(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년08월16일
(11) 등록번호 10-2694809
(24) 등록일자 2024년08월08일

(51) 국제특허분류(Int. Cl.)
C12Q 1/6883 (2018.01) *A61K 31/713* (2006.01)
A61P 1/16 (2006.01) *C12N 15/113* (2010.01)
(52) CPC특허분류
C12Q 1/6883 (2022.01)
A61K 31/713 (2013.01)
(21) 출원번호 10-2020-7013128
(22) 출원일자(국제) 2018년10월11일
심사청구일자 2021년10월08일
(85) 번역문제출일자 2020년05월07일
(65) 공개번호 10-2020-0062314
(43) 공개일자 2020년06월03일
(86) 국제출원번호 PCT/US2018/055377
(87) 국제공개번호 WO 2019/075181
국제공개일자 2019년04월18일
(30) 우선권주장
62/570,985 2017년10월11일 미국(US)
(56) 선행기술조사문헌
CN103520724 A*
(뒷면에 계속)

(73) 특허권자
리제너론 파마슈티칼스 인코포레이티드
미합중국 뉴욕주 10591 타리타운 울드 소우 밀 리버 로드 777
(72) 발명자
신, 여롱
미합중국 뉴욕주 10591 타리타운 울드 소우 밀 리버 로드 777
그로마다, 제스퍼
미합중국 뉴욕주 10591 타리타운 울드 소우 밀 리버 로드 777
(뒷면에 계속)
(74) 대리인
김진환, 박지하, 김민철

전체 청구항 수 : 총 42 항 심사관 : 최지다

(54) 발명의 명칭 **PNPLA3 I148M 변이를 발현하는 환자의 간 질환의 치료에서의 HSD17B13의 저해**

(57) 요 약

본 개시내용은 HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법을 제공한다. 본 개시내용은 또한 HSD17B13의 저해제를 투여함으로써 PNPLA3 Ile148Met+인 대상체를 치료하는 방법을 제공한다. 본 개시내용은 또한 대상체에서 PNPLA3 Ile148Met 변이체 및 기능성 HSD17B13을 검출하는 방법을 제공한다. 본 개시내용은 또한 간 질환에 대해서 보호 효과를 갖는 대상체를 식별하는 방법을 제공한다. 본 개시내용은 또한 간 질환의 치료에 사용하기 위한 HSD17B13의 저해제를 제공한다.

대 표 도

특징	GHS 발견 코호트 (N = 46,544)	GHS 비만 수술 코호트 (N = 2,644)	멀러스 심장 연구 (N = 1,357)	펜 메디슨 바이오뱅크 (N = 8,526)
연령(세) - 중위 (IQR)	63 (50 - 74)	53 (44 - 61)	46 (38 - 54)	68 (60 - 76)
여성 - 수 (%)	26,875 (58)	2,119 (80)	724 (53)	3,242 (38)
체질량 지수 - 중위 (IQR)	30 (25 - 45)	47 (42 - 54)	28 (25 - 32)	30 (25 - 32)
트랜스아미나제 수준(U/L) - 중위 (IQR)				
알라닌 아미노트랜스퍼라제 (ALT)	22.0 (17.0 - 29.0)	23.0 (17.5 - 29.5)	20.0 (15.0 - 27.0)	22.0 (17.0 - 30.0)
아스파테이트 아미노트랜스퍼라제(AST)	23.0 (20.0 - 27.5)	23.0 (20.0 - 27.0)	21.0 (18.0 - 25.0)	24.0 (20.0 - 30.5)
간 질환의 존재 - N (%)				
알코올성 간 질환	197 (0.4)	7 (0.3)	-	-
알코올성 간경변	130 (0.3)	3 (0.1)	-	-
비알코올성(비바이러스성) 간 질환	1,938 (4.2)	1,543 (58.4)	-	-
비알코올성 간경변	382 (0.8)	24 (0.9)	-	-
간세포 암종	76 (0.2)	1 (0.04)	-	-
간질환 없음	30,628 (65.8)	1 (0.04)	-	-

(52) CPC특허분류
A61P 1/16 (2018.01)
C12N 15/113 (2013.01)
C12Q 2600/106 (2013.01)
C12Q 2600/156 (2013.01)
C12Q 2600/158 (2013.01)

(72) 발명자
첸, 시핑
미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리버 로드 777
드웨이, 프레드릭
미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리버 로드 777
데슬로비취 도스탈, 탄야
미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리버 로드 777
스커르만, 클라우디아
미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리버 로드 777
바라스, 아리스
미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리버 로드 777
애블-허슨, 노우라
미국, 10128 뉴욕, 뉴욕, 아파트먼트 18이, 이스트 345 93번 스트리트

(56) 선행기술조사문헌
PLOS ONE. vol. 10, no. 7, e01 32640(2015.07.22.)
Journal of Lipid Research. vol. 56, no. 1, 167-175(2014.11.05.)
Nature genetics. vol. 43, no. 11, 1131-1138 (2011.10.16.)
Gastroenterology. vol. 150, no. 8, 1728-1744 (2016.02.10.)*
*는 심사관에 의하여 인용된 문헌

명 세 서

청구범위

청구항 1

비알코올성 또는 알코올성 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법으로서,

상기 대상체로부터의 샘플이,

i) I148M 변이를 포함하는 파타틴 유사 포스포리파제 도메인 함유 3(patatin like phospholipase domain containing 3: PNPLA3) 단백질을 암호화하는 제1 핵산 및 기능성 HSD17B13 단백질을 암호화하는 제2 핵산; 및

ii) I148M 변이를 포함하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질 중 적어도 하나를 포함하는지의 여부를 결정하는 단계; 및

i)에 정의된 바와 같은 상기 제1 핵산 및 제2 핵산 둘 다 및 ii)에 정의된 바와 같은 상기 단백질 둘 다가 검출되는 경우 상기 대상체를 HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 식별하는 단계

를 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 2

제1항에 있어서,

상기 제1 핵산 분자는 게놈 DNA, mRNA 또는 mRNA로부터 획득된 cDNA를 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 3

제2항에 있어서,

상기 게놈 DNA는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함하거나,

상기 mRNA는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함하거나;

상기 mRNA는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함하거나;

상기 cDNA는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함하거나; 또는

상기 cDNA는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 4

제2항 또는 제3항에 있어서,

상기 게놈 DNA는 서열번호 31에 따른 뉴클레오타이드 서열을 포함하거나;

상기 mRNA는 서열번호 34에 따른 뉴클레오타이드 서열을 포함하거나;

상기 mRNA는 서열번호 35에 따른 뉴클레오타이드 서열을 포함하거나;

상기 cDNA는 서열번호 38에 따른 뉴클레오타이드 서열을 포함하거나;

상기 cDNA는 서열번호 39에 따른 뉴클레오타이드 서열을 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 5

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 제1 핵산을 검출하는 단계는,

상기 제1 핵산의 적어도 일부를 서열분석하는 단계로서, 상기 일부는 상기 I148M 변이를 암호화하는 코돈을 포함하는, 상기 서열분석하는 단계; 또는

상기 제1 핵산을 상기 제1 핵산의 일부에 특이적으로 혼성화하는 프로브 또는 프라이머와 혼성화시키는 단계로서, 상기 일부는 상기 I148M 변이를 암호화하는 상기 코돈을 포함하는, 상기 혼성화시키는 단계

를 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 6

제5항에 있어서,

상기 프로브 또는 프라이머는 대립유전자-특이적(allele-specific) 프로브 또는 프라이머이되, 상기 프로브 또는 프라이머는 표지를 선택적으로 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 7

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 대상체가 상기 I148M 변이에 대해서 동형접합성(homozygous)인지 또는 이형접합성(heterozygous)인지를 결정하는 단계를 더 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 8

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 제2 핵산은 게놈 DNA, mRNA 또는 mRNA로부터 획득된 cDNA를 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 9

제8항에 있어서,

상기 게놈 DNA는 서열번호 1에 따른 12,667번 위치에 상응하는 위치에 아데닌을 포함하거나;

상기 게놈 DNA는 서열번호 1에 따른 뉴클레오타이드 서열을 포함하거나;

상기 mRNA는 서열번호 3에 따른 뉴클레오타이드 서열을 포함하거나;

상기 mRNA는 서열번호 4에 따른 뉴클레오타이드 서열을 포함하거나;

상기 mRNA는 서열번호 7에 따른 뉴클레오타이드 서열을 포함하거나;

상기 mRNA는 서열번호 11에 따른 뉴클레오타이드 서열을 포함하거나;

상기 cDNA는 서열번호 12에 따른 뉴클레오타이드 서열을 포함하거나;

상기 cDNA는 서열번호 13에 따른 뉴클레오타이드 서열을 포함하거나;

상기 cDNA는 서열번호 16에 따른 뉴클레오타이드 서열을 포함하거나; 또는

상기 cDNA는 서열번호 20에 따른 뉴클레오타이드 서열을 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 10

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 제2 핵산을 검출하는 단계는,

상기 제2 핵산을 서열분석하는 단계; 또는

상기 제2 핵산을 상기 제2 핵산의 일부에 특이적으로 혼성화하는 프로브 또는 프라이머와 혼성화시키는 단계로

서, 상기 일부는 서열번호 1에 따른 12,667번 위치에 상응하는 위치에 아데닌을 포함하는, 상기 혼성화시키는 단계

를 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 11

제10항에 있어서,

상기 프로브 또는 프라이머는 대립유전자-특이적 프로브 또는 프라이머이되, 상기 프로브 또는 프라이머는 표지를 선택적으로 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 12

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 대상체가 상기 샘플에서 기능성 HSD17B13 단백질을 암호화하는 상기 제2 핵산에 대해서 동형접합성인지 또는 이형접합성인지를 결정하는 단계를 더 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 13

삭제

청구항 14

삭제

청구항 15

제1항에 있어서,

상기 알코올성 간 질환은 알코올 섭취로 인한 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 16

삭제

청구항 17

제1항에 있어서,

상기 비알코올성 간 질환은 비알코올성 지방간 질환(nonalcoholic fatty liver disease: NAFLD) 또는 비알코올성 지방간염(non-alcoholic steatohepatitis: NASH)을 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 18

제1항에 있어서,

상기 비알코올성 간 질환은 알코올 섭취에 의해서 유발되지 않은 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함하는, 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법.

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

I148M 변이를 포함하는 PNPLA3 단백질을 갖고, 그리고 기능성 HSD17B13 단백질을 갖는 인간 대상체에서 비알코올성 또는 알코올성 간 질환을 치료하는데 사용하기 위한, 활성 성분으로 HSD17B13의 저해제를 포함하는 약제학적 조성물로서, 상기 HSD17B13의 저해제는 내인성 HSD17B13 게놈 DNA 또는 mRNA에 혼성화하고, 상기 대상체 내의 세포에서 HSD17B13 폴리펩티드의 발현을 감소시키는 안티센스 분자를 포함하는 것인, 약제학적 조성물.

청구항 58

제57항에 있어서,

상기 인간 대상체는 I148M 변이를 포함하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질에 대해서 양성인 것으로 시험된, 약제학적 조성물.

청구항 59

제57항 또는 제58항에 있어서,

상기 치료는 상기 인간 대상체가 I148M 변이를 포함하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질을 갖는지의 여부를 결정하는 단계를 포함하는, 약제학적 조성물.

청구항 60

제57항에 있어서,

상기 인간 대상체는, 제1항 내지 제3항 중 어느 한 항에 정의된 바와 같은 방법을 사용하여 HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자인 것으로 식별된, 약제학적 조성물.

청구항 61

제57항 또는 제58항에 있어서,

상기 변이체 PNPLA3 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하는, 약제학적 조성물.

청구항 62

제61항에 있어서,

상기 변이체 PNPLA3 단백질은 서열번호 42에 따른 아미노산 서열을 포함하는, 약제학적 조성물.

청구항 63

제57항 또는 제58항에 있어서,

상기 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 게놈 DNA인, 약제학적 조성물.

청구항 64

제63항에 있어서,

상기 게놈 DNA는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함하는, 약제학적 조성물.

청구항 65

제64항에 있어서,

상기 게놈 DNA는 서열번호 31에 따른 뉴클레오타이드 서열을 포함하는, 약제학적 조성물.

청구항 66

제57항 또는 제58항에 있어서,

상기 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 mRNA인, 약제학적 조성물.

청구항 67

제66항에 있어서,

상기 mRNA는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함하는, 약제학적 조성물.

청구항 68

제67항에 있어서,

상기 mRNA는 서열번호 34에 따른 뉴클레오타이드 서열을 포함하는, 약제학적 조성물.

청구항 69

제66항에 있어서,

상기 mRNA는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함하는, 약제학적 조성물.

청구항 70

제69항에 있어서,

상기 mRNA는 서열번호 35에 따른 뉴클레오타이드 서열을 포함하는, 약제학적 조성물.

청구항 71

제57항 또는 제58항에 있어서,

상기 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 cDNA인, 약제학적 조성물.

청구항 72

제71항에 있어서,

상기 cDNA는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함하는, 약제학적 조성물.

청구항 73

제72항에 있어서,

상기 cDNA는 서열번호 38에 따른 뉴클레오타이드 서열을 포함하는, 약제학적 조성물.

청구항 74

제71항에 있어서,

상기 cDNA는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함하는, 약제학적 조성물.

청구항 75

제74항에 있어서,

상기 cDNA는 서열번호 39에 따른 뉴클레오타이드 서열을 포함하는, 약제학적 조성물.

청구항 76

삭제

청구항 77

제57항에 있어서,

상기 알코올성 간 질환은 알코올 섭취로 인한 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함하는, 약제학적 조성물.

청구항 78

삭제

청구항 79

제57항에 있어서,

상기 비알코올성 간 질환은 비알코올성 지방간 질환(NAFLD) 또는 비알코올성 지방간염(NASH)을 포함하는, 약제학적 조성물.

청구항 80

제57항에 있어서,

상기 비알코올성 간 질환은 알코올 섭취에 의해서 유발되지 않은 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함하는, 약제학적 조성물.

청구항 81

제57항 또는 제58항에 있어서,

상기 인간 대상체는 기능성 HSD17B13에 대해서 동형접합성 또는 이형접합성인, 약제학적 조성물.

청구항 82

제57항 또는 제58항에 있어서,

상기 인간 대상체는 비만인, 약제학적 조성물.

청구항 83

제57항 또는 제58항에 있어서,

상기 인간 대상체는 지방간인, 약제학적 조성물.

청구항 84

제57항 또는 제58항에 있어서,

상기 안티센스 분자는 안티센스 RNA(antisense RNA)를 포함하는, 약제학적 조성물.

청구항 85

제57항 또는 제58항에 있어서,

상기 HSD17B13의 저해제는 siRNA(small interfering RNA)를 포함하는, 약제학적 조성물.

발명의 설명

기 술 분 야

[0001] **서열 목록에 대한 참고**

[0002] 본 출원은 2018년 10월 10일자로 생성되고, 238KB의 크기를 갖는, 파일명 18923801002SEQ의 텍스트 파일로서 전자적으로 제출된 서열 목록을 포함한다. 서열 목록은 참조에 의해 본 명세서에 포함된다.

[0003] **기술분야**

[0004] 본 개시내용은 일반적으로 정밀 의료의 분야에 관한 것이다. 보다 특별하게는, 본 개시내용은 파타틴 유사 포스포리파제 도메인 함유 3(patatin like phospholipase domain containing 3: PNPLA3) Ile148Met 양성이고, 간 질환 또는 간 질환에 대한 민감성을 갖는 대상체를 식별하는 방법 및 이러한 대상체를 하이드록시스테로이드 17-베타 데하이드로게나제 13(hydroxysteroid 17-beta dehydrogenase 13: HSD17B13)의 저해제로 치료하는 방법에 관한 것이다.

배 경 기 술

[0005] 특허, 특허 출원, 공개된 특허 출원, 등록 번호, 기술 논문 및 학술 논문을 비롯한 다양한 참고문헌이 본 명세서 전체에서 인용된다. 각각의 참고문헌은 본 명세서에서 전체적으로 그리고 모든 목적을 위해서 참조에 의해 포함된다.

[0006] 만성 간 질환 및 간경변은 미국에서 질병률 및 사망률의 선두적인 원인이며, 2014년에 사망자는 38,170명(전체 사망의 1.5%)을 차지한다(Kochanek et al., Nat'l. Vital Stat. Rep., 2016, 65, 1-122). 미국에서 간경변의 가장 일반적인 병인론은 알코올성 간 질환, 만성 C형 간염, 및 비알코올성 지방 간 질환(nonalcoholic fatty liver disease: NAFLD)이며, 2004년에서 2013년 사이에 간 이식을 기다리는 환자의 약 80%를 차지한다(Wong et al., Gastroenterology, 2015, 148, 547-555). 미국에서 NAFLD의 예측된 유병률은 19 내지 46%(Browning et al., Hepatology, 2004, 40, 1387-1395; Lazo et al., Am. J. Epidemiol., 2013, 178, 38-45; 및 Williams et al., Gastroenterology, 2011, 140, 124-131)이고, 아마도 이의 주요 인자인 비만율 증가(Cohen et al., Science, 2011, 332, 1519-1523)와 관련하여 점점 증가하고 있다(Younossi et al., Clin. Gastroenterol. Hepatol., 2011, 9, 524-530). C형 간염의 치료에서 상당한 진전이 이루어졌지만, 알코올성 또는 비알코올성 간 질환 및 간경변에 대해서는 현재 어떠한 증거 기반 치료법도 존재하지 않는다.

[0007] 이전의 전장 유전체 연관성 연구(genome wide association studies: GWAS)는 만성 간 질환의 위험성 증가와 연관된 서열 변이를 식별하였다. 가장 강력하게 검증된 연관성은 유전자 PNPLA3에 의해서 암호화된, 파타틴-유사 포스포리파제 도메인-함유 3에서의 일반적인 미스센스 변이체와의 연관성이다. 이러한 변이체(rs738409, p.Ile148Met)는 초기에 간 트라이글리세리드 수준의 증가와 연관되고(Romeo et al., Nat. Genet., 2008, 40, 1461-5), 그 다음 비알코올성 지방간염(NASH)(Rotman et al., Hepatology, 2010, 52, 894-903; Sookoian et al., J. Lipid Res., 2009, 50, 2111-2116) 및 간경변(Shen et al., J. Lipid Res., 2015, 56, 167-175)과 연관되는 것으로 밝혀졌다. 막관통 6 슈퍼패밀리 구성원 2를 암호화하는 TM6SF2에서의 미스센스 변이체가 또한 비알코올성 지방간 질환(NAFLD)의 위험을 증가시킨다(Kozlitina et al., Nat. Genet., 2014, 46, 352-6; Liu et al., Nat. Commun., 2014, 5, 4309; 및 Sookoian et al., Hepatology, 2015, 61, 515-25). PNPLA3 및 TM6SF2에서의 변이체가 정확히 어떻게 간 질환에 기여하는지는 아직 완전히 설명되어 있지는 않다(Smagris et al., J. Biol. Chem., 2016, 291, 10659-76; Mahdessian et al., Proc. Natl. Acad. Sci. USA, 2014, 111, 8913-8; Huang et al., J. Biol. Chem., 2011, 286, 37085-93; 및 Pirazzi et al., J. Hepatol., 2012, 57, 1276-82). 지금까지, 만성 간 질환을 보호하는 어떠한 유전자 변이체도 식별되어 있지 않다.

발명의 내용

[0008] 본 개시내용은 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법을 제공하며, 이 방법은, 대상체로부터의 샘플이 i) I148M 변이를 포함하는 파타틴 유사 포스포리파제 도메인 함유 3(PNPLA3) 단백질을 암호화하는 제1 핵산, 및 기능성 HSD17B13 단백질을 암호화하는 제2 핵산; 및/또는 ii) I148M 변이를 포함

하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질을 포함하는지의 여부를 결정하는 단계; 및 i)에 정의된 바와 같은 제1 핵산 및 제2 핵산 둘 다 및/또는 ii)에 정의된 바와 같은 단백질 둘 다가 검출되는 경우 대상체를 HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 식별하는 단계를 포함한다.

[0009] 일부 실시형태에서, 제1 핵산 분자는 게놈 DNA, mRNA 또는 mRNA로부터 획득된 cDNA를 포함한다.

[0010] 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함하거나; mRNA는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함하거나; mRNA는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함하거나; cDNA는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함하거나; 또는 cDNA는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함한다.

[0011] 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 뉴클레오타이드 서열 또는 서열번호 31과 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함하거나; mRNA는 서열번호 34에 따른 뉴클레오타이드 서열 또는 서열번호 34와 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함하거나; mRNA는 서열번호 35에 따른 뉴클레오타이드 서열 또는 서열번호 35와 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함하거나; cDNA는 서열번호 38에 따른 뉴클레오타이드 서열 또는 서열번호 38과 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함하거나; 또는 cDNA는 서열번호 39에 따른 뉴클레오타이드 서열 또는 서열번호 39와 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함한다.

[0012] 일부 실시형태에서, 제1 핵산을 검출하는 단계는, 제1 핵산의 적어도 일부를 서열분석하는 단계로서, 여기서 일부는 I148M 변이를 암호화하는 코돈을 포함하는, 상기 서열분석하는 단계; 또는 제1 핵산을 제1 핵산의 일부에 특이적으로 혼성화하는 프로브 또는 프라이머와 혼성화시키는 단계로서, 여기서 일부는 I148M 변이를 암호화하는 코돈을 포함하는, 상기 혼성화시키는 단계를 포함한다.

[0013] 일부 실시형태에서, 프로브 또는 프라이머는 대립유전자-특이적 프로브 또는 프라이머이고, 여기서 프로브 또는 프라이머는 표지를 선택적으로 포함한다.

[0014] 일부 실시형태에서, 방법은 대상체가 I148M 변이에 대해서 동형접합성(homozygous)인지 또는 이형접합성(heterozygous)인지를 결정하는 단계를 추가로 포함한다.

[0015] 일부 실시형태에서, 제2 핵산은 게놈 DNA, mRNA 또는 mRNA로부터 획득된 cDNA를 포함한다.

[0016] 일부 실시형태에서, 게놈 DNA는 서열번호 1에 따른 12,667번 위치에 상응하는 위치에 아데닌을 포함하거나; 게놈 DNA는 서열번호 1에 따른 뉴클레오타이드 서열 또는 서열번호 1과 적어도 90%의 서열 동일성을 갖고 기능성 HSD17B13 단백질을 암호화하는 뉴클레오타이드 서열을 포함하거나; mRNA는 서열번호 3에 따른 뉴클레오타이드 서열 또는 서열번호 3과 적어도 90%의 서열 동일성을 갖고 기능성 HSD17B13 단백질을 암호화하는 뉴클레오타이드 서열을 포함하거나; mRNA는 서열번호 4에 따른 뉴클레오타이드 서열 또는 서열번호 4와 적어도 90%의 서열 동일성을 갖고 기능성 HSD17B13 단백질을 암호화하는 뉴클레오타이드 서열을 포함하거나; mRNA는 서열번호 7에 따른 뉴클레오타이드 서열 또는 서열번호 7과 적어도 90%의 서열 동일성을 갖고 기능성 HSD17B13 단백질을 암호화하는 뉴클레오타이드 서열을 포함하거나; mRNA는 서열번호 11에 따른 뉴클레오타이드 서열 또는 서열번호 11과 적어도 90%의 서열 동일성을 갖고 기능성 HSD17B13 단백질을 암호화하는 뉴클레오타이드 서열을 포함하거나; cDNA는 서열번호 12에 따른 뉴클레오타이드 서열 또는 서열번호 12와 적어도 90%의 서열 동일성을 갖고 기능성 HSD17B13 단백질을 암호화하는 뉴클레오타이드 서열을 포함하거나; cDNA는 서열번호 13에 따른 뉴클레오타이드 서열 또는 서열번호 13과 적어도 90%의 서열 동일성을 갖고 기능성 HSD17B13 단백질을 암호화하는 뉴클레오타이드 서열을 포함하거나; cDNA는 서열번호 16에 따른 뉴클레오타이드 서열 또는 서열번호 16과 적어도 90%의 서열 동일성을 갖고 기능성 HSD17B13 단백질을 암호화하는 뉴클레오타이드 서열을 포함하거나; 또는 cDNA는 서열번호 20에 따른 뉴클레오타이드 서열 또는 서열번호 20과 적어도 90%의 서열 동일성을 갖고 기능성 HSD17B13 단백질을 암호화하는 뉴클레오타이드 서열을 포함한다.

[0017] 일부 실시형태에서, 제2 핵산을 검출하는 단계는, 제2 핵산을 서열분석하는 단계; 또는 제2 핵산을 제2 핵산의 일부에 특이적으로 혼성화하는 프로브 또는 프라이머와 혼성화시키는 단계로서, 여기서 일부는 서열번호 1에 따른 12,667번 위치에 상응하는 위치에 아데닌을 포함하는, 상기 혼성화시키는 단계를 포함한다.

[0018] 일부 실시형태에서, 프로브 또는 프라이머는 대립유전자-특이적 프로브 또는 프라이머이고, 여기서 프로브 또는

프라이머는 표지를 선택적으로 포함한다.

[0019] 일부 실시형태에서, 방법은 대상체가 샘플에서 기능성 HSD17B13 단백질을 암호화하는 제2 핵산에 대해서 동형접합성인지 또는 이형접합성인지를 결정하는 단계를 추가로 포함한다.

[0020] 일부 실시형태에서, 방법은 대상체에게 HSD17B13의 저해제를 투여하는 단계를 추가로 포함한다.

[0021] 일부 실시형태에서, 간 질환은 알코올성 간 질환이다. 일부 실시형태에서, 알코올성 간 질환은 알코올 섭취로 인한 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다.

[0022] 일부 실시형태에서, 간 질환은 비알코올성 간 질환이다. 일부 실시형태에서, 비알코올성 간 질환은 비알코올성 지방간 질환(NAFLD) 또는 비알코올성 지방간염(NASH)을 포함한다. 일부 실시형태에서, 비알코올성 간 질환은 알코올 섭취에 의해서 유발되지 않은 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다.

도면의 간단한 설명

[0023] 본 명세서의 부분에 포함되어 이를 구성하는 첨부된 도면은 몇몇 양상을 예시하며, 설명과 함께 본 개시내용의 원리를 설명하기 위해서 제공된다.

도 1은 발견 및 반복 코호트로부터의 서열분석된 유럽 혈통(European-ancestry) 개체의 기준선 특징을 나타낸 도면.

도 2는 발견 코호트에서 1.0×10^{-7} 미만의 P로 혈청 트랜스아미나제와 연관된 단일 뉴클레오타이드 변이체를 나타낸 도면.

도 3은 3개의 개별 유럽 혈통 코호트에서 발견 코호트로부터의 35개의 엑솜-와이드(exome-wide) 유의한 단일 뉴클레오타이드 변이체의 반복 및 공동 메타-분석(replication and joint meta-analysis)을 나타낸 도면.

도 4는 발견 코호트에서 간 질환 표현형을 갖는 13개의 엑솜-와이드 유의한 반복 단일 뉴클레오타이드 변이체의 연관성을 나타낸 도면.

도 5는 댈러스 간 연구(Dallas Liver Study) 및 소아 간 연구(Pediatric Liver Study)로부터의 유전자형분석된 다민족 사례 및 대조군의 기준선 특징을 나타낸 도면.

도 6(패널 A 및 패널 B)은 HSD17B13 주변 영역에서 GHS 발견 코호트에서의 알라닌 아미노트랜스퍼라제(ALT; A) 및 아스파테이트 아미노트랜스퍼라제(AST; B)에 대한 영역 연관성 플롯을 나타낸 도면.

도 7은 HSD17B13 rs72613567 스플라이스 변이체의 동형접합성 참조군(T/T), 이형접합성(T/TA) 및 동형접합성 대체(TA/TA) 보유자에서의 PNPLA3의 발현을 나타낸 도면.

도 8은 3개의 HSD17B13 rs72613567 유전자형(T/T, T/TA, TA/TA)에서의 63 PNPLA3 rs738409 보유자(C/C 및 C/G)의 발현 차이를 나타낸 도면.

도 9는 PNPLA3 rs738409(p.I148M)와 HSD17B13 rs72613567 간의 유전적 상호작용의 분석을 나타낸 도면.

도 10(패널 A 및 패널 B)은 HSD17B13 rs72613567:TA가 PNPLA3 p.I148M과 연관된 간 손상의 위험을 완화시킨다는 것을 나타낸 도면.

도 11(패널 A 내지 패널 F)은 PNPLA3 rs738409(p.I148M) 및 HSD17B13 rs72613567 유전자형에 의한 원시 및 잔차화(residualized) ALT 수준을 나타낸 도면.

도 12(패널 A 내지 패널 F)는 PNPLA3 rs738409(p.I148M) 및 HSD17B13 rs72613567 유전자형에 의한 원시 및 잔차화 AST 수준을 나타낸 도면.

도 13(패널 A 내지 패널 F)은 HSD17B13 스플라이스 변이체의 동형접합성 참조군(T/T), 이형접합성(T/TA) 및 동형접합성 대체(TA/TA) 보유자에서의 4개의 추가적인 신규 HSD17B13 전사물(E-H)의 mRNA 발현을 나타낸 도면.

도 14(패널 A 및 패널 B)는 GHS 발견 코호트에서 혈청 트랜스아미나제 수준과 단일 뉴클레오타이드 변이체 연관성의 맨허튼(Manhattan) 플롯(좌측) 및 분위수(quantile)-분위수 플롯(우측)을 나타낸 도면.

도 15(패널 A 및 패널 B)는 HSD17B13 rs72613567:TA가 알코올성 간 질환 표현형 및 비알코올성 간 질환 표현형의 위험 감소와 연관된다는 것을 나타낸 도면.

도 16(패널 A 및 패널 B)은 HSD17B13 rs72613567:TA가 PNPLA3 p.I148M과 연관된 간 손상의 위험을 완화시킨다는 것을 나타낸 도면.

도 17(패널 A 및 패널 B)은 HSD17B13 rs72613567:TA가 단순 지방증에서 지방간염 및 섬유증으로 진행되는 위험을 감소시키는 것과 연관된다는 것을 나타낸 도면.

도 18(패널 A 내지 패널 G)은 신규 HSD17B13 전사물의 발현, 세포 이하 국지화(subcellular localization) 및 효소 활성도를 나타낸 도면.

도 19(패널 A 및 패널 B)는 HSD17B13 rs72613567:TA가 PNPLA3 I148M과 연관된 알코올성 간 질환 및 비알코올성 간 질환의 위험을 완화시킨다는 것을 나타낸 도면. 각각의 막대 위의 번호는 대조군/사례를 나타낸다.

본 개시내용의 추가 이점은 하기 발명을 실시하기 위한 구체적인 내용에 부분적으로 제시될 것이며, 발명을 실시하기 위한 구체적인 내용으로부터 부분적으로 자명할 것이거나, 본 명세서에 개시된 실시형태의 실시에 의해서 학습될 수 있다. 본 개시내용의 이점은 첨부된 청구범위에 특별하게 제시된 요소 및 조합에 의해서 실현 및 달성될 것이다. 상기 일반적인 설명 및 하기 발명을 실시하기 위한 구체적인 내용 둘 다는 단지 예시 및 설명이며, 청구된 바와 같은 실시형태를 제한하지 않는 것으로 이해되어야 한다.

발명을 실시하기 위한 구체적인 내용

[0024] 본 개시내용의 양상과 관련된 다양한 용어가 명세서 및 청구범위 전체에서 사용된다. 이러한 용어는, 달리 제시되지 않는 한, 관련 기술 분야에서의 이의 통상의 의미로 제공되어야 한다. 다른 구체적으로 정의된 용어는 본 명세서에 제공된 정의와 일관된 방식으로 이해되어야 한다.

[0025] 달리 명확하게 표현되지 않는 한, 본 명세서에 제시된 임의의 방법 또는 양상은, 어떠한 방식으로도, 그 단계가 구체적인 순서로 수행되는 것을 요구하는 것으로 이해되도록 의도되지 않는다. 따라서, 방법 청구항이 청구범위 또는 설명에서 단계가 특정 순서로 제한되어야 한다는 것을 구체적으로 언급하지 않는 경우, 어떠한 경우에도 어떠한 식으로도 순서가 추론되는 것으로 의도되지 않는다. 이는, 단계 또는 작업 흐름의 배열과 관련된 로직의 문제, 문법적 구성 또는 구두점으로부터 유래된 일반적인 의미, 또는 본 명세서에 기재된 양상의 수 또는 양상을 비롯한, 해석에 대한 임의의 가능한 비표현 기반을 수용한다.

[0026] 본 명세서에서 사용되는 바와 같이, 단수 형태는 그 문맥이 달리 명확하게 언급하지 않는 한 복수 대상을 포함한다.

[0027] 본 명세서에서 사용되는 바와 같이, 용어 "대상체" 및 "환자"는 상호 교환 가능하게 사용된다. 대상체는 포유동물을 비롯한 임의의 동물을 포함한다. 포유동물은 비제한적으로 농장 동물(예를 들어, 말, 소, 돼지), 반려 동물(예를 들어, 개, 고양이), 실험 동물(예를 들어, 마우스, 래트, 토끼) 및 비인간 영장류를 포함한다. 일부 실시형태에서, 대상체는 인간이다.

[0028] 본 명세서에서 사용되는 바와 같이, "핵산", "핵산 분자", "핵산 서열", "폴리뉴클레오타이드", 또는 "올리고뉴클레오타이드"는 임의의 길이의 뉴클레오타이드의 중합체 형태를 포함할 수 있고, DNA 및/또는 RNA를 포함할 수 있고, 단일 가닥, 이중 가닥 또는 다중 가닥일 수 있다. 한 가닥의 핵산을 또한 이의 상보체라고 지칭한다.

[0029] 본 명세서에서 사용되는 바와 같이, 구 "에 상응하는" 또는 이의 문법적 변형은, 주어진 아미노산 또는 핵산 서열 또는 위치의 넘버링과 관련하여 사용되는 경우, 주어진 아미노산 서열 또는 핵산 서열을 참조 서열(예를 들어, 본 명세서에서 참조 서열은 예를 들어, (기능성 또는 기능성으로 거동하는 전사물) HSD17B13의 핵산 분자 또는 폴리펩타이드임)과 비교할 때 명시된 참조 서열의 넘버링을 지칭한다. 즉, 주어진 중합체의 잔기(예를 들어, 아미노산 또는 뉴클레오타이드) 번호 또는 잔기(예를 들어, 아미노산 또는 뉴클레오타이드) 위치는 주어진 아미노산 또는 핵산 서열 내의 잔기의 실제 번호 위치에 의하기 보다는 표준 서열에 대해서 지정된다. 예를 들어, 주어진 아미노산 서열은 두 서열 사이에 잔기 매치를 최적화하기 위한 갭을 도입함으로써 참조 서열에 정렬될 수 있다. 이러한 사례에서, 갭이 존재하긴 하지만, 주어진 아미노산 또는 핵산 서열에서 잔기의 넘버링은 이것이 정렬된 참조 서열에 대해서 수행된다.

[0030] 예를 들어, 구 "서열번호 2에 따른 12,667번 위치에 상응하는 위치에 티민을 포함하는 HSD17B13 기능 상실 변이체 단백질(loss-of-function variant protein)을 암호화하는 핵산 분자"(및 유사한 구)는, 조사될 HSD17B13 게놈 DNA의 핵산 서열이 서열번호 2에 따른 뉴클레오타이드 서열에 정렬되는 경우, 조사될 HSD17B13 게놈 DNA가 서열번호 2의 12,667번 위치에 상응하는 위치에 티민을 포함한다는 것을 의미한다.

[0031] 예를 들어, 서열번호 2에 따른 12,667번 위치에 상응하는 위치에 티민을 포함하는 HSD17B13 기능 상실 변이체 단백질을 암호화하는 핵산 분자는 주어진 HSD17B13 단백질과 서열번호 2의 핵산 서열 사이에 서열 정렬을 수행함으로써 쉽게 식별될 수 있다. 마찬가지로, 서열번호 42에 따른 148번 위치에 상응하는 위치 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 갖는 PNPLA3 Ile148Met 단백질은, 주어진 PNPLA3 단백질과 서열번호 42 또는 서열번호 43의 아미노산 서열 사이에 서열 정렬을 수행함으로써 쉽게 식별될 수 있다. 특정 서열번호의 위치에 상응하는 특정 위치에 특정 뉴클레오타이드 또는 아미노산을 갖는 특정 핵산 분자 및 단백질을 식별하기 위해서 서열 정렬을 수행하는 데 사용될 수 있는 다양한 컴퓨터 알고리즘이 존재한다. 예를 들어, 백분율 서열 동일성을 식별하기 위한 프로그램을 사용하여 서열 정렬을 수행할 수 있다. 핵산 내의 핵산 서열의 특정 스트레치 또는 폴리펩타이드 내의 아미노산 서열의 특정 스트레치 간의 백분율 동일성(또는 백분율 상보성)은 BLAST 프로그램(기본적인 국지 정렬 탐색 툴) 및 PowerBLAST 프로그램(Altschul *et al.*, J. Mol. Biol., 1990, 215, 403-410; Zhang and Madden, Genome Res., 1997, 7, 649-656) 또는 CLUSTALW 소프트웨어(Sievers et al., 2014, Methods Mol. Biol., 1079, 105-116)를 사용함으로써 또는 디폴트 설정을 사용하는 Gap 프로그램(위스콘신 서열 분석 패키지(Wisconsin Sequence Analysis Package), 유닉스용 버전 8, 제네틱스 컴퓨터 그룹(Genetics Computer Group), 미국 위스콘신주 매디슨 유니버시티 리서치 파크 소재))(이것은 스미쓰 및 워터만의 알고리즘(Adv. Appl. Math., 1981, 2, 482-489)을 사용함)을 사용함으로써 결정될 수 있다. 그러나, 서열은 또한 수동으로 정렬될 수 있다. 본 명세서에서, 백분율 서열 동일성에 대해서 언급되는 경우, 더 높은 서열 동일성 백분율이 더 낮은 것보다 바람직하다.

[0032] 본 개시내용은, HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법; HSD17B13의 저해제를 투여하는 것을 포함하는 간 질환을 치료 또는 저해하는 방법; 대상체에서 PNPLA3 Ile148Met(본 명세서에서 "I148M"이라고도 지칭됨) 및 기능성 HSD17B13을 검출하는 방법; 간 질환에 대해서 보호 효과를 갖는 대상체를 식별하는 방법; 및 간 질환의 치료에 사용하기 위한 HSD17B13의 저해제를 제공한다.

[0033] 본 개시내용은, HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 분류하는 방법; HSD17B13의 저해제를 투여하는 것을 포함하는 간 질환을 치료 또는 저해하는 방법; 대상체에서 PNPLA3 Ile148Met(본 명세서에서 "I148M"이라고도 지칭됨) 및 기능성 HSD17B13을 검출하는 방법; 간 질환에 대해서 보호 효과를 갖는 대상체를 분류하는 방법; 및 간 질환의 치료에 사용하기 위한 HSD17B13의 저해제를 제공한다.

[0034] 본 개시내용에 따라서, 간 지방 소립 단백질(hepatic lipid droplet protein)인 17-베타 하이드록시스테로이드 데하이드로게나제 13를 암호화하는, HSD17B13에서의 스플라이스 변이체(rs72613567:TA)가 감소된 ALT($P=4.2\times10^{-12}$) 및 AST($P=6.2\times10^{-10}$) 수준과 재현 가능하게 연관되어 있다는 것을 발견하였다. 또한 이러한 변이체는 대립유전자 용량-의존적 방식으로, 알코올성 간 질환 및 비알코올성 간 질환(각각의 rs72613567:TA 대립유전자에 대해서, 각각 38%, 95% 신뢰 구간(CI) 19% 내지 52%; 및 16%, 95% CI 9% 내지 22%) 및 간경변(각각의 rs72613567:TA 대립유전자에 대해서, 알코올성 간경변 및 알코올성 간경변에 대해서 각각 44%, 95% CI 22-59%; 및 26%, 95% CI 12% 내지 38%)의 감소된 위험과 연관되어 있다는 것을 발견하였다. 이러한 연관성은 2개의 독립적인 코호트에서 확인되었다. rs72613567:TA는 비알코올성 지방간염(NASH)의 조직학적 특징의 중증도 감소(지방간 질환을 갖는 개체 중에서 각각의 rs72613567:TA 대립유전자에 대해서 비알코올성 지방간염(NASH)에서 23% 감소, 95% CI 10% 내지 34%) 및 PNPLA3 p.I148M과 연관된 완화된 간 손상과 연관되어 있다. rs72613567:TA는 스테로이드 기질에 대한 효소 활성이 결핍된 절두된 아이소폼을 초래한다. 따라서, HSD17B13에서의 기능 상실 변이체는 알코올성 간 질환 및 비알코올성 간 질환 및 지방증에서 NASH로의 진행의 위험 감소와 연관되어 있다. 미국 특허 출원 공개 제US2018/0216084호(PCT 공개 제WO 2018/136702호에 상응함)는 이의 전문이 참조에 의해 본 명세서에 포함된다.

