4920
 aagaatgaga atagtttat agttcttct gatgtgcaga atgttttct gtccattata 4980
 attttctcgt cttcacatca agagtaagc agttagacat gattactccc actttccaga 5040
 tgaggagact gaggcttggg ggaagtgact tctcttggaa ggcagagggt gacatctaac 5100

 cctggtctct tgattccaag tacttagtat atcgagagag taaaagtga tccccctct 5160
 tgaagagggg agtgatgagg ggagagtgc atggcaagat ctggaagaat ggcaagagg 5220
 tccaagggtc tgcacctc caccaaggt caagacagaa cttttgctg ggtcacctca 5280

atctgccagc aatggaagat gagtagctgt ggggacattt cataaaagca agtgggtttt 5340
 ttgttttggt ttgttttggt tttgtttttt ttttctagaa caaggctgtg cacagtggct 5400
 cacctctgta atcccagcac tttgggaggc tgaggcggga ggatcacttg agctcaggag 5460
 ttcgagacca gccagagcaa tataaggaga ccccatctct acaaaaaatt taaaaattag 5520

ccaggtatag tgggtgtgtc ctatagtacc agctactctg aaggctgagg tgggaagatt 5580
 gcttcagccc aggagttcga ggatgcagtg agctatgaat gcaacactgc actacagcct 5640
 ggatgaaaga acaagactct gtctcaaac ataaataata agtaaaaaga ataaaagcaa 5700
 gagatgcact tgagaatctc cagccagatc tgtagccact gggcttctct ccaaggctaa 5760
 actattacag gaggggtggc ttgtgtctcg gtcaccacag accacagcgt tccattcact 5820
 cggggttggt ctggagctgg cttgtgagaa ctgactgtta gcttctctc ccaactccat 5880
 gtttgccagt gccacactga tagcttgaaa ttggttattg ccggagtgtt tacaccacia 5940

ggactagcaa acctacaaa tccgggcttt tgttcctgga gagcccgttg ttaacattca 6000
 ccagcacacc acagcattcg gcaatggctg gaccatggga tgcctacata tggggacatc 6060
 ctcttgggg atgaggtag agcagggcga tctttcacc tcttcttaa gggaggggac 6120
 aaaagtctg gtctgggaag cacacgtttt gctgatcagc gtaaccttgg gcaggctact 6180
 ccaccactcc gagcctcacc tgtaaagtgg gaatgatatc tccctccagg gcagatgtea 6240
 ggattcaatg gaatgagatc acagtaactg tgagagctcc cgttacatga ggagtacaag 6300
 tgaactcttc atgcgccct ttttagcgag aagttaacca ttaaactctc caggcttcag 6360

agcaccatt cgetgtctac ctgatcccta gggccctcc cgccttccc tgtgccttc 6420
 ctccactagt cagcaccagg aaatgtttc gataacgttg caacggaggc cttgttcattg 6480
 ctgccccat cggggacaag cgcggggggg ggggggtgga ggccagagga gactatttca 6540
 gtctaaatt gtgcttaata aacctatc aaaaccataa a 6581

<210> 34
 <211> 6588
 <212> DNA
 <213> Homo sapiens
 <400> 34

gcactaccg ccttctggt ggacggctc ctggtggctg tgctgctgct gtgagcgggc 60
 ccctgctct ccatgcccc agctctccgg ctgggtgggc ttggccatgg tcagcgtgaa 120

cgcgcccctc ggggctccag tggagagttc ttacggctcg cagtttccat ggagaccgce 180
 tcccaccatc cctgaaatct cattgtgacc agtattaagg gattaactct gccagcggc 240

cctcgcaggg tctgggggtgc tgagaggggc cagcaggagc tccctcaagg atgaagaggt 300
 tgtgactgt gggggacccc atagcagcag gaggaggatg tcggactggg gccagggtt 360
 cccccaggac ccaccagaca cgtcccccac agaaggcacc aacctcaacg cgcccaacag 420
 cctgggtgtc agcgcctgt gtgccatctg cggggaccgg gccacgggca aacactacgg 480
 tgccctgagc tgtgacggct gcaagggtt cttccggagg agcgtgcgga agaaccacat 540