[0035] 본 개시내용은 하이드록시스테로이드 17-베타 데하이드로게나제 13(HSD17B13)을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법을 제공하며, 이 방법은, 대상체로부터의 샘플이 I148M 변이를 포함하는 파타틴 유사 포스포리파제 도메인 함유 3(PNPLA3) 단백질을 암호화하는 제1 핵산, 및 기능성 HSD17B13 단백질을 암호화하는 제2 핵산, 및/또는 I148M 변이를 포함하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질을 포함하는지의 여부를 결정하는 단계; 및 제1 핵산 및 제2 핵산 둘 다 및/또는 단백질 둘 다가 검출되는 경우 대상체를 HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 식별하는 단계를 포함한다.

[0036] 본 개시내용은, 또한 HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 분류하는 방법; HSD17B13의 저해제를 투여하는 것을 포함하는 간 질환을 치료 또는 저해하는 방법; 대상체에서

PNPLA3 Ile148Met(본 명세서에서 "I148M"이라고도 지칭됨) 및 기능성 HSD17B13을 검출하는 방법; 간 질환에 대해서 보호 효과를 갖는 대상체를 분류하는 방법; 및 간 질환의 치료에 사용하기 위한 HSD17B13의 저해제를 제공한다.

[0037] 본 개시내용은 또한 간 질환을 치료 또는 저해하는 방법을 제공하며, 이 방법은 하이드록시스테로이드 17-베타 데하이드로게나제 13(HSD17B13)의 저해제를, I148M 변이를 포함하는 파타틴 유사 포스포리파제 도메인 함유 3(PNPLA3) 단백질을 발현하는 인간 간 질환 환자에게 투여하여 환자에서 간 질환을 치료 또는 저해하는 단계를 포함한다.

[0038] 본 명세서에 기재된 방법에서, 다양한 PNPLA3 및 HSD17B13 단백질, 및 이를 암호화하는 핵산 분자(예를 들어, 게놈 DNA, mRNA, 및 mRNA로부터 유래된 cDNA)가 검출, 발현 또는 사용된다. 이들 PNPLA3 및 HSD17B13 단백질, 및 이를 암호화하는 핵산 분자를 보다 상세하게 기재한다.

[0039] 두 야생형 PNPLA3 단백질에 대한 아미노산 서열을 서열번호 40 및 서열번호 41에 제시한다. 서열번호 40을 갖는 야생형 PNPLA3 단백질은 481개 아미노산 길이인 반면, 서열번호 41을 갖는 야생형 PNPLA3 단백질은 477개 아미노산 길이이다. 서열번호 40을 갖는 야생형 PNPLA3 단백질은 148번 위치에 아이소류신을 갖는다. 서열번호 41을 갖는 야생형 PNPLA3 단백질은 144번 위치에 아이소류신을 갖는다.

[0040] 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질은 서열번호 42에 따른 아미노산 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하고, 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질은 서열번호 42에 따른 아미노산 서열과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하고, 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질은 서열번호 42에 따른 아미노산 서열을 포함하거나 이것으로 이루어진다.

[0041] 일부 실시형태에서, 변이체 PNPLA3 Ile144Met 단백질은 서열번호 43에 따른 아미노산 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하고, 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile144Met 단백질은 서열번호 43에 따른 아미노산 서열과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하고, 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile144Met 단백질은 서열번호 43에 따른 아미노산 서열을 포함하거나 이것으로 이루어진다.

[0042] 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 및 변이체 PNPLA3 Ile144Met 단백질은 상기에 기재된 단백질의 단편이고, 여기서 단편은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 단편은 암호화된 폴리펩타이드(예컨대, 서열번호 42 또는 서열번호 43의 아미노산 서열을 갖는 폴리펩타이드)의 적어도 약 10, 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 55, 적어도 약 60, 적어도 약 65, 적어도 약 70, 적어도 약 75, 적어도 약 80, 적어도 약 85, 적어도 약 90, 적어도 약 95, 적어도 약 100, 적어도 약 150 또는 적어도 약 200개의 인접 아미노산 잔기를 포함한다. 이와 관련하여, 더 긴 단편이 더 짧은 단편에 비해서 바람직하다. 일부 실시형태에서, 단편은 암호화된 폴리펩타이드의 적어도 약 10, 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 55, 적어도 약 60, 적어도 약 65, 적어도 약 70, 적어도 약 75, 적어도 약 80, 적어도 약 85, 적어도 약 90, 적어도 약 95 또는 적어도 약 100개의 인접 아미노산 잔기를 포함한다. 이와 관련하여, 더 긴 단편이 더 짧은 단편에 비해서 바람직하다.

[0043] 야생형 PNPLA3 단백질을 암호화하는 게놈 DNA 분자에 대한 핵산 서열을 서열번호 30에 제시한다. 서열번호 30을 갖는 야생형 PNPLA3 게놈 DNA 분자는 5109번 위치에 사이토신을 포함한다. 서열번호 30을 갖는 야생형 PNPLA3 게놈 DNA 분자는 5107 내지 5109번 위치에 코돈 ATC를 포함한다.

[0044] 일부 실시형태에서, 변이체 PNPLA3 게놈 DNA 분자는 PNPLA3 Ile148Met 단백질, 또는 각각 서열번호 42 또는 서열번호 43과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%,

적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 게놈 DNA 분자는 PNPLA3 Ile148Met 단백질, 또는 각각 서열번호 42 또는 서열번호 43과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 게놈 DNA 분자는 PNPLA3 Ile148Met 단백질, 또는 각각 서열번호 42 또는 서열번호 43에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0045] 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질 또는 변이체 PNPLA3 Ile144Met 단백질을 암호화하는 변이체 PNPLA3 게놈 DNA 분자는, 서열번호 31과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 31에 따른 5109번 위치에 상응하는 위치에 구아닌을 포함하거나 또는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 코돈 ATG를 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질 또는 변이체 PNPLA3 Ile144Met 단백질을 암호화하는 변이체 PNPLA3 게놈 DNA 분자는, 서열번호 31과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 31에 따른 5109번 위치에 상응하는 위치에 구아닌을 포함하거나 또는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 코돈 ATG를 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질 또는 변이체 PNPLA3 Ile144Met 단백질을 암호화하는 변이체 PNPLA3 게놈 DNA 분자는 서열번호 31에 따른 뉴클레오타이드 서열을 포함하거나 이것으로 이루어진다.

[0046] 일부 실시형태에서, 변이체 PNPLA3 게놈 DNA 분자는 전체 게놈 DNA 서열보다 적은 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 게놈 DNA 분자는 서열번호 31의 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400, 적어도 약 500, 적어도 약 600, 적어도 약 700, 적어도 약 800, 적어도 약 900, 적어도 약 1000, 적어도 약 2000, 적어도 약 3000, 적어도 약 4000, 적어도 약 5000, 적어도 약 6000, 적어도 약 7000, 적어도 약 8000, 적어도 약 9000, 적어도 약 10000, 적어도 약 11000 또는 적어도 약 11500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 PNPLA3 게놈 DNA 분자는 서열번호 31의 적어도 약 1000 내지 적어도 약 2000개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다.

[0047] 일부 실시형태에서, 변이체 PNPLA3 게놈 DNA 분자는 서열번호 31의 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400, 적어도 약 500, 적어도 약 600, 적어도 약 700, 적어도 약 800, 적어도 약 900, 적어도 약 1000, 적어도 약 1100, 적어도 약 1200, 적어도 약 1300, 적어도 약 1400, 적어도 약 1500, 적어도 약 1600, 적어도 약 1700, 적어도 약 1800, 적어도 약 1900, 적어도 약 2000, 적어도 약 2100, 적어도 약 2200, 적어도 약 2300, 적어도 약 2400 또는 적어도 약 2500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다.

[0048] 두 야생형 PNPLA3 mRNA 분자의 핵산 서열을 서열번호 32 및 서열번호 33에 제시한다. 서열번호 32를 갖는 야생형 PNPLA3 mRNA 분자는 444번 위치에 사이토신을 포함한다. 서열번호 32를 갖는 야생형 PNPLA3 mRNA 분자는 442번 내지 444번 위치에 코돈 AUC를 포함한다. 서열번호 33을 갖는 야생형 PNPLA3 mRNA 분자는 432번 위치에 사이토신을 포함한다. 서열번호 33을 갖는 야생형 PNPLA3 mRNA 분자는 430번 내지 432번 위치에 코돈 AUC를 포함한다.

[0049] 일부 실시형태에서, 변이체 PNPLA3 mRNA 분자는 PNPLA3 Ile148Met 단백질, 또는 각각 서열번호 42 또는 서열번호 43과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메

티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 mRNA 분자는 PNPLA3 Ile148Met 단백질, 또는 각각 서열번호 42 또는 서열번호 43과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 mRNA 분자는 PNPLA3 Ile148Met 단백질, 또는 각각 서열번호 42 또는 서열번호 43에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0050] 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질을 암호화하는 변이체 PNPLA3 mRNA 분자는 서열번호 34와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 34에 따른 444번 위치에 상응하는 위치에 구아닌을 포함하거나 또는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 코돈 AUG를 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질을 암호화하는 변이체 PNPLA3 mRNA 분자는 서열번호 34와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 34에 따른 444번 위치에 상응하는 위치에 구아닌을 포함하거나 또는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 코돈 AUG를 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질을 암호화하는 변이체 PNPLA3 mRNA 분자는 서열번호 34에 따른 뉴클레오타이드 서열을 포함하거나 이것으로 이루어진다.

[0051] 일부 실시형태에서, 변이체 PNPLA3 Ile144Met 단백질을 암호화하는 변이체 PNPLA3 mRNA 분자는 서열번호 35와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 35에 따른 432번 위치에 상응하는 위치에 구아닌을 포함하거나, 또는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 코돈 AUG를 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile144Met 단백질을 암호화하는 변이체 PNPLA3 mRNA 분자는 서열번호 35와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 35에 따른 432번 위치에 상응하는 위치에 구아닌을 포함하거나 또는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 코돈 AUG를 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile144Met 단백질을 암호화하는 변이체 PNPLA3 mRNA 분자는 서열번호 35에 따른 뉴클레오타이드 서열을 포함하거나 이것으로 이루어진다.

[0052] 일부 실시형태에서, 변이체 PNPLA3 mRNA 분자는 전체 변이체 PNPLA3 mRNA 서열보다 더 적은 뉴클레오타이드를 포함한다. 일부 실시형태에서, 변이체 PNPLA3 mRNA 분자는 서열번호 34 또는 서열번호 35의 적어도 약 5, 적어도 약 8, 적어도 약 10, 적어도 약 12, 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400, 적어도 약 500 또는 적어도 약 600개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 PNPLA3 mRNA 분자는 서열번호 34 또는 서열번호 35의 적어도 약 200 내지 적어도 약 500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 이와 관련하여, 더 긴 mRNA 분자가 더 짧은 것에 비해서 바람직하다. 일부 실시형태에서, 변이체 PNPLA3 mRNA 분자는 서열번호 34 또는 서열번호 35의 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400 또는 적어도 약 500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 이와 관련하여, 더 긴 mRNA 분자가 더 짧은 것에 비해서 바람직하다. 일부 실시형태에서, 이러한 변이체 PNPLA3 mRNA 분자는 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 암호화하는 코돈 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 암호화하는 코돈을 포함한다. 일부 실시형태에서, 이러한 변이체 PNPLA3 mRNA 분자는 서열번호 34에 따른 444번 위치에 상응하는 위치에 구아닌을 포함하거나 또는 서열번호 35에 따른 432번 위치에 사응하는 위치에 구아닌을 포함한다. 일부 실시형태에서, 이러한 변이체 PNPLA3 mRNA 분자는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 코돈 AUG를 포함하거나 또는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 코돈 AUG를 포함한다.

[0053] 두 야생형 PNPLA3 cDNA 분자의 핵산 서열을 서열번호 36 및 서열번호 37에 제시한다. 서열번호 36을 갖는 야생형 PNPLA3 cDNA 분자는 444번 위치에 사이토신을 포함한다. 서열번호 36을 갖는 야생형 PNPLA3 cDNA 분자는 442 내지 444번 위치에 코돈 ATC를 포함한다. 서열번호 37을 갖는 야생형 PNPLA3 cDNA 분자는 432번 위치에 사이토

신을 포함한다. 서열번호 37을 갖는 야생형 PNPLA3 cDNA 분자는 430 내지 432번 위치에 코돈 ATC를 포함한다.

[0054] 일부 실시형태에서, 변이체 PNPLA3 cDNA 분자는 PNPLA3 Ile148Met 단백질, 또는 각각 서열번호 42 또는 서열번호 43과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 cDNA 분자는 PNPLA3 Ile148Met 단백질, 또는 각각 서열번호 42 또는 서열번호 43과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 cDNA 분자는 PNPLA3 Ile148Met 단백질, 또는 각각 서열번호 42 또는 서열번호 43에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0055] 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질을 암호화하는 변이체 PNPLA3 cDNA 분자는 서열번호 38과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 38에 따른 444번 위치에 상응하는 위치에 구아닌을 포함하거나 또는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 코돈 ATG를 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질을 암호화하는 변이체 PNPLA3 cDNA 분자는 서열번호 38와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 38에 따른 444번 위치에 상응하는 위치에 구아닌을 포함하거나 또는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 코돈 ATG를 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질을 암호화하는 변이체 PNPLA3 cDNA 분자는 서열번호 38에 따른 뉴클레오타이드 서열을 포함하거나 이것으로 이루어진다.

[0056] 일부 실시형태에서, 변이체 PNPLA3 Ile144Met 단백질을 암호화하는 변이체 PNPLA3 cDNA 분자는 서열번호 39와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 39에 따른 432번 위치에 상응하는 위치에 구아닌을 포함하거나 또는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 코돈 ATG를 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile144Met 단백질을 암호화하는 변이체 PNPLA3 cDNA 분자는 서열번호 39와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어지고, 서열번호 39에 따른 432번 위치에 상응하는 위치에 구아닌을 포함하거나 또는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 코돈 ATG를 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile144Met 단백질을 암호화하는 변이체 PNPLA3 cDNA 분자는 서열번호 39에 따른 뉴클레오타이드 서열을 포함하거나 이것으로 이루어진다.

[0057] 일부 실시형태에서, 변이체 PNPLA3 cDNA 분자는 전체 변이체 PNPLA3 cDNA 서열보다 더 적은 뉴클레오타이드를 포함한다. 일부 실시형태에서, 변이체 PNPLA3 cDNA 분자는 서열번호 38 또는 서열번호 39의 적어도 약 5, 적어도 약 8, 적어도 약 10, 적어도 약 12, 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400, 적어도 약 500 또는 적어도 약 600개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 PNPLA3 cDNA 분자는 서열번호 38 또는 서열번호 39의 적어도 약 200 내지 적어도 약 500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 이와 관련하여, 더 긴 cDNA 분자가 더 짧은 것에 비해서 바람직하다. 일부 실시형태에서, 변이체 PNPLA3 cDNA 분자는 서열번호 38 또는 서열번호 39의 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400 또는 적어도 약 500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 이와 관련하여, 더 긴 cDNA 분자가 더 짧은 것에 비해서 바람직하다. 일부 실시형태에서, 이러한 변이체 PNPLA3 cDNA 분자는 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 암호화하는 코돈 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 암호화하는 코돈을 포함한다. 일부 실시형태에서, 이러한 변이체 PNPLA3 cDNA 분자는 서열번호 38에 따른 444번 위치에 상응하는 위치에 구아닌을 포함하거나 또는 서열번호 39에 따른 432번 위치에 사응하는 위치에 구아닌을 포함한다. 일부 실시형태에서, 이러한 변이체 PNPLA3 cDNA 분자는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위

치에 코돈 ATG를 포함하거나 또는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 코돈 ATG를 포함한다.

[0058] 기능성 HSD17B13 단백질과 연관된 4개의 HSD17B13 아이소폼 단백질에 대한 아미노산 서열을 서열번호 21(아이소폼 A), 서열번호 22(아이소폼 B), 서열번호 25(아이소폼 E) 및 서열번호 29(아이소폼 I)에 제시한다. 서열번호 21(아이소폼 A)을 갖는 HSD17B13 단백질은 300개 아미노산 길이이다. 서열번호 22(아이소폼 B)를 갖는 HSD17B13 단백질은 264개 아미노산 길이이다. 서열번호 25(아이소폼 E)를 갖는 HSD17B13 단백질은 324개 아미노산 길이이다. 서열번호 29(아이소폼 I)를 갖는 HSD17B13 단백질은 271개 아미노산 길이이다.

[0059] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 HSD17B13 아이소폼 단백질은 서열번호 21(아이소폼 A)에 따른 아미노산 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 HSD17B13 아이소폼 단백질은 서열번호 21(아이소폼 A)에 따른 아미노산 서열과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 HSD17B13 아이소폼 단백질은 서열번호 21(아이소폼 A)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진다.

[0060] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 HSD17B13 아이소폼 단백질은 서열번호 22(아이소폼 B)에 따른 아미노산 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 HSD17B13 아이소폼 단백질은 서열번호 22(아이소폼 B)에 따른 아미노산 서열과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 HSD17B13 아이소폼 단백질은 서열번호 22(아이소폼 B)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진다.

[0061] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 HSD17B13 아이소폼 단백질은 서열번호 25(아이소폼 E)에 따른 아미노산 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 HSD17B13 아이소폼 단백질은 서열번호 25(아이소폼 E)에 따른 아미노산 서열과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 HSD17B13 아이소폼 단백질은 서열번호 25(아이소폼 E)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진다.

[0062] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 HSD17B13 아이소폼 단백질은 서열번호 29(아이소폼 I)에 따른 아미노산 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 HSD17B13 아이소폼 단백질은 서열번호 29(아이소폼 I)에 따른 아미노산 서열과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 HSD17B13 아이소폼 단백질은 서열번호 29(아이소폼 I)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진다.

[0063] 기능 상실 rs72613567 HSD17B13 단백질(서열번호 2)과 연관된 5개의 HSD17B13 아이소폼 단백질에 대한 아미노산 서열을 서열번호 23(아이소폼 C), 서열번호 24(아이소폼 D), 서열번호 26(아이소폼 F), 서열번호 27(아이소폼 G) 및 서열번호 28(아이소폼 H)에 제시한다. 서열번호 23(아이소폼 C)을 갖는 HSD17B13 단백질은 261개 아미노산 길이이다. 서열번호 24(아이소폼 D)를 갖는 HSD17B13 단백질은 274개 아미노산 길이이다. 서열번호 26(아이소폼 F)을 갖는 HSD17B13 단백질은 284개 아미노산 길이이다. 서열번호 27(아이소폼 G)을 갖는 HSD17B13 단백질은 238개 아미노산 길이이다. 서열번호 28(아이소폼 H)을 갖는 HSD17B13 단백질은 298개 아미노산 길이이다.

[0064] 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 23(아이소폼 C)에 따른 아미노산 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다.

일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 23(아이소폼 C)에 따른 아미노산 서열과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 23(아이소폼 C)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진다.

[0065] 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 24(아이소폼 D)에 따른 아미노산 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 24(아이소폼 D)에 따른 아미노산 서열과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 24(아이소폼 D)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 아미노산 서열을 포함한다.

[0066] 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 26(아이소폼 F)에 따른 아미노산 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 26(아이소폼 F)에 따른 아미노산 서열과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 26(아이소폼 F)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진다.

[0067] 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 27(아이소폼 G)에 따른 아미노산 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 27(아이소폼 G)에 따른 아미노산 서열과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 27(아이소폼 G)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진다.

[0068] 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 28(아이소폼 H)에 따른 아미노산 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 28(아이소폼 H)에 따른 아미노산 서열과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 기능 상실과 연관된 HSD17B13 변이체 단백질은 서열번호 28(아이소폼 H)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진다.

[0069] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 HSD17B13 아이소폼 단백질 및 기능 상실과 연관된 HSD17B13 변이체 단백질은 상기에 기재된 단백질의 단편이다. 일부 실시형태에서, 단편은 암호화된 폴리펩타이드(예컨대, 서열번호 21, 서열번호 22, 서열번호 23, 서열번호 24, 서열번호 25, 서열번호 26, 서열번호 27, 서열번호 28 또는 서열번호 29의 아미노산 서열을 갖는 폴리펩타이드)의 적어도 약 10, 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 55, 적어도 약 60, 적어도 약 65, 적어도 약 70, 적어도 약 75, 적어도 약 80, 적어도 약 85, 적어도 약 90, 적어도 약 95, 적어도 약 100, 적어도 약 150 또는 적어도 약 200개의 인접 아미노산 잔기를 포함한다. 이와 관련하여, 더 긴 단편이 더 짧은 단편에 비해서 바람직하다. 일부 실시형태에서, 단편은 암호화된 폴리펩타이드의 적어도 약 10, 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 55, 적어도 약 60, 적어도 약 65, 적어도 약 70, 적어도 약 75, 적어도 약 80, 적어도 약 85, 적어도 약 90, 적어도 약 95 또는 적어도 약 100개의 인접 아미노산 잔기를 포함한다. 이와 관련하여, 더 긴 단편이 더 짧은 단편에 비해서 바람직하다.

[0070] 기능성 HSD17B13 게놈 DNA 분자에 대한 핵산 서열을 서열번호 1에 제시한다. 서열번호 1을 갖는 기능성 HSD17B13 게놈 DNA 분자는 12,667번 위치에 아데닌을 포함한다.

[0071] 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 21(아이소폼 A)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도

약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 21(아이소폼 A)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 21(아이소폼 A)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0072] 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 22(아이소폼 B)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 22(아이소폼 B)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 22(아이소폼 B)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0073] 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 25(아이소폼 E)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 25(아이소폼 E)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 25(아이소폼 E)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0074] 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 29(아이소폼 I)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 29(아이소폼 I)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 29(아이소폼 I)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0075] 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 1과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 1과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 분자는 서열번호 21에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0076] 기능 상실과 연관된 HSD17B13 변이체 단백질을 암호화하는 변이체 HSD17B13 게놈 DNA 분자에 대한 핵산 서열을 서열번호 2에 제시한다. 서열번호 2를 갖는 변이체 HSD17B13 게놈 DNA 분자는 12,667번 위치에 티민을 포함한다.

[0077] 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 23(아이소폼 C)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 23(아이소폼 C)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의

서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 23(아이소폼 C)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0078] 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 24(아이소폼 D)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 24(아이소폼 D)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 24(아이소폼 D)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0079] 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 26(아이소폼 F)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 26(아이소폼 F)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 26(아이소폼 F)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0080] 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 27(아이소폼 G)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 27(아이소폼 G)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 27(아이소폼 G)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0081] 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 28(아이소폼 H)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 28(아이소폼 H)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 28(아이소폼 H)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0082] 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 2와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 2와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 변이체 HSD17B13 게놈 DNA 분자는 서열번호 2에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0083] 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 및 변이체 HSD17B13 게놈 DNA 분자는 전체 게놈 DNA 서열보다 적은 서열을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 및 변이체 HSD17B13 게놈 DNA 분자는 서열

번호 1(기능성 HSD17B13 게놈 DNA) 또는 서열번호 2(변이체 HSD17B13 게놈 DNA)의 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400, 적어도 약 500, 적어도 약 600, 적어도 약 700, 적어도 약 800, 적어도 약 900, 적어도 약 1000, 적어도 약 2000, 적어도 약 3000, 적어도 약 4000, 적어도 약 5000, 적어도 약 6000, 적어도 약 7000, 적어도 약 8000, 적어도 약 9000, 적어도 약 10000, 적어도 약 11000 또는 적어도 약 11500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 및 변이체 HSD17B13 게놈 DNA 분자는 서열번호 1(기능성 HSD17B13 게놈 DNA) 또는 서열번호 2(변이체 HSD17B13 게놈 DNA)의 적어도 약 1000 내지 적어도 약 2000개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다.

[0084] 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA 및 변이체 HSD17B13 게놈 DNA 분자는 서열번호 1(기능성 HSD17B13 게놈 DNA) 또는 서열번호 2(변이체 HSD17B13 게놈 DNA)의 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400, 적어도 약 500, 적어도 약 600, 적어도 약 700, 적어도 약 800, 적어도 약 900, 적어도 약 1000, 적어도 약 1100, 적어도 약 1200, 적어도 약 1300, 적어도 약 1400, 적어도 약 1500, 적어도 약 1600, 적어도 약 1700, 적어도 약 1800, 적어도 약 1900, 적어도 약 2000, 적어도 약 2100, 적어도 약 2200, 적어도 약 2300, 적어도 약 2400 또는 적어도 약 2500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다.

[0085] 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 4개의 HSD17B13 RNA 전사물에 대한 핵산 서열을 서열번호 44(전사물 A), 서열번호 45(전사물 B), 서열번호 48(전사물 E) 및 서열번호 52(전사물 I)에 제시한다.

[0086] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 21(아이소폼 A)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 21(아이소폼 A)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 21(아이소폼 A)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0087] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 22(아이소폼 B)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 22(아이소폼 B)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 22(아이소폼 B)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0088] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 25(아이소폼 E)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 25(아이소폼 E)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 25(아이소폼 E)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진

HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0089] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 29(아이소폼 I)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 29(아이소폼 I)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 29(아이소폼 I)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0090] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 44(전사물 A)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 44(전사물 A)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 44(전사물 A)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0091] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 45(전사물 B)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 45(전사물 B)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 45(전사물 B)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0092] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 48(전사물 E)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 48(전사물 E)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 48(전사물 E)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0093] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 52(전사물 I)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 52(전사물 I)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 52(전사물 I)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0094] 기능 상실과 연관된 아이소폼 단백질을 암호화하는 5개의 HSD17B13 RNA 전사물에 대한 핵산 서열을 서열번호 46(전사물 C), 서열번호 47(전사물 D), 서열번호 49(전사물 F), 서열번호 50(전사물 G) 및 서열번호 51(전사물 H)에 제시한다.

[0095] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 23(아이소폼 C)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%,

적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 23(아이소폼 C)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 23(아이소폼 C)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0096] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 24(아이소폼 D)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 24(아이소폼 D)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 24(아이소폼 D)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0097] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 26(아이소폼 F)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 26(아이소폼 F)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 26(아이소폼 F)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0098] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 27(아이소폼 G)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 27(아이소폼 G)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 27(아이소폼 G)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0099] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 28(아이소폼 H)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 28(아이소폼 H)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 28(아이소폼 H)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0100] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 46(전

사물 C)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 46(전사물 C)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 46(전사물 C)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0101] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 47(전사물 D)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 47(전사물 D)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 47(전사물 D)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0102] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 49(전사물 F)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 49(전사물 F)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 49(전사물 F)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0103] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 50(전사물 G)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 50(전사물 G)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 50(전사물 G)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0104] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 51(전사물 H)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 51(전사물 H)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 RNA 전사물은 서열번호 51(전사물 H)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0105] 일부 실시형태에서, 기능성 HSD17B13 RNA 전사물 및 변이체 HSD17B13 RNA 전사물은 RNA 전사물 서열보다 적은 서열을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 RNA 전사물 및 변이체 HSD17B13 RNA 전사물은 서열번호 44, 서열번호 45, 서열번호 48 또는 서열번호 52(기능성 HSD17B13 RNA 전사물) 또는 서열번호 46, 서열번호 47, 서열번호 49, 서열번호 50 또는 서열번호 51(변이체 HSD17B13 RNA 전사물)의 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400, 적어도 약 500, 적어도 약 600, 적어도 약 700, 적어도 약 800, 적어도 약 900, 적어도 약 1000, 적어도 약 2000 또는 적어도 약 2500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 RNA 전사물 및 변이체 HSD17B13 RNA 전사물은 RNA 전사물 서열보다 적은 서열을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 RNA 전사물 및 변이체 HSD17B13 RNA 전사물은 서열번호 44, 서열번호 45, 서열번호

48 또는 서열번호 52(기능성 HSD17B13 RNA 전사물) 또는 서열번호 46, 서열번호 47, 서열번호 49, 서열번호 50 또는 서열번호 51(변이체 HSD17B13 RNA 전사물)의 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400 또는 적어도 약 500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 RNA 전사물 및 변이체 HSD17B13 RNA 전사물은 서열번호 44, 서열번호 45, 서열번호 48 또는 서열번호 52(기능성 HSD17B13 RNA 전사물) 또는 서열번호 46, 서열번호 47, 서열번호 49, 서열번호 50 또는 서열번호 51(변이체 HSD17B13 RNA 전사물)의 적어도 약 1000 내지 적어도 약 2000개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다.

[0106] 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 4개의 HSD17B13 cDNA 전사물에 대한 핵산 서열을 서열번호53(전사물 A), 서열번호 54(전사물 B), 서열번호 57(전사물 E) 및 서열번호 61(전사물 I)에 제시한다.

[0107] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 21(아이소폼 A)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 21(아이소폼 A)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 21(아이소폼 A)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0108] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 22(아이소폼 B)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 22(아이소폼 B)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 22(아이소폼 B)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0109] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 25(아이소폼 E)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 25(아이소폼 E)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 25(아이소폼 E)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0110] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 29(아이소폼 I)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 29(아이소폼 I)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는

HSD17B13 cDNA 전사물은 서열번호 29(아이소폼 I)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0111] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 53(전사물 A)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 53(전사물 A)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 53(전사물 A)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0112] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 54(전사물 B)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 54(전사물 B)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 54(전사물 B)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0113] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 57(전사물 E)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 57(전사물 E)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 57(전사물 E)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0114] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 61(전사물 I)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 61(전사물 I)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 61(전사물 I)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0115] 기능 상실과 연관된 아이소폼 단백질을 암호화하는 5개의 HSD17B13 cDNA 전사물에 대한 핵산 서열을 서열번호 55(전사물 C), 서열번호 56(전사물 D), 서열번호 58(전사물 F), 서열번호 59(전사물 G) 및 서열번호 60(전사물 H)에 제시한다.

[0116] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 23(아이소폼 C)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 23(아이소폼 C)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 23(아이소폼 C)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0117] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 24(아

이소폼 D)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 24(아이소폼 D)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 24(아이소폼 D)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0118] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 26(아이소폼 F)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 26(아이소폼 F)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 26(아이소폼 F)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0119] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 27(아이소폼 G)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 27(아이소폼 G)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 27(아이소폼 G)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0120] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 28(아이소폼 H)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 28(아이소폼 H)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 28(아이소폼 H)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0121] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 55(전사물 C)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 55(전사물 C)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 55(전사물 C)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0122] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 56(전사물 D)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적

어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 56(전사물 D)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 56(전사물 D)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0123] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 58(전사물 F)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 58(전사물 F)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 58(전사물 F)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0124] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 59(전사물 G)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 59(전사물 G)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 59(전사물 G)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0125] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 60(전사물 H)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 60(전사물 H)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 전사물은 서열번호 60(전사물 H)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0126] 일부 실시형태에서, HSD17B13 cDNA 전사물은 cDNA 전사물 서열보다 적은 서열을 포함한다. 일부 실시형태에서, HSD17B13 cDNA 전사물은 서열번호 53, 서열번호 54, 서열번호 57 또는 서열번호 61 또는 서열번호 55, 서열번호 56, 서열번호 58, 서열번호 59 또는 서열번호 60의 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400, 적어도 약 500, 적어도 약 600, 적어도 약 700, 적어도 약 800, 적어도 약 900, 적어도 약 1000, 적어도 약 2000 또는 적어도 약 2500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 일부 실시형태에서, HSD17B13 cDNA 전사물은 cDNA 전사물 서열보다 적은 서열을 포함한다. 일부 실시형태에서, HSD17B13 cDNA 전사물은 서열번호 53, 서열번호 54, 서열번호 57 또는 서열번호 61 또는 서열번호 55, 서열번호 56, 서열번호 58, 서열번호 59 또는 서열번호 60의 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400 또는 적어도 약 500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 일부 실시형태에서, HSD17B13 cDNA 전사물은 서열번호 53, 서열번호 54, 서열번호 57 또는 서열번호 61 또는 서열번호 55, 서열번호 56, 서열번호 58, 서열번호 59 또는 서열번호 60의 적어도 약 1000 내지 적어도 약 2000개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다.

[0127] 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 4개의 HSD17B13 mRNA 분자에 대한 핵산 서열을 서열번호 3(전사물 A), 서열번호 4(전사물 B), 서열번호 7(전사물 E) 및 서열번호 11(전사물 I)에 제시한다.

[0128] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서

열번호 21(아이소폼 A)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 21(아이소폼 A)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 21(아이소폼 A)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0129] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 22(아이소폼 B)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 22(아이소폼 B)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 22(아이소폼 B)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0130] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 25(아이소폼 E)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 25(아이소폼 E)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 25(아이소폼 E)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0131] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 29(아이소폼 I)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 29(아이소폼 I)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 29(아이소폼 I)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0132] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 3(전사물 A)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 3(전사물 A)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 3(전사물 A)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0133] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 4(전사물 B)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도

약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 4(전사물 B)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 4(전사물 B)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0134] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 7(전사물 E)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 7(전사물 E)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 7(전사물 E)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0135] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 11(전사물 I)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 11(전사물 I)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 11(전사물 I)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0136] 기능 상실과 연관된 아이소폼 단백질을 암호화하는 5개의 HSD17B13 mRNA 분자에 대한 핵산 서열을 서열번호 5(전사물 C), 서열번호 6(전사물 D), 서열번호 8(전사물 F), 서열번호 9(전사물 G) 및 서열번호 10(전사물 H)에 제시한다.

[0137] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 23(아이소폼 C)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 23(아이소폼 C)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 23(아이소폼 C)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0138] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 24(아이소폼 D)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 24(아이소폼 D)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 24(아이소폼 D)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0139] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 26(아이소폼 F)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하

는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 26(아이소폼 F)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 26(아이소폼 F)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0140] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 27(아이소폼 G)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 27(아이소폼 G)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 27(아이소폼 G)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0141] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 28(아이소폼 H)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 28(아이소폼 H)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 28(아이소폼 H)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0142] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 5(전사물 C)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 5(전사물 C)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 5(전사물 C)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0143] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 6(전사물 D)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 6(전사물 D)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 6(전사물 D)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0144] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 8(전사물 F)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 8(전사물 F)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관

된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 8(전사물 F)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0145] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 9(전사물 G)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 9(전사물 G)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 9(전사물 G)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0146] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 10(전사물 H)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 10(전사물 H)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 mRNA 분자는 서열번호 10(전사물 H)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0147] 일부 실시형태에서, HSD17B13 mRNA 분자는 전체 mRNA 서열보다 더 적은 뉴클레오타이드를 포함한다. 일부 실시형태에서, HSD17B13 mRNA 분자는 서열번호 3, 서열번호 4, 서열번호 7 또는 서열번호 11 또는 서열번호 5, 서열번호 6, 서열번호 8, 서열번호 9 또는 서열번호 10의 적어도 약 5, 적어도 약 8, 적어도 약 10, 적어도 약 12, 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400, 적어도 약 500, 적어도 약 600, 적어도 약 700, 적어도 약 800 또는 적어도 약 900개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 일부 실시형태에서, HSD17B13 mRNA 분자는 서열번호 3, 서열번호 4, 서열번호 7 또는 서열번호 11 또는 서열번호 5, 서열번호 6, 서열번호 8, 서열번호 9 또는 서열번호 10의 적어도 약 200 내지 적어도 약 500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 이와 관련하여, 더 긴 mRNA 분자가 더 짧은 것에 비해서 바람직하다. 일부 실시형태에서, HSD17B13 mRNA 분자는 서열번호 3, 서열번호 4, 서열번호 7 또는 서열번호 11 또는 서열번호 5, 서열번호 6, 서열번호 8, 서열번호 9 또는 서열번호 10의 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400 또는 적어도 약 500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 이와 관련하여, 더 긴 mRNA 분자가 더 짧은 것에 비해서 바람직하다.

[0148] 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 4개의 HSD17B13 cDNA 분자에 대한 핵산 서열을 서열번호 12(전사물 A), 서열번호 13(전사물 B), 서열번호 16(전사물 E) 및 서열번호 20(전사물 I)에 제시한다.

[0149] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 21(아이소폼 A)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 21(아이소폼 A)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 21(아이소폼 A)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0150] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 22(아이소폼 B)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부

실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 22(아이소폼 B)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 22(아이소폼 B)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0151] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 25(아이소폼 E)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 25(아이소폼 E)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 25(아이소폼 E)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0152] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 29(아이소폼 I)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 29(아이소폼 I)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 29(아이소폼 I)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0153] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 12(전사물 A)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 12(전사물 A)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 12(전사물 A)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0154] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 13(전사물 B)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 13(전사물 B)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 13(전사물 B)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0155] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 16(전사물 E)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 16(전사물 E)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 16

(전사물 E)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0156] 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 20(전사물 I)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 20(전사물 I)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능성 HSD17B13 단백질과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 20(전사물 I)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0157] 기능 상실과 연관된 아이소폼 단백질을 암호화하는 5개의 HSD17B13 cDNA 분자에 대한 핵산 서열을 서열번호 14(전사물 C), 서열번호 15(전사물 D), 서열번호 17(전사물 F), 서열번호 18(전사물 G) 및 서열번호 19(전사물 H)에 제시한다.

[0158] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 23(아이소폼 C)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 23(아이소폼 C)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 23(아이소폼 C)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0159] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 24(아이소폼 D)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 24(아이소폼 D)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 24(아이소폼 D)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0160] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 26(아이소폼 F)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 26(아이소폼 F)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 26(아이소폼 F)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0161] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 27(아이소폼 G)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 27(아이소폼 G)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을

포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 27(아이소폼 G)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0162] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 28(아이소폼 H)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 28(아이소폼 H)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 28(아이소폼 H)에 따른 아미노산 서열을 포함하거나 이것으로 이루어진 HSD17B13 아이소폼 단백질을 암호화하는 핵산 서열을 포함하거나 이것으로 이루어진다.

[0163] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 14(전사물 C)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 14(전사물 C)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 14(전사물 C)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0164] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 15(전사물 D)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 15(전사물 D)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 15(전사물 D)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0165] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 17(전사물 F)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 17(전사물 F)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 17(전사물 F)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0166] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 18(전사물 G)과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 18(전사물 G)과 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 18(전사물 G)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0167] 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 19(전사물 H)와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어

도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 19(전사물 H)와 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 이것으로 이루어진다. 일부 실시형태에서, 기능 상실과 연관된 아이소폼 단백질을 암호화하는 HSD17B13 cDNA 분자는 서열번호 19(전사물 H)에 따른 핵산 서열을 포함하거나 이것으로 이루어진다.

[0168] 일부 실시형태에서, HSD17B13 cDNA 분자는 전체 cDNA 서열보다 더 적은 뉴클레오타이드를 포함한다. 일부 실시형태에서, HSD17B13 cDNA 분자는 서열번호 12, 서열번호 13, 서열번호 16 또는 서열번호 20 또는 서열번호 14, 서열번호 15, 서열번호 17, 서열번호 18 또는 서열번호 19의 적어도 약 5, 적어도 약 8, 적어도 약 10, 적어도 약 12, 적어도 약 15, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400, 적어도 약 500, 적어도 약 600, 적어도 약 700, 적어도 약 800 또는 적어도 약 900개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 일부 실시형태에서, HSD17B13 cDNA 분자는 서열번호 12, 서열번호 13, 서열번호 16 또는 서열번호 20 또는 서열번호 14, 서열번호 15, 서열번호 17, 서열번호 18 또는 서열번호 19의 적어도 약 200 내지 적어도 약 500개의 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 이와 관련하여, 더 긴 cDNA 분자가 더 짧은 것에 비해서 바람직하다. 일부 실시형태에서, HSD17B13 cDNA 분자는 서열번호 12, 서열번호 13, 서열번호 16 또는 서열번호 20 또는 서열번호 14, 서열번호 15, 서열번호 17, 서열번호 18 또는 서열번호 19의 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90, 적어도 약 100, 적어도 약 200, 적어도 약 300, 적어도 약 400 또는 적어도 약 500 인접 뉴클레오타이드를 포함하거나 이것으로 이루어진다. 이와 관련하여, 더 긴 cDNA 분자가 더 짧은 것에 비해서 바람직하다.

[0169] 본 명세서에 기재된 프로브 및 프라이머를 사용하여 본 명세서에 기재된 기능성 또는 변이체 PNPLA3 게놈 DNA 분자, mRNA 분자 또는 mRNA 분자로부터 유래된 cDNA 분자 중 임의의 것에 혼성화시킬 수 있다. 프라이머를 사용하여 예를 들어, 본 명세서에 기재된 기능성 또는 변이체 PNPLA3 게놈 DNA 분자, mRNA 분자 또는 mRNA 분자로부터 유래된 cDNA 분자 중 임의의 것의 부분을 증폭시킬 수 있어서, 증폭 산물을 예를 들어, 검출 또는 서열분석할 수 있다.

[0170] 예를 들어, 프로브 및 프라이머를 사용하여 서열번호 30을 포함하는 야생형 PNPLA3 게놈 DNA 분자를 비롯한, 본 명세서에 기재된 야생형 PNPLA3 게놈 DNA 분자 중 임의의 것에 혼성화시킬 수 있다. 프로브 및 프라이머를 또한 사용하여 서열번호 32 또는 서열번호 33을 포함하는 야생형 PNPLA3 mRNA 분자를 비롯한, 본 명세서에 기재된 야생형 PNPLA3 mRNA 분자 중 임의의 것에 혼성화시킬 수 있다. 프로브 및 프라이머를 또한 사용하여 서열번호 36 또는 서열번호 37을 포함하는 야생형 PNPLA3 cDNA 분자를 비롯한, 본 명세서에 기재된 야생형 PNPLA3 cDNA 분자 중 임의의 것에 혼성화시킬 수 있다.