 gtactcctgc agatttagcc ggcagtgcgt ggtggacaaa gacaagagga accagtgccg 600
 ctactgcagg ctcaagaaat gcttccgggc tggcatgaag aaggaagccg tccagaatga 660
 gggggaccgg atcagcactc gaaggtcaag ctatgaggac agcagcctgc cctccatcaa 720
 tgcgctcctg caggcggagg tccgtccccg acagatcacc tccccgtct cgggatcaa 780
 cggcgacatt cgggcgaaga agattgccag catcgagat gtgtgtgagt ccatgaagga 840
 gcagctgctg gttctcgttg agtgggcaa gtacatccca gctttctgcg agctccccct 900
 ggacgaccag gtggccctgc tcagagccca tgctggcgag cacctgctgc tcggagccac 960

 caagagatcc atggtgttca aggacgtgct gctcctaggc aatgactaca ttgtccctcg 1020
 gcactgcccc gagctggcgg agatgagccg ggtgtccata cgcatccttg acgagctggt 1080
 gctgcccctc caggagctgc agatcgatga caatgagtat gcctacctca aagccatcat 1140
 cttctttgac ccagatgcca aggggctgag cgatccaggg aagatcaagc ggctgcgttc 1200
 ccagggtcag gtgagcttgg aggactacat caacgaccgc cagtatgact cgcgtggccg 1260
 ctttgagag ctgctgctgc tgctgcccac cttgcagagc atcacctggc agatgatcga 1320
 gcagatccag ttcataaagc tcttcggcat ggccaagatt gacaacctgt tgcaggagat 1380

 gctgctggga gggccccca gcgatgcacc ccatgcccac caccctcgc accctcacct 1440
 gatgcaggaa catatgggaa ccaactcat cgttgccaac acaatgcca ctcacctcag 1500
 caacggacag atgtccacc ctgagacccc acagccctca ccgccaggtg getcagggtc 1560
 tgagccctat aagctcctgc cgggagccgt cgccacaatc gtcaagcccc tctctgcat 1620
 cccccagcc accatcacca agcaggaagt tatctagcaa gccgctgggg cttgggggt 1680
 ccaactggctc cccccagccc cctaagagag cacctggtga tcactggtc acggcaaagg 1740
 aagacgtgat gccaggacca gtcccagagc aggaatggga aggatgaagg gcccgagaac 1800

 atggcctaag ggccacatcc cactgccacc cttgacgccc tgctctggat aacaagactt 1860
 tgacttgggg agaccttac tgccctggac aacttttctc atgttgaagc cactgccttc 1920
 accttcacct tcatcaatgt ccaacccccg acttcatccc aaaggacagc cgcttgaga 1980
 tgacttgagg ccttacttaa acccagctcc cttcttcctt agcctggtgc ttctctctc 2040
 ctagcccctg tcatggtgtc cagacagagc cctgtgaggc tgggtccaat tgtggcactt 2100

ggggcacctt gctcctcctt ctgctgctgc ccccacctct gctgcctccc tctgctgtca 2160
 ccttgctcag ccatcccgtc ttctccaaca ccacctctcc agaggccaag gaggccttgg 2220

 aaacgattcc cccagtcatt ctgggaacat gttgtaagca ctgactggga ccaggcacca 2280
 ggcagggtct agaaggctgt ggtgaggga gacgcctttc tcctccaacc caacctcatc 2340
 ctcttcttc agggacttgg gtgggtactt gggtaggat ccctgaagc cttcaaccg 2400
 agaaaacaaa cccaggttgg cgactgcaac aggaacttgg agtggagagg aaaagcatca 2460
 gaaagaggca gaccatccac caggcctttg agaaaggta gaattctggc tggtagagca 2520
 ggtgagatgg gacattcaa agaacagcct gagccaaggc ctagtgttag taagaatcta 2580
 gcaagaattg aggaagaatg gtgtgggaga gggatgatga agagagagag ggcctgctgg 2640