[0171] 프로브 및 프라이머를 또한 사용하여 서열번호 31을 포함하는 변이체 PNPLA3 게놈 DNA 분자를 비롯한, 본 명세서에 기재된 변이체 PNPLA3 게놈 DNA 분자 중 임의의 것에 혼성화시킬 수 있다. 프로브 및 프라이머를 또한 사용하여 서열번호 34 또는 서열번호 35를 포함하는 변이체 PNPLA3 mRNA 분자를 비롯한, 본 명세서에 기재된 변이체 PNPLA3 mRNA 분자 중 임의의 것에 혼성화시킬 수 있다. 프로브 및 프라이머를 또한 사용하여 서열번호 38 또는 서열번호 39를 포함하는 변이체 PNPLA3 cDNA 분자를 비롯한, 본 명세서에 기재된 변이체 PNPLA3 cDNA 분자 중 임의의 것에 혼성화시킬 수 있다.

[0172] 프로브를 사용하여 예를 들어, 본 명세서에 기재된 기능성 또는 변이체 HSD17B13 게놈 DNA 분자, mRNA 분자 또는 mRNA 분자로부터 유래된 cDNA 분자 중 임의의 것을 검출할 수 있다. 프라이머를 사용하여 예를 들어, 본 명세서에 기재된 기능성 또는 변이체 HSD17B13 게놈 DNA 분자, mRNA 분자 또는 mRNA 분자로부터 유래된 cDNA 분자 중 임의의 것의 부분을 증폭시킬 수 있어서, 증폭 산물을 예를 들어, 검출 또는 서열분석할 수 있다.

[0173] 예를 들어, 프로브 및 프라이머를 사용하여 서열번호 1을 포함하는 기능성 HSD17B13 게놈 DNA 분자를 비롯한, 본 명세서에 기재된 기능성 HSD17B13 게놈 DNA 분자 중 임의의 것에 혼성화시킬 수 있다. 프로브 및 프라이머를 또한 사용하여 서열번호 44, 서열번호 45, 서열번호 48 또는 서열번호 52를 포함하는 기능성 HSD17B13 RNA 전사물을 비롯한, 본 명세서에 기재된 기능성 HSD17B13 RNA 전사물 중 임의의 것에 혼성화시킬 수 있다. 프로브 및 프라이머를 또한 사용하여 서열번호 53, 서열번호 54, 서열번호 57 또는 서열번호 61을 포함하는 기능성 HSD17B13 DNA 전사물을 비롯한, 본 명세서에 기재된 기능성 HSD17B13 DNA 전사물 중 임의의 것에 혼성화시킬 수 있다. 프로브 및 프라이머를 또한 사용하여 서열번호 3, 서열번호 4, 서열번호 7 또는 서열번호 11을 포함하는

기능성 HSD17B13 mRNA 분자를 비롯한 본 명세서에 기재된 기능성 HSD17B13 mRNA 분자 중 임의의 것에 혼성화시킬 수 있다. 프로브 및 프라이머를 또한 사용하여 서열번호 12, 서열번호 13, 서열번호 16 또는 서열번호 20을 포함하는 기능성 HSD17B13 cDNA 분자를 비롯한 본 명세서에 기재된 기능성 HSD17B13 cDNA 분자 중 임의의 것에 혼성화시킬 수 있다.

[0174] 프로브 및 프라이머를 또한 사용하여 서열번호 2를 포함하는 변이체 HSD17B13 게놈 DNA 분자를 비롯한, 본 명세서에 기재된 변이체 HSD17B13 게놈 DNA 분자 중 임의의 것에 혼성화시킬 수 있다. 프로브 및 프라이머를 또한 사용하여 서열번호 46, 서열번호 47, 서열번호 49, 서열번호 50 또는 서열번호 51을 포함하는 변이체 HSD17B13 RNA 전사물을 비롯한, 본 명세서에 기재된 변이체 HSD17B13 RNA 전사물 중 임의의 것에 혼성화시킬 수 있다. 프로브 및 프라이머를 또한 사용하여 서열번호 55, 서열번호 56, 서열번호 58, 서열번호 59 또는 서열번호 60을 포함하는 변이체 HSD17B13 DNA 전사물을 비롯한, 본 명세서에 기재된 변이체 HSD17B13 DNA 전사물 중 임의의 것에 혼성화시킬 수 있다. 프로브 및 프라이머를 또한 사용하여 서열번호 5, 서열번호 6, 서열번호 8, 서열번호 9 또는 서열번호 10을 포함하는 변이체 HSD17B13 mRNA 분자를 비롯한, 본 명세서에 기재된 변이체 HSD17B13 mRNA 분자 중 임의의 것에 혼성화시킬 수 있다. 프로브 및 프라이머를 또한 사용하여 서열번호 14, 서열번호 15, 서열번호 17, 서열번호 18 또는 서열번호 19를 포함하는 HSD17B13 cDNA 분자를 비롯한 본 명세서에 기재된 HSD17B13 cDNA 분자 중 임의의 것에 혼성화시킬 수 있다.

[0175] 일부 실시형태에서, 본 명세서에 기재된 프로브 및/또는 프라이머는 본 명세서에 기재된 핵산 분자 또는 이의 상보체 중 임의의 것에 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 엄격한(stringent) 조건 하에서 본 명세서에 개시된 핵산 분자 중 임의의 것에 특이적으로 혼성화한다. 본 개시내용은 또한 온화한 조건 하에서 본 명세서에 개시된 핵산 분자 또는 이의 상보체 중 임의의 것에 혼성화하는 핵산 서열을 갖는 핵산 분자를 제공한다.

[0176] DNA 혼성화를 촉진시키는 적절한 엄격성 조건, 예를 들어, 약 45℃에서 6× 염화나트륨/시트르산나트륨(SSC), 그 다음 50℃에서 2× SSC의 세척을 포함한다(또한 문헌[*Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6)] 참고). 전형적으로, 혼성화 및 검출을 위한 엄격한 조건은 염 농도가 pH 7.0 내지 8.3에서 약 1.5M Na 이온보다 낮고, 전형적으로 약 0.01 내지 1.0M Na 이온 농도(또는 다른 염)이고, 온도가, 짧은 프로브(예를 들어, 10 내지 50개 뉴클레오타이드)의 경우 적어도 약 30℃이고, 보다 긴 프로브(예를 들어, 50개 초과의 뉴클레오타이드)의 경우 적어도 약 60℃인 것일 것이다. 엄격한 조건은 또한 탈안정화제, 예컨대, 폼아마이드의 첨가로 달성될 수 있다. 예시적인 저 엄격성 조건은 37℃에서 30 내지 35% 폼아마이드, 1M NaCl, 1% SDS(황산도데실나트륨)의 완충액으로 혼성화 및 50 내지 55℃에서 1× 내지 2× SSC(20× SSC = 3.0M NaCl/0.3M 구연산삼나트륨)에서의 세척을 포함한다. 예시적인 중간 엄격성 조건은 37℃에서 40 내지 45% 폼아마이드, 1.0M NaCl, 1% SDS에서의 혼성화 및 55 내지 60℃에서 0.5× 내지 1× SSC에서의 세척을 포함한다. 예시적인 고 엄격성 조건은 37℃에서 50% 폼아마이드, 1M NaCl, 1% SDS에서의 혼성화 및 60 내지 65℃에서 0.1× SSC에서의 세척을 포함한다. 선택적으로, 세척 완충액은 약 0.1% 내지 약 1%의 SDS를 포함할 수 있다. 혼성화의 기간은 일반적으로 약 24시간 미만, 통상적으로 약 4 내지 약 12시간이다. 세척 시간의 기간은 적어도 평형에 도달하는데 충분한 시간 기간일 것이다.

[0177] 혼성화 반응에서, 특이성은 전형적으로 혼성화 후 세척의 함수이고, 결정적인 인자는 이온 강도 및 최종 세척 용액의 온도이다. DNA-DNA 혼성체의 경우, T_m은 문헌[Meinkoth and Wahl, *Anal. Biochem.*, 1984, 138, 267-284]의 식: T_m = 81.5℃ + 16.6(log M) + 0.41(GC%) - 0.61(폼%) - 500/L로부터 근사될 수 있고; 식 중 M은 일가 양이온의 몰 농도이고, GC%는 DNA 내의 구아노신 및 사이토신 뉴클레오타이드의 백분율이고, 폼%는 혼성화 용액에서 폼아마이드의 백분율이고, L은 염기쌍에서 혼성체의 길이이다. T_m은 상보적인 표적 서열의 50%가 완벽하게 매치되는 프로브에 혼성화되는 온도(명시된 이온 강도 및 pH 하에서)이다. T_m은 각각 1%의 미스매치에 대해 약 1℃만큼 감소되고; 따라서, T_m, 혼성화 및/또는 세척 조건은 목적하는 동일성의 서열에 혼성화되도록 조정될 수 있다. 예를 들어, 90% 이상의 동일성을 갖는 서열이 추구되면, T_m은 10℃ 감소될 수 있다. 일반적으로, 엄격한 조건은 정의된 이온 강도 및 pH에서 특정 서열 및 이의 상보체에 대한 열 융점(T_m)보다 약 5℃ 낮도록 선택된다. 그러나, 극심한 엄격한 조건은 열 융점(T_m)보다 1℃, 2℃, 3℃ 또는 4℃ 더 낮은 온도에서의 혼성화 및/또는 세척을 활용할 수 있고; 중간 정도의 엄격한 조건은 열 융점(T_m)보다 6℃, 7℃, 8℃, 9℃ 또는 10℃ 더 낮은 온도에서의 혼성화 및/또는 세척을 활용할 수 있고; 저 엄격성 조건은 열 융점(T_m)보다 11℃, 12℃, 13℃,

14℃, 15℃ 또는 20℃ 더 낮은 온도에서의 혼성화 및/또는 세척을 활용할 수 있다. 상기 식, 혼성화 및 세척 조성물 및 목적하는 T_m을 사용하여, 당업자는 혼성화의 엄격성 및/또는 세척 용액에서 변동이 내재적으로 설명된다고 이해할 것이다. 목적하는 정도의 미스매치가 45℃(수성 용액) 또는 32℃(폼아마이드 용액)보다 낮은 T_m을 유발하면, 보다 더 높은 온도가 사용될 수 있도록 SSC 농도를 증가시키는 것이 최적이다.

[0178] 본 명세서에 기재된 프로브는 검출을 보조하기 위해서 표지에 연결 또는 융합될 수 있다. 표지는 직접 검출 가능(예를 들어, 형광단)하거나 또는 간접적으로 검출 가능(예를 들어, 합텐, 효소 또는 형광단 켄처)할 수 있다. 이러한 표지는 분광학, 광화학, 생화학, 면역화학 또는 화학 수단에 의해 검출 가능할 수 있다. 이러한 표지는 예를 들어, 방사선-계수 장치로 측정될 수 있는 방사성 표지; 시각적으로 관찰되거나 또는 분광광도계로 측정될 수 있는 안료, 염료 또는 다른 색소원; 스핀 표지 분석기로 측정될 수 있는 스핀 표지; 및 형광 표지(예를 들어, 형광단)를 포함하고, 여기서 출력 신호는 적합한 분자 부가물의 여기에 의해 생성되고, 염료에 의해 흡수되는 광으로의 여기에 의해 가시화될 수 있거나, 또는 표준 형광측정기 또는 영상화 시스템으로 측정될 수 있다. 표지는 또한, 예로서 화학발광 물질(여기서 출력 신호는 신호 화합물의 화학적 변형에 의해서 생성됨); 금속-함유 물질; 또는 효소(여기서 신호의 효소-의존성 이차 생성, 예컨대, 무색 기질로부터 착색된 산물의 형성이 발생함)일 수 있다. 용어 "표지"는 또한 접합된 분자가 나중에 기질과 함께 첨가될 때, 검출 가능한 신호를 산출하는데 사용되도록, 접합된 분자에 선택적으로 결합할 수 있는 "태그" 또는 합텐을 지칭할 수 있다. 예를 들어, 바이오틴을 태그로서 사용하고, 이어서, 호스래디쉬 퍼옥시데이트(HRP)의 아비딘 또는 스트렙타비딘 접합체를 사용하여 태그에 결합시키고, 이어서 열량측정 기질(예를 들어, 테트라메틸 벤지딘(TMB)) 또는 형광원 기질을 사용하여 HRP의 존재를 검출할 수 있다. 예시적인 표지는 myc, HA, FLAG 또는 3XFLAG, 6XHis 또는 폴리히스티딘, 글루타티온-S-트랜스퍼라제(GST), 말토스 결합 단백질, 에피토프 태그, 또는 면역글로불린의 Fc 부분을 포함하지만 이로 제한되지 않는다. 다수의 표지는 예를 들어, 입자, 형광단, 합텐, 효소 및 이들의 열량측정, 형광원 및 화학발광 기질 및 다른 표지를 포함한다.

[0179] 프로브 또는 프라이머는 임의의 적합한 길이를 포함할 수 있고, 이의 비제한적인 예는 적어도 약 5, 적어도 약 8, 적어도 약 10, 적어도 약 11, 적어도 약 12, 적어도 약 13, 적어도 약 14, 적어도 약 15, 적어도 약 16, 적어도 약 17, 적어도 약 18, 적어도 약 19, 적어도 약 20, 적어도 약 21, 적어도 약 22, 적어도 약 23, 적어도 약 24 또는 적어도 약 25개 뉴클레오타이드 길이를 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 적어도 약 18개 뉴클레오타이드 길이 내지 약 25개 뉴클레오타이드 길이를 포함할 수 있다. 프로브 또는 프라이머는 약 10 내지 약 35, 약 10 내지 약 30, 약 10 내지 약 25, 약 12 내지 약 30, 약 12 내지 약 28, 약 12 내지 약 24, 약 15 내지 약 30, 약 15 내지 약 25, 약 18 내지 약 30, 약 18 내지 약 25, 약 18 내지 약 24, 또는 약 18 내지 약 22개 뉴클레오타이드 길이를 포함할 수 있다. 일부 실시형태에서, 프로브 또는 프라이머는 약 18 내지 약 30개 뉴클레오타이드 길이이다. 대체적으로, 일부 실시형태에서, 프로브는 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 55, 적어도 약 60, 적어도 약 65, 적어도 약 70, 적어도 약 75, 적어도 약 80, 적어도 약 85, 적어도 약 90, 적어도 약 95 또는 적어도 약 100개의 뉴클레오타이드를 포함하거나 이것으로 이루어진다.

[0180] 일부 실시형태에서, 프로브 및/또는 프라이머는 본 명세서에 기재된 야생형 PNPLA3 또는 HSD17B13 핵산 분자 또는 변이체 PNPLA3 또는 HSD17B13 핵산 분자 중 임의의 것과 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99% 또는 100% 동일한 핵산 분자의 적어도 약 15개의 인접 뉴클레오타이드에 혼성화할 수 있다.

[0181] 일부 실시형태에서, 프로브 또는 프라이머는 DNA를 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 RNA를 포함한다.

[0182] 본 명세서에 기재된 프로브 및 프라이는 또한 변경(alteration)-특이적 프로브 및 변경-특이적 프라이머일 수 있다. 변경-특이적 프로브 또는 변경-특이적 프라이머는 변이체 PNPLA3 단백질을 암호화하는 핵산 서열에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하지만, 야생형 PNPLA3 단백질을 암호화하는 핵산 서열에 상보성이 아니고/아니거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함할 수 있다. 이와 관련하여, "특이적으로 혼성화한다"는 프로브 또는 프라이머(예를 들어, 변경-특이적 프로브 또는 변경-특이적 프라이머)가 야생형 PNPLA3 단백질을 암호화하는 핵산 분자에 혼성화하지 않는다는 것을 의미한다. 본 명세서에서, 용어 "특이적으로 혼성화한다"는 제시된 핵산 분자에만 혼성화하고, 또 다른 핵산 분자에 혼성화하지 않는 프로브 또는 프라이머를 의미한다. 따라서, I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 핵산 분자에 특이적으로 혼성화하는 프로브 또는 프라이머는 I148M 변이를 포함하지 않는 PNPLA3 단백질을 암호화하는 핵산 분자에 혼성화하지

않는다. 변경-특이적 프로브 또는 변경-특이적 프라이머는 또한 야생형 PNPLA3 단백질을 암호화하는 핵산 서열에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하지만, 변이체 PNPLA3 단백질을 암호화하는 핵산 서열에 상보성이 아니고/아니거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함할 수 있다. 이와 관련하여, "특이적으로 혼성화한다"는 프로브 또는 프라이머(예를 들어, 변경-특이적 프로브 또는 변경-특이적 프라이머)가 변이체 PNPLA3 단백질을 암호화하는 핵산 분자에 혼성화하지 않는다는 것을 의미한다.

[0183] 변경-특이적 프로브 또는 변경-특이적 프라이머는 또한 변이체 HSD17B13 단백질을 암호화하는 핵산 서열에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하지만, 기능성 HSD17B13 단백질을 암호화하는 핵산 서열에 상보성이 아니고/아니거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함할 수 있다. 이와 관련하여, "특이적으로 혼성화한다"는 프로브 또는 프라이머(예를 들어, 변경-특이적 프로브 또는 변경-특이적 프라이머)가 기능성 HSD17B13 단백질을 암호화하는 핵산 분자에 혼성화하지 않는다는 것을 의미한다. 예를 들어, 이와 관련하여, "특이적으로 혼성화한다"는 프로브 또는 프라이머가 비활성/기능 상실 HSD17B13 단백질을 암호화하는 핵산 분자에 혼성화하지 않는다는 것을 의미한다. 변경-특이적 프로브 또는 변경-특이적 프라이머는 또한 가능성 HSD17B13 단백질을 암호화하는 핵산 서열에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하지만, 변이체 HSD17B13 단백질을 암호화하는 핵산 서열에 상보성이 아니고/아니거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함할 수 있다. 이와 관련하여, "특이적으로 혼성화한다"는 프로브 또는 프라이머(예를 들어, 변경-특이적 프로브 또는 변경-특이적 프라이머)가 변이체 HSD17B13 단백질을 암호화하는 핵산 분자에 혼성화하지 않는다는 것을 의미한다.

[0184] 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함하는 PNPLA3 핵산 서열의 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 게놈 DNA의 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함하는 변이체 PNPLA3 mRNA의 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함하는 변이체 PNPLA3 mRNA의 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 cDNA의 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 cDNA의 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다.

[0185] 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 1에 따른 12,667번 위치에 상응하는 위치에 아데닌을 포함하는 HSD17B13 게놈 DNA의 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 2에 따른 12,667번 위치에 상응하는 위치에 티민을 포함하는 HSD17B13 게놈 DNA의 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다.

[0186] 일부 실시형태에서, 프로브 또는 프라이머가 혼성화되는 핵산 분자의 일부는 약 10 내지 약 200, 약 10 내지 약 150, 약 10 내지 약 100, 약 10 내지 약 50, 약 10 내지 약 40, 약 10 내지 약 30, 또는 약 10 내지 약 20개의 뉴클레오타이드를 포함하고, 특정 변이를 암호화하는 코돈을 함유하는 위치에 상응하는 코돈(예를 들어, PNPLA3의 I148M 또는 상응하는 야생형 HSD17B13 단백질과 상이한 변이체 HSD17B13 단백질의 일부)을 포함한다. 일부 바람직한 실시형태에서, 프로브 또는 프라이머가 혼성화되는 핵산 분자의 일부는 약 10 내지 약 50, 약 10 내지 약 40, 약 10 내지 약 30, 또는 약 10 내지 약 20개의 뉴클레오타이드를 포함하고, 특정 변이를 암호화하는 코돈을 함유하는 위치에 상응하는 코돈(예를 들어, PNPLA3의 I148M 또는 상응하는 야생형 HSD17B13 단백질과 상이한 변이체 HSD17B13 단백질의 일부)을 포함한다.

[0187] 본 명세서에 기재된 키트는 본 명세서에 기재된 야생형 PNPLA3 및/또는 HSD17B13 핵산 분자 중 임의의 것 및/또는 본 명세서에 기재된 변이체 PNPLA3 및/또는 HSD17B13 핵산 분자 중 임의의 것을 검출 및/또는 증폭시키는 데 사용될 수 있는 검출 및/또는 증폭 검정 시약을 포함할 수 있다. 일부 실시형태에서, 이러한 검출 및/또는 증폭을 위한 키트는 본 명세서에 기재된 시약(예를 들어, 프로브 및 프라이머) 중 임의의 것을 함유할 수 있다. 일

부 실시형태에서, 베이직 키트는 본 명세서에 개시된 핵산 분자 중 임의의 것 내의 유전자좌에 대해서, 적어도 하나의 프로브 또는 프라이머 또는 적어도 2개의 프로브 또는 프라이머, 예컨대, 변경-특이적 프로브 또는 변경-특이적 프라이머를 갖는 용기를 포함할 수 있다. 키트는 또한 사용 설명서를 선택적으로 포함할 수 있다. 키트는 또한 다른 선택적인 키트 성분, 예를 들어, 증폭된 유전자좌 각각에 안내되는 대립유전자 래더, 충분한 품질의 증폭용 효소, 증폭을 용이하게 하기 위한 증폭 완충제, 효소 활성을 용이하게 하기 위한 2가 양이온 용액, 증폭 동안 가닥 연장을 위한 dNTP, 전기영동법을 위한 증폭된 재료의 제조를 위한 로딩 용액, 주형 대조군으로서의 게놈 DNA, 분리 배지에서 기대되는 바와 같이 물질이 이동하는 것을 보장하기 위한 크기 마커 및 사용자를 교육하고, 사용시 오류를 제한하기 위한 프로토콜 및 매뉴얼 중 하나 이상을 포함할 수 있다. 키트 중의 다양한 시약의 양은 또한 다수의 인자, 예컨대, 공정의 최적의 민감도에 따라서 달라질 수 있다. 수동 응용에 사용하기 위한 시험 키트 또는 자동화 샘플 제조, 반응 설정, 검출기 또는 분석기와 함께 사용하기 위한 시험 키트를 제공하는 것은 이러한 개시내용의 범주에 포함된다. 일부 실시형태에서, 키트는 검출을 위한 적어도 하나의 표지된 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 본 명세서에 개시된 키트 중 임의의 것은 어닐링 반응을 수행하는 데 필요한 시약 및 생성물, 및 설명서를 추가로 포함할 수 있다.

[0188] 본 개시내용은 본 명세서에 기재된 야생형 PNPLA3 단백질 중 임의의 것의 존재를 검출하기 위한 방법을 제공한다. 본 개시내용은 또한 본 명세서에 기재된 변이체 PNPLA3 단백질 중 임의의 것의 존재를 검출하기 위한 방법을 제공한다. 본 개시내용은 또한 본 명세서에 기재된 야생형 PNPLA3 핵산 분자(예를 들어, 본 명세서에 기재된 게놈 DNA 분자, mRNA 분자, 및 cDNA 분자) 중 임의의 것을 검출하기 위한 방법을 제공한다. 본 개시내용은 또한 본 명세서에 기재된 변이체 PNPLA3 핵산 분자(예를 들어, 본 명세서에 기재된 게놈 DNA 분자, mRNA 분자, 및 cDNA 분자) 중 임의의 것을 검출하기 위한 방법을 제공한다.

[0189] 본 개시내용은 또한 본 명세서에 기재된 기능성 HSD17B13 단백질 중 임의의 것의 존재를 검출하기 위한 방법을 제공한다. 본 개시내용은 또한 본 명세서에 기재된 변이체 HSD17B13 단백질 중 임의의 것의 존재를 검출하기 위한 방법을 제공한다. 본 개시내용은 또한 본 명세서에 기재된 기능성 HSD17B13 핵산 분자(예를 들어, 본 명세서에 기재된 게놈 DNA 분자, RNA 전사물, cDNA 전사물, mRNA 분자, 및 cDNA 분자) 중 임의의 것을 검출하기 위한 방법을 제공한다. 본 개시내용은 또한 본 명세서에 기재된 변이체 HSD17B13 핵산 분자(예를 들어, 본 명세서에 기재된 게놈 DNA 분자, RNA 전사물, cDNA 전사물, mRNA 분자, 및 cDNA 분자) 중 임의의 것을 검출하기 위한 방법을 제공한다.

[0190] 본 명세서에 기재된 방법 중 임의의 것의 일부 실시형태에서, 기능성 HSD17B13 단백질 또는 이를 암호화하는 핵산 분자는 대상체 또는 환자에서 검출되거나 검출이 추구된다. 일부 실시형태에서, 대상체 또는 환자는 기능성 HSD17B13 단백질을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 본 명세서에 기재된 기능성 HSD17B13 단백질(이를 암호화하는 본 명세서에 기재된 핵산 분자 중 임의의 것에 의해서 암호화될 수 있음) 중 하나이다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40에 따른 아미노산 서열을 갖는 HSD17B13 단백질의 생물학적 활성도의 적어도 90%, 적어도 80%, 적어도 70%, 적어도 60%, 적어도 50%, 적어도 40%, 적어도 30%, 적어도 20%, 적어도 10%, 적어도 5% 또는 적어도 1%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40에 따른 아미노산 서열을 갖는 HSD17B13 단백질의 생물학적 활성도의 적어도 90%, 적어도 80%, 적어도 70%, 적어도 60% 또는 적어도 50%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 적어도 90%, 적어도 80%, 적어도 70%, 적어도 60%, 적어도 50%, 적어도 40%, 적어도 30%, 적어도 20%, 적어도 10%, 적어도 5% 또는 적어도 1%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 적어도 70%, 적어도 60%, 적어도 50%, 적어도 40%, 적어도 30%, 적어도 20%, 적어도 10%, 적어도 5% 또는 적어도 1%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 적어도 50%, 적어도 40%, 적어도 30%, 적어도 20%, 적어도 10%, 적어도 5% 또는 적어도 1%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 적어도 30%, 적어도 20%, 적어도 10%, 적어도 5% 또는 적어도 1%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 90%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 80%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 70%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 60%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 50%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 40%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적

활성도의 30%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 20%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 10%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 5%를 갖는다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 40을 갖는 HSD17B13 단백질의 생물학적 활성도의 1%를 갖는다. 일부 실시형태에서, HSD17B13 단백질의 활성도(예를 들어, 기능성)는 예를 들어, 옥시도리덕타제 활성도 검정을 수행함으로써 결정될 수 있다.

[0191] 집단 내에서 유전자 서열, 및 이러한 유전자에 의해 암호화된 mRNA 및 단백질은 다형성, 예컨대, 단일-뉴클레오타이드 다형성으로 인해 달라질 수 있다는 것이 이해된다. 본 명세서에 제공된 서열은 단지 예시적인 서열이다. 변이체 PNPLA3 및 HSD17B13 게놈 DNA, mRNA, cDNA 및 폴리펩타이드의 다른 서열이 또한 가능하다.

[0192] 생물학적 샘플은 대상체로부터의 임의의 세포, 조직 또는 생물학적 유체로부터 유래될 수 있다. 샘플은 임의의 임상적으로 관련된 조직, 예컨대, 골수 샘플, 종양 생검, 미세 바늘 흡인물 또는 체액, 예컨대, 혈액, 치은 열구액(gingival crevicular fluid), 혈장, 혈청, 림프, 복수, 낭성 유체 또는 소변을 포함할 수 있다. 일부 실시형태에서, 샘플은 협측 스왑(swab)을 포함한다. 본 명세서에 개시된 방법에서 사용되는 샘플은 검정 포맷, 검출 방법의 특성, 및 샘플로서 사용되는 조직, 세포 또는 추출물을 기초로 달라질 것이다. 생물학적 샘플은 사용되는 검정에 따라서 상이하게 가공될 수 있다. 예를 들어, 변이체 PNPLA3 핵산 분자를 검출하는 경우, 게놈 DNA에 대한 샘플을 단리시키거나 또는 농축시키도록 설계된 예비 가공이 사용될 수 있다. 다양한 기술이 이러한 목적으로 사용될 수 있다. 변이체 PNPLA3 mRNA의 수준을 검출하는 경우, 상이한 기술이 mRNA를 갖는 생물학적 샘플을 농축시키는 데 사용될 수 있다. mRNA의 존재 또는 수준 또는 특정 변이체 게놈 DNA 유전자좌의 존재를 검출하는 다양한 방법이 사용될 수 있다.

[0193] 일부 실시형태에서, 특정 PNPLA3 단백질 또는 HSD17B13 단백질(예를 들어, 기능성 또는 변이체)의 존재 또는 부재는, 단백질의 적어도 일부를 서열분석하여 단백질이 본 명세서에 기재된 변이체 PNPLA3 단백질 또는 HSD17B13 단백질(예를 들어, 기능성 또는 변이체) 중 임의의 것을 암호화하는 아미노산을 포함하는지의 여부를 결정함으로써 검출된다. 일부 실시형태에서, 특정 PNPLA3 단백질 또는 HSD17B13 단백질(예를 들어, 기능성 또는 변이체)의 존재 또는 부재는, 면역검정, 예컨대, ELISA를 수행하여 본 명세서에 기재된 변이체 PNPLA3 단백질 또는 HSD17B13 단백질(예를 들어, 기능성 또는 변이체) 중 임의의 것이 샘플 중에 존재하는지의 여부를 결정함으로써 검출된다.

[0194] 일부 실시형태에서, 서열분석된 단백질 중 일부는 약 5 내지 약 100, 약 5 내지 약 50, 약 5 내지 약 40, 약 5 내지 약 30, 약 5 내지 약 20 또는 약 5 내지 약 10개의 아미노산을 포함하고, 변이(예를 들어, PNPLA3의 I148M 또는 상응하는 야생형 HSD17B13 단백질과 상이한 변이체 HSD17B13 단백질의 일부)를 함유하는 위치에 상응하는 위치를 포함한다. 일부 바람직한 실시형태에서, 서열분석된 단백질 중 일부는 약 5 내지 약 20 또는 약 5 내지 약 10개의 아미노산을 포함하고, 변이(예를 들어, PNPLA3의 I148M 또는 상응하는 야생형 HSD17B13 단백질과 상이한 변이체 HSD17B13 단백질의 일부)를 함유하는 위치에 상응하는 위치를 포함한다.

[0195] 단백질 서열분석 기술의 예시적인 비제한적인 예는 질량 분광분석법 및 에드만 분해를 포함하지만 이들로 제한되지 않는다. 면역검정의 예시적인 면역침전, 웨스턴 블롯, 면역조직화학, ELISA, 면역세포화학, 유세포 분석법 및 면역-PCR을 포함하지만 이들로 제한되지 않는다. 다양한 기술(예를 들어, 비색정량, 형광, 화학발광 또는 방사성)을 이용하여 검출 가능하게 표지된 다클론성 또는 단클론성 항체가 면역검정에서의 사용에 적합하다.

[0196] 일부 실시형태에서, 특정 PNPLA3 핵산 분자 또는 HSD17B13 핵산 분자(예를 들어, 기능성 또는 변이체 게놈 DNA, mRNA, cDNA, RNA 전사물 또는 cDNA 전사물)의 존재 또는 부재는 핵산 분자의 적어도 일부를 서열분석하여 핵산 분자가 본 명세서에 기재된 변이체 PNPLA3 핵산 분자 또는 HSD17B13 핵산 분자(예를 들어, 기능성 또는 변이체) 중 임의의 것에 따른 핵산 서열를 포함을 포함하는지의 여부를 결정함으로써 검출된다.

[0197] 일부 실시형태에서, 서열분석되는 핵산 분자의 일부는 약 10 내지 약 200, 약 10 내지 약 150, 약 10 내지 약 100, 약 10 내지 약 50, 약 10 내지 약 40, 약 10 내지 약 30, 또는 약 10 내지 약 20개의 뉴클레오타이드를 포함하고, 특정 변이를 암호화하는 코돈을 함유하는 위치에 상응하는 코돈(예를 들어, PNPLA3의 I148M 또는 상응하는 야생형 HSD17B13 단백질과 상이한 변이체 HSD17B13 단백질의 일부)을 포함한다. 일부 바람직한 실시형태에서, 서열분석되는 핵산 분자의 일부는 약 10 내지 약 50, 약 10 내지 약 40, 약 10 내지 약 30, 또는 약 10 내지 약 20개의 뉴클레오타이드를 포함하고, 특정 변이를 암호화하는 코돈을 함유하는 위치에 상응하는 코돈(예를 들어, PNPLA3의 I148M 또는 상응하는 야생형 HSD17B13 단백질과 상이한 변이체 HSD17B13 단백질의 일부)을 포함한다.

[0198] 일부 실시형태에서, 대상체에서 본 명세서에 기재된 특정 PNPLA3 핵산 분자 또는 HSD17B13 핵산 분자 중 임의의 것(예를 들어, 기능성 또는 변이체 게놈 DNA 분자, mRNA 분자, cDNA 분자, RNA 전사물 또는 cDNA 전사물 중 임의의 것)의 존재 또는 부재를 검출하는 방법은, 대상체로부터 획득된 생물학적 샘플에 대해서 검정을 수행하는 것을 포함하는데, 이 검정은 생물학적 샘플 중의 핵산 분자가 본 명세서에 기재된 특정 PNPLA3 핵산 분자 또는 HSD17B13 핵산 분자 중 임의의 것(예를 들어, 기능성 또는 변이체 게놈 DNA 분자, mRNA 분자, cDNA 분자, RNA 전사물 또는 cDNA 전사물 중 임의의 것)을 포함하는지의 여부를 결정한다. 일부 실시형태에서, 생물학적 샘플은 세포 또는 세포 용해물을 포함한다. 이러한 방법은 예를 들어, 대상체로부터 생물학적 샘플을 획득하는 단계, 선택적으로 mRNA를 cDNA로 역전사시키는 단계, 및 검정을 수행하는 단계를 추가로 포함할 수 있다. 이러한 검정은 예를 들어, 본 명세서에 기재된 특정 핵산 분자의 특정 위치의 동일성을 결정하는 단계를 포함할 수 있다.

[0199] 예를 들어, 검정은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함하는 PNPLA3 핵산 서열의 일부에 또는 이에 인접한 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함하는 변경-특이적 프로브 또는 변경-특이적 프라이머를 사용하는 것을 포함할 수 있다. 일부 실시형태에서, 검정은 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 게놈 DNA의 일부에, 또는 이에 인접한 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함하는 변경-특이적 프로브 또는 변경-특이적 프라이머를 사용하는 것을 포함할 수 있다. 일부 실시형태에서, 검정은 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함하는 변이체 PNPLA3 mRNA의 일부 또는 이에 인접한 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함하는 변경-특이적 프로브 또는 변경-특이적 프라이머를 사용하는 것을 포함할 수 있다. 일부 실시형태에서, 검정은 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함하는 변이체 PNPLA3 mRNA의 일부 또는 이에 인접한 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함하는 변경-특이적 프로브 또는 변경-특이적 프라이머를 사용하는 것을 포함할 수 있다. 일부 실시형태에서, 검정은 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 cDNA의 일부 또는 이에 인접한 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함하는 변경-특이적 프로브 또는 변경-특이적 프라이머를 사용하는 것을 포함할 수 있다. 일부 실시형태에서, 검정은 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 cDNA의 일부 또는 이에 인접한 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함하는 변경-특이적 프로브 또는 변경-특이적 프라이머를 사용하는 것을 포함할 수 있다.

[0200] 일부 실시형태에서, 검정은 서열번호 1에 따른 12,667번 위치에 상응하는 위치에 아데닌을 포함하는 HSD17B13 게놈 DNA의 일부에, 또는 이에 인접한 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함하는 변경-특이적 프로브 또는 변경-특이적 프라이머를 사용하는 것을 포함할 수 있다. 일부 실시형태에서, 검정은 서열번호 2에 따른 12,667번 위치에 상응하는 위치에 티민을 포함하는 HSD17B13 게놈 DNA의 일부에, 또는 이에 인접한 일부에 상보성이고/이거나 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함하는 변경-특이적 프로브 또는 변경-특이적 프라이머를 사용하는 것을 포함할 수 있다.

[0201] 일부 실시형태에서, 검정은 대상체로부터의 생물학적 샘플 중에 존재하는 본 명세서에 기재된 핵산 분자의 적어도 일부를 서열분석하는 것을 포함하며, 여기서 서열분석되는 일부는 본 명세서에 개시된 위치를 포함한다. 예를 들어, 서열분석되는 일부는 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나, 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함하는 PNPLA3 핵산 서열의 일부일 수 있다. 일부 실시형태에서, 서열분석되는 일부는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 게놈 DNA의 일부일 수 있다. 일부 실시형태에서, 서열분석되는 일부는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함하는 변이체 PNPLA3 mRNA의 일부일 수 있다. 일부 실시형태에서, 서열분석되는 일부는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함하는 변이체 PNPLA3 mRNA의 일부일 수 있다. 일부 실시형태에서, 서열분석되는 일부는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 cDNA의 일부일 수 있다. 일부 실시형태에서, 서열분석되는 일부는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 cDNA의 일부일 수 있다.

[0202] 일부 실시형태에서, 서열분석되는 일부는 서열번호 1에 따른 12,667번 위치에 상응하는 위치에 아데닌을 포함하는 HSD17B13 게놈 DNA의 일부일 수 있다. 일부 실시형태에서, 서열분석되는 일부는 서열번호 2에 따른 12,667번 위치에 상응하는 위치에 티민을 포함하는 HSD17B13 게놈 DNA의 일부일 수 있다.

[0203] 일부 실시형태에서, 검정은 a) 생물학적 샘플을, 본 명세서에서 식별된 핵산 분자의 일부에 인접한(예를 들어, 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나, 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함하는 PNPLA3 핵산 서열의 일부에 인접하거나; 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 게놈 DNA의 일부에 인접하거나; 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함하는 변이체 PNPLA3 mRNA의 일부에 인접하거나; 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함하는 변이체 PNPLA3 mRNA의 일부에 인접하거나; 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 cDNA의 일부에 인접하거나; 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 cDNA의 일부에 인접하거나; 서열번호 1에 따른 12,667번 위치에 상응하는 위치에 아데닌을 포함하는 HSD17B13 게놈 DNA의 일부에 인접하거나; 또는 서열번호 2에 따른 12,667번 위치에 상응하는 위치에 티민을 포함하는 HSD17B13 게놈 DNA의 일부에 인접한) 영역에 혼성화하는 프라이머(또는 변경-특이적 프라이머)와 접촉시키는 단계; b) 프라이머를 적어도 뉴클레오타이드 위치에 상응하는 핵산 분자의 위치를 통해서 변경된 부위(예를 들어, 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나, 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함하는 PNPLA3 핵산 서열의 일부; 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 게놈 DNA의 일부; 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함하는 변이체 PNPLA3 mRNA의 일부; 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함하는 변이체 PNPLA3 mRNA의 일부; 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 cDNA의 일부; 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 cDNA의 일부; 서열번호 1에 따른 12,667번 위치에 상응하는 위치에 아데닌을 포함하는 HSD17B13 게놈 DNA의 일부; 또는 서열번호 2에 따른 12,667번 위치에 상응하는 위치에 티민을 포함하는 HSD17B13 게놈 DNA의 일부)를 지나서 연장시키는 단계; 및 c) 프라이머의 연장 산물이 본 명세서에 기재된 변이체 또는 야생형 PNPLA3 또는 HSD17B13 핵산 분자 중 임의의 것의 핵산 서열을 포함하는지의 여부를 결정하는 단계를 포함한다.

[0204] 일부 실시형태에서, PNPLA3 게놈 DNA만 분석된다. 일부 실시형태에서, PNPLA3 mRNA만 분석된다. 일부 실시형태에서, PNPLA3 mRNA로부터 획득된 PNPLA3 cDNA만 분석된다. 일부 실시형태에서, HSD17B13 게놈 DNA만 분석된다. 일부 실시형태에서, HSD17B13 mRNA만 분석된다. 일부 실시형태에서, HSD17B13 mRNA로부터 획득된 HSD17B13 cDNA만 분석된다. 일부 실시형태에서, HSD17B13 RNA 전사물만 분석된다. 일부 실시형태에서, HSD17B13 RNA 전사물로부터 획득된 HSD17B13 cDNA만 분석된다.

[0205] 일부 실시형태에서, 검정은 생물학적 샘플을 엄격한 조건 하에서 상응하는 기능성 핵산 분자가 아닌 본 명세서에 기재된 특정 변이체 PNPLA3 핵산 분자 또는 변이체 HSD17B13 핵산 분자 중 임의의 것(예를 들어, 변이체 게놈 DNA 분자, mRNA 분자, cDNA 분자, RNA 전사물 또는 cDNA 전사물 중 임의의 것)에 특이적으로 혼성화하는 프라이머 또는 프로브와 접촉시키는 단계, 및 혼성화가 일어났는지의 여부를 결정하는 단계를 포함한다.

[0206] 일부 실시형태에서, 검정은 생물학적 샘플을 엄격한 조건 하에서 각각 야생형 PNPLA3 또는 변이체 HSD17B13을 암호화하는 상응하는 핵산 분자가 아닌 본 명세서에 기재된 특정 변이체 PNPLA3 핵산 분자 중 임의의 것(예를 들어, 변이체 게놈 DNA 분자, mRNA 분자, cDNA 분자, RNA 전사물 또는 cDNA 전사물 중 임의의 것) 또는 기능성 HSD17B13 단백질을 암호화하는 핵산 분자 중 임의의 것(예를 들어, 기능성 HSD17B13 단백질을 암호화하는 게놈 DNA 분자, mRNA 분자, cDNA 분자, RNA 전사물 또는 cDNA 전사물 중 임의의 것)에 특이적으로 혼성화하는 프라이머 또는 프로브와 접촉시키는 단계, 및 혼성화가 일어났는지의 여부를 결정하는 단계를 포함한다.

[0207] 일부 실시형태에서, 검정은 RNA 서열분석(RNA-Seq)을 포함한다. 일부 실시형태에서, 검정은 또한 역 전사효소 중합효소 연쇄 반응(RT-PCR)을 통해서 mRNA를 cDNA로 역전사시키는 단계를 포함한다.

[0208] 이러한 프로브 및 프라이머는 고도로 엄격한 혼성화 조건 하에 표적 서열에 특이적으로 혼성화할 수 있다. 비록 표적 핵산 서열과 상이하지만 표적 핵산 서열을 특이적으로 검출 및/또는 식별하는 능력을 보유하는 프로브가 종래의 방법에 의해 설계될 수 있지만, 프로브 및 프라이머는 인접 뉴클레오타이드와 표적 서열의 완전한 핵산 서열 동일성을 가질 수 있다. 따라서, 프로브 및 프라이머는 표적 핵산 분자와 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99% 또는 100%의 서열 동일성 또는 상보성을 공유할 수 있다.

[0209] 핵산 분자에 대한 프로브의 결합을 허용하는 조건 하에 생물학적 샘플에서 프로브가 핵산 분자와 혼성화되는 경우, 이러한 결합은 검출되고, 생물학적 샘플에서 특정 변이체 또는 야생형 PNPLA3 또는 변이체 또는 기능성

HSD17B13 유전자좌의 존재 또는 특정 변이체 또는 야생형 PNPLA3 또는 변이체 또는 기능성 HSD17B13 mRNA 또는 cDNA의 존재 또는 수준의 표시를 허용할 수 있다. 결합된 프로브의 이러한 식별은 기재되어 있다. 특정 프로브는 변이체 또는 야생형 PNPLA3 또는 변이체 또는 기능성 HSD17B13 유전자의 특정 영역과 적어도 약 80%, 약 80% 내지 약 85%, 약 85% 내지 약 90%, 약 90% 내지 약 95%, 및 약 95% 내지 약 100% 동일한(또는 상보적인) 서열을 포함할 수 있다. 특정 프로브는 변이체 또는 야생형 PNPLA3 또는 변이체 또는 기능성 HSD17B13 mRNA의 특정 영역과 적어도 약 80%, 약 80% 내지 약 85%, 약 85% 내지 약 90%, 약 90% 내지 약 95%, 및 약 95% 내지 약 100% 동일한(또는 상보적인) 서열을 포함할 수 있다. 특정 프로브는 변이체 또는 야생형 PNPLA3 또는 변이체 또는 기능성 HSD17B13 cDNA의 특정 영역과 적어도 약 80%, 약 80% 내지 약 85%, 약 85% 내지 약 90%, 약 90% 내지 약 95%, 및 약 95% 내지 약 100% 동일한(또는 상보적인) 서열을 포함할 수 있다.

[0210] 일부 실시형태에서, 생물학적 샘플의 특정 핵산 상보체가 특정 기능성 또는 변이체 PNPLA3 단백질 또는 HSD17B13 단백질을 암호화하는 핵산 서열을 포함하는지의 여부를 결정하기 위해서, 생물학적 샘플은 관심대상 부위를 암호화하는 위치(예를 들어, 본 명세서에 기재된 위치 중 임의의 것)에 인접한 5' 측접 서열로부터 유래된 제1 프라이머, 및 관심대상의 동일한 부위를 암호화하는 위치에 인접한 3' 측접 서열로부터 유래된 제2 프라이머를 포함하는 프라이머 쌍을 사용하여 핵산 증폭 방법에 적용되어, 특정 기능성 또는 변이체 PNPLA3 단백질 또는 HSD17B13 단백질의 존재를 진단하는 앰플리콘을 생성시킬 수 있다. 예를 들어, PNPLA3과 관련하여 앰플리콘은 서열번호 42에 따른 148번 위치에 상응하는 위치를 암호화하는 뉴클레오타이드 서열을 포함할 수 있다. HSD17B13과 관련하여 앰플리콘은 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 뉴클레오타이드 서열을 포함할 수 있다. 일부 실시형태에서, 앰플리콘은 길이가 프라이머 쌍과 하나의 뉴클레오타이드 염기쌍을 합한 길이에서부터, DNA 증폭 프로토콜에 의해 생산 가능한 앰플리콘의 임의의 길이까지의 범위일 수 있다. 이러한 거리는 하나의 뉴클레오타이드 염기쌍에서부터, 증폭 반응의 한계치, 또는 약 2만 개 뉴클레오타이드 염기쌍까지의 범위일 수 있다. 선택적으로, 프라이머 쌍은 관심대상 부위를 암호화하는 위치 및 관심대상 부위를 암호화하는 위치의 각각의 측면 상의 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 그 초과의 뉴클레오타이드를 포함하는 영역에 측접한다. 유사한 앰플리콘은 mRNA 및/또는 cDNA 서열로부터 생성될 수 있다.

[0211] 프로브 및 프라이머를 제조 및 사용하는 대표적인 방법은 예를 들어, 문헌[*Molecular Cloning: A Laboratory Manual,* 2nd Ed., Vol. 1-3, ed. Sambrook et al., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. 1989(이하, "Sambrook et al., 1989"); *Current Protocols in Molecular Biology*, ed. Ausubel et al., Greene Publishing and Wiley-Interscience, New York, 1992(정기적으로 업데이트됨) (이하, "Ausubel et al., 1992"); 및 Innis et al., PCR Protocols: A Guide to Methods and Applications, Academic Press: San Diego, 1990)]에 기재되어 있다. PCR 프라이머 쌍은 예를 들어, 이러한 목적을 위해서 의도된 컴퓨터 프로그램, 예컨대, Vector NTI 버전 10(인포맥스사(Informax Inc.), 미국 메릴랜드주 베데스다주 소재)에서의 PCR 프라이머 분석 툴; PrimerSelect(디엔에이스타사(DNASTAR Inc.), 미국 매디슨주 위스콘신 소재); 및 Primer3(Version 0.4.0.COPYRGT., 1991, 화이트헤드 생의학 연구소(Whitehead Institute for Biomedical Research), 미국 매사추세츠주 캠브리지 소재)을 이용함으로써, 공지된 서열로부터 유래될 수 있다. 추가로, 서열은 시각적으로 스캔될 수 있고, 프라이머는 가이드라인을 사용하여 수동으로 식별될 수 있다.

[0212] 임의의 핵산 혼성화 또는 증폭 또는 서열분석 방법을 사용하여 기능성 또는 변이체 PNPLA3 또는 HSD17B13 유전자 유전자좌의 존재 및/또는 기능성 또는 변이체 PNPLA3 또는 HSD17B13 mRNA 또는 mRNA로부터 생산된 cDNA의 수준을 특이적으로 검출할 수 있다. 일부 실시형태에서, 핵산 분자를 프라이머로서 사용하여 기능성 또는 변이체 PNPLA3 또는 HSD17B13 핵산의 영역을 증폭시킬 수 있거나, 또는 핵산 분자를 예를 들어, 엄격한 조건 하에서, 기능성 또는 변이체 PNPLA3 또는 HSD17B13 유전자 유전자좌를 포함하는 핵산 분자 또는 기능성 또는 변이체 PNPLA3 또는 HSD17B13 mRNA 또는 mRNA로부터 생산된 cDNA를 포함하는 핵산 분자에 특이적으로 혼성화하는 프로브로서 사용할 수 있다.

[0213] 예를 들어, 핵산 서열분석, 핵산 혼성화 및 핵산 증폭을 비롯한 다양한 기술이 입수 가능하다. 핵산 서열분석 기술의 예시적인 예는 쇄 종결인자(생어(Sanger)) 서열분석 및 염료 종결인자 서열분석을 포함하지만 이들로 제한되지 않는다.

[0214] 다른 방법은 정제된 DNA, 증폭된 DNA 및 고정된 세포 제제에 대해서 지향되는 표지화된 프라이머 또는 프로브(형광 동일계 혼성화(fluorescence *in situ* hybridization: FISH))를 사용하는 것을 비롯하여, 서열분석 이외의 핵산 혼성화 방법을 포함한다. 일부 방법에서, 표적 핵산은 검출 이전에 또는 검출과 동시에 증폭될 수 있다. 핵산 증폭 기술의 예시적인 예는 중합효소 연쇄 반응(PCR), 리가제 연쇄 반응(LCR), 가닥 대체 증폭(strand displacement amplification: SDA) 및 핵산 서열 기반 증폭(NASBA)을 포함하지만 이들로 제한되지 않

는다. 다른 방법은 리가제 연쇄 반응, 가닥 대체 증폭 및 고온성 SDA(thermophilic SDA: tSDA)를 포함하지만 이들로 제한되지 않는다.

[0215] 예를 들어, 혼성화 보호 검정(Hybridization Protection Assay: HPA), 실시간 증폭 공정의 정량적 평가, 및 샘플 중에 초기에 존재하지만 실시간 증폭을 기초로 하지 않고 표적 서열의 양을 결정하는 것을 비롯한, 임의의 방법이 비-증폭된 또는 증폭된 폴리뉴클레오타이드를 검출하는데 사용될 수 있다.

[0216] 서열 증폭을 반드시 필요로 하지는 않고, 예를 들어, 서던(DNA:DNA) 블롯 혼성화, 동일계 혼성화(ISH) 및 형광 동일계 혼성화(FISH) 방법을 기반으로 하는 핵산을 식별하기 위한 방법이 또한 제공된다. 서던 블로팅은 특정 핵산 서열을 검출하는 데 사용될 수 있다. 이러한 방법에서, 샘플로부터 추출되는 핵산은 단편화되고, 매트릭스 겔에서 전기이동적으로 분리되고, 막 필터로 전달된다. 필터 결합된 핵산은 관심대상 서열에 상보적인 표지된 프로브와의 혼성화에 적용된다. 필터에 결합된 혼성화된 프로브가 검출된다. 임의의 이러한 방법에서, 방법은 본 명세서에 기재되거나 또는 예시된 프로브 중 임의의 것을 사용한 혼성화를 포함할 수 있다.

[0217] 혼성화 기술에서, 프로브 또는 프라이머가 이의 표적에 특이적으로 혼성화할 것이도록 엄격한 조건이 사용될 수 있다. 일부 실시형태에서, 엄격한 조건 하에서, 폴리뉴클레오타이드 프라이머 또는 프로브는 다른 서열(예를 들어, 상응하는 기능성 또는 변이체 PNPLA3 또는 HSD17B13 유전자좌, 기능성 또는 변이체 PNPLA3 또는 HSD17B13 mRNA, 또는 기능성 또는 변이체 PNPLA3 또는 HSD17B13 cDNA)보다 검출 가능하게 더 큰 정도로, 배경보다 10배를 초과하는 것을 비롯하여, 배경보다, 예컨대, 적어도 2배, 적어도 3배, 적어도 4배 또는 그 초과의 더 큰 정도로, 이의 표적 서열(예를 들어, 기능성 또는 변이체 PNPLA3 또는 HSD17B13 유전자좌, 기능성 또는 변이체 PNPLA3 또는 HSD17B13 mRNA, 또는 기능성 또는 변이체 PNPLA3 또는 HSD17B13 cDNA)에 혼성화할 것이다. 일부 실시형태에서, 엄격한 조건 하에서 폴리뉴클레오타이드 프라이머 또는 프로브는 다른 서열에 대한 것보다 적어도 2배만큼 검출 가능하게 큰 정도로 이의 표적 서열에 혼성화할 것이다. 일부 실시형태에서, 엄격한 조건 하에서 폴리뉴클레오타이드 프라이머 또는 프로브는 다른 서열에 대한 것보다 적어도 3배만큼 검출 가능하게 큰 정도로 이의 표적 서열에 혼성화할 것이다. 일부 실시형태에서, 엄격한 조건 하에서 폴리뉴클레오타이드 프라이머 또는 프로브는 다른 서열에 대한 것보다 적어도 4배만큼 검출 가능하게 큰 정도로 이의 표적 서열에 혼성화할 것이다. 일부 실시형태에서, 엄격한 조건 하에서 폴리뉴클레오타이드 프라이머 또는 프로브는 배경보다 10배를 초과하게 검출 가능하게 큰 정도로 다른 서열에 대한 것보다 이의 표적 서열에 혼성화할 것이다. 엄격한 조건은 서열 의존적이고, 상이한 상황에서 상이할 것이다. 혼성화의 엄격성 및/또는 세척 조건을 제어함으로써, 프로브에 100% 상보적인 표적 서열이 식별될 수 있다(상동성 프로빙). 대안으로, 엄격성 조건은 더 낮은 정도의 동일성이 검출되도록, 서열에서 일부 미스매치를 허용하도록 조정될 수 있다(이종성 프로빙).

[0218] 일부 실시형태에서, 검출 단계는, 관심대상 부위(예를 들어, 본 명세서에 기재된 위치 중 임의의 것)를 암호화하는 핵산 분자의 적어도 일부를 증폭시키는 단계; 핵산 분자를 검출 가능한 표지로 표지하는 단계; 표지된 핵산을 프로브를 포함하는 지지체와 접촉시키는 단계로서, 프로브는 엄격한 조건 하에서 관심대상 부위(예를 들어, 본 명세서에 기재된 위치 중 임의의 것)를 암호화하는 핵산에 혼성화하는 핵산 서열을 포함하는, 상기 단계; 및 검출 가능한 표지를 검출하는 단계를 포함한다.

[0219] 일부 실시형태에서, 검출 단계는, PNPLA3 또는 HSD17B13 단백질을 암호화하는 핵산 분자의 적어도 일부를 증폭시키는 단계로서, 증폭된 핵산 분자는 관심대상 부위(예를 들어, 본 명세서에 기재된 위치 중 임의의 것)를 포함하는 아미노산 서열을 암호화하는, 상기 단계; 핵산 분자를 검출 가능한 표지로 표지하는 단계; 표지된 핵산을 프로브를 포함하는 지지체와 접촉시키는 단계로서, 프로브는 엄격한 조건 하에서 관심대상 부위(예를 들어, 본 명세서에 기재된 위치 중 임의의 것)를 암호화하는 핵산에 혼성화하는 핵산 서열을 포함하는, 상기 단계; 및 검출 가능한 표지를 검출하는 단계를 포함한다. 본 명세서에 개시된 핵산 분자 중 임의의 것은 증폭될 수 있다. 예를 들어, 본 명세서에 개시된 게놈 DNA, cDNA 또는 mRNA 분자는 증폭될 수 있다. 일부 실시형태에서, 핵산 분자는 mRNA이고, 방법은 증폭 단계 이전에 mRNA를 cDNA로 역전사시키는 단계를 추가로 포함한다.

[0220] 일부 실시형태에서, 검출 단계는 PNPLA3 또는 HSD17B13 단백질을 암호화하는 핵산 분자를 검출 가능한 표지를 포함하는 프로브와 접촉시키는 단계로서, 프로브는 엄격한 조건 하에서 변이체 PNPLA3 또는 HSD17B13 단백질을 암호화하는 핵산 서열에 혼성화하는 핵산 서열을 포함하는, 상기 단계, 및 검출 가능한 표지를 검출하는 단계를 포함한다. 일부 실시형태에서, 검출 단계는 PNPLA3 또는 HSD17B13 단백질을 암호화하는 핵산 분자를 검출 가능한 표지를 포함하는 프로브와 접촉시키는 단계로서, 프로브는 엄격한 조건 하에서 관심대상 부위(예를 들어, 본 명세서에 기재된 위치 중 임의의 것)를 암호화하는 핵산 서열에 혼성화하는 핵산 서열을 포함하는, 상기 단계, 및 검출 가능한 표지를 검출하는 단계를 포함한다. 일부 실시형태에서, 핵산 분자는 인간 대상체로부터 획득된

세포 내에 존재하여, 검출은 동일계 혼성화 기술에 따른다.

[0221] 본 명세서에 개시된 방법에 사용될 수 있는 다른 검정은 예를 들어, 역전사 중합효소 연쇄 반응(RT-PCR) 또는 정량적 RT-PCR(qRT-PCR)을 포함한다. 본 명세서에 개시된 방법에 사용될 수 있는 추가의 다른 검정은 예를 들어, RNA 서열분석(RNA-Seq)에 이어서 생물학적 샘플에서 변이체 mRNA 또는 cDNA의 존재 또는 양의 검출을 포함한다.

[0222] 일부 실시형태에서, 검출 단계는 특정 PNPLA3 또는 HSD17B13 단백질을 암호화하는 핵산 분자의 적어도 일부를 증폭시키는 단계, 증폭된 핵산 분자를 검출 가능한 표지로 표지하는 단계, 표지된 핵산 분자를 프로브를 포함하는 지지체와 접촉시키는 단계로서, 프로브는 예를 들어, 엄격한 조건 하에서 특정 PNPLA3 또는 HSD17B13 단백질을 암호화하는 핵산 서열에 특이적으로 혼성화하는 핵산 서열을 포함하는, 상기 단계, 및 검출 가능한 표지를 검출하는 단계를 포함한다. 일부 실시형태에서, 검출 단계는 특정 PNPLA3 또는 HSD17B13 단백질을 암호화하는 핵산 분자의 적어도 일부를 증폭시키는 단계, 증폭된 핵산 분자를 검출 가능한 표지로 표지하는 단계, 표지된 핵산 분자를 프로브를 포함하는 지지체와 접촉시키는 단계로서, 프로브는 예를 들어, 엄격한 조건 하에서, 관심 대상 부위(예를 들어, 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 암호화하거나, 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 암호화하는 PNPLA3 핵산 서열의 일부; 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 게놈 DNA의 일부; 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함하는 변이체 PNPLA3 mRNA의 일부; 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함하는 변이체 PNPLA3 mRNA의 일부; 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 cDNA의 일부; 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함하는 변이체 PNPLA3 cDNA의 일부; 서열번호 1에 따른 12,667번 위치에 상응하는 위치에 아데닌을 포함하는 HSD17B13 게놈 DNA의 일부; 또는 서열번호 2에 따른 12,667번 위치에 상응하는 위치에 티민을 포함하는 HSD17B13 게놈 DNA의 일부)를 암호화하는 핵산 서열에 특이적으로 혼성화하는 핵산 서열을 포함하는, 상기 단계, 및 검출 가능한 표지를 검출하는 단계를 포함한다. 핵산이 mRNA를 포함하는 경우, 방법은 증폭 단계 이전에 mRNA를 cDNA로 역전사시키는 단계를 추가로 포함할 수 있다. 일부 실시형태에서, 결정 단계는 특정 PNPLA3 또는 HSD17B13 단백질을 암호화하는 핵산 분자를 검출 가능한 표지를 포함하는 프로브와 접촉시키는 단계 및 검출 가능한 표지를 검출하는 단계를 포함한다.

[0223] 본 개시내용은 하이드록시스테로이드 17-베타 데하이드로게나제 13(HSD17B13)을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법을 제공하며, 이 방법은, 대상체로부터의 샘플이 I148M 변이를 포함하는 파타틴 유사 포스포리파제 도메인 함유 3(PNPLA3) 단백질을 암호화하는 제1 핵산, 및 기능성 HSD17B13 단백질을 암호화하는 제2 핵산, 및/또는 I148M 변이를 포함하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질을 포함하는지의 여부를 결정하는 단계; 및 제1 핵산 및 제2 핵산 둘 다 및/또는 단백질 둘 다가 검출되는 경우 대상체를 HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 식별하는 단계를 포함한다. 일부 실시형태에서, 대상체는 비만이다. 일부 실시형태에서, 대상체는 지방간을 갖는다. 일부 실시형태에서, 제1 핵산 분자는 게놈 DNA를 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 제1 핵산 분자는 mRNA를 포함한다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 제1 핵산 분자는 mRNA로부터 획득된 cDNA를 포함한다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 제1 핵산을 검출하는 단계는 제1 핵산의 적어도 일부를 서열분석하는 단계를 포함하고, 일부는 I148M 변이를 암호화하는 코돈을 포함한다. 일부 실시형태에서, 제1 핵산을 검출하는 단계는 제1 핵산을 제1 핵산의 일부에 특이적으로 혼성화하는 프로브 또는 프라이머와 혼성화시키는 단계를 포함하며, 여기서 일부는 I148M 변이를 암호화하는 코돈을 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 대립유전자-특이적 프로브 또는 프라이머이다. 일부 실시형태에서, 프로브 또는 프라이머는 표지를 포함한다. 일부 실시형태에서, 방법은 대상체가 I148M 변이에 대해서 동형접합성인지 또는 이형접합성인지를 결정하는 단계를 추가로 포함한다.

[0224] 일부 실시형태에서, 제2 핵산은 게놈 DNA를 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 1에 따른 12,667번 위치에 상응하는 위치에 아데닌을 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 1을 포함한다. 일부 실시형태에서, 제2 핵산 분자는 mRNA를 포함한다. 일부 실시형태에서, mRNA는 서열번호 3을 포함한다. 일부 실시형태에서, mRNA는 서열번호 4를 포함한다. 일부 실시형태에서, mRNA는 서열번호 7을 포함한다. 일부 실시형태에서, mRNA는 서열번호 11을 포함한다. 일부 실시형태에서, 제2 핵산 분자는 mRNA로부터 획득된 cDNA를 포함한다. 일부 실시형태에서, cDNA는 서열번호 12를 포함한다. 일부 실시형태에서, cDNA는 서열번호 13을 포함한다. 일부 실시형태에서, cDNA는 서열번호 16을 포함한다. 일부 실시형태에서, cDNA는 서열번호 20을 포함한다. 일부 실시형태에서, 제2 핵산을 검출하는 단계는 제2 핵산을 서열분석하는 단계를 포함한다. 일부 실시형태에서, 제2 핵산을 검출하는 단계는 제2 핵산을 제2 핵산에 특이적으로 혼성화하는 프로브 또는 프라이머와 혼성화시키는 단계를 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 대립유전자-특이적 프로브 또는 프라이머이다. 일부 실시형태에서, 프로브 또는 프라이머는 표지를 포함한다. 일부 실시형태에서, 방법은 대상체가 샘플에서 기능성 HSD17B13 단백질을 암호화하는 제2 핵산에 대해서 동형접합성인지 또는 이형접합성인지를 결정하는 단계를 추가로 포함한다.

[0225] 본 개시내용은 HSD17B13 저해를 위한 후보자인 대상체를 식별하는 방법을 제공하며, 이 방법은 대상체로부터의 샘플이 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체를 암호화하는 핵산을 포함하는지의 여부를 결정하는 단계를 포함한다. 본 개시내용은 또한 HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 식별하는 방법을 제공하며, 이 방법은, 대상체로부터의 샘플이 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 제1 핵산, 및 기능성 HSD17B13 단백질을 암호화하는 제2 핵산, 및/또는 I148M 변이를 포함하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질을 포함하는지의 여부를 결정하는 단계; 및 제1 핵산 및 제2 핵산 둘 다 및/또는 단백질 둘 다가 검출되는 경우 대상체를 HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 식별하는 단계를 포함한다.

[0226] 본 개시내용은 또한 HSD17B13 저해를 위한 후보자인 대상체를 분류하는 방법을 제공하며, 이 방법은 대상체로부터의 샘플이 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체를 암호화하는 핵산을 포함하는지의 여부를 결정하는 단계를 포함한다. 본 개시내용은 또한 HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 인간 대상체를 분류하는 방법을 제공하며, 이 방법은, 대상체로부터의 샘플이 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 제1 핵산, 및 기능성 HSD17B13 단백질을 암호화하는 제2 핵산, 및/또는 I148M 변이를 포함하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질을 포함하는지의 여부를 결정하는 단계; 및 제1 핵산 및 제2 핵산 둘 다 및/또는 단백질 둘 다가 검출되는 경우 대상체를 HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자로서 분류하는 단계를 포함한다.

[0227] 변이체 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체는 본 명세서에 기재된 변이체 PNPLA3 Ile148Met 변이체 및 PNPLA3 Ile144Met 변이체 중 임의의 것일 수 있다. 변이체 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체는 본 명세서에 기재된 방법 중 임의의 것에 의해서 검출될 수 있다. 일부 실시형태에서, 방법은 대상체가 변이체 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체에 대해서 동형접합성인지 또는 이형접합성인지를 결정하는 단계를 추가로 포함한다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체에 대해서 이형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile148Met 변이체에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile148Met 변이체에 대해서 이형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile144Met 변이체에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile144Met 변이체에 대해서 이형접합성이다.

[0228] 바람직한 실시형태에서, 대상체는 HSD17B13 단백질 내에 기능 상실 변이를 암호화하는 어떠한 유전자도 포함하지 않는다. 본 명세서에 기재되거나 예시된 것을 비롯한 HSD17B13 단백질 내의 기능 상실 변이는 간 질환-보호 효과를 부여한다고 여겨지고, 추가로 이러한 보호 효과는 변이체 PNPLA3 Ile148M 변이의 존재 하에서 향상된다고 여겨진다. 따라서, HSD17B13 단백질을 암호화하는 (각각의 염색체로부터의) 유전자의 카피 모두가 기능 상실 변이를 암호화하는 대상체(예를 들어, PNPLA3 내에 I148M 변이를 포함하는 대상체)는 HSD17B13 저해 요법으로부터 이득을 얻지 못할 것이라고 여겨진다. 그럼에도 불구하고, 적어도 부분적인 기능성 HSD17B13 단백질을 발현하는 대상체는 HSD17B13 저해 요법으로부터 이득을 얻을 것이라고 여겨진다. 따라서, 방법은 유전자가 HSD17B13 단백질에서 기능 상실 변이를 암호화하는지의 여부, 뿐만 아니라, 대상체가 동형접합성인지 또는 이형접합성인지의 여부를 비롯하여, (하나 또는 두 염색체에서) HSD17B13을 암호화하는 유전자의 상태를 분류하는 단계를 포

함할 수 있다.

[0229] 일부 실시형태에서, 방법은 대상체로부터의 샘플에서 기능성 HSD17B13 단백질을 암호화하는 핵산 분자 또는 유전자의 존재를 검출하는 단계를 추가로 포함한다. 핵산 분자는 본 명세서에 기재된 기능성 HSD17B13 단백질 중 임의의 것을 암호화할 수 있다. HSD17B13 핵산 분자는 본 명세서에 기재된 방법 중 임의의 것에 의해서 검출될 수 있다. 일부 실시형태에서, 방법은 대상체가 기능성 HSD17B13 단백질을 암호화하는 유전자에 대해서 동형접합성인지 또는 이형접합성인지를 결정하는 단계를 추가로 포함한다. 일부 실시형태에서, 대상체는 기능성 HSD17B13 단백질을 암호화하는 유전자에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 기능성 HSD17B13을 암호화하는 유전자에 대해서 이형접합성이다.

[0230] 본 개시내용은 또한 본 명세서에 개시된 프로브 중 임의의 하나 이상이 부착된 기질을 포함하는 지지체를 제공한다. 고체 지지체는 분자, 예컨대, 본 명세서에 개시된 프로브 중 임의의 것이 회합될 수 있는 고체-상태 기질 또는 지지체이다. 고체 지지체의 형태는 어레이이다. 고체 지지체의 또 다른 형태는 어레이 검출기이다. 어레이 검출기는 다수의 상이한 프로브가 어레이, 격자 또는 다른 조직화된 패턴으로 커플링된 고체 지지체이다.

[0231] 고체 지지체에서 사용하기 위한 고체-상태 기질은 분자가 커플링될 수 있는 임의의 고체 물질을 포함할 수 있다. 이것은 아크릴아마이드, 아가로스, 셀룰로스, 나이트로셀룰로스, 유리, 폴리스타이렌, 폴리에틸렌 비닐 아세테이트, 폴리프로필렌, 플리메타크릴레이트, 폴리에틸렌, 폴리에틸렌 옥사이드, 폴리실리케이트, 폴리카보네이트, 테플론, 플루오로카본, 나일론, 규소(silicon) 고무, 폴리산무수물(polyanhydride), 폴리글리콜산, 폴리락트산, 폴리오쏘에스터, 폴리프로필퓨머레이트, 콜라겐, 글리코사미노글리칸 및 폴리아미노산과 같은 물질을 포함한다. 고체-상태 기질은 박막, 막, 병, 접시, 섬유, 제직 섬유, 성형 중합체, 입자, 비드, 마이크로입자 또는 조합물을 비롯한 임의의 유용한 형태를 가질 수 있다. 고체-상태 기질 및 고체 지지체는 다공성 또는 비다공성일 수 있다. 고체-상태 기질을 위한 형태는 마이크로역가 접시, 예컨대, 표준 96-웰 유형이다. 일부 실시형태에서, 일반적으로 웰당 하나의 어레이를 함유하는 멀티-웰 유리 슬라이드가 사용될 수 있다. 이러한 특징은 검정 재현성의 더 큰 제어, 증가된 산출량 및 샘플 조작 및 자동화 용이성을 가능하게 한다. 일부 실시형태에서, 지지체는 마이크로어레이이다.

[0232] 일부 실시형태에서, 방법은 대상체가 비만인지의 여부를 결정하는 단계를 추가로 포함한다. 일부 실시형태에서, 체질량 지수(BMI)가 30kg/㎡를 초과하면 대상체는 비만이다. 비만은 간 질환을 갖거나 또는 이의 발달 위험이 있는 대상체의 특징일 수 있다. 일부 실시형태에서, 방법은 대상체가 지방간을 갖는지의 여부를 결정하는 단계를 추가로 포함한다. 지방간은 간 질환을 갖거나 또는 이의 발달 위험이 있는 대상체의 특징일 수 있다. 일부 실시형태에서, 방법은 대상체가 비만이고, 지방간을 갖는지의 여부를 결정하는 단계를 추가로 포함한다.

[0233] 일부 실시형태에서, 방법은 대상체에게 HSD17B13의 저해제를 투여하는 단계를 추가로 포함한다. HSD17B13의 저해제를 대상체에게 투여하는 방법을 본 명세서에 기재한다.

[0234] 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 분자는 본 명세서에 기재된 핵산 분자 중 임의의 것이다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 분자는 게놈 DNA이다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 게놈 DNA는 본 명세서에 기재된 게놈 DNA 분자 중 임의의 것이다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질을 암호화하는 핵산 분자는 게놈 DNA이다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 게놈 DNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다. 일부 실시형태에서, ATG 코돈을 포함하는 변이체 PNPLA3 단백질을 암호화하는 게놈 DNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다.

[0235] 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 분자는 mRNA이다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 mRNA는 본 명세서에 기재된 mRNA 분자 중 임의의 것이다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질을 암호화하는 핵산 분자는 mRNA이다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 뉴클레오타이드 서열을 포함한다. 일부

실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 mRNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다. 일부 실시형태에서, AUG 코돈을 포함하는 변이체 PNPLA3 단백질을 암호화하는 mRNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 식별된다.

[0236] 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 분자는 cDNA이다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 cDNA는 본 명세서에 기재된 cDNA 분자 중 임의의 것이다. 일부 실시형태에서, PNPLA3 Ile148Met 단백질을 암호화하는 핵산 분자는 cDNA이다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 cDNA는 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 식별된다. 일부 실시형태에서, ATG 코돈을 포함하는 변이체 PNPLA3 단백질을 암호화하는 cDNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 식별된다.

[0237] 일부 실시형태에서, 기능성 HSD17B13 단백질을 암호화하는 핵산 분자는 본 명세서에 기재된 핵산 분자 중 임의의 것이다. 일부 실시형태에서, 기능성 HSD17B13 단백질을 암호화하는 핵산 분자는 게놈 DNA이다. 일부 실시형태에서, 기능성 HSD17B13 단백질을 암호화하는 게놈 DNA는 본 명세서에 기재된 게놈 DNA 분자 중 임의의 것이다. 일부 실시형태에서, 게놈 DNA는 서열번호 1에 따른 12,667번 위치에 상응하는 위치에 아데닌을 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 1을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 게놈 DNA의 존재는 명세서에 기재된 바와 같은, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다.

[0238] 일부 실시형태에서, 기능성 HSD17B13 단백질을 암호화하는 핵산 분자는 mRNA이다. 일부 실시형태에서, 기능성 HSD17B13 단백질을 암호화하는 mRNA는 본 명세서에 기재된 mRNA 분자 중 임의의 것이다. 일부 실시형태에서, 기능성 HSD17B13 핵산 분자는 mRNA이다. 일부 실시형태에서, mRNA는 서열번호 3을 포함한다. 일부 실시형태에서, mRNA는 서열번호 4를 포함한다. 일부 실시형태에서, mRNA는 서열번호 7을 포함한다. 일부 실시형태에서, mRNA는 서열번호 11을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 mRNA의 존재는 명세서에 기재된 바와 같은, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다.

[0239] 일부 실시형태에서, 기능성 HSD17B13 단백질을 암호화하는 핵산 분자는 cDNA이다. 일부 실시형태에서, 기능성 HSD17B13 단백질을 암호화하는 cDNA는 본 명세서에 기재된 cDNA 분자 중 임의의 것이다. 일부 실시형태에서, cDNA는 서열번호 12를 포함한다. 일부 실시형태에서, cDNA는 서열번호 13을 포함한다. 일부 실시형태에서, cDNA는 서열번호 16을 포함한다. 일부 실시형태에서, cDNA는 서열번호 20을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 cDNA의 존재는 명세서에 기재된 바와 같은, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다.

[0240] 일부 실시형태에서, 방법은 대상체로부터 샘플을 획득하는 단계를 추가로 포함한다. 일부 실시형태에서, HSD17B13 저해에 대한 후보자인 대상체는 간 질환을 갖거나 또는 간 질환의 발달에 민감성이다. 일부 실시형태에서, 간 질환은 만성 간 질환이다. 일부 바람직한 실시형태에서, 만성 간 질환은 비알코올성 지방간 질환(NAFLD), 알코올성 간 질환(ALD), 비알코올성 지방간염(NASH), 간경변, 지방증 또는 간세포 암종이다. 일부 바람직한 실시형태에서, 만성 간 질환은 비알코올성 지방간 질환(NAFLD), 알코올성 간 질환(ALD), 비알코올성 지방간염(NASH), 간경변 또는 지방증이다. 일부 실시형태에서, 간 질환은 알코올성 간 질환이다. 일부 실시형태에서, 알코올성 간 질환은 알코올 섭취로 인한 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다. 일부 실시형태에서, 간 질환은 비알코올성 간 질환이다. 일부 실시형태에서, 비알코올성 간 질환은 비알코올성 지방간 질환(NAFLD) 또는 비알코올성 지방간염(NASH)을 포함한다. 일부 실시형태에서, 비알코올성 간 질환은 알코올 섭취에 의해서 유발되지 않은 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다.

[0241] 본 개시내용은 또한 대상체에서 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체 및 기능성 HSD17B13을 검출하는 방법을 제공하며, 이 방법은 대상체로부터의 샘플에서, PNPLA3 Ile148Met 단백질, 또는 PNPLA3 Ile148Met 단백질을 암호화하는 핵산 분자의 존재를 검출하는 단계, 또는 대상체로부터의 샘플에서, PNPLA3 Ile144Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 분자의 존재를 검출하는 단계; 및 대상체로부터의 샘플에서 기능성 HSD17B13 단백질 또는 기능성 HSD17B13 단백질을 암호화하는 핵산 분자의 존재를 검출하는 단계를 포함한다. 변이체 PNPLA3 Ile148Met 변이체 단백질 또는 핵산 분자는 본 명세서에 기재된 변이체

PNPLA3 Ile148Met 변이체 단백질 또는 핵산 분자 중 임의의 것일 수 있다. 변이체 PNPLA3 Ile144Met 변이체 단백질 또는 핵산 분자는 본 명세서에 기재된 변이체 PNPLA3 Ile144Met 변이체 단백질 또는 핵산 분자 중 임의의 것일 수 있다. 기능성 HSD17B13 단백질 또는 핵산 분자는 본 명세서에 기재된 기능성 HSD17B13 단백질 또는 핵산 분자 중 임의의 것일 수 있다.

[0242] 일부 실시형태에서, 방법은 대상체가 변이체 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체에 대해서 동형접합성인지 또는 이형접합성인지를 결정하는 단계를 추가로 포함한다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체에 대해서 이형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile148Met 변이체에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile148Met 변이체에 대해서 이형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile144Met 변이체에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile144Met 변이체에 대해서 이형접합성이다.

[0243] 일부 실시형태에서, 방법은 대상체가 기능성 HSD17B13에 대해서 동형접합성인지 또는 이형접합성인지를 결정하는 단계를 추가로 포함한다. 일부 실시형태에서, 대상체는 기능성 HSD17B13에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 기능성 HSD17B13에 대해서 이형접합성이다.

[0244] 일부 실시형태에서, 기능성 HSD17B13 단백질의 존재를 샘플에서 검출한다. 기능성 HSD17B13 단백질은 본 명세서에 기재된 기능성 HSD17B13 단백질 중 임의의 것일 수 있다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 서열번호 21, 서열번호 22, 서열번호 25 또는 서열번호 29에 따른 아미노산 서열을 포함한다. 일부 실시형태에서, 기능성 HSD17B13 단백질은 본 명세서에 기재된 바와 같은 아미노산 서열분석 또는 면역검정에 의해서 검출된다.

[0245] 일부 실시형태에서, 기능성 HSD17B13 핵산 분자의 존재를 샘플에서 검출한다. 기능성 HSD17B13 핵산 분자는 본 명세서에 기재된 기능성 HSD17B13 핵산 분자 중 임의의 것일 수 있다. 일부 실시형태에서, 기능성 HSD17B13 핵산 분자는 게놈 DNA이다. 기능성 HSD17B13 게놈 DNA 분자는 본 명세서에 기재된 기능성 HSD17B13 게놈 DNA 분자 중 임의의 것일 수 있다. 일부 실시형태에서, 게놈 DNA는 서열번호 1에 따른 12,667번 위치에 상응하는 위치에 아데닌을 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 1을 포함한다. 일부 실시형태에서, 게놈 DNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다.

[0246] 일부 실시형태에서, 기능성 HSD17B13 핵산 분자는 mRNA이다. 기능성 HSD17B13 mRNA 분자는 본 명세서에 기재된 기능성 HSD17B13 mRNA 분자 중 임의의 것일 수 있다. 일부 실시형태에서, mRNA는 서열번호 3을 포함한다. 일부 실시형태에서, mRNA는 서열번호 4를 포함한다. 일부 실시형태에서, mRNA는 서열번호 7을 포함한다. 일부 실시형태에서, mRNA는 서열번호 11을 포함한다. 일부 실시형태에서, mRNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다.

[0247] 일부 실시형태에서, 기능성 HSD17B13 핵산 분자는 cDNA이다. 기능성 HSD17B13 cDNA 분자는 본 명세서에 기재된 기능성 HSD17B13 cDNA 분자 중 임의의 것일 수 있다. 일부 실시형태에서, cDNA는 서열번호 12를 포함한다. 일부 실시형태에서, cDNA는 서열번호 13을 포함한다. 일부 실시형태에서, cDNA는 서열번호 16을 포함한다. 일부 실시형태에서, cDNA는 서열번호 20을 포함한다. 일부 실시형태에서, cDNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다.

[0248] 일부 실시형태에서, PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질의 존재를 샘플에서 검출한다. 변이체 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질은 본 명세서에 기재된 변이체 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질 중 임의의 것일 수 있다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나, 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 샘플 중의 변이체 PNPLA3 단백질은 서열번호 42에 따른 아미노산 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질은 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 샘플 중의 변이체 PNPLA3 단백질은 서열번호 43에 따른 아미노산 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질은 본 명세서에 기재된 바와 같은 아미노산 서열분석 또는 면역검정에 의해서 검출된다.

[0249] 일부 실시형태에서, PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile148Met 단백질을 암호화하는 핵산 분자의 존재가

샘플에서 검출된다. PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile148Met 단백질을 암호화하는 핵산 분자는 본 명세서에 기재된 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile148Met 단백질을 암호화하는 핵산 분자 중 임의의 것일 수 있다.

[0250] 일부 실시형태에서, PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 분자는 게놈 DNA이다. PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 게놈 DNA는 본 명세서에 기재된 게놈 DNA 분자 중 임의의 것일 수 있다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, ATG 코돈을 포함하는 변이체 PNPLA3 단백질을 암호화하는 게놈 DNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다.

[0251] 일부 실시형태에서, PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile148Met 단백질을 암호화하는 핵산 분자는 mRNA이다. PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile148Met 단백질을 암호화하는 mRNA 분자는 본 명세서에 기재된 mRNA 분자 중 임의의 것일 수 있다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, AUG 코돈을 포함하는 변이체 PNPLA3 단백질을 암호화하는 mRNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 식별된다.

[0252] 일부 실시형태에서, PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 분자는 cDNA이다. PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 cDNA는 본 명세서에 기재된 cDNA 분자 중 임의의 것일 수 있다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, ATG 코돈을 포함하는 변이체 PNPLA3 단백질을 암호화하는 cDNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 식별된다.

[0253] 일부 실시형태에서, 방법은 대상체로부터 샘플을 획득하는 단계를 추가로 포함한다.