 agagcatagg gtctggaaca ccagctgag gtctgatca gcttcaagga gtatgcaggg 2700
 agctgggctt ccagaaaatg aacacagcag ttctgcagag gacgggaggc tggaaactgg 2760
 gaggtcaggt ggggtggatg atataatgcg ggtgagagta atgaggcttg gggctggaga 2820
 ggacaagatg ggtaaaccct cacatcagag tgacatccag gaggaataag ctcccagggc 2880
 ctgtctcaag ctcttctta ctcccaggca ctgtcttaag gcatctgaca tgcacatct 2940
 catttaatcc tccttctct cctattaacc tagagattgt ttttgtttt tatttctctc 3000
 ctccctccc gccctcacc gccccactcc ctccctaacct agagattgtt acagaagctg 3060

 aaattgcgtt ctaagagggt aagtgatttt ttttctgaaa ctcacacaac taggaagtgg 3120
 ctgagtcagg acttgaacc aggtctcct ggatcagaac aggagctctt aactacagtg 3180
 gctgaatagc ttctccaaag gctccctgtg ttctcaccgt gatcaagttg aggggcttcc 3240
 ggctcccttc tacagcctca gaaaccagac tcgttcttct gggaacctg cccactccca 3300
 ggaccaagat tggcctgagg ctgcaactaa attcacttag ggtcgagcat cctgtttgct 3360
 gataaatatt aaggagaatt catgactctt gacagctttt ctctcttcac tcccgaagtc 3420
 aaggggaggg gtggcagggg tctgtttcct ggaagtcagg ctcatctggc ctgttggcat 3480

 ggggggtggga cagtgtgcac agtgtggggg caggggaggg ctaagcaggc ctgggtttga 3540
 gggctgctcc ggagaccgtc actccaggig cattctgaa gcattagacc ccaggatgga 3600
 gcgaccagca tgtcatccat gtggaactt ggtggctttg aggacattct ggaanaatgcc 3660
 actgaccagt gtgaacaaaa gggatgtgtt atgggctgg aggtgtgatt aggtaggagg 3720
 gaaactgttg gaccgactcc tgccccctgc tcaaacctga cccctctgag tggttggagg 3780
 cagtccccca gtgccagaa atcccacat tagtgattgt ttttatgag aaagagcgt 3840

ggagaagtat tggggcaatg tgtcagggag gaatcaccac atccctacgg cagtcccagc 3900

 caagcccca atcccagcgg agactgtgcc ctgctcagag ctccaagcc ttccccacc 3960
 acctcactca agtgcccctg aaatecctgc cagacggctc agcctggtct gcggtaaaggc 4020
 agggaggctg gaaccatttc tgggcattgt ggtcattccc actgtgttcc tccacctcct 4080
 ccctccagcg ttgctcagac ctctgtcttg ggagaaaggt tgagataaga atgtcccatg 4140
 gagtgccgtg ggcaacagtg gcccttcatg ggaacaatct gttggagcag ggggtcagtt 4200
 ctctgtctgg aatctacccc tttctggagg agaaacccat tccaccttaa taactttatt 4260
 gtaatgtgag aaacacaaaa caaagtttac ttttttgact ctaagctgac atgatattag 4320

 aaaatctctc gcctctcttt ttttttttt ttttttttt ggctacttga gttgtgttcc 4380
 taaaacataa aatctgatgg acaaacagag ggttgctggg gggacaagcg tgggcacaat 4440
 ttccccacca agacacctg atcttcaggc ggttctcagg agcttctaaa aatccgcatg 4500
 gctctcctga gagtggacag aggagaggag agggtcagaa atgaacgctc ttctatttct 4560
 tgtcattacc aagccaatta cttttgcaa atttttctgt gatctgccct gattaagatg 4620
 aattgtgaaa tttacatcaa gcaattatca aagcgggctg ggtcccatca gaacgaccea 4680
 catctttctg tgggtgtgaa tgtcattagg tcttgcgctg acccctgagc ccccatcact 4740