[0254] 본 개시내용은 또한 간 질환에 대해서 보호 효과를 갖는 대상체를 식별하는 방법을 제공하며, 이 방법은 대상체로부터의 샘플에서 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체의 존재를 검출하는 단계; 및 대상체로부터의 샘플에서 HSD17B13 기능 상실 변이체의 존재를 검출하는 단계를 포함한다. 본 개시내용은 또한 간 질환에 대해서 보호 효과를 갖는 대상체를 분류하는 방법을 제공하며, 이 방법은 대상체로부터의 샘플에서 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체의 존재를 검출하는 단계; 및 대상체로부터의 샘플에서 HSD17B13 기능 상실 변이체의 존재를 검출하는 단계를 포함한다. 변이체 PNPLA3 Ile148Met 변이체 및 PNPLA3 Ile144Met 변이체는 본 명세서에 기재된 변이체 PNPLA3 Ile148Met 변이체 및 PNPLA3 Ile144Met 변이체 중 임의의 것일 수 있다. HSD17B13 기능 상실 변이체는 본 명세서에 기재된 HSD17B13 기능 상실 변이체 중 임의의 것일 수 있다.

[0255] 일부 실시형태에서, 방법은 대상체가 변이체 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체에 대해서 동형접합성인지 또는 이형접합성인지를 결정하는 단계를 추가로 포함한다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체에 대해서 이형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile148Met 변이체에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile148Met 변이체에 대해서 이형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile144Met 변이체에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 변이체 PNPLA3 Ile144Met 변이체에 대해서 이형접합성이다.

[0256] 일부 실시형태에서, 방법은 대상체가 기능성 HSD17B13에 대해서 동형접합성인지 또는 이형접합성인지를 결정하는 단계를 추가로 포함한다. 일부 실시형태에서, 대상체는 기능성 HSD17B13에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 기능성 HSD17B13에 대해서 이형접합성이다.

[0257] 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체는 대상체로부터의 샘플에서

PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile148Met 단백질을 암호화하는 핵산 분자를 검출하거나 또는 대상체로부터의 샘플에서 PNPLA3 Ile144Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 분자를 검출함으로써 대상체에서 검출되고; HSD17B13 기능 상실은 대상체로부터의 샘플에서 HSD17B13 기능 상실 변이체 단백질 또는 HSD17B13 기능 상실 변이체 단백질을 암호화하는 핵산 분자를 검출함으로써 대상체에서 검출된다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체는 대상체로부터의 샘플에서 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile148Met 단백질을 암호화하는 핵산 분자를 검출함으로써 대상체에서 검출되고; HSD17B13 기능 상실은 대상체로부터의 샘플에서 HSD17B13 기능 상실 변이체 단백질 또는 HSD17B13 기능 상실 변이체 단백질을 암호화하는 핵산 분자를 검출함으로써 대상체에서 검출된다.

[0258] 일부 실시형태에서, PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질의 존재를 샘플에서 검출한다. 변이체 PNPLA3 Ile148Met 단백질 및 PNPLA3 Ile144Met 단백질은 본 명세서에 기재된 변이체 PNPLA3 Ile148Met 단백질 및 PNPLA3 Ile144Met 단백질 중 임의의 것일 수 있다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나, 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 샘플 중의 변이체 PNPLA3 단백질은 서열번호 42에 따른 아미노산 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질은 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 샘플 중의 변이체 PNPLA3 단백질은 서열번호 43에 따른 아미노산 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질은 본 명세서에 기재된 바와 같은 아미노산 서열분석 또는 면역검정에 의해서 검출된다.

[0259] 일부 실시형태에서, PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 분자를 샘플에서 검출한다. 변이체 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 분자는 본 명세서에 기재된 변이체 PNPLA3 Ile148Met 단백질 및 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 분자 중 임의의 것일 수 있다.

[0260] 일부 실시형태에서, 여기서 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 분자는 게놈 DNA이다. 변이체 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 게놈 DNA는 본 명세서에 기재된 변이체 PNPLA3 Ile148Met 단백질 및 PNPLA3 Ile144Met 단백질 중 임의의 것일 수 있다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, ATG 코돈을 포함하는 변이체 PNPLA3 단백질을 암호화하는 게놈 DNA는 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다.

[0261] 일부 실시형태에서, 변이체 PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile144Met 단백질을 암호화하는 핵산 분자는 mRNA이다. 변이체 PNPLA3 Ile148Met 단백질 및 PNPLA3 Ile144Met 단백질을 암호화하는 mRNA 분자는 본 명세서에 기재된 변이체 PNPLA3 Ile148Met 단백질 및 PNPLA3 Ile144Met 단백질 중 임의의 것일 수 있다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, AUG 코돈을 포함하는 변이체 PNPLA3 단백질을 암호화하는 mRNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 식별된다.

[0262] 일부 실시형태에서, PNPLA3 Ile148Met 단백질 또는 PNPLA3 Ile148Met 단백질을 암호화하는 핵산 분자는 cDNA이다. 변이체 PNPLA3 Ile148Met 단백질 및 PNPLA3 Ile148Met 단백질을 암호화하는 cDNA는 본 명세서에 기재된 변이체 PNPLA3 Ile148Met 단백질 및 PNPLA3 Ile148Met 단백질 중 임의의 것일 수 있다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, ATG 코돈을 포함하는 변이체 PNPLA3 단백질을 암호화하는 cDNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 식별된다.

[0263] 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질의 존재를 샘플에서 검출한다. HSD17B13 기능 상실 변이체는 본 명세서에 기재된 HSD17B13 기능 상실 변이체 단백질 중 임의의 것일 수 있다. 일부 실시형태에서,

HSD17B13 기능 상실 변이체 단백질은 서열번호 23에 따른 아미노산 서열을 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질은 서열번호 24에 따른 아미노산 서열을 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질은 서열번호 26에 따른 아미노산 서열을 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질은 서열번호 27에 따른 아미노산 서열을 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질은 서열번호 28에 따른 아미노산 서열을 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질은 본 명세서에 기재된 바와 같은 아미노산 서열분석 또는 면역검정에 의해서 검출된다.

[0264] 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 핵산 분자의 존재를 샘플에서 검출한다. HSD17B13 기능 상실 변이체 단백질을 암호화하는 핵산 분자는 본 명세서에 기재된 HSD17B13 기능 상실 변이체 단백질을 암호화하는 핵산 분자 중 임의의 것일 수 있다.

[0265] 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 핵산 분자는 게놈 DNA이다. HSD17B13 기능 상실 변이체 단백질을 암호화하는 게놈 DNA 분자는 본 명세서에 기재된 HSD17B13 기능 상실 변이체 단백질 중 임의의 것일 수 있다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 게놈 DNA는 서열번호 2에 따른 12,667번 위치에 상응하는 위치에 티민을 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 게놈 DNA는 서열번호 2를 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 게놈 DNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다.

[0266] 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 핵산 분자는 mRNA이다. HSD17B13 기능 상실 변이체 단백질을 암호화하는 mRNA 분자는 본 명세서에 기재된 HSD17B13 기능 상실 변이체 단백질을 암호화하는 mRNA 분자 중 임의의 것일 수 있다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 mRNA는 서열번호 5를 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 mRNA는 서열번호 6을 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 mRNA는 서열번호 8을 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 mRNA는 서열번호 9를 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 mRNA는 서열번호 10을 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 mRNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다.

[0267] 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 핵산 분자는 cDNA이다. HSD17B13 기능 상실 변이체 단백질을 암호화하는 cDNA 분자는 본 명세서에 기재된 HSD17B13 기능 상실 변이체 단백질을 암호화하는 cDNA 분자 중 임의의 것일 수 있다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 cDNA는 서열번호 5를 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 cDNA는 서열번호 6을 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 cDNA는 서열번호 8을 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 cDNA는 서열번호 9를 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 cDNA는 서열번호 10을 포함한다. 일부 실시형태에서, HSD17B13 기능 상실 변이체 단백질을 암호화하는 cDNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 검출된다.

[0268] 일부 실시형태에서, 방법은 대상체로부터 샘플을 획득하는 단계를 추가로 포함한다. 일부 실시형태에서, 간 질환은 만성 간 질환이다. 일부 실시형태에서, 만성 간 질환은 비알코올성 지방간 질환(NAFLD), 알코올성 간 질환(ALD), 비알코올성 지방간염(NASH), 간경변, 지방증 또는 간세포 암종이다. 일부 실시형태에서, 간 질환은 알코올성 간 질환이다. 일부 실시형태에서, 알코올성 간 질환은 알코올 섭취로 인한 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다. 일부 실시형태에서, 간 질환은 비알코올성 간 질환이다. 일부 실시형태에서, 비알코올성 간 질환은 비알코올성 지방간 질환(NAFLD) 또는 비알코올성 지방간염(NASH)을 포함한다. 일부 실시형태에서, 비알코올성 간 질환은 알코올 섭취에 의해서 유발되지 않은 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다.

[0269] 본 개시내용은 또한 대상체에게 HSD17B13의 저해제를 투여하는 단계를 추가로 포함하는 본 명세서에 기재된 방법 중 임의의 것을 제공한다. 일부 실시형태에서, HSD17B13의 저해제는 내인성 HSD17B13 게놈 DNA 또는 mRNA에 혼성화하고, 대상체 내의 세포에서 HSD17B13 폴리펩타이드의 발현을 감소시키는 기능성 폴리펩타이드, 안티센스 DNA, RNA, siRNA, 또는 shRNA를 포함한다. 일부 실시형태에서, HSD17B13 저해제는 또한 단쇄 데하이드로게나제/리덕타제(SDR) 패밀리 중 1종 이상의 추가 구성원을 저해할 수 있고, 이들 중에서 HSD17B13이 구성원이다. 이러

한 다른 구성원은 HSD17B1, HSD17B2, HSD17B3, HSD17B4, HSD17B6, HSD17B7, HSD17B8, HSD17B10, HSD17B11, HSD17B12, HSD17B13, HSD17B14, HSD11B1, HSD11B2, HSD3B1, HSD3B2 및 HSD3B7, 뿐만 아니라 유사한 유사체 데하이드로게나제/리덕타제 3(DHRS3) 및 레티놀 데하이드로게나제10(RDH10)을 포함하지만 이들로 제한되지 않는다. 일부 실시형태에서, HSD17B13의 저해제는 대상체에서 간 질환을 저해하기 위해서 투여된다. 일부 실시형태에서, HSD17B13의 저해제는 대상체에서 간 질환을 치료하기 위해서 투여된다. 일부 실시형태에서, 간 질환은 만성 간 질환이다. 일부 실시형태에서, 만성 간 질환은 비알코올성 지방간 질환(NAFLD), 알코올성 간 질환(ALD), 비알코올성 지방간염(NASH), 간경변, 지방증 또는 간세포 암종 중 1종 이상이다. 일부 실시형태에서, 간 질환은 알코올성 간 질환이다. 일부 실시형태에서, 알코올성 간 질환은 알코올 섭취로 인한 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다. 일부 실시형태에서, 간 질환은 비알코올성 간 질환이다. 일부 실시형태에서, 비알코올성 간 질환은 비알코올성 지방간 질환(NAFLD) 또는 비알코올성 지방간염(NASH)을 포함한다. 일부 실시형태에서, 비알코올성 간 질환은 알코올 섭취에 의해서 유발되지 않은 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다. 일부 실시형태에서, 대상체는 I148M 변이를 암호화하는 유전자에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 I148M 변이를 암호화하는 유전자에 대해서 이형접합성이다. 일부 실시형태에서, 대상체는 추가로 기능성 HSD17B13 단백질을 암호화하는 유전자에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 추가로 기능성 HSD17B13 단백질을 암호화하는 유전자 및 HSD17B13의 기능 상실 변이체를 암호화하는 유전자에 대해서 이형접합성이다.

[0270] 본 개시내용은 또한 간 질환을 치료 또는 저해하는 방법을 제공하며, 이 방법은 하이드록시스테로이드 17-베타 데하이드로게나제 13(HSD17B13)의 저해제를, I148M 변이를 포함하는 파타틴 유사 포스포리파제 도메인 함유 3(PNPLA3) 단백질을 발현하는 인간 간 질환 환자에게 투여하여 환자에서 간 질환을 치료 또는 저해하는 단계를 포함한다. 일부 실시형태에서, 간 질환은 만성 간 질환이다. 일부 실시형태에서, 간 질환은 비알코올성 지방간 질환(NAFLD), 알코올성 간 질환(ALD), 비알코올성 지방간염(NASH), 간경변, 지방증 또는 간세포 암종 중 1종 이상이다. 일부 실시형태에서, 간 질환은 알코올성 간 질환이다. 일부 실시형태에서, 알코올성 간 질환은 알코올 섭취로 인한 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다. 일부 실시형태에서, 간 질환은 비알코올성 간 질환이다. 일부 실시형태에서, 비알코올성 간 질환은 비알코올성 지방간 질환(NAFLD) 또는 비알코올성 지방간염(NASH)을 포함한다. 일부 실시형태에서, 비알코올성 간 질환은 알코올 섭취에 의해서 유발되지 않은 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다. 일부 실시형태에서, 환자는 비만이다. 일부 실시형태에서, 환자는 지방간을 갖는다. 일부 실시형태에서, 환자는 대상체로부터의 샘플에서 변이체 PNPLA3 단백질을 검출함으로써 변이체 PNPLA3 단백질(예를 들어, I148M 또는 I144M 변이를 포함하는 PNPLA3 단백질)을 발현하는 것으로 결정되었다. 일부 실시형태에서, 변이체 PNPLA3 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 샘플 중의 변이체 PNPLA3 단백질은 서열번호 42에 따른 아미노산 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질은 아미노산 서열분석에 의해서 또는 면역검정에 의해서 검출된다. 일부 실시형태에서, 대상체는 대상체로부터의 샘플에서 변이체 PNPLA3 단백질을 암호화하는 핵산 분자(예를 들어, I148M 또는 I144M 변이를 포함하는 PNPLA3 단백질을 암호화하는 변이체 PNPLA3 핵산 분자)를 검출함으로써 변이체 PNPLA3 단백질을 발현하는 것으로 결정되었다. 일부 실시형태에서, 변이체 PNPLA3 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질은 서열번호 42에 따른 아미노산 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 게놈 DNA, mRNA 또는 mRNA로부터 획득된 cDNA를 포함한다. 일부 실시형태에서, 핵산 분자는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함하는 게놈 DNA를 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 핵산 분자는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함하는 mRNA를 포함한다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 핵산 분자는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함하는 mRNA를 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 핵산 분자는 mRNA로부터 획득된 cDNA를 포함하고, cDNA는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 핵산 분자는 mRNA로부터 획득된 cDNA를 포함하고, cDNA는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 핵산은 핵산의 적어도 일부를 서열분석함으로써 검출되고, 일부는 I148M 변이를 암호화한다. 일부 실시형태에서, 핵산은 핵산의 일부에 특이적으로 혼성화하는 프로브 또는 프라이머의 혼성화에 의해서 검출되며, 여기서 일부는 I148M 변이를 암호화하는 코돈을 포함한다. 일부 실시형태에서, 프로브 또

는 프라이머는 대립유전자-특이적 프로브 또는 프라이머이다. 일부 실시형태에서, 프로브 또는 프라이머는 표지를 포함한다. 일부 실시형태에서, 환자는 변이체 PNPLA3 단백질을 암호화하는 유전자에 대해서 동형접합성이다. 일부 실시형태에서, 환자는 변이체 PNPLA3 단백질을 암호화하는 유전자에 대해서 이형접합성이다. 일부 실시형태에서, 환자는 기능성 HSD17B13 단백질을 암호화하는 유전자에 대해서 동형접합성이다. 일부 실시형태에서, 환자는 기능성 HSD17B13 단백질을 암호화하는 유전자에 대해서 이형접합성이다. 일부 실시형태에서, 환자는 기능성 HSD17B13 단백질을 암호화하는 유전자 및 HSD17B13의 기능 상실 변이체를 암호화하는 유전자에 대해서 이형접합성이다.

[0271] HSD17B13의 저해제는 I148M 변이를 포함하는 PNPLA3 단백질을 갖고, 기능성 HSD17B13 단백질을 갖는 인간 대상체에서 간 질환의 치료를 위해서 본 명세서에 기재된 바와 같이 사용될 수 있다. 일부 실시형태에서, 인간 대상체는 I148M 변이를 포함하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질에 대해서 양성인 것으로 시험되었다. 일부 실시형태에서, 치료는 상기 인간 대상체가 I148M 변이를 포함하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질을 갖는지의 여부를 결정하는 단계를 포함한다. 일부 실시형태에서, 인간 대상체는, 본 명세서에 정의된 바와 같은 방법 중 임의의 것을 사용하여 HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자인 것으로 식별되었다. 일부 실시형태에서, 변이체 PNPLA3 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질은 서열번호 42에 따른 아미노산 서열, 또는 서열번호 42와 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 아미노산 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질은 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질은 서열번호 43에 따른 아미노산 서열, 또는 서열번호 43과 적어도 90%의 서열 동일성을 갖고 I144M 변이를 포함하는 아미노산 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 게놈 DNA이다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 뉴클레오타이드 서열, 또는 서열번호 31과 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 mRNA이다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 뉴클레오타이드 서열, 또는 서열번호 34와 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 뉴클레오타이드 서열, 또는 서열번호 35과 적어도 90%의 서열 동일성을 갖고 I144M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 cDNA이다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 뉴클레오타이드 서열, 또는 서열번호 38과 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 뉴클레오타이드 서열, 또는 서열번호 39와 적어도 90%의 서열 동일성을 갖고 I144M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 간 질환은 만성 간 질환이다. 일부 실시형태에서, 만성 간 질환은 비알코올성 지방간 질환(NAFLD), 알코올성 간 질환(ALD), 비알코올성 지방간염(NASH), 간경변, 지방증 또는 간세포 암종이다. 일부 실시형태에서, 간 질환은 알코올성 간 질환이다. 일부 실시형태에서, 알코올성 간 질환은 알코올 섭취로 인한 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다. 일부 실시형태에서, 간 질환은 비알코올성 간 질환이다. 일부 실시형태에서, 비알코올성 간 질환은 비알코올성 지방간 질환(NAFLD) 또는 비알코올성 지방간염(NASH)을 포함한다. 일부 실시형태에서, 비알코올성 간 질환은 알코올 섭취에 의해서 유발되지 않은 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다. 일부 실시형태에서, 인간 대상체는 기능성 HSD17B13에 대해서 동형접합성 또는 이형접합성이다.

[0272] 일부 실시형태에서, HSD17B13의 저해제는 HSD17B13 유전자 발현 또는 HSD17B13 단백질의 기능을 감소 또는 저해한다. HSD17B13의 저해제는 자연 발생 리간드 및 합성 리간드, 길항제, 효능제, 항체, 펩타이드, 환식 펩타이드, 핵산, 기능성 폴리뉴클레오타이드, 작은 유기 분자 등을 포함하지만 이들로 제한되지 않는다. 기능성 폴리뉴클레오타이드는 특정 기능을 갖는, 예컨대, 표적 분자에 결합하거나 또는 특정 반응을 촉매작용하는 핵산 분자이다. 기능성 폴리뉴클레오타이드의 예는 안티센스 분자, 압타머, 리보자임 및 트라이플렉스 형성 분자를 포함하지만 이들로 제한되지 않는다. 기능성 폴리뉴클레오타이드는 표적 분자에 의해서 보유된 특이적 활성의 저해제로서 작용할 수 있다. 안티센스 분자는 정규 또는 비정규 염기 짝지움을 통해서 표적 핵산 분자와 상호작

용하도록 설계된다. 안티센스 분자 및 표적 분자의 상호작용은 예를 들어, RNase-H-매개된 RNA-DNA 혼성 분해를 통해서 표적 분자의 파괴를 촉진시키도록 설계된다. 대체적으로, 안티센스 분자는 표적 분자 상에서 정상적으로 발생할 가공 기능, 예컨대, 전사 또는 복제를 방해하도록 설계된다. 안티센스 분자는 표적 분자의 서열을 기초로 설계될 수 있다. 표적 분자의 가장 접근 가능한 영역을 식별함으로써 안티센스 효율의 최적화를 위한 다양한 방법이 존재한다. 예시적인 방법은 시험관내 선택 실험, 및 DMS 및 DEPC를 사용한 DNA 변형 연구를 포함하지만 이들로 제한되지 않는다. 안티센스 분자는 일반적으로 약 10^{-6} 이하, 약 10^{-8} 이하, 약 10^{-10} 이하 또는 약 10^{-12} 이하의 해리 상수(k_d)로 표적 분자에 결합한다. 안티센스 분자의 설계 및 사용을 보조하는 방법 및 기술의 대표적인 샘플, 및 안티센스 분자는 하기 비제한적인 미국 특허 목록에서 찾아볼 수 있다: 5,135,917; 5,294,533; 5,627,158; 5,641,754; 5,691,317; 5,780,607; 5,786,138; 5,849,903; 5,856,103; 5,919,772; 5,955,590; 5,990,088; 5,994,320; 5,998,602; 6,005,095; 6,007,995; 6,013,522; 6,017,898; 6,018,042; 6,025,198; 6,033,910; 6,040,296; 6,046,004; 6,046,319; 6,057,437; 및 미국 시리얼 번호 제62/645,941호(출원일 2018년 3월 21일)(이들 각각은 이들의 전문이 참조에 의해 본 명세서에 포함됨). 안티센스 분자의 예는 안티센스 RNA, 짧은 간섭 RNA(siRNA) 및 짧은 헤어핀 RNA(shRNA)를 포함하지만 이들로 제한되지 않는다. 예를 들어, 안티센스 RNA, siRNA 또는 shRNA는 HSD17B13 게놈 DNA 또는 mRNA의 고유한 영역을 표적화하도록 설계될 수 있다. 일부 실시형태에서, HSD17B13의 저해제는 안티센스 분자이다. 일부 실시형태에서, HSD17B13의 저해제는 shRNA 분자이다. 일부 실시형태에서, HSD17B13의 저해제는 siRNA 분자이다.

[0273] 본 명세서에 기재된 방법 중 임의의 것에서, HSD17B13의 저해제의 투여는 간 질환의 특정 특징의 감소 또는 제거를 초래할 수 있다. 일부 실시형태에서, 간 질환의 특징은 염증 및 섬유증을 포함하지만 이들로 제한되지 않는다.

[0274] 본 개시내용은 또한 PNPLA3 Ile148Met 양성(즉, "PNPLA3 Ile148Met+") 또는 PNPLA3 Ile144Met 양성(즉, "PNPLA3 Ile144Met+")인 대상체를 치료하는 방법을 제공하며, 이 방법은 HSD17B13의 저해제를 대상체에게 투여하는 단계를 포함한다. 본 개시내용은 또한 간 질환을 치료 또는 저해하는 방법을 제공하며, 이 방법은 HSD17B13의 저해제를, I148M 변이를 포함하는 PNPLA3 단백질을 발현하는 인간 간 질환 환자에게 투여하여 환자에서 간 질환을 치료 또는 저해하는 단계를 포함한다.

[0275] 변이체 PNPLA3 Ile148Met 양성 또는 PNPLA3 Ile144Met 양성 대상체는 본 명세서에 기재된 변이체 PNPLA3 단백질 중 임의의 것을 가질 수 있다. 일부 실시형태에서, 대상체는 또한 기능성 HSD17B13에 대해서 동형접합성이거나 이형접합성이다. 일부 실시형태에서, 대상체는 기능성 HSD17B13에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 기능성 HSD17B13에 대해서 이형접합성이다. 일부 실시형태에서, 대상체는 HSD17B13 기능 상실 변이체에 대해서 동형접합성이다. 대상체는 본 명세서에 기재된 기능성 HSD17B13 단백질 중 임의의 것을 가질 수 있다.

[0276] 일부 실시형태에서, PNPLA3 Ile148Met+ 또는 PNPLA3 Ile144Met+인 대상체는 대상체로부터의 샘플에서 PNPLA3 단백질을 검출함으로써 PNPLA3 Ile148Met+ 또는 PNPLA3 Ile144Met+인 것으로 결정되었고, 여기서 변이체 PNPLA3 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나, 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 변이체 PNPLA3 Ile148Met 양성 또는 PNPLA3 Ile144Met 양성 대상체는 본 명세서에 기재된 변이체 PNPLA3 단백질 중 임의의 것을 가질 수 있다.

[0277] 일부 실시형태에서, PNPLA3 Ile148Met+인 대상체는 대상체로부터의 샘플에서 PNPLA3 단백질을 검출함으로써 PNPLA3 Ile148Met+인 것으로 결정되었고, 여기서 변이체 PNPLA3 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 샘플 중의 변이체 PNPLA3 단백질은 서열번호 42에 따른 아미노산 서열을 포함한다.

[0278] 일부 실시형태에서, PNPLA3 Ile144Met+인 대상체는 대상체로부터의 샘플에서 PNPLA3 단백질을 검출함으로써 PNPLA3 Ile144Met+인 것으로 결정되었고, 여기서 변이체 PNPLA3 단백질은 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 샘플 중의 변이체 PNPLA3 단백질은 서열번호 43에 따른 아미노산 서열을 포함한다.

[0279] 일부 실시형태에서, 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나, 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함하는 변이체 PNPLA3 단백질은 본 명세서에 기재된 바와 같은 아미노산 서열분석 또는 면역검정에 의해서 식별된다.

[0280] 일부 실시형태에서, PNPLA3 Ile148Met+ 또는 PNPLA3 Ile144Met+인 대상체는 대상체로부터의 샘플에서 PNPLA3

단백질을 암호화하는 핵산을 검출함으로써 PNPLA3 Ile148Met+ 또는 PNPLA3 Ile144Met+인 것으로 결정되었고, 여기서 변이체 PNPLA3 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나, 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 변이체 PNPLA3 Ile148Met 양성 또는 PNPLA3 Ile144Met 양성 대상체는 본 명세서에 기재된 변이체 PNPLA3 핵산 분자 중 임의의 것을 가질 수 있다.

[0281] 일부 실시형태에서, PNPLA3 Ile148Met+인 대상체는 대상체로부터의 샘플에서 PNPLA3 단백질을 암호화하는 핵산 분자를 검출함으로써 PNPLA3 Ile148Met+인 것으로 결정되었고, 여기서 변이체 PNPLA3 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 샘플 중의 변이체 PNPLA3 단백질은 서열번호 42에 따른 아미노산 서열을 포함한다.

[0282] 일부 실시형태에서, PNPLA3 Ile144Met+인 대상체는 대상체로부터의 샘플에서 PNPLA3 단백질을 암호화하는 핵산 분자를 검출함으로써 PNPLA3 Ile144Met+인 것으로 결정되었고, 여기서 변이체 PNPLA3 단백질은 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 샘플 중의 변이체 PNPLA3 단백질은 서열번호 43에 따른 아미노산 서열을 포함한다.

[0283] 일부 실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 게놈 DNA, mRNA 또는 mRNA로부터 유래된 cDNA이다.

[0284] 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, ATG 코돈을 포함하는 변이체 PNPLA3 단백질을 암호화하는 게놈 DNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 식별된다.

[0285] 일부 실시형태에서, mRNA는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, AUG 코돈을 포함하는 변이체 PNPLA3 단백질을 암호화하는 mRNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 식별된다.

[0286] 일부 실시형태에서, cDNA는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, ATG 코돈을 포함하는 변이체 PNPLA3 단백질을 암호화하는 cDNA는 본 명세서에 기재된 바와 같이, 핵산 서열분석에 의해서 또는 프로브의 혼성화에 의해서 식별된다.

[0287] HSD17B13의 저해제의 투여는, 예를 들어, 비경구, 정맥내, 경구, 피하, 동맥내, 두개내, 척수강내, 복강내, 국소, 비강내, 또는 근육내를 포함하지만 이들로 제한되지 않는 임의의 적합한 경로에 의해서 일어날 수 있다. 약제학적 조성물은 바람직하게 무균성이고, 실질적으로 등장성이고, GMP 조건 하에 제조된다. 약제학적 조성물은 단위 투여 형태(즉, 단일 투여를 위한 투여량)로 제공될 수 있다. 약제학적 조성물은 1종 이상의 생리학적으로 그리고 약제학적으로 허용 가능한 담체, 희석제, 부형제 또는 보조제를 사용하여 제형화될 수 있다. 제형은 선택된 투여 경로에 좌우된다. 용어 "약제학적으로 허용 가능한"은 담체, 희석제, 부형제 또는 보조제가 제형의 다른 성분과 상용성이고, 이의 수용자에게 실질적으로 유해하지 않다는 것을 의미한다.

[0288] 일부 실시형태에서, 대상체는 간 질환을 갖거나 간 질환의 발달에 민감성이다. 일부 실시형태에서, 간 질환은 만성 간 질환이다. 일부 실시형태에서, 만성 간 질환은 비알코올성 지방간 질환(NAFLD), 알코올성 간 질환(ALD), 비알코올성 지방간염(NASH), 간경변, 지방증 또는 간세포 암종이다. 일부 실시형태에서, 간 질환은 알코올성 간 질환이다. 일부 실시형태에서, 알코올성 간 질환은 알코올 섭취로 인한 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다. 일부 실시형태에서, 간 질환은 비알코올성 간 질환이다. 일부 실시형태에서, 비알코올성 간 질환은 비알코올성 지방간 질환(NAFLD) 또는 비알코올성 지방간염(NASH)을 포함한다. 일부 실시형태에서, 비알코올성 간 질환은 알코올 섭취에 의해서 유발되지 않은 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다.

[0289] 본 개시내용은 또한 간 질환 치료제로 환자를 치료하는 방법을 제공하며, 여기서 환자는 간 질환을 앓고 있다. 방법은 대상체로부터의 샘플이 i) I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 제1 핵산, 및 기능성

HSD17B13 단백질을 암호화하는 제2 핵산, 및/또는 ii) I148M 변이를 포함하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질을 포함하는지의 여부를 결정하는 단계를 포함한다. 일부 실시형태에서, 방법은 대상체로부터의 샘플이 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 제1 핵산, 및 기능성 HSD17B13 단백질을 암호화하는 제2 핵산을 포함하는지의 여부를 결정하는 단계를 포함한다. 일부 실시형태에서, 방법은 대상체로부터의 샘플이 I148M 변이를 포함하는 PNPLA3 단백질, 및 기능성 HSD17B13 단백질을 포함하는지의 여부를 결정하는 단계를 포함한다.

[0290] 일부 실시형태에서, 이러한 결정은 환자로부터 생물학적 샘플을 획득하거나 획득함으로써 수행된다. 일부 실시형태에서, 방법은 생물학적 샘플에 대한 유전자형분석 검정을 수행하거나 수행하였어서 환자가 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 제1 핵산 및 기능성 HSD17B13 단백질을 암호화하는 제2 핵산을 갖는지를 결정하는 단계를 추가로 포함한다. 일부 실시형태에서, 방법은 생물학적 샘플에 대한 검정을 수행하거나 수행하였어서 환자가 I148M 변이를 포함하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질을 갖는지를 결정하는 단계를 추가로 포함한다.

[0291] 일부 실시형태에서, 환자가 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 핵산을 갖고/갖거나 I148M 변이를 포함하는 PNPLA3 단백질을 갖고, 기능성 HSD17B13 단백질을 암호화하는 핵산을 갖고/갖거나 기능성 HSD17B13 단백질을 갖는 경우, 방법은 환자에게 HSD17B13의 저해제를 투여하는 단계를 추가로 포함한다. 일부 실시형태에서, 환자가 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 핵산을 갖고/갖거나 I148M 변이를 포함하는 PNPLA3 단백질을 갖고, 기능성 HSD17B13 단백질을 암호화하는 핵산을 갖고/갖거나 기능성 HSD17B13 단백질을 갖는 경우, 방법은 환자에게 HSD17B13의 저해제를 투여하는 단계 및 환자에게 간 질환 치료제를 투여하는 단계를 추가로 포함한다. 일부 실시형태에서, 환자가 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 핵산을 갖고/갖거나 I148M 변이를 포함하는 PNPLA3 단백질을 갖지만, 기능성 HSD17B13 단백질을 암호화하는 핵산을 갖지 않고/않거나 기능성 HSD17B13 단백질을 갖는 경우, 방법은 환자에게 간 질환 치료제를 투여하는 단계를 추가로 포함한다.

[0292] 간 질환 치료제의 예는 다이설피람, 날트렉손, 아캄프로세이트, 프레드니손, 아자티오프린, 페니실라민, 트라이엔틴, 데페록사민, 시프로플록사신, 노로플록사신, 세프트리악손, 오플록사신, 아목시실린-클라불라네이트, 파이토나다이온, 부메타나이드, 푸로세마이드, 하이드로클로로티아자이드, 클로로 티아자이드, 아밀로라이드, 트라이암테렌, 스피로놀락톤, 옥트레오타이드, 아테놀롤, 메토프롤롤, 나돌롤, 프로프라놀롤, 티몰롤 및 카르베딜롤을 포함하지만 이들로 제한되지 않는다.

[0293] (예를 들어, 만성 C형 간염의 치료에 사용하기 위한) 간 질환 치료제의 추가 예는 리바비린, 파리타프레비어, 시메프레비어(Olysio), 그라조프레비어, 레디파스비어, 옴비타스비어, 엘바스비어, 다클라타스비어(Daklinza), 다사부비어, 리토나비어, 소포스부비어, 벨파타스비어, 복실라프레비어, 글레카프레비어, 피브렌타스비어, 페그인터페론 알파-2a, 페그인터페론 알파-2b 및 인터페론 알파-2b를 포함하지만 이들로 제한되지 않는다.

[0294] (예를 들어, 비알코올성 지방간 질환에 사용하기 위한) 간 질환 치료제의 추가 예는 체중 감소 유도제, 예컨대, 오를리스타트 또는 시부트라민; 인슐린 반응 개선제(insulin sensitizing agent), 예컨대, 티아졸리딘다이온(TZD), 메트포민 및 메글리티나이드; 지질 저하제, 예컨대, 스타틴, 피브레이트 및 오메가-3 지방산; 항산화제, 예컨대, 비타민 E, 베타인, N-아세틸-시스테인, 레시틴, 실리마린 및 베타-카로텐; 항 TNF 작용제, 예컨대, 펜톡시필린; 프로바이오틱스, 예컨대, VSL#3; 및 세포보호제(cytoprotective agent), 예컨대, 우르소데옥시콜산(UDCA)을 포함하지만 이들로 제한되지 않는다. 다른 적합한 치료제는 ACE 저해제/ARB, 올리고프룩토스 및 인크레틴 유사체를 포함한다.

[0295] (예를 들어, NASH에 사용하기 위한) 간 질환 치료제의 예는 오베티콜산(Ocaliva(등록상표)), 셀론세르팁, 엘라피브라노어, 세니크리비로크, GR_MD_02, MGL_3196, IMM124E, 아라키딜 아미도 콜란산(Aramchol(상표명)), GS0976, 엠리카산, 볼릭시바트, NGM282, GS9674, 트로피펙서, MN_001, LMB763, BI_1467335, MSDC_0602, PF_05221304, DF102, 사로글리타자, BMS986036, 라니피브라노어(Lanifibranor), 세마글루타이드, 니타족사나이드, GRI_0621, EYP001, VK2809, 날메펜, LIK066, MT_3995, 엘로빅시바트, 나모데노손, 포랄루맙, SAR425899, 소타글리플로진, EDP_305, 아이소사부테이트, 젬카벤, TERN_101, KBP_042, PF_06865571, DUR928, PF_06835919, NGM313, BMS_986171, 나마시주맙, CER_209, ND_L02_s0201, RTU_1096, DRX_065, IONIS_DGAT2Rx, INT_767, NC_001, 셀라데파, PXL770, TERN_201, NV556, AZD2693, SP_1373, VK0214, 헤파스템, TGFTX4, RLBN1127, GKT_137831, RYI_018, CB4209-CB4211 및 JH_0920을 포함하지만 이들로 제한되지 않는다.

[0296] 본 개시내용은 또한 PNPLA3 Ile148Met 양성 또는 PNPLA3 Ile144Met 양성이고, 또한 기능성 HSD17B13에 대해서

동형접합성이거나 이형접합성인 인간 대상체에서 간 질환을 치료하기 위한 의약의 제조에 사용하기 위한 HSD17B13의 저해제를 제공한다. 일부 실시형태에서, 대상체는 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 PNPLA3 Ile148Met 변이체 또는 PNPLA3 Ile144Met 변이체에 대해서 이형접합성이다. 일부 실시형태에서, 대상체는 기능성 HSD17B13에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 기능성 HSD17B13에 대해서 이형접합성이다.

[0297] 일부 실시형태에서, HSD17B13의 저해제는 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하는 PNPLA3 단백질, 또는 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하거나 또는 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함하는 PNPLA3 단백질을 암호화하는 핵산 분자, 또는 서열번호 43에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함하는 PNPLA3 단백질을 암호화하는 핵산 분자를 갖는 인간 대상체에서 간 질환을 치료하는 데 사용된다. 일부 실시형태에서, 변이체 PNPLA3 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질은 서열번호 42에 따른 아미노산 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질은 서열번호 43에 따른 144번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질은 서열번호 43에 따른 아미노산 서열을 포함한다.

[0298] 일부 실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 게놈 DNA이다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 뉴클레오타이드 서열을 포함한다.

[0299] 일부 실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 mRNA이다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 뉴클레오타이드 서열을 포함한다.

[0300] 일부 실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 cDNA이다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 뉴클레오타이드 서열을 포함한다.

[0301] 일부 실시형태에서, 간 질환은 만성 간 질환이다. 일부 실시형태에서, 만성 간 질환은 비알코올성 지방간 질환(NAFLD), 알코올성 간 질환(ALD), 비알코올성 지방간염(NASH), 간경변, 지방증 또는 간세포 암종이다. 일부 실시형태에서, 간 질환은 알코올성 간 질환이다. 일부 실시형태에서, 알코올성 간 질환은 알코올 섭취로 인한 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다. 일부 실시형태에서, 간 질환은 비알코올성 간 질환이다. 일부 실시형태에서, 비알코올성 간 질환은 비알코올성 지방간 질환(NAFLD) 또는 비알코올성 지방간염(NASH)을 포함한다. 일부 실시형태에서, 비알코올성 간 질환은 알코올 섭취에 의해서 유발되지 않은 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다.

[0302] 일부 실시형태에서, 인간 대상체는 기능성 HSD17B13에 대해서 동형접합성 또는 이형접합성이다. 일부 실시형태에서, 대상체는 기능성 HSD17B13에 대해서 동형접합성이다. 일부 실시형태에서, 대상체는 기능성 HSD17B13에 대해서 이형접합성이다.

[0303] 본 명세서에 기재된 방법 중 임의의 것에서, 프로브 또는 프라이머 또는 변경-특이적 프로브 또는 변경-특이적 프라이머는 단일 핵산 종에 특이적으로 상보성이거나 또는 단일 핵산 종과 특이적으로 혼성화할 수 있다. 예를 들어, HSD17B13 전사물 A, 전사물 B, 전사물 E 또는 전사물 I에 대한 핵산(예를 들어, 본 명세서에 기재된 기능성 HSD17B13에 대한 mRNA, cDNA, RNA 전사물 또는 cDNA 전사물 중 임의의 것)에 특이적으로 상보성이거나 이것에 특이적으로 혼성화하는 프로브 또는 프라이머 또는 변경-특이적 프로브 또는 변경-특이적 프라이머는 변이체 HSD17B13에 대한 핵산 분자 중 임의의 것(예를 들어, HSD17B13의 변이체 C, D, F, G, H에 대한 mRNA, cDNA, RNA 전사물 또는 cDNA 전사물 중 임의의 것)에 상보성이 아니거나 이에 혼성화한다.

[0304] 본 개시내용은 또한 I148M 변이를 포함하는 PNPLA3 단백질을 갖고, 그리고 기능성 HSD17B13 단백질을 갖는 인간 대상체에서 간 질환을 치료하는데 사용하기 위한 HSD17B13의 저해제를 제공한다. 일부 실시형태에서, 인간 대상체는 I148M 변이를 포함하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질에 대해서 양성인 것으로 시험되었다. 일

부 실시형태에서, 치료는 상기 인간 대상체가 I148M 변이를 포함하는 PNPLA3 단백질 및 기능성 HSD17B13 단백질을 갖는지의 여부를 결정하는 단계를 포함한다. 일부 실시형태에서, 인간 대상체는, 본 명세서에 정의된 바와 같은 방법 중 임의의 것을 사용하여 HSD17B13을 저해함으로써 간 질환을 치료 또는 저해하기 위한 후보자인 것으로 식별되었다.

[0305] 일부 실시형태에서, 변이체 PNPLA3 단백질은 서열번호 42에 따른 148번 위치에 상응하는 위치에 메티오닌을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질은 서열번호 42에 따른 아미노산 서열, 또는 서열번호 42와 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 아미노산 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 게놈 DNA이다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 5107 내지 5109번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, 게놈 DNA는 서열번호 31에 따른 뉴클레오타이드 서열, 또는 서열번호 31과 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 mRNA이다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 442 내지 444번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 34에 따른 뉴클레오타이드 서열, 또는 서열번호 34와 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 430 내지 432번 위치에 상응하는 위치에 AUG 코돈을 포함한다. 일부 실시형태에서, mRNA는 서열번호 35에 따른 뉴클레오타이드 서열, 또는 서열번호 35과 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 변이체 PNPLA3 단백질을 암호화하는 핵산 분자는 cDNA이다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 442 내지 444번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 38에 따른 뉴클레오타이드 서열, 또는 서열번호 38과 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 430 내지 432번 위치에 상응하는 위치에 ATG 코돈을 포함한다. 일부 실시형태에서, cDNA는 서열번호 39에 따른 뉴클레오타이드 서열, 또는 서열번호 39와 적어도 90%의 서열 동일성을 갖고 I148M 변이를 포함하는 PNPLA3 단백질을 암호화하는 뉴클레오타이드 서열을 포함한다.

[0306] 일부 실시형태에서, 간 질환은 만성 간 질환이다. 일부 실시형태에서, 만성 간 질환은 비알코올성 지방간 질환(NAFLD), 알코올성 간 질환(ALD), 비알코올성 지방간염(NASH), 간경변, 지방증 또는 간세포 암종이다. 일부 실시형태에서, 간 질환은 알코올성 간 질환이다. 일부 실시형태에서, 알코올성 간 질환은 알코올 섭취로 인한 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다. 일부 실시형태에서, 간 질환은 비알코올성 간 질환이다. 일부 실시형태에서, 비알코올성 간 질환은 비알코올성 지방간 질환(NAFLD) 또는 비알코올성 지방간염(NASH)을 포함한다. 일부 실시형태에서, 비알코올성 간 질환은 알코올 섭취에 의해서 유발되지 않은 간경변, 지방증 또는 간세포 암종 중 하나 이상을 포함한다.

[0307] 일부 실시형태에서, 인간 대상체는 기능성 HSD17B13에 대해서 동형접합성 또는 이형접합성이다.

[0308] 상기 또는 하기에서 인용된 모든 특허 출원, 웹사이트, 기타 간행물, 등록 번호 등은, 각각의 개별적인 항목이 구체적으로 및 개별적으로 참고로 포함된 것으로 제시된 것과 동일한 정도로 모든 목적을 위해서 전문이 참조에 의해 포함된다. 상이한 버전의 서열이 상이한 시점에서의 등록 번호와 연관되는 경우, 본 출원의 유효한 출원일의 등록 번호와 연관된 버전이 의미가 있다. 유효한 출원일은 실제 출원일 이전이거나 해당되는 경우 등록 번호를 언급한 우선권 출원의 출원일을 의미한다. 마찬가지로, 상이한 버전의 간행물, 웹사이트 등이 상이한 시점에 공개되면, 달리 제시되지 않는 한 본 출원의 유효한 출원일에 가장 가깝게 공개된 버전을 의미한다. 본 개시내용의 임의의 특징, 단계, 요소, 실시형태 또는 양상은 구체적으로 달리 제시되지 않는 한 임의의 다른 것과 조합하여 사용될 수 있다. 본 발명이 명료성 및 이해를 목적으로 어느 정도 상세하게 설명 및 예에 의해서 기재되었지만, 특정 변화 및 변형이 첨부된 청구범위의 범주 내에서 실시될 수 있는 것이 명백할 것이다.

[0309] 본 명세서에서 언급된 뉴클레오타이드 및 아미노산 서열은 뉴클레오타이드 염기에 대한 표준 문자 약어 및 아미노산에 대한 1-문자 암호를 사용하여 제시된다. 뉴클레오타이드 서열은 서열의 5' 단부에서 시작하여 3' 단부를 향해 진행하는(즉, 각 라인에서 좌측에서 우측으로) 표준 관례를 따른다. 각각의 뉴클레오타이드 서열의 단지 하나의 가닥이 제시되지만, 상보성 가닥은 표시된 가닥의 임의의 언급에 의해서 포함되는 것으로 이해된다. 아미노산 서열은 서열의 아미노 말단에서 시작하여 카복시 말단을 향해 진행하는(즉, 각각의 라인에서 좌측에서 우측으로) 표준 관례를 따른다.

[0310] 하기 실시예는 실시형태를 보다 상세하게 기재하기 위해서 제공된다. 이것은 청구된 실시형태를 제한하는 것이

아니라, 설명하도록 의도된다.

[0311] **실시예**

[0312] 하기 실시예는, 본 명세서에 청구된 화합물, 조성물, 물품, 디바이스 및/또는 방법이 제조 및 평가되는 방법의 완전한 개시 및 설명을 당업자에게 제공하도록 제시되며, 순수하게 예시적인 것으로 의도되며, 본 발명자들이 발명 주제로서 간주하는 것의 범주를 제한하도록 의도되지 않는다. 수치(예를 들어, 양, 온도 등)와 관련하여 정확성을 보장하기 위해서 노력하였지만, 일부 오류 및 편차가 고려되어야 한다. 달리 제시되지 않는 한, 부는 중량부이고, 온도는 ℃ 단위 또는 주변 온도이고, 압력은 대기압 또는 그 근처이다.

[0313] **실시예 1: PNPLA3 rs738409(p.I148M)와 HSD17B13 rs72613567 간의 유전적 상호작용 - 연구 설계**

[0314] 본 연구에서, 엑솜 서열분석을 사용하여 혈청 알라닌 아미노트랜스퍼라제(ALT) 및 아스파테이트 아미노트랜스퍼라제(AST) 수준과 연관된 변이체를 식별하였는데, 이것은 DiscovEHR 인간 유전학 연구, 엑솜 서열 데이터를 전자 건강 기록(electronic health record: EHR)과 연결하는 코호트 그리고 3개의 추가 연구에서 간세포 손상의 마커이다. DiscovEHR 및 2개의 독립적인 코호트에서 연관된 유전자 변이체와 만성 간 질환의 임상적 진단 간의 연관성을 또한 연구하였다. 간 생검을 진행한 비만 수술 환자의 독립적인 코호트에서 이들 변이체 중 하나와 간 질환의 조직병리학적 중증도를 또한 연구하였다.

[0315] *연구 설계 및 참여자*

[0316] 게놈 DNA 샘플 및 6개의 코호트로부터의 데이터를 사용하여 인간 유전학 연구를 수행하였다. 이들 연구는 GHS20의 MyCode(등록상표) 지역 건강 계획(Community Health Initiative)으로부터의 첫 번째 50,726명의 동의한 성인 참여자로부터 기원한 2개의 리제네런 유전학 센터(Regeneron Genetics Center) 및 가이싱거 헬쓰 시스템(Geisinger Health System: GHS) DiscovEHR 연구 집단을 포함하였다. GHS 발견 코호트는 비만 수술 코호트에 모집된 모든 사람을 제외하고, 2007년 내지 2016년 사이에 외래환자 1차 의료 및 전문 의료로부터 모집된 46,544명의 유럽인 개체로 이루어졌다. GHS 비만 수술 코호트는 비만 수술에 대해서 조회된 2,644명의 유럽인 개체로 이루어졌다. 간 트랜스아미나제와의 연관성의 반복 연구를 댈러스 심장 연구(Dallas Heart Study) 및 펜 메디슨 바이오뱅크(Penn Medicine Biobank)에서 수행하였는데, 이것은 각각 1,357명 및 8,527명의 유럽 혈통 개체를 포함하였다. 만성 간 질환과의 연관성의 반복 연구는 댈러스 간 연구(Dallas Liver Study: DLS)로부터의 517명의 개체 및 댈러스 소아 간 연구(Dallas Pediatric Liver Study: DPLS)로부터의 439명의 개체를 포함하였다. 전체 연구 설명 및 임상 표현형 및 질환 정의는 보충 부록의 방법 섹션에 기재되어 있다.

[0317] 댈러스 간 연구 및 소아 간 연구로부터의 유전자형분석된 다민족 사례 및 대조군의 기준선 특징을 도 5에 나타낸다.

[0318] *샘플 제조, 서열분석 및 유전자형분석*

[0319] DiscovEHR 연구, 댈러스 심장 연구 및 펜 매디슨 바이오뱅크에서의 참여자에 대한 DNA 샘플 제조 및 전체 엑솜 서열분석을 이미 기재된 바와 같이 리제네런 유전학에서 수행하였다(Dewey et al., Science, 2016, In Press). HSD17B13 rs72613567을 댈러스 간 연구 및 델러스 소아 간 연구에서 태크맨(Taqman) 검정에 의해서 유전자형분석하였다(그리고 각각의 유전자형 중 5명의 개체에서 생어 서열분석에 의해서 검증하였다).

[0320] *발견 코호트에서의 임상 측정치 및 만성 간 질환 정의*

[0321] 알라닌 아미노트랜스퍼라제(ALT) 및 아스파테이트 아미노트랜스퍼라제(AST)에 대한 임상 실험 측정치를 GHS 발견 코호트 및 비만 수술 코호트로부터의 참여자의 EHR로부터 얻었다. 중위 ALT 및 AST 값을 2회 이상의 측정치를 사용하여 모든 참여자에 대해서 계산하였고, 연관성 분석 이전에 $\log_{10}$-변환하여 분포를 정규화하였다.

[0322] 국제 질환 분류, 제9차 개정(International Classification of Diseases, Ninth Revision)(ICD-9) 질환 진단 코드를 EHR로부터 얻었고, 비바이러스성 비알코올성(ICD-9 571.40, 571.41, 571.49, 571.5, 571.8, 571.9) 또는 알코올성(ICD-9 571.0, 571.1, 571.2, 571.3) 간 질환 사례 정의에 대한 임상 질환 카테고리로 나누었다. 단일 진단 코드에 기초한 추가 사례 정의는 알코올성 간경변(ICD-9 571.2), 비알코올성 간경변(ICD-9 571.5) 및 HCC(ICD-9 155.0)를 포함하였다. 이러한 사례 정의의 경우, 간 질환이 없는 일반적인 대조군("간 질환 없음")을 임의의 유형의 간 질환을 나타내는 사례 기준 또는 단일-충돌 또는 문제-목록 진단 코드가 없는 참여자로서 정의하였다.

[0323] HSD17B13 주변 영역에서 GHS 발견 코호트에서의 알라닌 아미노트랜스퍼라제(ALT; A) 및 아스파테이트 아미노트

랜스퍼라제(AST; B)에 대한 영역 연관성 플롯을 도 6(패널 A 및 패널 B)에 나타낸다. 자색 다이아몬드는 스플라이스 변이체 rs72613567을 나타낸다. 각각의 원은 단일 뉴클레오타이드 변이체를 나타내는데, 원의 색상은 그 변이체와 rs72613567 간의 연관 비평형(linkage disequilibrium)(DiscovEHR 코호트에서 계산된 r2)을 나타낸다. 청색 선은 HapMap에서 예측된 재조합율을 나타낸다. 패널의 하단 부분은 유전자좌 내의 각각의 유전자의 상대적인 위치 및 전사 가닥을 나타낸다. 이웃하는 유전자 HSD17B11(ALT 및 AST에 대해서 각각 가장 유의한 P-값 1.4×10^{-1} 및 4.3×10^{-2})에서 AST 또는 ALT와 암호 또는 스플라이스 영역 변이체 간에는 유의한 연관성이 존재하지 않았다.

[0324] *비만 수술 코호트에서의 간 조직병리학적 표현형 정의*

[0325] GHS 비만 수술 코호트는 2,644명의 유럽 혈통(European descent) 개체로 이루어졌으며, 이들 개체 중 2,391명으로부터 수술 중 간 생검 시편을 입수하였다. 간 생검 시편을 이미 기재된 바와 같이, 폼알린-고정시키고, 조직학을 위해서 헤마톡실린 및 에오신으로 염색하고, 섬유증의 평가를 위해서 매슨 트라이크롬 염색제(Masson's trichrome stain)로 염색하였다(Gerhard et al., Patient Saf. Surg., 2011, 5, 1). 이미 확립된 기준을 사용하여 간병리학자에 의해서 조직학 진단을 결정하였다(Brunt et al., Am. J. Gastroenterol., 1999, 94, 2467-74). 조직학 진단을 사용하여 하기 표현형을 정의하였다: 1) 정상: 지방증, NASH 또는 섬유증 증거 없음; 2) 단순 지방증: NASH 또는 섬유증의 증거가 없는 지방증(등급 무관); 3) NASH: 소엽 염증 또는 간세포 팽창(등급 무관)의 임의의 존재, 또는 섬유증(병기 무관)의 임의의 존재.

[0326] 서열분석된 유럽 혈통 개체 발견 및 반복 코호트의 기준선 특징을 도 1에 나타낸다. 발견 코호트에서 1.0×10^{-7} 미만의 P로 혈청 트랜스아미나제와 연관된 단일 뉴클레오타이드 변이체를 도 2에 나타낸다.

[0327] *DNA 샘플 제조 및 서열분석*

[0328] 간략하면, 제조사의 제안된 프로토콜(로슈 님블젠사(Roche NimbleGen))에 따라서 NimbleGen 프로브를 사용하여 엑솜 캡처를 수행하였다. 캡처된 DNA를 PCR 증폭시키고, qRT-PCR(카파 바이오시스템즈사(Kapa Biosystems))에 의해서 정량하였다. 멀티플렉싱된 샘플을 Illumina v4 HiSeq 2500 상에서 75bp 페어드-엔드 서열분석을 사용하여 샘플의 96%에서 표적 염기의 85% 초과의 20x 초과의 반수체 판독 깊이를 제공하기에 충분한 커버리지 깊이(표적 염기의 대략 80x 평균 반수체 판독 깊이)까지 서열분석하였다. 각각의 Illumina Hiseq 2500 실시로부터의 원시 서열 데이터를 서열 판독 정렬 및 변이체 식별을 위해서 DNAnexus 플랫폼에 업로딩하였다(Reid et al., BMC Bioinformatics, 2014, 15, 30). 간략하면, 원시 서열 데이터를 BCL 파일로부터 샘플-특이적 FASTQ-파일로 변환하고, 이것을 BWA-mem을 사용하여 인간 참조 빌드 GRCh37.p13에 정렬하였다(Li et al., Bioinformatics, 2009, 25, 1754-60). 단일 뉴클레오타이드 변이체(SNV) 및 삽입/결실(indel) 서열 변이체를 Genome Analysis Toolkit를 사용하여 식별하였다(McKenna et al., Genome Res., 2010, 20, 1297-303).

[0329] *간 효소 및 만성 간 질환 표현형의 엑솜-와이드 연관성 분석*

[0330] 1% 미만의 누락 데이터율(missing data rate), 1.0×10^{-6} 초과의 하디-바인버그 평형(Hardy-Weinberg equilibrium) P-값 및 0.1% 초과의 부 대립유전자 빈도를 갖는 502,219개의 이대립유전자 변이체를 트랜스아미나제 수준과의 연관성에 대해서 조사하였다. Log_{10}-변환된 중위 ALT 및 AST를 연령, 연령^2, 성별, BMI 및 혈통의 제1의 4개의 주성분에 대해서 조정하였다. 연구 참여자에서의 관련성을 설명하기 위해서, 유전적 관련성 매트릭스는 무작위-효과 공변량(random-effects covariate)으로서 적합하였다. 주성분 및 유전적 관련성 매트릭스 둘 다를 대략적인 연관 평형(linkage equilibrium)에서 0.1% 초과의 부 대립 유전자 빈도로 39,858개의 비-MHC 마커로부터 구축하였다. 선형 혼합 모델을 GCTA 패키지에서 구현되는 바와 같이 사용하여(Yang et al., Am. J. Hum. Genet., 2011, 88, 76-82) 특질 잔차(trait residual)와 단일 뉴클레오타이드 변이체 간의 연관성에 대해서 시험하였다. 본문에 기록된 모든 P-값은 대립유전자 모델에 상응한다.

[0331] GHS 발견 코호트에서의 연관성 반복을 3개의 개별 유럽 혈통 코호트: GHS 비만 수술 코호트, 댈러스 심장 연구 및 펜 매디슨 바이오뱅크(상기에 기술됨)에서 시도하였다. GHS 비만 수술 코호트 및 펜 매디슨 바이오뱅크로부터의 ALT 및 AST 측정치를 log_{10}-변환시키고, 연령, 연령^2, 성별, BMI 및 혈통의 제1의 4개의 주성분에 대해서 조정하였다. 유전적 관련성 매트릭스를 무작위-효과 공변량으로서 포함시켰고, GCTA에서 선형 혼합 모델을 사용하여 분석을 수행하였다. 댈러스 심장 연구에서, log_{10}-변환된 ALT 및 AST 측정치를 ALT 및 AST를 연령, 연령^2,

성별, BMI 및 혈통의 제1의 10개의 주성분에 대해서 조정하고, PLINK로 구현되는 선형 회귀를 사용하여 분석을 수행하였다. 3개의 반복 코호트에 대한 통계 요약을 METAL을 사용하여 메타 분석하였다(Willer et al., Bioinformatics, 2010, 26, 2190-1)(반복 메타-분석). 발견 코호트 및 3개의 반복 코호트에 대한 통계 요약을 유사하게 메타 분석하였다(공동 메타-분석).

[0332] 3개의 개별 유럽 혈통 코호트에서 발견 코호트로부터의 35개의 엑솜-와이드 유의한 뉴클레오타이드 변이체의 반복 및 공동 메타-분석을 도 3에 나타낸다.

[0333] GHS 발견 코호트에서 트랜스아미나제와 엑솜 와이드 유의한 연관성($p<1\times10^{-7}$)을 갖는 변이체에 대해서, 본 명세서에 기재된 유럽 혈통 반복 연구로 본 명세서에 기재된 바와 같이 연관성 분석 및 메타 분석을 수행하였다. 시험된 변이체의 수에 의해서 결정된 본페로니 유의성 역치(Bonferroni significance threshold)를 사용하여 반복 연관성을 정의하였다. 발견 및 반복 연구의 메타 분석을 또한 수행하였다. 본문에 기록된 모든 P-값은 대립유전자 모델에 상응한다.

[0334] 트랜스아미나제-연관된 단일 뉴클레오타이드 변이체를 또한 만성 간 질환 표현형과의 연관성에 대해서 조사하였다(본 명세서에 기재된 바와 같이 정의 및 분석됨). 시험된 변이체의 수 및 넓은 만성 간 질환 카테고리에 의해서 결정된 본페로니 유의성 역치를 사용하여 연관 유의성을 결정하였다. 반복된 신규 변이체를 GHS 비만 수술 코호트로부터 조직병리학적으로 정의된 간 표현형과의 연관성에 대해서 또한 조사하였다.

[0335] *만성 간 질환 표현형과의 연관성 분석*

[0336] 간 효소 ExWAS로부터의 13개의 유의한 반복 단일 뉴클레오타이드 변이체를 상기에 기재된 바와 같이, GHS 발견 코호트로부터 정의된 만성 간 질환 표현형과의 연관성에 대해서 분석하였다. 0.05/26 미만의 P(1.92×10^{-3} 미만의 P)의 본페로니 유의성 역치를 사용하여 13개의 변이체 및 시험된 2개의 넓은 만성 간 질환 카테고리(알코올성 및 비알코올성)를 설명하였다. *HSD17B13* rs72613567 변이체를 상기에 기재된 바와 같이 GHS 비만 수술 코호트로부터 조직병리학적으로 정의된 간 표현형과의 연관성에 대해서 추가로 시험하였다. 연령, 연령^2, 성별, BMI 및 혈통의 제1의 4개의 주성분에 대해서 조정한 후 승산비(odds ratio)를 퍼스의 로지스틱 회귀의 벌점 가능성 방법(Firth's penalized likelihood method of logistic regression)을 사용하여 예측하였다. 유전자형 승산비를 동일한 공변량을 사용하여 *HSD17B13* rs72613567에 대해서 예측하였다.

[0337] 연령, 연령^2, 성별, BMI 및 자가-보고된 민족성에 대해서 조정된, 로지스틱 회귀에 의해서 DLS에서 간 질환에 대한 승산비를 예측하였다. 입수 가능한 rs72613567 유전자형을 갖는 댈러스 심장 연구로부터의 참여자를 정상 대조군으로서 사용하였다(n=4,279). DPLS에서의 승산비를 로지스틱 회귀에 의해서 예측하였다.

[0338] 발견 코호트에서 간 질환 표현형을 갖는 13개의 엑솜-와이드의 유의한 반복 단일 뉴클레오타이드 변이체의 연관성을 도 4에 나타낸다.

[0339] *PNPLA3 rs738409(p.I148M)와 HSD17B13 rs72613567 간의 유전적 상호작용 - 분석*

[0340] *PNPLA3* rs738409와 *HSD17B13* rs72613567의 조합 효과를 평가하기 위해서, 각각 선형 회귀 및 로지스틱 회귀를 사용하여 정량적(ALT 및 AST) 및 2항(비알코올성 간 질환 및 알코올성 간 질환) 특질에 대한 연관성 분석을 수행하고, 상가적 유전자 모델(additive genetic model)을 가정하여 두 유전자 변이체뿐만 아니라 상호작용항(interaction term)에 대한 주요 효과를 모델링하였다. 모든 모델을 연령, 연령^2, 성별, BMI 및 혈통의 제1의 4개의 주성분에 대해서 조정하였다. 염기 R에서 glm 함수를 사용하여 통계학적 분석을 수행하였다.

[0341] *소프트웨어*

[0342] GCTA 소프트웨어, 버전 1.25.0(Yang et al., Am. J. Hum. Genet., 2011, 88, 76-82) 및 PLINK, 버전 1.9.0을 사용하여 유전적 연관성 분석을 수행하였다. 분위수-분위수 및 맨허튼 플롯을 R 소프트웨어, 버전 3.2.1(R Project for Statistical Computing)을 사용하여 생성시켰다. LocusZoom(Pruim et al., Bioinformatics, 2010, 26, 2336-7)을 사용하여 영역 연관성 플롯을 생성시켰다.

[0343] *RNA 서열분석 연구*

[0344] Agilent RNA Nano Bioanalyzer 칩 상에서 전체 RNA를 시험함으로써 RNA 품질 및 농도를 평가하였는데; 모든 샘플은 8 초과의 RNA 온전성 값(RNA integrity number: RIN)을 가졌다. 폴리아데닐화된 RNA 전사물을 올리고(dT)25 비드(써모 피셔 사이언티픽사(Thermo Fisher Scientific))를 사용하여 2 라운드의 풍부화를 사용하여 단

리시켰다. 샘플을 RNAclean XP 비드(베크만 쿨터사(Beckman Coulter))를 사용하여 정제 및 농축시키고, 대략 140 염기쌍으로 열 단편화시켰다. 무작위 육량체를 사용하여 SuperScript III 역전사효소(써모 피셔 사이언티픽사)로 제1-가닥 합성을 완결하였고; dTTP를 제2-가닥 합성 동안 dUTP로 대체하였다. 우라실 DNA-글리코실라제 단계를 추가하여 엑솜에 대해서 상기에 언급된 표준 DNA 라이브러리 제조 방법에 따라서 샘플을 처리하여 가닥-특이적 서열분석 라이브러리를 생성시켰다. 샘플을 풀링(pooling)시키고, Illumina v4 HiSeq 2500 상에서 75bp 페어드-엔드 서열분석을 사용하여 서열분석하였다.

[0345] *신규 HSD17B13 전사물의 식별 및 검증*

[0346] 판독물을 2개의 미스매치를 허용하는 ArrayStudio(등록상표) 소프트웨어(OmicSoft(등록상표), 미국 노쓰캐롤라이나주 캐리 소재)를 사용하여 Human.B38에 대해서 맵핑하였다. 두 접근법을 사용하여 신규 *HSD17B13* 전사물을 식별하였다. 신규 엑손 접합부를 ArrayStudio를 사용하여 Gencode v24에 기초하여 발견하였다. Trinity(v2.2.0)를 디폴트 설정으로 사용하여 신생 전사물 조립을 수행하였다. 커스텀 유전자 모델을 구축하여 *HSD17B13*의 신규 전사물을 혼입하였고, 전사물 정량을 커스텀 유전자 모델에 대한 판독물 정렬에 의해서 예측하였다. 모든 식별된 HSD17B13 아이소폼의 단백질 서열 정렬을 결정하였다. Platinum(상표명) Taq DNA Polymerase(써모피셔사(Thermofisher))를 사용하는 SuperScript(상표명) One-Step RT-PCR을 사용하여 인간 간 샘플로부터의 전체 RNA에 대해서 RT-PCR을 수행하였다. 각각의 50㎕의 RT-PCR 반응은 1X 반응 믹스, 500nM(각각 정방향 프라이머 및 역방향 프라이머(PST516: ATGAACATCATCCTAGAAATCCTTC; 서열번호 62) 및 PST517: ATCATGCATACATCTCTGGCTGGAG; 서열번호 63)), 1㎕의 RT/Platinum Taq 및 75ng의 RNA를 함유하였다. 주기 조건은, 30분 동안 45℃의 1회 사이클; 2분 동안 94℃의 1회 사이클; 20초 동안 94℃, 30초 동안 53℃ 및 90초 동안 72℃의 40회 사이클; 5분 동안 72℃의 1회 사이클; 이어서 10℃ 유지였다. 생성물을 QIAquick PCR 정제 키트(퀴아젠사)를 사용하여 정제시키고, 프라이머 DE002(ATCAGAACTTC AGGCCTTGG; 서열번호 64)를 사용하여 직접 생어 서열분석에 적용하였다. B 전사물 및 C 전사물을 식별하기 위해서, RT-PCR 생성물을 SYBR GoldSYBR(등록상표) Gold Nucleic Acid Gel Stain(써모피셔사)로 염색된 2% 아가로스 겔 상에서 전개시키고, 예측된 분자량의 밴드를 절제하고, QIAquick Gel Extraction Kit(퀴아젠사)를 사용하여 정제시키고, 이어서 TOPO(등록상표) TA Cloning Kit(써모피셔사)를 사용하여 클로닝에 적용하였다. M13F 및 M13R 서열분석 프라이머를 사용하여 TOPO 클론의 서열분석을 수행하였다. Sequencher DNA 분석 소프트웨어(진 코즈사(Gene Codes Corporation))를 사용하여 서열 분석을 수행하였다.

[0347] 제1 엑손(GCAAAGCCATGAACATCATCC; 서열번호 65) 및 마지막 엑손(TCTTGATGTAGTGGGAGTCGGATT; 서열번호 66)에서 유전자-특이적 프라이머를 사용하여 Platinum Taq High Fidelity가 구비된 SuperScript III One-step RT-PCR System(써모 피셔 사이언티픽사)으로 50ng의 전체 RNA로부터 직접 전장 HSD17B13 전사물을 증폭시켜 약 2.2kb의 앰플리콘(최대 예측된 크기의 전사물)을 생성시켰다. 앰플리콘을 Agilent Bioanalyzer 상에서 검증하였다. PacBio-호환성 바코드 어댑터를 앰플리콘에 결찰시키고, PacBio PB 비드(퍼시픽 바이오사이언시스사(Pacific Biosciences))로 세정하였다. 라이브러리를 동일한 양으로 풀링시키고, 하나의 SMRT 세포에 대해서 180분 동안 PacBio RSII 플랫폼 상에서 서열분석하였다. PacBio 소프트웨어 smrtanalysis v2.3 tool labelzmw을 사용하여 디멀티플렉싱(demultiplexing)하고, 이어서 ConsensusTools AmpliconAnalysis로 분석하였다. 생성된 앰플리콘을 HSD17B13 RefSeq 유전자와 비교하여 아이소폼 및 유전자형 상태를 결정하였다.

[0348] *HSD17B13 아이소폼의 세포 이하 국지화*

[0349] HepG2 세포에 HSD17B13 A 전사물 및 D 전사물을 보유하는 렌티바이러스를 감염시키고, 안정적인 세포주를 선택하고, HSD17B13 아이소폼, 지방 소립 및 소포체를 면역형광을 사용하여 가시화하였다. 간략하면, HepG2 세포를 10% 우태아 혈청이 보충된 이글 최소 필수 배지(Eagle's Minimum Essential Medium) 중에서 배양하였다. *HSD17B13* 전사물 A 및 D를 Myc-DDK 골격 렌티바이러스 작제물 내에 서브클로닝시키고, 렌티바이러스를 생성시켰다. HepG2 세포에 *HSD17B13* 전사물을 보유하는 렌티바이러스를 감염시켰다. 각각의 *HSD17B13* 전사물을 발현하는 안정적인 세포주를 완전 배양 배지 중에서 1 내지 3㎎/㎖의 Geneticin G-418 설페이트로 2주 동안 선택하였다. 고정 이후에, HSD17B13 아이소폼을 마우스 항-Myc 항체를 사용하여 검출하였다. 지방 소립을 BODIPY FL 염료(시그마사)로 표지하였다. 지질 코트 단백질 및 소포체를 토끼 항-PLIN 항체(시그마사) 및 토끼 항-칼넥신 항체(셀 시그널링 테크놀로지사)로 표지하였다. 면역형광을 위한 2차 항체는 Alexa Fluor 488 당나귀 항-토끼 IgG 및 Alexa Fluor 594 당나귀 항-마우스 IgG(잭슨 이뮤노리서치사)였다.

[0350] *인간 간 생검 조직에서의 HSD171B3 단백질 발현의 정량*

[0351] 인간 간 및 세포 펠릿 샘플을 프로테아제 및 포스파타제 저해제 혼합물(써모피셔사)의 존재 하에서 얼음 냉각된

1× RIPA 용해 완충액(EMD 밀리포어사(EMD Millipore)) 중에서 균질화시켰다. 상청액을 수집하고, BCA 단백질 검정(써모피셔사)을 사용하여 단백질 농축을 위해서 사용하였다. 인간 조직 용해물을 30㎍/웰로 로딩하고, 안정적인 세포주를 9㎍/웰로 로딩하고, SDS/PAGE 겔(바이오래드사) 상에서 분리시키고, PVDF 막(바이오래드사)에 전달하였다. 막을 1시간 동안 0.1% Tween20(바이오래드사)이 보충된 1x TBS 중의 5%(wt/vol) 우유로 차단시켰다. 막을 항체와 함께 4℃에서 밤새 HSD17B13(1:200, 써모피셔사) 및 B-액틴(1:500, 셀 시그널링 테크놀로지사)에 대해서 인큐베이션시켰다. 결합된 항체를 HRP-접합된 항-토끼 항체(1:10,000, 잭슨 이뮤노리서치사)를 사용하여 검출하고, 화학발광 시약(써모피셔사)을 사용하여 증강시켰다. Image J 소프트웨어를 사용하여 밴드 강도를 정량하였다.

[0352] *HSD17B13 효소 활성도의 시험관내 및 세포 특징규명*

[0353] 재조합 인간 HSD17B13 단백질을 HSD17B13 전사물 A 또는 전사물 D를 보유하는 플라스미드 DNA로 형질전환된 이. 콜라이(*E. coli*)(진스크립트사(Genscript))로부터 정제시켰다. HSD17B13 변이체는 C 말단에 10×His 태그를 함유하였고, Ni^{2+} 친화도 정제를 사용하여 가용성 분획으로부터 정제되었다. NAD(P)H-Glo 검출 시스템(프로메가사(Promega))을 사용한 NADH 생산의 측정을 통해서 효소 활성도를 결정하였다. 반응을 3시간 동안 25℃에서 0.2M tris-HCl, pH 7.5, 0.5mM NAD^+, 75μM의 기질(시그마사) 및 500ng의 정제된 효소 중에서 100㎕의 최종 부피로 수행하였다. 인큐베이션 후, 20㎕의 반응을 20㎕의 루시퍼라제 시약(프로메가사)과 합하고, 1시간 동안 실온에서 인큐베이션시키고, Envision Plate Reader(퍼킨 엘머사) 상에서 판독하였다.

[0354] HSD17B13 전사물 A, 전사물 D 또는 녹색 형광 단백질(GFP, 대조군)을 과발현하는 HEK293 세포를 사용하여 세포-기반 검정에서 에스트라다이올에 대한 HSD17B13의 활성도를 조사하였다. 에스트라다이올을 각각의 세포 유형에 공급하였다. 48시간 후에, 배지를 수집하고, 에스트라다이올의 농도 및 이의 전환 생성물인 에스트론을 식별하고, LC-MS에 의해서 정량하였다. 하이드록시에스트라다이올(에스트라다이올로부터의 대사산물) 및 하이드록시에스트론(에스트론의 대사산물)을 LC-MS에 의해서 식별하였다.

[0355] **실시예 2: 66개 인간 간 샘플에서 HSD17B13 및 PNPLA3의 유전자 발현 분석**

[0356] HSD17B13 및 PNPLA3의 유전자 발현을 66개의 인간 간 샘플을 사용하여 분석하였다. 모든 샘플은 지방증, 소엽염증 또는 섬유증이 없는 대조군 공여자로부터 유래되었다. HSD17B13 rs72613567(T/T, T/TA 및 TA/TA) 및 PNPLA3 rs738409(C/C, C/G 및 G/G) 유전자형의 분포를 표 1에 나타낸다.

표 1

[0357]

유전자형	C/C	C/G	G/G	ND
T/T	12	8	1	0
T/TA	15	12	0	2
TA/TA	12	4	0	0

[0358] PNPLA3의 발현은 HSD17B13 rs72613567 스플라이스 변이체의 동형접합성 대체 보유자에서 상당히 감소되었다(도 7 참고). mRNA 발현은 FPKM 단위로 나타낸다. FDR 0.0071을 갖는 T/T와 비교할 때 1.6배 감소가 관찰되었다. HSD17B13 TA/TA 유전자형을 갖는 변이체 PNPLA3 C/C 보유자는 HSD17B13 T/T 보유자와 비교할 때 상당히 감소된 발현을 가졌다: 1.7배(FDR 0.017) 감소. TA/TA 유전자형을 갖는 변이체 PNPLA3 C/G 보유자는 발현 감소를 나타내었지만, 통계학적으로 유의하지 않았다(1.4배, FDR 1). 도 8은 3개의 HSD17B13 rs72613567 유전자형(T/T, T/TA, TA/TA)에서의 63 PNPLA3 rs738409 보유자(C/C 및 C/G, 표 1 참고)의 발현 차이를 나타낸다.

[0359] 변이체 PNPLA3 p.I148M 변이체는 NAFLD에 대해서 가장 널리 검증된 유전적 위험 인자이며, 148M 대립유전자는 혈통에 따라서 개체의 5 내지 25%에서 동형접합성 상태로 존재한다. HSD17B13 rs72613567:TA가 PNPLA3 p.I148M과 연관된 간 손상의 위험을 변경시키는지의 여부를 이해하기 위해서, DiscovEHR에서 ALT, AST 및 만성 간 질환 표현형과의 연관성에서 두 변이체 간의 상호작용의 분석을 수행하였다. 이러한 분석은 모든 참여자에서 그리고 비만(체질량 지수(BMI) > 30kg/㎡) 및 비-비만(BMI ≤ 30kg/㎡) 하위 집단에서 수행하였다. ALT(상호작용에 대해서 $P=1.8\times10^{-3}$) 및 AST(상호작용에 대해서 $P=4.5\times10^{-3}$) 수준의 연관성 분석에서 HSD17B13 rs72613567:TA와

PNPLA3 p.I148M 사이에 명목상 유의한 상호작용이 존재하였고; 이러한 연관성은 비만 개체에서의 연관성에 의해서 주로 기인된다(도 9 참고). 이러한 분석에서, rs72613567:TA 대립유전자는 ALT 및 AST 증가와 PNPLA3 148M 대립유전자의 대립유전자 용량-의존적 연관성을 완화시켰다(도 10 참고). 도 10을 참고하면, 패널 A는 각각의 PNPLA3 p.I148M 유전자형을 갖는 개체에서 HSD17B13 rs72613567와 ALT의 연관성을 나타내고, 패널 B는 각각의 PNPLA3 p.I148M 유전자형을 갖는 개체에서 HSD17B13 rs72613567과 AST의 연관성을 나타낸다. 효과 추정치(베타 및 95% CI)를 선형 회귀를 사용하여 계산하였고, 연령, 연령^2, 성별, BMI 및 혈통의 4개의 주성분에 대해서 조정하였다. 도 11(패널 A 내지 패널 F)은 PNPLA3 rs738409(p.I148M) 및 HSD17B13 rs72613567 유전자형에 의한 원시 및 잔차 ALT 수준을 나타낸다. 연령, 연령^2, 성별, BMI 및 4개의 주성분 1 내지 4에 대해서 조정된 선형 회귀에 의해서 잔차를 계산하였다. 도 12(패널 A 내지 패널 F)는 PNPLA3 rs738409(p.I148M) 및 HSD17B13 rs72613567 유전자형에 의한 원시 및 잔차 AST 수준을 나타낸다. 연령, 연령^2, 성별, BMI 및 4개의 주성분 1 내지 4에 대해서 조정된 선현 회귀에 의해서 잔차를 계산하였다. 도 13(패널 A 내지 패널 F)은 HSD17B13 스플라이스 변이체의 동형접합성 참조군(T/T), 이형접합성(T/TA) 및 동형접합성 대체(TA/TA) 보유자에서의 4개의 추가적인 신규 HSD17B13 전사물(E-H)의 mRNA 발현을 나타낸다. 유전자 모델에서 암호 영역은 적색으로 나타내고, 비번역 영역은 흑색으로 나타낸다. 전사물 E 및 H는 엑손 3과 4 사이에 추가 엑손을 함유한다. 전사물 F는 엑손 6에서 인트론 6까지 리드-쓰루(read-through)를 포함한다. 청색 화살표는 rs72613567로부터의 A 삽입을 나타낸다. 전사물 G는 엑손 2가 없다. 전사물 G 및 H에서의 별표는 엑손 6의 3'-단부에서의 G의 삽입을 도시하며, 이것은 단백질의 조기 절두로 이어진다(전사물 D와 유사함). 전사물은 박스 플롯에 도시된 바와 같이 HSD17B13 유전자형에 따라서 차별적으로 발현된다. mRNA 발현을 FPKM 단위로 나타낸다.

[0360] 이러한 데이터는, HSD17B13 rs72613567:TA 변이체가 변이체 PNPLA3 p.I148M 변이체에 의해서 지방증 간 질환에 유전적으로 취약한 개체에서 간 손상의 위험을 완화시킨다는 것을 시사한다. 이러한 발견은, HSD17B13의 치료 조절에 대한 중요한 하위 집단 - 변이체 PNPLA3 148M 대립유전자에 대해서 이형접합성이거나 동형접합성인 개체를 시사할 수 있다.

[0361] **실시예 3: 66개 인간 간 샘플에서 HSD17B13 및 PNPLA3의 유전자 발현 분석**

[0362] *엑손 변이체와 아스파테이트 및 알라닌 아미노트랜스퍼라제의 연관성*

[0363] 502,219개의 이대립유전자 단일 유전자 변이체를 DiscovEHR 연구("GHS 발견 코호트"; 도 1의 기본 인구통계학)로부터의 유럽 혈통의 46,544명의 개체에서 혈청 ALT 또는 AST 수준과의 연관성에 대해서 조사하였다. 19개의 유전자에서 총 35개의 변이체가 1.0×10^{-7} 미만의 P에서 ALT 또는 AST와 연관된 것을 발견하였다(도 14 및 도 2 참고). 도 14를 참고하면, GHS 발견 코호트에서 혈청 트랜스아미나제 수준과 단일 뉴클레오타이드 변이체 연관성의 맨허튼 플롯(좌측) 및 분위수-분위수 플롯(우측)을 나타낸다. 1.0×10^{-7} 미만의 P에서 알라닌 아미노트랜스퍼라제(ALT) 수준과 상당히 연관된 16개 유전자에서의 31개 변이체가 존재하였다(패널 A 참고). 1.0×10^{-7} 미만의 P에서 아스파테이트 아미노트랜스퍼라제(AST) 수준과 상당히 연관된 10개 유전자에서의 12개 변이체가 존재하였다(패널 B 참고). 모든 유의한 연관성을 도 2에 나타낸다. 3개의 개별 유럽 혈통 코호트의 반복 메타 분석에서 ALT 또는 AST와 유의하게 연관된, HSD17B13을 비롯한, 9개의 유전자(유전자 명칭에 의해서 본 명세서에 나타냄)에서 13개의 변이체가 존재하였다(도 3 참고). 엑솜-와이드 분위수-분위수 플롯 및 게놈 대조군 람다 값에 의해서 제시된 바와 같이, 연관성 시험을 널리 보정하였다.

[0364] 유럽 혈통 개체의 3개의 코호트에서 반복 연구를 수행하였다: 1) DiscovEHR로부터의 비만 수술 환자(n=2,644)("GHS 비만 수술 코호트"); 2) 댈러스 심장 연구로부터의 1,357명의 개체; 및 3) 펜 매디슨 바이오뱅크로부터의 8,526명의 개체. 반복 코호트의 메타 분석에서, 9개의 유전자에서의 13개의 변이체가 ALT 또는 AST의 혈청 수준과 유의하게 연관되었다(시험된 35개 변이체에 대해서 1.43×10^{-3} 미만의 P의 본페로니 유의성 역치)(도 3 참고). 이들은 증가된 트랜스아미나제 수준과 연관되어 있다고 이미 보고된 변이체, 예컨대, *PNPLA3*(Romeo et al., Nat. Genet., 2008, 40, 1461-5), *TM6SF2*(Kozlitina et al., Nat. Genet. 2014, 46, 352-6), *SERPINA1*(Brantly et al., Am. J. Med., 1988, 84, 13-31), *SAMM50*(Kitamoto et al., Hum. Genet., 2013, 132, 783-92) 및 *ERLIN1*(Feitosa et al., Atherosclerosis, 2013, 228, 175-80)을 포함하였다. *SERPINA1*은 알파-1-항트립신을 암호화하는데, 이의 기능성 결핍은 간 질환을 유발하며; *SAMM50*과의 연관성은 *PNPLA3*에서의 변이와의 연관 비평형을 통해서 매개되고, *ERLIN1*은 간 지방 침착에 연관되었다. 간 질환과 연관되어 있다고 이전에 보고되어 있지 않은 변이체를 또한 식별하였다. 이들은 각각 ALT 및 AST를 암호화하는 유전

자인 *GPT* 및 *GOT1*, 및 용질 운반체 패밀리(solute carrier family) 39 구성원 12를 암호화하는 *SLC39A12*에서의 몇몇 변이체를 포함하였다.