 gccgcctgat ggggcaaaga aacaaaaaac atttcttact cttctgtgtt ttaaaaaag 4800
 tttataaac aaaataaatg gcgcatatgt tttctaagtc cttggataag tatcttttct 4860
 ttcaggtatc agaataaga ctgaatcttc tggttctact tgggggttaa aaaatTTTT 4920
 ttaaaggaag aatgagaata gttttatagt tcttttgat gtgcagaatg tttttgtgtc 4980
 cattataatt ttcagtctt cacatcaaga ggtaagcagt tagacatgat tactcccact 5040
 ttccagatga ggagactgag gcttggggga agtgacttct cttggaagc agaggtggac 5100
 atctaacctt ggtctcttga ttccaagtac ttagtatatc gagagagtga aagttgatec 5160

 cccttcttga agaggggagt gatgagggga gaggcaatg gcaagatctg gaagaatggc 5220
 aagaggttcc aagggtctgt catctccac caaggttcaa gacagaacct tttgctgggt 5280
 cacctcaatc tgccagcaat ggaagatgag tagctgtggg gacatttcat aaaagcaagt 5340
 ggttttttg tttgttttg tttgttttt tgttttttt tctagaacaa ggctgtgcac 5400
 agtggctcac ctctgtaate ccagcacttt gggaggctga ggcgggagga tcaacttgac 5460
 tcaggagttc gagaccagcc agagcaatat aaggagacc catctctaca aaaaatttaa 5520
 aaattagcca ggtatagtg tgtgtgccta tagtaccagc tactctgaag gctgaggtgg 5580

| | |
|--|------|
| gaagattgct tcagcccagg agttcgagga tgcagtgagc tatgaatgca aactgcact | 5640 |
| acagcctgga tgaagaaca agactctgtc tcaaacata aataataagt aaaaagaata | 5700 |
| aaagcaagag atgcacttga gaatctccag ccagatctgt agccactggg cttctctcca | 5760 |
| aggctaaact attacaggag ggtggccttg tgtctcggtc accacagacc acagcgttcc | 5820 |
| attcactcgg ggttgtgctg gagctggctt gtgagaactg actgttagct tctcttcca | 5880 |
| actccatggt tgccagtgcc aactgatag cttgaaattg gttattgccg gagtgtttac | 5940 |
| accacaagga ctagcaaaact ctacaaatcc gggcttttgt tcctggagag cccgttgta | 6000 |
| | |
| acattcacca gcacaccaca gcattcggca atggctggac catgggatgc ctacatatgg | 6060 |
| ggacatcctc cttggggatg agggtagagc agggcgatcc tttcacctct tccttaaggg | 6120 |
| aggggacaaa agttctggtc tgggaagcac acgttttct gatcagcgta accttgggca | 6180 |
| ggtcactcca ccaactccgag cctcatctgt aaagtgggaa tgatatctcc ctccagggca | 6240 |
| gatgtcagga ttcaatggaa tgagatcaca gtaactgtga gagctcccgt tacatgagga | 6300 |
| gtacaagtga actcttcatg cgccccttt tagcgagaag ttaaccatta aactctccag | 6360 |
| gcttcagagc acccattcgc tgtctactg atccctaggg ccgtcccgc ctcccctgt | 6420 |
| | |
| gccttcctc cactagtcag caccagaaa tgttttcgat aacgttgcaa cggaggcctt | 6480 |
| gttcatgctg ccgccatcgg ggacaagcgc gggggggggg gggaggagc cagaggagac | 6540 |
| tatttcagtc ctaaattgtg ctttaataaac ccatatcaaa accataaa | 6588 |