[0365] 17-베타 하이드록시스테로이드 데하이드로게나제 패밀리의 특징규명되지 않은 구성원인 하이드록시스테로이드 17-베타 데하이드로게나제 13을 암호화하는 유전자인 *HSD17B13*에서의 변이체와 ALT(발견 $P=4.2\times10^{-12}$, 반복 $P=1.7\times10^{-4}$) 및 AST(발견 $P=6.2\times10^{-10}$, 반복 $P=1.7\times10^{-4}$, 도 3 참고)의 감소된 수준 간의 재현 가능한 연관성이 또한 식별되었다. 연관된 변이체, rs72613567는 엑손 6(TA 대립유전자)의 공여자 스플라이스 부위에 인접한 아데닌의 삽입이고, GHS 발견 코호트에서 26.0%의 대립유전자 빈도를 가졌다. 이전에, Chambers 등은 ALT 수준과 연관된 4q22(rs6834314)의 인근 유전자좌를 식별하였고(Chambers et al., Nat. Genet., 2011, 43, 1131-8); rs72613567은 지금까지 트랜스아미나제 수준과 연관된다고 보고되지 않았다. *HSD17B13*은 동일한 유전자 패밀리의 또 다른 구성원인 *HSD17B11*의 30kb 상류에 존재한다. *HSD17B11*에서의 암호 또는 스플라이스 변이체와, 발견 코호트(도 6 참고) 또는 발견 코호트 및 3개의 반복 코호트의 공동 메타-분석에서의 트랜스아미나제 수준 사이에는 어떠한 엑솜-와이드 유의한 연관성도 관찰되지 않았다. 추가로, rs72613567과 *HSD17B11*에서의 변이체의 연관 비평형은 모든 혈통군에서 보통이었다(모든 혈통군의 *HSD17B11*에서의 모든 확인된 변이체에서 $r^2<0.4$). 전체적으로, 이러한 발견은 *HSD17B13*이 트랜스아미나제 수준과 기능적으로 관련될 가능성이 가장 높은 게놈 영역의 유전자임을 시사한다.

[0366] *엑손 변이체와 만성 간 질환의 임상적 진단의 연관성*

[0367] 발견 및 반복 코호트에서 발견되는 9개의 유전자 내의 13개의 트랜스아미나제-연관된 변이체와, 알코올성 간질환 및 비알코올성(비바이러스성) 간 질환을 비롯한 만성 간 질환, 뿐만 아니라 가장 진행된 형태의 만성 간 질환: 알코올성 간경변, 비알코올성 간경변 및 간세포 암종(HCC)을 비롯한 만성 간 질환 간의 관계를 또한 분석하였다. 시험된 13개의 변이체에 대해서 1.92×10^{-3} 미만의 P의 본페로니 유의성 역치를 사용하여, 5개의 유전자(*HSD17B13*, *SERPINA1*, *TM6SF2*, *PNPLA3* 및 *SAMM50*) 내의 6개의 변이체와 만성 간 질환 표현형 간의 유의한 연관성을 발견하였다(도 4 참고). *SERPINA1*, *TM6SF2*, *PNPLA3* 및 *SAMM50* 연관성은 이미 보고된 연관성을 확인한다. 발견 코호트에서, *HSD17B13* rs72613567:TA는 대립유전자 용량-의존적 방식으로 알코올성 간 질환 및 비알코올성 간 질환의 모든 EHR-유래된 카테고리의 더 낮은 승산과 연관되었다(도 15, 패널 A 참고): 알코올성 간 질환의 모든 카테고리, 이형접합성 승산비(OR_{het})(95% 신뢰 구간) 0.58(0.42 내지 0.80), 동형접합성 OR(OR_{hom}) 0.47(0.23 내지 0.97), 대립유전자 OR($OR_{대립유전자}$) 0.62(0.48 내지 0.81), $P=1.8\times10^{-4}$; 비알코올성 간 질환의 모든 카테고리, OR_{het} 0.83(0.75 내지 0.92), OR_{hom} 0.70(0.57 내지 0.87), $OR_{대립유전자}$ 0.84(0.78 내지 0.91), $P=1.3\times10^{-5}$. *HSD17B13* rs72613567:TA는 또한 알코올성 간경변 및 비알코올성 간경변의 더 낮은 승산과 연관되었는데, 각각 이형접합체 및 동형접합체에 대해서 알코올성 간경변의 더 낮은 승산은 42% 및 73%(OR_{het} 0.58(0.39 내지 0.86), OR_{hom} 0.27(0.09 내지 0.85), $OR_{대립유전자}$ 0.56(0.41 내지 0.78), $P=3.4\times10^{-4}$)이고, 각각 이형접합체 및 동형접합체에 대해서 비알코올성 간경변의 더 낮은 승산은 26% 및 49%이다(OR_{het} 0.74 (0.60 내지 0.93), OR_{hom} 0.51(0.31 내지 0.85), $OR_{대립유전자}$ 0.74(0.62 내지 0.88), $P=4.5\times10^{-4}$). *HSD17B13* rs72613567:TA는 또한 명목상 HCC의 더 낮은 승산과 연관되었다.

[0368] 이러한 발견은 다민족 델러스 간 연구(DLS) 및 델러스 소아 간 연구(DPLS)에서 확인 및 확장되었다(도 5 참고). DLS에서, TA 대립유전자는 대립유전자-용량 의존적인 방식으로 임의의 간 질환의 더 낮은 승산과 연관되었다(OR_{het} 0.74(0.57 내지 0.97), OR_{hom} 0.41(0.21 내지 0.83), $OR_{대립유전자}$ 0.70(0.5 내지 0.88), $P=1.8\times10^{-3}$, 도 15, 패널 B 참고). 후기 간경변 형태의 알코올성($OR_{대립유전자}$ 0.72(0.53 내지 0.99), $P=4.4\times10^{-2}$) 및 비알코올성($OR_{대립유전자}$ 0.65(0.40 내지 1.07), $P=9.0\times10^{-2}$) 간 질환과의 보호 연관성을 비롯한 EHR-유래된 간 질환 하위유형 전체에서 유사한 효과가 관찰되었다. 자가 보고된 민족성에 의해서 군집화된 개체의 하위세트 분석에서, 간 질환과의 연관성은 DLS의 히스패닉계 미국인(Hispanic Americans)에서 유의하였고(n=326 사례 및 722개 대조군, $OR_{대립유전자}$ 0.51(0.35 내지 0.74), $P=4.0\times10^{-4}$); 통계학적 유의성을 달성하지 않은 유사한 수치 경향이 또한 아프리카

계 미국인(African American)(n=33 사례 및 2,291개의 대조군, $OR_{대립유전자}$ 0.74(0.25 내지 2.47), P=0.67) 및 유럽계 미국인(European American)(n=158 사례 및 1,266개 대조군, $OR_{대립유전자}$ 0.87(0.65 내지 1.15), P=0.32) 하위세트에서 또한 주목되었다. DPLS에서, 히스패닉계 미국 소아 간 질환 환자 및 비만 대조군의 개별 연구에서, TA 대립유전자는 또한 간 질환의 더 낮은 승산과 연관되었다($OR_{대립유전자}$ 0.61(0.37 내지 0.99), $P=4.6\times10^{-2}$). 따라서, *HSD17B13* rs72613567:TA는, 3개의 독립적인 집단에서, 성인 및 유아에서, 간경변을 비롯한 만성 간 질환의 다양한 형태의 감소된 승산과 연관되었다.

[0369] 도 15를 참고하면, HSD17B13 rs72613567:TA는 도시된 바와 같이, 알코올성 간 질환 표현형 및 비알코올성 간 질환 표현형의 위험 감소와 연관된다. GHS 발견 코호트에서, HSD17B13 rs72613567은 대립유전자 용량-의존적인 방식으로 비알코올성 및 알코올성 간 질환, 간경변, 및 간세포 암종의 더 낮은 승산과 연관되었다(패널 A 참고). 승산비를 로지스틱 회귀를 사용하여 계산하였고, 연령, 연령^2, 성별, BMI 및 혈통의 주성분에 대해서 조정하였다. 이형접합성(Het OR) 보유자 및 동형접합성(Hom OR) 보유자에 대한 유전자형 승산비를 또한 나타낸다. 댈러스 간 연구에서, HSD17B13 rs72613567은 대립유전자 용량-의존적인 방식으로 임의의 간 질환의 더 낮은 승산과 연관되었다(패널 B 참고). 유사한 대립유전자 용량-의존적 효과를 간 질환 하위유형 전체에서 관찰하였다. 승산비를 로지스틱 회귀를 사용하여 계산하였고, 연령, 연령^2, 성별, BMI 및 자가 보고된 민족성에 대해서 조정하였다.

[0370] *PNPLA3 rs738409(p.I148M)와 HSD17B13 rs72613567 간의 유전적 상호작용*

[0371] 변이체 PNPLA3 p.I148M 변이체는 NAFLD에 대해서 가장 널리 검증된 유전적 위험 인자이며, 148M 대립유전자는 혈통에 따라서 개체의 5 내지 25%에서 동형접합성 상태로 존재한다. HSD17B13 rs72613567:TA가 PNPLA3 p.I148M과 연관된 간 손상의 위험을 변경시키는지의 여부를 이해하기 위해서, DiscovEHR에서 ALT, AST 및 만성 간 질환 표현형과의 연관성에서 두 변이체 간의 상호작용의 분석을 수행하였다. 이러한 분석은 모든 참여자에서 그리고 비만(체질량 지수[BMI] ≥ 30kg/㎡) 및 비-비만(BMI < 30kg/㎡) 하위 집단에서 수행하였다. ALT(상호작용에 대해서 $P=1.8\times10^{-3}$) 및 AST(상호작용에 대해서 $P=4.5\times10^{-3}$) 수준의 연관성 분석에서 HSD17B13 rs72613567:TA와 PNPLA3 p.I148M 사이에 명목상 유의한 상호작용이 존재하였고; 이러한 연관성은 비만 개체에서의 연관성에 의해서 주로 기인된다(도 9 참고). 이러한 분석에서, rs72613567:TA 대립유전자는 ALT 및 AST 증가와 PNPLA3 148M 대립유전자의 대립유전자 용량-의존적 연관성을 완화시켰다(도 16, 도 11 및 도 12참고). RNA 서열분석-기반 발현 분석은, *HSD17B13* rs72613567:TA가 대립유전자 용량-의존적인 방식으로 감소된 *PNPLA3* mRNA 발현과 연관되었음을 나타내었다(도 7 참고). 이러한 데이터는, HSD17B13 rs72613567:TA 변이체가 변이체 PNPLA3 p.I148M 변이체에 의해서 지방증 간 질환에 유전적으로 취약한 개체에서 간 손상의 위험을 완화시킨다는 것을 시사한다.

[0372] 도 16을 참고하면, HSD17B13 rs72613567:TA가 도시된 바와 같이, PNPLA3 p.I148M과 연관된 간 손상의 위험을 완화시킨다. 각각 PNPLA3 p.I148M 유전자형을 갖는 개체에서 HSD17B13 rs72613567과 ALT의 연관성(패널 A 참고). 각각 PNPLA3 p.I148M 유전자형을 갖는 개체에서 HSD17B13 rs72613567과 AST의 연관성(패널 B 참고). 효과 추정치(베타 및 95% CI)를 선형 회귀를 사용하여 계산하였고, 연령, 연령^2, 성별, BMI 및 혈통의 4개의 주성분에 대해서 조정하였다. ALT 및 AST 수준의 연관성 분석에서 HSD17B13 rs72613567:TA와 PNPLA3 p.I148M 간의 상호작용에 대한 P 값은 각각 $P=1.8\times10^{-3}$ 및 $P=4.5\times10^{-3}$이었다.

[0373] *HSD17B13 rs72613567:TA와 간 병리학의 연관성*

[0374] NAFLD는 상당한 염증의 증거가 없는 간 지방 축적(단순 지방증)에서부터 임상적으로 영향을 미치는 NASH까지의 범위의 질환 스펙트럼을 설명한다. *HSD17B13* rs72613567:TA와 EHR-유래된 간 질환 진단 코드 간의 연관성을 확인하고, 지방증에서 NASH로의 조직병리학적 진행과의 연관성을 추가로 이해하기 위해서, GHS 비만 수술 코호트에서 연관성의 시험을 수행하였다. 비만 수술 시에 간 생검에 의해서 평가된 전체 엑솜 서열분석된 개체 중 2,391명의 이러한 코호트에서, 총 555명(23%)의 개체는 지방증, 지방간염 또는 섬유증의 증거가 없었고("정상"), 830명(35%)은 단순 지방증을 가졌고, 1006명(42%)은 NASH를 가졌다. 유전자형에 의한 정상 간, 단순 지방증 및 NASH의 유병률을 비교하는 경우, 각각의 TA 대립유전자에 대해서, 정상 간의 유병률은 유전자형에 의해서 달라지는 것으로 보이지 않았지만(T/T, T/TA 및 TA/TA 보유자에 대해서 각각 23%, 24%, 및 23%, 비율 경향에 대한 카이 제곱 검정(Chi-squared test)에 의해서 P = 0.5), NASH의 유병률은 감소되었고(T/T, T/TA 및 TA/TA 보유자에 대해서 각각 45%, 40% 및 31%, $P = 1.6\times10^{-4}$), 단순 지방증의 유병률은 증가되는 것을 관찰하였다(T/T,

T/TA 및 TA/TA 보유자에 대해서 각각 33%, 35% 및 47%, P = 1.1×10^{-3})(도 17, 패널 A 참고). 지방증을 갖는 개체 중에서, TA 대립유전자는, 단순 지방증과 비교할 때, 대립유전자 용량-의존적인 방식으로, NASH의 유의하게 더 낮은 승산과 연관되었다(OR_{het} 0.87(0.71 내지 1.06), OR_{hom} 0.48(0.33 내지 0.70), $OR_{대립유전자}$ 0.77(0.66 내지 0.90), P=6.5×10^{-4})(도 17, 패널 B 참고). 전체적으로, 이러한 데이터는, 단순 지방증에서 더 후기 병기의 NASH 및 섬유증으로의 NAFLD 진행을 매개하는 데 있어서의 *HSD17B13*에 대한 역할을 시사한다.

[0375] 도 17을 참고하면, HSD17B13 rs72613567:TA는 도시된 바와 같이, 단순 지방증에서 지방간염 및 섬유증으로 진행되는 위험을 감소시키는 것과 연관된다. GHS 비만 수술 코호트로부터의 간 생검 진행된 2,391명의 개체에서 HSD17B13 rs72613567 유전자형에 따른 조직병리학적-특징규명된 간 질환의 유병률(패널 A 참고). 각각의 TA 대립유전자에 대해서, 정상 간의 유병률은 유전자형에 의해서 달라지는 것으로 보이지 않았지만(비율 경향에 대한 카이 제곱 검정에 의해서 P = 0.5), NASH의 유병률은 감소되었고(P = 1.6×10^{-4}), 단순 지방증의 유병률은 증가되었다(P = 1.1×10^{-3}). GHS 비만 수술 코호트에서, HSD17B13 rs72613567은 이형접합성 TA 보유자 및 동형접합성 TA 보유자 각각에서 13% 및 52%의 더 낮은 승산과 연관되었다(패널 B 참고). 승산비를 로지스틱 회귀를 사용하여 계산하였고, 연령, 연령^2, 성별, BMI 및 혈통의 주성분에 대해서 조정하였다. 이형접합성(Het OR) 보유자 및 동형접합성(Hom OR) 보유자에 대한 유전자형 승산비를 또한 나타낸다.

[0376] *HSD17B13 mRNA 및 HSD17B13 단백질 발현에 대한 rs72613567:TA의 효과*

[0377] 유전자의 공지된 전사물 및 신규 전사물의 발현에 대한 *HSD17B13* rs72613567:TA 대립유전자의 효과를 조사하였다. RNA 서열분석을 사용하여 *HSD17B13* rs72613567 스플라이스 변이체의 22명의 T/T 동형접합성 보유자, 30명의 T/TA 이형접합성 보유자 및 17명의 TA/TA 동형접합성 보유자로부터의 조직학적 정상 간 샘플에서 *HSD17B13* mRNA 발현을 평가하였다. 2개의 *HSD17B13* 전사물, A 및 B에 더하여, 2개의 신규 전사물: 엑손 6이 결여된 전사물 C 및 엑손 6의 3' 단부에 구아닌 뉴클레오타이드의 삽입을 함유한 전사물 D를 식별하였는데, 이것은 단백질의 조기 절두를 초래한다고 예측될 것이다. 4개의 추가 전사물(E 내지 H)은 매우 낮은 수준으로 발현되었다(도 13 참고). 전사물을 RT-PCR 및 생어 서열분석에 의해서 검증하였다(데이터 나타내지 않음). D 전사물을 또한 긴 판독물 cDNA 서열분석을 사용하여 검증하였다. 이러한 전사물의 발현 수준은 *HSD17B13* rs72613567 유전자형에 따라서 달라졌고; 각각의 TA 대립유전자에 대해서, 전사물 A 및 B의 수준은 감소되었지만, 전사물 C 및 D의 수준은 대립유전자 용량-의존적인 방식으로 증가하였다(도 18, 패널 A 및 패널 B). 전장 300개의 아미노산 단백질을 암호화하는 전사물 A는 T/T 동형접합체에서 우세한 전사물인 반면, 조기 절두된 단백질을 암호화하는 전사물 D는 TA/TA 동형접합체에서 우세한 전사물이었다. 인간 간 생검 조직에서, 절두된 아이소폼 D 단백질은 이형접합체 및 TA/TA 동형접합체에서 최소한으로 존재하였고, 아이소폼 A 단백질 존재비(abundance)는 대립유전자 용량-의존적인 방식으로 감소되었다(도 18, 패널 C 및 패널 D 참고). 이러한 데이터는 mRNA 스플라이스를 변경하여, 인간 간에서 실질적으로 감소된 발현을 갖는 절두된 단백질 형태의 합성을 초래하는 *HSD17B13* rs72613567과 일치한다.

[0378] 도 18을 참고하면, 신규 HSD17B13 전사물의 발현, 세포 이하 국지화 및 효소 활성도를 나타낸다. HSD17B13 rs72613567 스플라이스 변이체의 동형접합성 참조군(T/T), 이형접합성(T/TA) 및 동형접합성 대체(TA/TA) 보유자에서의 HSD17B13 전사물 A 및 D의 발현(패널 A 참고). HSD17B13 유전자 내의 암호 영역을 적색으로 나타내고, 비번역 영역을 두꺼운 흑색선으로 나타내고, 인트론을 얇은 흑색선으로 나타낸다. 전사물 D에서의 별표는 rs72613567로부터의 A 삽입을 나타낸다. mRNA 발현은 FPKM 단위로 나타낸다(단편/킬로베이스의 전사물/백만개의 맵핑된 판독물 단편). HSD17B13 전사물 A 및 D를 과발현하는 HepG2 세포로부터의 웨스턴 블롯. HSD17B13 전사물 D는 HSD17B13 전사물 A에 비해서 더 낮은 분자량을 갖는 절두된 단백질로 번역되었다(패널 B 참고). 새로 동결된 인간 간 및 HEK293 세포 샘플로부터의 HSD17B13 웨스턴 블롯(패널 C 참고). 인간 간 샘플은 HSD17B13 rs72613567 스플라이스 변이체의 동형접합성 참조군(T/T), 이형접합성(T/TA) 및 동형접합성 대체(TA/TA) 보유자로부터 기원한다. 세포 샘플은 비-태그화된 HSD17B13 전사물 A 및 D를 과발현하는 HEK293 세포로부터 기원한다. HSD17B13 전사물 D는 HSD17B13 아이소(Iso)A에 비해서 더 낮은 분자량을 갖는 절두된 단백질 아이소D로 번역되었다. HSD17B13 아이소D 단백질 수준은 인간 간(좌측) 및 세포(우측) 샘플로부터의 아이소A 단백질 수준보다 더 낮았다(패널 D 참고). 액틴에 정규화된 단백질 수준을 막대 칼럼으로 나타낸다; ** P<0.001, *P<0.05. HSD17B13 아이소폼 A 및 D 둘 다는 지방 소립막에 국지화되었다(패널 E 참고). HSD17B13 전사물 A 또는 D를 안정적으로 과발현하는 HepG2를, 지방 소립을 나타내기 위해서 BODIPY로 표지하고, HSD17B13 국지화를 나타내기 위해서 항-Myc로 표지하였다. 모든 도면을 동일한 정도로 확대한다. 척도 막대는 10㎛를 나타낸다. 삽도는 본래 영상의 4

× 확대를 나타낸다. 17-베타 에스트라다이올(에스트라다이올), 류코트라이엔 B4(LTB4) 및 13-하이드록시옥타데카다이엔산(13(S)-HODE에 대한 HSD17B13 아이소폼 A 및 D의 효소 활성도(패널 F 참고). HSD17B13 아이소폼 D는 아이소폼 A에 대한 상응하는 값의 10% 미만의 효소 활성도를 나타낸다. G, HEK293 세포에서 과발현되는 경우 HSD17B13 아이소폼 D는 배양 배지에서 측정되는 경우 에스트라다이올(기질)에서 에스트론(산물)으로의 전환율이 크지 않았지만, 과발현된 HSD17B13 아이소폼 A는 강한 전환을 나타내었다.

[0379] *인간 간 세포에서의 HSD17B13의 발현*

[0380] HSD17B13은 간에서 주로 발현되고(Liu et al., Acta Biochim. Pol., 2007, 54, 213-8), 여기서 그것은 지방 소립에 국지화되는데(Su et al., Proc. Natl. Acad. Sci. USA, 2014, 111, 11437-42), 이는 지방간 질환의 발병기전에서의 역할과 일치한다. HSD171B3의 발현 및 이의 국지화를, *HSD17B13* 전사물 A 및 D를 발현하는 렌티바이러스가 안정적으로 형질도입된 불멸화된 인간 간 세포주에서 평가하였다. HSD17B13 아이소폼 A는 BODIPY-표지된 지방 소립 주변의 막 상에서 주로 검출되었다(도 18, 패널 E 참고). 유사한 세포 이하 국지화가 지방 소립 표면에서 HSD17B13 아이소폼 D에 대해서 관찰되었다(도 18, 패널 F 참고).

[0381] *시험관내 및 세포 모델에서 HSD17B13 활성도에 대한 rs72613567:TA의 효과의 효과*

[0382] rs72613567:TA로 인한 HSD17B13 단백질의 조기 절두의 기능성 결과를 이해하기 위해서, 아이소폼 A 및 D의 효소 활성도를 재조합 단백질을 사용하여 시험관내에서 평가하였다. 300개 초과의 추정 기질을 조사하였고, 이들 중 에스트라다이올, 류코트라이엔 B4, 및 13-하이드록시옥타데카다이엔산은 HSD17B13에 의해서 효소적으로 전환되어, 하이드록실의 케톤기로의 산화를 초래하였다. HSD17B13 아이소폼 D는 3개의 기질에 대해서 상당히 감소된 활성도를 나타내었다(도 18, 패널 F 참고).

[0383] GFP 대조군과 비교할 때, HSD17B13 전사물 A 과발현 세포는 세포 배양 배지 중에서 더 낮은 농도의 에스트라다이올 및 더 높은 농도의 에스트론을 가졌는데, 이는 에스트라다이올에 대한 효소 활성도를 시사한다(도 18, 패널 G 참고). HSD17B13 전사물 D 과발현 세포는 GFP 대조군 세포와 유사한 에스트론/에스트라다이올의 비를 가졌는데, 이는 HSD17B13 전사물 D가 유의한 기능 상실을 가짐을 시사한다. 질량 스펙 분석은, 소모된 에스트라다이올에 비해서 에스트론의 더 적은 축적을 고려할 때 에스트론의 하이드록시에스트론 및 다른 산물로의 신속한 전환을 나타내었다.

[0384] 대규모 엑솜 서열분석을 통해서, *HSD17B13*에서의 스플라이스 변이체와, 감소된 혈청 트랜스아미나제 수준, 뿐만 아니라 후기 간경변 형태의 간 질환 및 HCC를 비롯한, 비알코올성 및 알코올성 형태의 간 질환 사이의 신규 연관성을 식별하였다. 본 발명자들이 아는 한, 이것은 간 질환과의 보호 연관성을 갖는 단백질-변경 변이체의 최초 보고이다. *HSD17B13* rs72613567:TA 대립유전자는 단순 지방증과는 연관되지 않았고, NASH에 대한 진행 위험 감소와 연관되었다. 몇몇 상이한 간 질환 카테고리 및 민족성에 걸친 4개의 독립적인 코호트(DiscovEHR, DiscovEHR에서의 독립적인 비만 수술 코호트, DLS 및 DPLS)에서의 용량-의존적인 보호 연관성의 일관성은, 보고된 *HSD17B13* 변이체가 만성 간 질환의 보다 임상적으로 진전된 병기로 진행하는 것을 보호한다는 개념을 뒷받침한다. 관찰된 대립유전자 용량-의존성은 또한 HSD17B13 기능의 보다 강한 조절이 질환 위험 및 진행에 보다 강한 효과를 초래할 수 있다는 것을 입증한다. HSD17B13 rs72613567:TA 대립유전자는 또한 변이체 PNPLA3 p.I148M 변이체에 의해서 지방증 간 질환에 유전적으로 취약한 개체에서 간 손상의 위험을 완화시켰다. 이러한 발견은, HSD17B13의 치료 조절에 대한 중요한 하위 집단 - 변이체 PNPLA3 148M 대립유전자에 대해서 이형접합성이거나 동형접합성인 개체를 시사할 수 있다.

[0385] 본 명세서에 기재된 연관성 발견은 주로 BMI가 높은 유럽계 및 히스패닉계 미국인에서의 관찰에 기초하였다. *HSD17B13*은 *HSD17B13*과 높은 서열 유사성을 갖지만 더 넓은 조직 분포를 갖는 동일한 유전자 패밀리의 구성원인 *HSD17B11*과 근접한다. 종합하면, 본 명세서에 제시된 데이터는 *HSD17B13*이 인간에서 지방간 질환의 예방 및 치료를 위한 잠재적인 치료 표적이라는 지위를 뒷받침한다. 본 명세서에 제시된 데이터는, HSD17B13의 표적화가 간 질환이 지방증에서 더 후기 병기의 NASH, 섬유증 및 간경변(이것은 심각한 이환율 및 사망률과 연관되며, 현재 효과적인 치료법이 없음)으로 진행하는 것을 감소시킬 수 있다는 것을 나타낸다.

[0386] **실시예 4: HSD17B13 rs72613567:TA는 PNPLA3 I148M과 연관된 알코올성 간 질환 및 비알코올성 간 질환의 위험을 완화시킨다**

[0387] 지방증, 소엽 염증 또는 섬유증이 없는 대조군 공여자로부터의 29,928개의 인간 간 샘플의 HSD17B13 및 PNPLA3 유전자형을, 알코올성 간 질환을 갖는 환자로부터의 190개의 샘플 또는 비알코올성 간 질환을 갖는 1857명의 환자와 비교함으로써 HSD17B13 및 PNPLA3 유전자형과 간 질환의 연관성을 분석하였다. 95% 신뢰 구간을 갖는

HSD17B13 및 PNPLA3 유전자형의 각각의 조합에 대해서 (질환을 갖는 군의 발생률)/(대조군의 발생률)의 식에 의해서 승산비를 계산하였다. 도 19를 참고하면, 패널 A는 각각의 PNPLA3 p.I148M 유전자형을 갖는 개체에서 HSD17B13 rs72613567와 알코올성 간 질환의 연관성을 나타내고, 패널 B는 각각의 PNPLA3 p.I148M 유전자형을 갖는 개체에서 HSD17B13 rs72613567과 비알코올성 간 질환의 연관성을 나타낸다. 이러한 데이터는, PNPLA3 p.I148M이 알코올성 간 질환 및 비알코올성 간 질환 둘 다의 더 높은 발생률과 용량-의존적인 방식으로 연관됨을 입증한다. HSD17B13 rs72613567:TA 유전자형은 알코올성 간 질환 및 비알코올성 간 질환 둘 다에 대한 감소된 위험과 대립유전자 용량-의존적인 방식으로 연관되었다.

도면

도면1

특징	GHS 발견 코호트 (N = 46,544)	GHS 비만 수술 코호트 (N = 2,644)	밸러스 심장 연구 (N = 1,357)	펜 메디슨 바이오뱅크 (N = 8,526)
연령(세) - 중위 (IQR)	63 (50 - 74)	53 (44 - 61)	46 (38 - 54)	68 (60 - 76)
여성 - 수 (%)	26,875 (58)	2,119 (80)	724 (53)	3,242 (38)
체질량 지수 - 중위 (IQR)	30 (25 - 45)	47 (42 - 54)	28 (25 - 32)	30 (25 - 32)
트랜스아미나제 수준(U/L) – 중위 (IQR)				
알라닌 아미노트랜스퍼라제 (ALT)	22.0 (17.0 - 29.0)	23.0 (17.5 - 29.5)	20.0 (15.0 - 27.0)	22.0 (17.0 - 30.0)
아스파테이트 아미노트랜스퍼라제(AST)	23.0 (20.0 - 27.5)	23.0 (20.0 - 27.0)	21.0 (18.0 - 25.0)	24.0 (20.0 - 30.5)
간 질환의 존재 – N (%)				
알코올성 간 질환	197 (0.4)	7 (0.3)	-	-
알코올성 간경변	130 (0.3)	3 (0.1)	-	-
비알코올성(비바이러스성) 간 질환	1,938 (4.2)	1,543 (58.4)	-	-
비알코올성 간경변	382 (0.8)	24 (0.9)	-	-
간세포 암종	76 (0.2)	1 (0.04)	-	-
간질환 없음	30,628 (65.8)	1 (0.04)	-	-

도면2a

특질	Chr	BP	Ref	Alt	rsID	유전자	주석	AA 치환	베타 (SE)	P	AAF	N
ALT	1	220970028	A	G	rs2642438	*MARC1*	미스센스	p.Thr165Ala	0.008 (0.001)	4.67E-08	0.7067	40,834
	4	88231392	T	TA	rs72613567	*HSD17B13*	스플라이스 공여자		-0.009 (0.001)	4.16E-12	0.2634	40,834
	8	144997604	C	T	rs371119003	*PLEC*	미스센스	p.Ala2302Thr	-0.160 (0.026)	1.30E-09	0.0005	40,833
	8	145008502	G	A		*PLEC*	미스센스	p.Arg522Cys	-0.268 (0.032)	3.26E-17	0.0003	40,834
	8	145692918	G	A	rs35968570	*KIFC2*	미스센스	p.Glu174Lys	-0.033 (0.005)	1.40E-11	0.0139	40,834
	8	145730072	G	A	rs143408057	*GPT*	미스센스	p.Arg83His	-0.314 (0.036)	3.28E-18	0.0003	40,834
	8	145730161	C	T	rs201815297	*GPT*	미스센스	p.Ala87Val	-0.224 (0.014)	6.28E-59	0.0018	40,834
	8	145730221	G	A	rs112574791	*GPT*	미스센스	p.Arg107Lys	-0.033 (0.005)	4.25E-11	0.0136	40,834
	8	145731636	T	G	rs145155876	*GPT*	중단 얻음	p.Tyr326*	-0.235 (0.031)	1.76E-14	0.0004	40,814
	8	145732114	G	C	rs141505249	*GPT*	미스센스	p.Glu430Gln	-0.224 (0.013)	8.84E-64	0.0019	40,795
	8	145732151	G	A	rs143462595	*GPT*	미스센스	p.Arg442His	-0.077 (0.013)	1.18E-09	0.0021	40,826
	8	145732180	G	C	rs147998249	*GPT*	미스센스	p.Val452Leu	-0.225 (0.013)	8.19E-65	0.0019	40,833
	8	145732305	G	GC		*GPT*	프레임시프트	p.Glu475fs	-0.271 (0.031)	1.00E-18	0.0004	40,834
	8	145748532	A	G	rs567402720	*LRRC24*	미스센스	p.Leu290Ser	-0.185 (0.028)	3.42E-11	0.0004	40,813
	9	117122202	C	T	rs3748177	*AKNA*	동의	p.Glu755Glu	-0.007 (0.001)	9.51E-09	0.5232	40,834
	9	117124731	G	A	rs3748176	*AKNA*	미스센스	p.Pro624Leu	-0.007 (0.001)	4.31E-09	0.5230	40,832
	10	101595996	T	A	rs17222723	*ABCC2*	미스센스	p.Val1188Glu	-0.015 (0.003)	2.97E-08	0.0608	40,834
	10	101606861	G	T	rs1137968	*ABCC2*	동의	p.Val1430Val	-0.015 (0.003)	2.71E-08	0.0608	40,834
	10	101610533	C	T	rs8187707	*ABCC2*	동의	p.His1496His	-0.015 (0.003)	2.77E-08	0.0608	40,834
	10	101611294	G	A	rs8187710	*ABCC2*	미스센스	p.Cys1515Tyr	-0.015 (0.003)	2.15E-08	0.0611	40,834
	10	101912064	T	C	rs2862954	*ERLIN1*	미스센스	p.Ile291Val	-0.012 (0.001)	2.43E-21	0.4755	40,834

회색 음영은 ALT 및 AST 둘 다와 엑솜-와이드-유의한 연관성을 갖는 변이체를 나타냄

약어 : AAF, 대체 대립유전자 빈도 ; Alt,대체 대립유전자; ALT,알라닌 아미노트랜스퍼라제;
AST, 아스파테이트 아미노트랜스퍼라제 ; Ref, 참조 대립유전자;SE, 표준오차;

									N			평균 ALT 또는 AST수준 (U/L)		
특질	Chr	BP	Ref	Alt	rsID	유전자	주석	AA 치환	Ref/Ref	Ref/Alt	Alt/Alt	Ref/Ref	Ref/Alt	Alt/Alt
ALT	1	2.2E+08	A	G	rs2642438	*MARC1*	미스센스	p.Thr165Ala	3,515	17,262	20,637	23.88	24.52	24.92
	4	8.8E+07	T	TA	rs72613567	*HSD17B13*	스플라이스 공여자		22,441	16,130	2,843	25.02	24.26	24.1
	8	1.4E+08	C	T	rs371119003	*PLEC*	미스센스	p.Ala2302Thr	41,373	40	0	24.67	18.1	NA
	8	1.5E+08	G	A		*PLEC*	미스센스	p.Arg522Cys	41,387	27	0	24.67	13.8	NA
	8	1.5E+08	G	A	rs35968570	*KIFC2*	미스센스	p.Glu174Lys	40,271	1,133	10	24.67	12.07	NA
	8	1.5E+08	G	A	rs143408057	*GPT*	미스센스	p.Arg83His	41,393	21	0	24.67	12.07	NA
	8	1.5E+08	C	T	rs201815297	*GPT*	미스센스	p.Ala87Val	41,270	144	0	24.7	14.68	NA
	8	1.5E+08	G	A	rs112574791	*GPT*	미스센스	p.Arg107Lys	40,293	1,111	10	24.71	23.09	18.35
	8	1.5E+08	T	G	rs145155876	*GPT*	중단 얻음	p.Tyr326*	41,364	30	0	24.67	14.07	NA
	8	1.5E+08	G	C	rs141505249	*GPT*	미스센스	p.Glu430Gln	41,223	150	2	24.7	14.48	13.75
	8	1.5E+08	G	A	rs143462595	*GPT*	미스센스	p.Arg442His	41,232	174	0	24.68	20.87	NA
	8	1.5E+08	G	C	rs147998249	*GPT*	미스센스	p.Val452Leu	41,254	159	0	24.7	14.74	NA
	8	1.5E+08	G	GC		*GPT*	프레임시프트	p.Glu475fs	41,385	29	0	24.67	14.24	NA
	8	1.5E+08	A	G	rs567402720	*LRRC24*	미스센스	p.Leu290Ser	41,358	35	0	24.67	17.71	NA
	9	1.2E+08	C	T	rs3748177	*AKNA*	동의	p.Glu755Glu	9,414	20,645	11,355	25.12	24.72	24.18
	9	1.2E+08	G	A	rs3748176	*AKNA*	미스센스	p.Pro624Leu	9,427	20,634	11,351	25.12	24.73	24.17
	10	1E+08	T	A	rs17222723	*ABCC2*	미스센스	p.Val1188Glu	36,543	4,704	167	24.77	23.97	22.12
	10	1E+08	G	T	rs1137968	*ABCC2*	동의	p.Val1430Val	36,543	4,704	167	24.77	23.97	22.04
	10	1E+08	C	T	rs8187707	*ABCC2*	동의	p.His1496His	36,542	4,706	166	24.77	23.97	22.03
	10	1E+08	G	A	rs8187710	*ABCC2*	미스센스	p.Cys1515Tyr	36,519	4,726	169	24.77	23.97	21.99
	10	1E+08	T	C	rs2862954	*ERLIN1*	미스센스	p.Ile291Val	11,318	20,819	9,277	25.32	24.71	23.77

회색 음영은 ALT 및 AST 둘 다와 엑솜-와이드-유의한 연관성을 갖는 변이체를 나타냄

약어 : AAF, 대체 대립유전자 빈도 ; Alt,대체 대립유전자; ALT,알라닌 아미노트랜스퍼라제;
AST, 아스파테이트 아미노트랜스퍼라제 ; Ref, 참조 대립유전자;SE, 표준오차;

도면2b

특질	Chr	BP	Ref	Alt	rsID	유전자	주석	AA 치환	베타 (SE)	P	AAF	N
	10	101977883	C	T	rs2230804	*CHUK*	미스센스	p.Val268Ile	-0.009 (0.001)	1.93E-13	0.5072	40,834
	10	113917085	T	A	rs2254537	*GPAM*	동의	p.Pro681Pro	-0.008 (0.001)	4.61E-10	0.7073	40,834
	10	113940329	T	C	rs2792751	*GPAM*	미스센스	p.Ile43Val	-0.008 (0.001)	2.54E-10	0.7097	40,832
	14	94844947	C	T	rs28929474	*SERPINA1*	미스센스	p.Glu366Lys	0.042 (0.005)	9.28E-21	0.0171	40,834
	19	19379549	C	T	rs58542926	*TM6SF2*	미스센스	p.Glu167Lys	0.014 (0.002)	4.76E-09	0.0759	40,833
	22	44324727	C	G	rs738409	*PNPLA3*	미스센스	p.Ile148Met	0.023 (0.002)	1.34E-50	0.2351	40,834
	22	44324730	C	T	rs738408	*PNPLA3*	동의	p.Pro149Pro	0.023 (0.002)	1.11E-50	0.2349	40,834
	22	44342116	A	G	rs2294918	*PNPLA3*	미스센스	p.Lys434Glu	0.007 (0.001)	8.26E-08	0.5986	40,832
	22	44368122	A	G	rs3761472	*SAMM50*	미스센스	p.Asp110Gly	0.019 (0.002)	8.85E-30	0.1682	40,833
	22	44395451	T	C	rs1007863	*PARVB*	미스센스	p.Trp37Arg	0.011 (0.001)	7.98E-16	0.3963	40,834
AST	4	88231392	T	TA	rs72613567	*HSD17B13*	스플라이스 공여자		-0.005 (0.001)	6.24E-10	0.2638	40,193
	10	18242311	A	G	rs10764176	*SLC39A12*	미스센스	p.Ser36Gly	-0.006 (0.001)	1.09E-10	0.2881	40,193
	10	101157378	CGTT	C		*GOT1*	인프레임 인델	p.Asn389del	-0.221 (0.024)	1.96E-20	0.0002	40,193
	10	101165533	G	C	rs374966349	*GOT1*	미스센스	p.Gln208Glu	0.271 (0.027)	2.43E-24	0.0002	40,193
	10	101912064	T	C	rs2862954	*ERLIN1*	미스센스	p.Ile291Val	-0.005 (0.001)	4.82E-09	0.4754	40,193
	11	22271870	A	T	rs7481951	*ANO5*	미스센스	p.Leu322Phe	0.004 (0.001)	9.61E-08	0.5833	40,162
	14	94844947	C	T	rs28929474	*SERPINA1*	미스센스	p.Glu366Lys	0.027 (0.003)	2.44E-20	0.0172	40,193
	19	19379549	C	T	rs58542926	*TM6SF2*	미스센스	p.Glu167Lys	0.008 (0.002)	6.54E-08	0.0760	40,192
	22	44324727	C	G	rs738409	*PNPLA3*	미스센스	p.Ile148Met	0.014 (0.001)	8.31E-46	0.2343	40,193
	22	44324730	C	T	rs738408	*PNPLA3*	동의	p.Pro149Pro	0.014 (0.001)	8.93E-46	0.2341	40,193
	22	44368122	A	G	rs3761472	*SAMM50*	미스센스	p.Asp110Gly	0.011 (0.001)	1.22E-22	0.1680	40,192
	22	44395451	T	C	rs1007863	*PARVB*	미스센스	p.Trp37Arg	0.006 (0.001)	1.31E-13	0.3961	40,193

회색 음영은 ALT 및 AST 둘 다와 엑솜-와이드-유의한 연관성을 갖는 변이체를 나타냄

약어 : AAF, 대체 대립유전자 빈도 ; Alt,대체 대립유전자; ALT,알라닌 아미노트랜스퍼라제;
AST, 아스파테이트 아미노트랜스퍼라제 ; Ref, 참조 대립유전자 ; SE, 표준오차 ;

									N			평균 ALT 또는 AST수준 (U/L)		
특질	Chr	BP	Ref	Alt	rsID	유전자	주석	AA 치환	Ref/Ref	Ref/Alt	Alt/Alt	Ref/Ref	Ref/Alt	Alt/Alt
	10	101977883	C	T	rs2230804	*CHUK*	미스센스	p.Val268Ile	10,048	20,733	10,633	25.18	24.75	24.01
	10	113917085	T	A	rs2254537	*GPAM*	동의	p.Pro681Pro	3,627	16,984	20,803	25	24.97	24.36
	10	113940329	T	C	rs2792751	*GPAM*	미스센스	p.Ile43Val	3,567	16,910	20,935	25	24.98	24.35
	14	94844947	C	T	rs28929474	*SERPINA1*	미스센스	p.Glu366Lys	40,006	1,399	9	24.58	26.91	43.89
	19	19379549	C	T	rs58542926	*TM6SF2*	미스센스	p.Glu167Lys	35,388	5,780	245	24.52	25.46	26.84
	22	44324727	C	G	rs738409	*PNPLA3*	미스센스	p.Ile148Met	24,257	14,837	2,320	24.06	24.99	28.91
	22	44324730	C	T	rs738408	*PNPLA3*	동의	p.Pro149Pro	24,273	14,824	2,317	24.06	24.98	28.92
	22	44342116	A	G	rs2294918	*PNPLA3*	미스센스	p.Lys434Glu	6,691	19,833	14,888	24.15	24.47	25.15
	22	44368122	A	G	rs3761472	*SAMM50*	미스센스	p.Asp110Gly	28,626	11,618	1,169	24.23	25.36	28.45
	22	44395451	T	C	rs1007863	*PARVB*	미스센스	p.Trp37Arg	15,036	19,920	6,458	24.15	24.6	26.09
AST	4	88231392	T	TA	rs72613567	*HSD17B13*	스플라이스 공여자		22,068	15,870	2,815	24.47	24.1	23.96
	10	18242311	A	G	rs10764176	*SLC39A12*	미스센스	p.Ser36Gly	20,645	16,738	3,370	24.47	24.15	23.85
	10	101157378	CGTT	C		*GOT1*	인프레임 인델	p.Asn389del	40,733	20	0	24.29	14.7	NA
	10	101165533	G	C	rs374966349	*GOT1*	미스센스	p.Gln208Glu	40,736	17	0	24.28	44.5	NA
	10	101912064	T	C	rs2862954	*ERLIN1*	미스센스	p.Ile291Val	11,138	20,486	9,129	24.59	24.26	23.99
	11	22271870	A	T	rs7481951	*ANO5*	미스센스	p.Leu322Phe	7,123	19,686	13,913	24.03	24.22	24.53
	14	94844947	C	T	rs28929474	*SERPINA1*	미스센스	p.Glu366Lys	39,361	1,384	8	24.24	25.76	34.5
	19	19379549	C	T	rs58542926	*TM6SF2*	미스센스	p.Glu167Lys	34,811	5,698	243	24.21	24.74	25.43
	22	44324727	C	G	rs738409	*PNPLA3*	미스센스	p.Ile148Met	23,889	14,622	2,242	23.96	24.48	26.62
	22	44324730	C	T	rs738408	*PNPLA3*	동의	p.Pro149Pro	23,905	14,609	2,239	23.96	24.47	26.63
	22	44368122	A	G	rs3761472	*SAMM50*	미스센스	p.Asp110Gly	28,170	11,450	1,132	24.07	24.64	26.24
	22	44395451	T	C	rs1007863	*PARVB*	미스센스	p.Trp37Arg	14,761	19,678	6,314	24.02	24.23	25.1

회색 음영은 ALT 및 AST 둘 다와 엑솜-와이드-유의한 연관성을 갖는 변이체를 나타냄

약어 : AAF, 대체 대립유전자 빈도 ; Alt,대체 대립유전자; ALT,알라닌 아미노트랜스퍼라제;
AST, 아스파테이트 아미노트랜스퍼라제 ; Ref, 참조 대립유전자 ; SE, 표준오차 ;

도면3a

특질	Chr	BP	Ref	Alt	RSID	유전자	주석	AA 치환	GHS 발견 코호트		
									메타 (SE)	P 발견	N
ALT	1	220970028	A	G	rs2642438	*MARC1*	미스센스	p.Thr165Ala	0.008 (0.001)	4.67E-08	40834
	4	88231392	T	TA	rs72613567	*HSD17B13*	스플라이스 공여자		-0.009 (0.001)	4.16E-12	40834
	8	144997604	C	T	rs371119003	*PLEC*	미스센스	p.Ala2302Thr	-0.160 (0.026)	1.30E-09	40833
	8	145008502	G	A		*PLEC*	미스센스	p.Arg522Cys	-0.268 (0.032)	3.26E-17	40834
	8	145692918	G	A	rs35968570	*KIFC2*	미스센스	p.Glu174Lys	-0.033 (0.005)	1.40E-11	40834
	8	145730072	G	A	rs143408057	*GPT*	미스센스	p.Arg83His	-0.314 (0.036)	3.28E-18	40834
	8	145730161	C	T	rs201815297	*GPT*	미스센스	p.Ala87Val	-0.224 (0.014)	6.28E-59	40834
	8	145730221	G	A	rs112574791	*GPT*	미스센스	p.Arg107Lys	-0.033 (0.005)	4.25E-11	40834
	8	145731636	T	G	rs145155876	*GPT*	중단 얻음	p.Tyr326*	-0.235 (0.031)	1.76E-14	40814
	8	145732114	G	C	rs141505249	*GPT*	미스센스	p.Glu430Gln	-0.224 (0.013)	8.84E-64	40795
	8	145732151	G	A	rs143462595	*GPT*	미스센스	p.Arg442His	-0.077 (0.013)	1.18E-09	40826
	8	145732180	G	C	rs147998249	*GPT*	미스센스	p.Val452Leu	-0.225 (0.013)	8.19E-65	40833
	8	145732305	G	GC		*GPT*	프레임시프트	p.Glu475fs	-0.271 (0.031)	1.00E-18	40834
	8	145748532	A	G	rs567402720	*LRRC24*	미스센스	p.Leu290Ser	-0.185 (0.028)	3.42E-11	40813
	9	117122202	C	T	rs3748177	*AKNA*	동의	p.Glu755Glu	-0.007 (0.001)	9.51E-09	40834
	9	117124731	G	A	rs3748176	*AKNA*	미스센스	p.Pro624Leu	-0.007 (0.001)	4.31E-09	40832
	10	101595996	T	A	rs17222723	*ABCC2*	미스센스	p.Val1188Glu	-0.015 (0.003)	2.97E-08	40834

회색 음영은 P < 1.43x10-3 의 본페로니 유의성 역치를 역치를 충족시키는 P-값을 나타냄

* 반복 메타-분석은 3개의 반복 코호트: GHS 비만 수술 코호트, 델러스 심장 연구 및 펜 메디슨 바이오뱅크를 포함함

* 공동 메타-분석은 발견 코호트 및 3개의 반복 코호트: GHS 발견 코호트, GHS 비만 수술 코호트, 델러스 심장 연구 및 펜 메디슨 바이오뱅크를 포함함

약어 : AAF, 대체 대립유전자 빈도; Alt, 대체 대립유전자; ALT, 알라닌 아미노트랜스퍼라제; AST, 아스파테이트 아미노트랜스퍼타제; Ref, 참조 대립유전자; SE, 표준 오차;

특질	Chr	BP	Ref	Alt	RSID	유전자	주석	AA 치환	GHS 발견 코호트		
									메타 (SE)	P 발견	N
	10	101606861	G	T	rs113796	*ABCC2*	동의	p.Val1430Val	-0.015 (0.003)	2.71E-08	40834
	10	101610533	C	T	rs818770	*ABCC2*	동의	p.His1496His	-0.015 (0.003)	2.77E-08	40834
	10	101611294	G	A	rs818771	*ABCC2*	미스센스	p.Cys1515Tyr	-0.015 (0.003)	2.15E-08	40834
	10	101912064	T	C	rs286295	*ERLIN1*	미스센스	p.Ile291Val	-0.012 (0.001)	2.43E-21	40834
	10	101977883	C	T	rs223080	*CHUK*	미스센스	p.Val268Ile	-0.009 (0.001)	1.93E-13	40834
	10	113917085	T	A	rs225453	*GPAM*	동의	p.Pro681Pro	-0.008 (0.001)	4.61E-10	40834
	10	113940329	T	C	rs279275	*GPAM*	미스센스	p.Ile43Val	-0.008 (0.001)	2.54E-10	40832
	14	94844947	C	T	rs289294	*SERPINA1*	미스센스	p.Glu366Lys	0.042 (0.005)	9.28E-21	40834
	19	19379549	C	T	rs585429	*TM6SF2*	미스센스	p.Glu167Lys	0.014 (0.002)	4.76E-09	40833
	22	44324727	C	G	rs738409	*PNPLA3*	미스센스	p.Ile148Met	0.023 (0.002)	1.34E-50	40834
	22	44324730	C	T	rs738408	*PNPLA3*	동의	p.Pro149Pro	0.023 (0.002)	1.11E-50	40834
	22	44342116	A	G	rs229491	*PNPLA3*	미스센스	p.Lys434Glu	0.007 (0.001)	8.26E-08	40832
	22	44368122	A	G	rs376147	*SAMM50*	미스센스	p.Asp110Gly	0.019 (0.002)	8.85E-30	40833
	22	44395451	T	C	rs100786	*PARVB*	미스센스	p.Trp37Arg	0.011 (0.001)	7.98E-16	40834

회색 음영은 P < 1.43x10-3 의 본페로니 유의성 역치를 역치를 충족시키는 P-값을 나타냄

* 반복 메타-분석은 3개의 반복 코호트: GHS 비만 수술 코호트, 델러스 심장 연구 및 펜 메디슨 바이오뱅크를 포함함

* 공동 메타-분석은 발견 코호트 및 3개의 반복 코호트: GHS 발견 코호트, GHS 비만 수술 코호트, 델러스 심장 연구 및 펜 메디슨 바이오뱅크를 포함함

약어 : AAF, 대체 대립유전자 빈도; Alt, 대체 대립유전자; ALT, 알라닌 아미노트랜스퍼라제; AST, 아스파테이트 아미노트랜스퍼타제; Ref, 참조 대립유전자; SE, 표준 오차;

도면3b

특질	Chr	BP	Ref	Alt	RSID	유전자	주석	AA 치환	GHS 발견 코호트		
									베타 (SE)	P discovery	N
AST	4	88231392	T	TA	rs72613567	*HSD17B13*	스플라이스 공여자		-0.005 (0.001)	6.24E-10	40193
	10	18242311	A	G	rs10764176	*SLC39A12*	미스센스	p.Ser36Gly	-0.006 (0.001)	1.09E-10	40193
	10	101157378	CGTT	C		*GOT1*	인프레임 인델	p.Asn389del	-0.221 (0.024)	1.96E-20	40193
	10	101165533	G	C	rs374966349	*GOT1*	미스센스	p.Gln208Glu	0.271 (0.027)	2.43E-24	40193
	10	101912064	T	C	rs2862954	*ERLIN1*	미스센스	p.Ile291Val	-0.005 (0.001)	4.82E-09	40193
	11	22271870	A	T	rs7481951	*ANO5*	미스센스	p.Leu322Phe	0.004 (0.001)	9.61E-08	40162
	14	94844947	C	T	rs28929474	*SERPINA1*	미스센스	p.Glu366Lys	0.027 (0.003)	2.44E-20	40193
	19	19379549	C	T	rs58542926	*TM6SF2*	미스센스	p.Glu167Lys	0.008 (0.002)	6.54E-08	40192
	22	44324727	C	G	rs738409	*PNPLA3*	미스센스	p.Ile148Met	0.014 (0.001)	8.31E-46	40193
	22	44324730	C	T	rs738408	*PNPLA3*	동의	p.Pro149Pro	0.014 (0.001)	8.93E-46	40193
	22	44368122	A	G	rs3761472	*SAMM50*	미스센스	p.Asp110Gly	0.011 (0.001)	1.22E-22	40192
	22	44395451	T	C	rs1007863	*PARVB*	미스센스	p.Trp37Arg	0.006 (0.001)	1.31E-13	40193

회색 음영은 P < 1.43x10-3 의 본페로니 유의성 역치를 역치를 충족시키는 P-값을 나타냄

* 반복 메타-분석은 3개의 반복 코호트: GHS 비만 수술 코호트, 델러스 심장 연구 및 펜 메디슨 바이오뱅크를 포함함

* 공동 메타-분석은 발견 코호트 및 3개의 반복 코호트: GHS 발견 코호트, GHS 비만 수술 코호트, 델러스 심장 연구 및 펜 메디슨 바이오뱅크를 포함함

약어 : AAF, 대체 대립유전자 빈도; Alt, 대체 대립유전자; ALT, 알라닌 아미노트랜스퍼라제; AST, 아스파테이트 아미노트랜스퍼타제; Ref, 참조 대립유전자; SE, 표준 오차;

특질	Chr	반복 코호트								
		GHS 비만 수술 코호트			델러스 심장 연구			펜 메디슨 바이오뱅크		
		Beta (SE)	P	N	Beta (SE)	P	N	베타 (SE)	P	N
ALT	1	0.005 (0.005)	3.10E-01	2475	0.009 (0.008)	2.58E-01	1356	0.006 (0.004)	1.81E-01	6158
	4	-0.010 (0.005)	5.57E-02	2475	-0.014 (0.008)	9.68E-02	1356	-0.012 (0.004)	4.85E-03	6158
	8	-0.492 (0.165)	2.84E-03	2475	NA (NA)	NA	NA	-0.054 (0.071)	4.46E-01	6158
	8	-0.161 (0.165)	3.29E-01	2475	NA (NA)	NA	NA	-0.259 (0.143)	6.90E-02	6158
	8	-0.009 (0.02)	6.48E-01	2475	0.027 (0.035)	4.48E-01	1355	-0.051 (0.019)	7.52E-03	6158
	8	-0.189 (0.165)	2.50E-01	2475	NA (NA)	NA	NA	-0.305 (0.101)	2.54E-03	6158
	8	-0.341 (0.074)	3.64E-06	2475	NA (NA)	NA	NA	-0.144 (0.054)	7.67E-03	6158
	8	-0.009 (0.02)	6.45E-01	2475	0.024 (0.035)	5.01E-01	1356	-0.059 (0.018)	1.13E-03	6158
	8	-0.314 (0.165)	5.71E-02	2475	-0.334 (0.137)	1.49E-02	1355	-0.151 (0.143)	2.90E-01	6157
	8	-0.273 (0.048)	9.83E-09	2474	-0.244 (0.073)	8.91E-04	1356	-0.188 (0.041)	5.52E-06	6157
	8	-0.115 (0.058)	4.82E-02	2475	-0.092 (0.097)	3.43E-01	1355	-0.042 (0.043)	3.36E-01	6157
	8	-0.273 (0.050)	4.26E-08	2475	-0.198 (0.068)	3.90E-03	1356	-0.188 (0.041)	5.52E-06	6158
	8	-0.161 (0.165)	3.29E-01	2475	NA (NA)	NA	NA	-0.506 (0.202)	1.22E-02	NA
	8	-0.161 (0.165)	3.29E-01	2475	NA (NA)	NA	NA	-0.303 (0.143)	3.37E-02	NA
	9	-0.004 (0.005)	4.09E-01	2475	0.003 (0.008)	6.46E-01	1356	-0.007 (0.004)	6.38E-02	6158

회색 음영은 P < 1.43x10-3 의 본페로니 유의성 역치를 역치를 충족시키는 P-값을 나타냄

* 반복 메타-분석은 3개의 반복 코호트: GHS 비만 수술 코호트, 델러스 심장 연구 및 펜 메디슨 바이오뱅크를 포함함

* 공동 메타-분석은 발견 코호트 및 3개의 반복 코호트: GHS 발견 코호트, GHS 비만 수술 코호트, 델러스 심장 연구 및 펜 메디슨 바이오뱅크를 포함함

약어 : AAF, 대체 대립유전자 빈도; Alt, 대체 대립유전자; ALT, 알라닌 아미노트랜스퍼라제; AST, 아스파테이트 아미노트랜스퍼타제; Ref, 참조 대립유전자; SE, 표준 오차;

도면3c

특질	Chr	반복 코호트								
		GHS 비만 수술 코호트			멜러스 심장 연구			팬 메디슨 바이오뱅크		
		베타 (SE)	P	N	베타 (SE)	P	N	베타 (SE)	P	N
	9	-0.004 (0.005)	3.90E-01	2475	0.002 (0.008)	7.69E-01	1355	-0.007 (0.004)	5.29E-02	6158
	10	-0.002 (0.010)	8.01E-01	2475	-0.003 (0.017)	8.37E-01	1356	-0.015 (0.007)	4.49E-02	6158
	10	-0.003 (0.010)	7.74E-01	2475	-0.005 (0.017)	7.49E-01	1356	-0.014 (0.007)	4.86E-02	6158
	10	-0.003 (0.010)	7.93E-01	2475	-0.005 (0.017)	7.49E-01	1356	-0.014 (0.007)	5.02E-02	6158
	10	-0.001 (0.010)	9.11E-01	2475	-0.008 (0.016)	6.41E-01	1356	-0.013 (0.007)	7.46E-02	6158
	10	-0.01 (0.005)	2.91E-02	2475	-0.006 (0.007)	4.02E-01	1356	-0.009 (0.004)	2.06E-02	6158
	10	-0.006 (0.005)	2.05E-01	2475	-0.001 (0.007)	9.07E-01	1356	-0.011 (0.004)	5.26E-03	6158
	10	-0.003 (0.005)	5.80E-01	2475	-0.014 (0.008)	8.25E-02	1356	-0.007 (0.004)	7.45E-02	6158
	10	-0.003 (0.005)	5.61E-01	2475	-0.014 (0.008)	9.08E-02	1356	-0.008 (0.004)	6.34E-02	6158
	14	0.035 (0.020)	7.97E-02	2475	0.044 (0.032)	1.63E-01	1356	0.056 (0.013)	1.38E-05	6158
	19	0.040 (0.010)	2.40E-05	2475	0.025 (0.014)	7.24E-02	1356	0.013 (0.008)	1.07E-01	6158
	22	0.019 (0.006)	5.54E-04	2475	0.005 (0.009)	5.75E-01	1356	0.018 (0.005)	5.51E-05	6158
	22	0.019 (0.006)	5.51E-04	2475	0.005 (0.009)	5.75E-01	1356	0.018 (0.005)	5.78E-05	6158
	22	0.001 (0.005)	7.77E-01	2475	0.004 (0.008)	6.26E-01	1356	0.005 (0.004)	2.00E-01	6158
	22	0.009 (0.006)	1.66E-01	2475	-0.002 (0.01)	8.80E-01	1356	0.021 (0.005)	5.29E-05	6158
	22	0.003 (0.005)	5.22E-01	2475	0.007 (0.008)	3.31E-01	1356	0.008 (0.004)	3.82E-02	6158

회색 음영은 P < 1.43x10-3 의 본페로니 유의성 역치를 역치를 충족시키는 P-값을 나타냄
*반복 메타-분석은 3개의 반복 코호트: GHS 비만 수술 코호트, 멜러스 심장 연구
및 팬 메디슨 바이오뱅크를 포함함
*공동 메타-분석은 발견 코호트 및 3개의 반복 코호트: GHS 발견 코호트,
GHS 비만 수술 코호트, 멜러스 심장 연구 및 팬 메디슨 바이오뱅크를 포함함
약어: AAF, 대체 대립유전자 빈도; Alt, 대체 대립유전자; ALT, 알라닌 아미노트랜스퍼라제; AST,
아스파테이트 아미노트랜스퍼타제; Ref, 참조 대립유전자; SE, 표준 오차;

특질	Chr	반복 메타-분석* (N=3)		공동 메타-분석** (N=4)	
		베타 (SE)	P 반복	베타 (SE)	P 공동
ALT	1	0.006 (0.003)	4.77E-02	0.007 (0.001)	6.91E-09
	4	-0.012 (0.003)	1.67E-04	-0.010 (0.001)	3.85E-15
	8	-0.124 (0.066)	5.92E-02	-0.155 (0.024)	2.41E-10
	8	-0.259 (0.143)	6.90E-02	-0.264 (0.03)	4.65E-18
	8	-0.024 (0.013)	6.89E-02	-0.032 (0.005)	3.36E-12
	8	-0.305 (0.101)	2.54E-03	-0.308 (0.033)	2.21E-20
	8	-0.213 (0.044)	1.01E-06	-0.223 (0.013)	4.00E-64
	8	-0.029 (0.013)	2.09E-02	-0.032 (0.005)	2.89E-12
	8	-0.264 (0.085)	1.84E-03	-0.238 (0.029)	1.35E-16
	8	-0.227 (0.029)	2.39E-15	-0.224 (0.012)	1.94E-77
	8	-0.070 (0.033)	3.12E-02	-0.076 (0.012)	1.12E-10
	8	-0.218 (0.029)	3.94E-14	-0.224 (0.012)	2.92E-77
	8	-0.506 (0.202)	1.22E-02	-0.272 (0.030)	6.42E-20
	8	-0.303 (0.143)	3.37E-02	-0.189 (0.027)	2.99E-12
	9	-0.005 (0.003)	9.79E-02	-0.007 (0.001)	3.57E-09

회색 음영은 P < 1.43x10-3 의 본페로니 유의성 역치를 역치를 충족시키는 P-값을 나타냄
*반복 메타-분석은 3개의 반복 코호트: GHS 비만 수술 코호트, 멜러스 심장 연구
및 팬 메디슨 바이오뱅크를 포함함
*공동 메타-분석은 발견 코호트 및 3개의 반복 코호트: GHS 발견 코호트,
GHS 비만 수술 코호트, 멜러스 심장 연구 및 팬 메디슨 바이오뱅크를 포함함
약어: AAF, 대체 대립유전자 빈도; Alt, 대체 대립유전자; ALT, 알라닌 아미노트랜스퍼라제; AST,
아스파테이트 아미노트랜스퍼타제; Ref, 참조 대립유전자; SE, 표준 오차;

도면3d

특질	Chr	반복 메타-분석* (N=3)		공동 메타-분석** (N=4)	
		메타 (SE)	P 반복	메타 (SE)	P 공동
	9	-0.005 (0.003)	7.27E-02	-0.007 (0.001)	1.19E-09
	10	-0.01 (0.006)	8.55E-02	-0.014 (0.002)	9.49E-09
	10	-0.01 (0.006)	7.96E-02	-0.014 (0.002)	8.03E-09
	10	-0.01 (0.006)	8.38E-02	-0.014 (0.002)	8.70E-09
	10	-0.009 (0.006)	1.19E-01	-0.014 (0.002)	1.07E-08
	10	-0.009 (0.003)	1.14E-03	-0.011 (0.001)	1.76E-23
	10	-0.008 (0.003)	5.11E-03	-0.009 (0.001)	4.17E-15
	10	-0.007 (0.003)	2.44E-02	-0.008 (0.001)	3.82E-11
	10	-0.007 (0.003)	2.13E-02	-0.008 (0.001)	1.85E-11
	14	0.049 (0.010)	1.51E-06	0.044 (0.004)	9.55E-26
	19	0.024 (0.006)	1.58E-05	0.016 (0.002)	1.35E-12
	22	0.017 (0.003)	2.60E-07	0.022 (0.001)	7.75E-56
	22	0.017 (0.003)	2.71E-07	0.022 (0.001)	6.77E-56
	22	0.004 (0.003)	2.03E-01	0.006 (0.001)	6.49E-08
	22	0.014 (0.004)	2.57E-04	0.018 (0.002)	2.36E-32
	22	0.006 (0.003)	2.63E-02	0.01 (0.001)	1.83E-16

회색 음영은 P < 1.43x10-3 의 본페로니 유의성 역치를 역치를 충족시키는 P-값을 나타냄
*반복 메타-분석은 3개의 반복 코호트: GHS 비만 수술 코호트, 멜러스 심장 연구 및 팬 메디슨 바이오뱅크를 포함함
* 공동 메타-분석은 발견 코호트 및 3개의 반복 코호트: GHS 발견 코호트, GHS 비만 수술 코호트, 멜러스 심장 연구 및 팬 메디슨 바이오뱅크를 포함함
약어 : AAF, 대체 대립유전자 빈도; Alt, 대체 대립유전자; ALT, 알라닌 아미노트랜스퍼라제; AST, 아스파테이트 아미노트랜스퍼타제; Ref, 참조 대립유전자; SE, 표준 오차;

특질	Chr	반복 코호트								
		GHS 비만 수술 코호트			멜러스 심장 연구			팬 메디슨 바이오뱅크		
		메타 (SE)	P	N	메타 (SE)	P	N	메타 (SE)	P	N
AST	4	-0.010 (0.003)	3.12E-03	2469	-0.012 (0.006)	5.87E-02	1356	-0.006 (0.004)	1.02E-01	6166
	10	-0.010 (0.003)	2.91E-03	2469	-0.003 (0.006)	6.25E-01	1356	-0.010 (0.004)	6.75E-03	6166
	10	-0.205 (0.062)	8.57E-04	2469	NA (NA)	NA	NA	-0.244 (0.089)	5.90E-03	6165
	10	NA (NA)	NA	2469	NA (NA)	NA	NA	0.339 (0.079)	1.85E-05	6166
	10	-0.004 (0.003)	1.54E-01	2469	-0.007 (0.006)	2.18E-01	1356	-0.003 (0.003)	3.13E-01	6166
	11	-0.001 (0.003)	7.85E-01	2466	0.006 (0.006)	2.80E-01	1356	-0.003 (0.003)	3.54E-01	6165
	14	0.023 (0.013)	7.79E-02	2469	0.046 (0.024)	6.09E-02	1356	0.052 (0.011)	4.75E-06	6166
	19	0.023 (0.006)	1.99E-04	2469	0.010 (0.011)	3.42E-01	1356	0.004 (0.007)	5.94E-01	6166
	22	0.014 (0.004)	1.27E-04	2469	0.004 (0.007)	5.53E-01	1356	0.017 (0.004)	1.16E-05	6166
	22	0.014 (0.004)	1.32E-04	2469	0.004 (0.007)	5.53E-01	1356	0.017 (0.004)	1.17E-05	6166
	22	0.008 (0.004)	6.03E-02	2469	-0.001 (0.008)	9.33E-01	1356	0.018 (0.005)	6.47E-05	6166
	22	0.003 (0.003)	4.12E-01	2469	0.006 (0.006)	3.03E-01	1356	0.009 (0.003)	1.37E-02	6166

회색 음영은 P < 1.43x10-3 의 본페로니 유의성 역치를 역치를 충족시키는 P-값을 나타냄
*반복 메타-분석은 3개의 반복 코호트: GHS 비만 수술 코호트, 멜러스 심장 연구 및 팬 메디슨 바이오뱅크를 포함함
* 공동 메타-분석은 발견 코호트 및 3개의 반복 코호트: GHS 발견 코호트, GHS 비만 수술 코호트, 멜러스 심장 연구 및 팬 메디슨 바이오뱅크를 포함함
약어 : AAF, 대체 대립유전자 빈도; Alt, 대체 대립유전자; ALT, 알라닌 아미노트랜스퍼라제; AST, 아스파테이트 아미노트랜스퍼타제; Ref, 참조 대립유전자; SE, 표준 오차;

도면3e

특질	Chr	반복 메타-분석* (N=3)		공동 메타-분석** (N=4)	
		베타 (SE)	P 반복	베타 (SE)	P 공동
AST	4	-0.009 (0.002)	1.69E-04	-0.006 (0.001)	1.13E-12
	10	-0.009 (0.002)	8.86E-05	-0.006 (0.001)	9.66E-14
	10	-0.205 (0.062)	8.57E-04	-0.220 (0.022)	1.66E-24
	10	0.339 (0.079)	1.85E-05	0.271 (0.027)	2.43E-24
	10	-0.004 (0.002)	3.92E-02	-0.005 (0.001)	5.52E-10
	11	-0.001 (0.002)	7.03E-01	0.004 (0.001)	1.48E-06
	14	0.040 (0.008)	6.56E-07	0.029 (0.003)	2.78E-25
	19	0.014 (0.004)	1.20E-03	0.009 (0.002)	5.92E-10
	22	0.014 (0.002)	2.00E-08	0.014 (0.001)	1.12E-52
	22	0.014 (0.002)	2.10E-08	0.014 (0.001)	1.26E-52
	22	0.011 (0.003)	1.77E-04	0.011 (0.001)	1.01E-25
	22	0.005 (0.002)	1.34E-02	0.006 (0.001)	6.61E-15

회색 음영은 P < 1.43x10-3 의 본페로니 유의성 역치를 역치를 충족시키는 P-값을 나타냄
*반복 메타-분석은 3개의 반복 코호트: GHS 비만 수술 코호트, 밸러스 심장 연구 및 팬 메디슨 바이오뱅크를 포함함
* 공동 메타-분석은 발견 코호트 및 3개의 반복 코호트: GHS 발견 코호트, GHS 비만 수술 코호트, 밸러스 심장 연구 및 팬 메디슨 바이오뱅크를 포함함
약어 : AAF, 대체 대립유전자 빈도; Alt, 대체 대립유전자; ALT, 알라닌 아미노트랜스퍼라제; AST, 아스파테이트 아미노트랜스퍼타제; Ref, 참조 대립유전자; SE, 표준 오차;

도면4

CHR:BP:Ref:Alt	유전자	rsID	알코올성 간 질환		알코올성 간경변	
			OR (95% CI)	P	OR (95% CI)	P
4:88231392:T:TA	*HSD17B13*	rs72613567	0.62 (0.48-0.81)	1.82E-04	0.56 (0.41-0.78)	3.35E-04
8:145730161:C:T	*GPT*	rs201815297	3.83 (1.05-13.94)	8.88E-02	6.33 (1.71-23.43)	2.88E-02
8:145732114:G:C	*GPT*	rs141505249	0.77 (0.06-10.73)	8.43E-01	1.13 (0.08-15.39)	9.30E-01
8:145732180:G:C	*GPT*	rs147998249	0.73 (0.05-11.76)	8.17E-01	1.07 (0.07-17.16)	9.60E-01
10:18242311:A:G	*SLC39A12*	rs10764176	0.85 (0.68-1.07)	1.64E-01	0.92 (0.70-1.22)	5.80E-01
10:101157378:CGTT:C	*GOT1*		4.60 (0.25-86.41)	3.93E-01	7.11 (0.38-133.19)	3.00E-01
10:101165533:G:C	*GOT1*	rs374966349	2.20 (0.13-37.68)	6.24E-01	3.47 (0.20 - 59.04)	4.70E-01
10:101912064:T:C	*ERLIN1*	rs2862954	0.92 (0.75-1.12)	4.05E-01	1.05 (0.82-1.34)	7.13E-01
14:94844947:C:T	*SERPINA1*	rs28929474	2.49 (1.49-4.17)	2.30E-03	3.35 (1.93-5.83)	3.01E-04
19:19379549:C:T	*TM6SF2*	rs58542926	1.47 (1.06-2.04)	2.76E-02	1.35 (0.89-2.04)	1.80E-01
22:44324727:C:G	*PNPLA3*	rs738409	1.76 (1.43-2.18)	4.98E-07	2.07 (1.60-2.67)	1.08E-07
22:44324730:C:T	*PNPLA3*	rs738408	1.77 (1.43-2.18)	4.70E-07	2.07 (1.61-2.67)	1.03E-07
22:44368122:A:G	*SAMM50*	rs3761472	1.90 (1.52-2.38)	1.36E-07	2.28 (1.75-2.98)	1.83E-08

회색 음영은 P < 1.92x10-3 의 본페로니 유의성 역치를 충족시키는 P-값을 나타냄

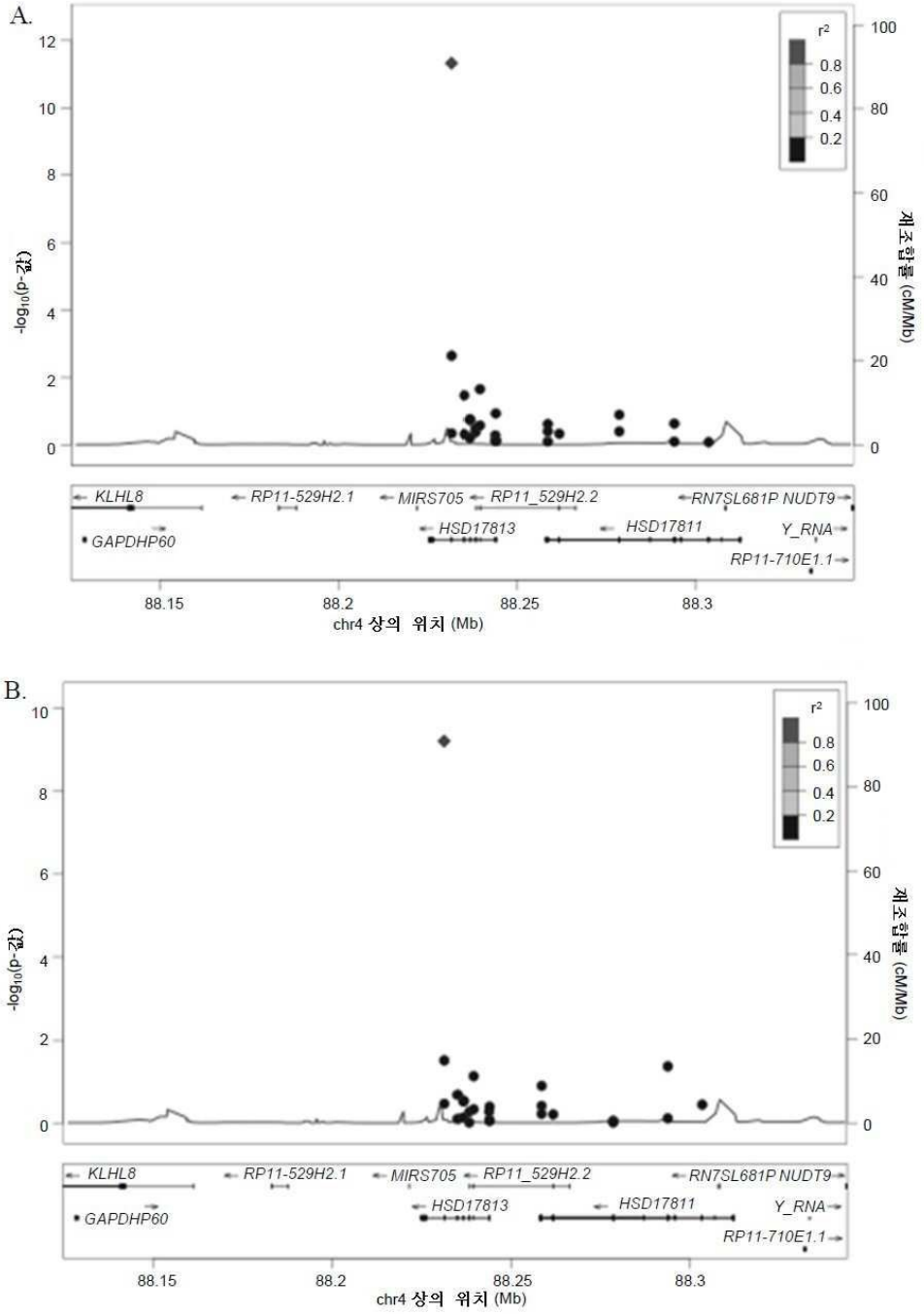
CHR:BP:Ref:Alt	유전자	rsID	비알코올성 간 질환		비알코올성 간경변		간세포 암종	
			OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
4:88231392:T:TA	*HSD17B13*	rs72613567	0.84 (0.78-0.91)	1.31E-05	0.74 (0.62-0.88)	4.48E-04	0.67 (0.45-1.00)	4.66E-02
8:145730161:C:T	*GPT*	rs201815297	0.23 (0.04-1.14)	1.86E-02	1.25 (0.24-6.38)	7.98E-01	3.66 (0.70-19.01)	2.01E-01
8:145732114:G:C	*GPT*	rs141505249	1.02 (0.49-2.11)	9.70E-01	0.36 (0.02-5.37)	3.82E-01	1.84 (0.15-23.25)	6.88E-01
8:145732180:G:C	*GPT*	rs147998249	1.03 (0.49-2.17)	9.30E-01	0.34 (0.02-5.59)	3.67E-01	1.74 (0.11-27.05)	7.21E-01
10:18242311:A:G	*SLC39A12*	rs10764176	0.92 (0.86 (0.99)	3.43E-02	1.03 (0.88-1.21)	7.15E-01	1.29 (0.93-1.79)	1.37E-01
10:101157378:CGTT:C	*GOT1*		2.37 (0.61-9.27)	2.50E-01	8.27 (1.44-47.49)	5.92E-02	9.81 (0.52-183.54	2.43E-01
10:101165533:G:C	*GOT1*	rs374966349	1.63 (0.53-4.96)	4.20E-01	1.17 (0.07-20.09)	9.13E-01	5.37 (0.32-91.12)	3.55E-01
10:101912064:T:C	*ERLIN1*	rs2862954	0.98 (0.91-1.04)	4.61E-01	1.13 (0.98-1.31)	9.90E-02	0.94 (0.69-1.28)	6.94E-01
14:94844947:C:T	*SERPINA1*	rs28929474	1.50 (1.21-1.87)	5.29E-04	2.99 (2.11-4.24)	9.08E-08	1.86 (0.74-4.67)	2.40E-01
19:19379549:C:T	*TM6SF2*	rs58542926	1.36 (1.21-1.52)	2.42E-07	1.64 (1.31-2.05)	6.04E-05	1.93 (1.22-3.04)	1.08E-02
22:44324727:C:G	*PNPLA3*	rs738409	1.65 (1.54-1.78)	1.31E-41	2.05 (1.76-2.38)	1.70E-19	2.20 (1.60-3.02)	5.59E-06
22:44324730:C:T	*PNPLA3*	rs738408	1.65 (1.54-1.78)	1.42E-41	2.05 (1.77-2.38)	1.45E-19	2.20 (1.60-3.03)	5.41E-06
22:44368122:A:G	*SAMM50*	rs3761472	1.52 (1.41-1.65)	7.33E-24	1.86 (1.58-2.19)	1.81E-12	1.66 (1.16-2.39)	1.05E-02

회색 음영은 P < 1.92x10-3 의 본페로니 유의성 역치를 충족시키는 P-값을 나타냄

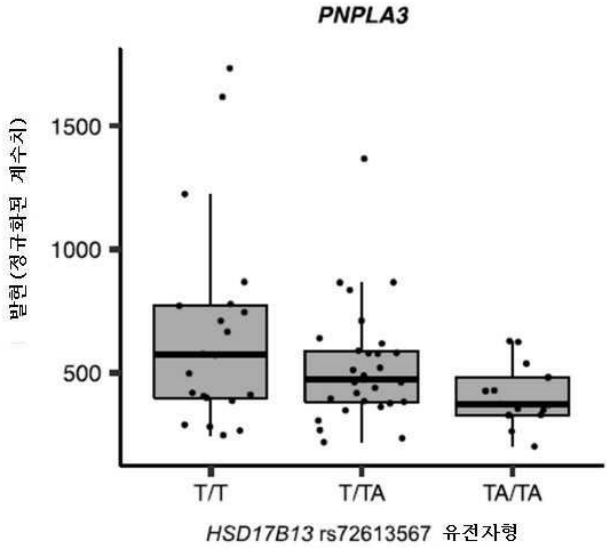
도면5

특징	멜러스 간 연구 사례 (N = 517)	멜러스 간 연구 대조군 (N = 4,279)	멜러스 소아 간 연구 사례 (N = 203)	멜러스 소아 간 연구 대조군 (N = 244)
연령(세) - 중위 (IQR)	55 (48 - 60)	44 (36 - 53)	12 (10 - 15)	12 (11 - 14)
여성 - 수 (%)	277 (54)	2,494 (58)	65 (32)	126 (52)
체질량 지수 - 중위 (IQR)	30 (27 - 35)	30 (26 - 35)	30 (27 - 34)	31 (28 - 35)
자가-보고된 민족성				
아프리카계 미국인	33 (6)	2,291 (54)	-	-
유럽계 미국인	158 (31)	1,266 (30)	-	-
히스패닉계 미국인	326 (63)	722 (17)	203 (100)	244 (100)
간 질환의 존재 - N (%)				
알코올성 간 질환	223 (43)	-	-	-
알코올성 간경변	215 (42)	-	-	-
비알코올성(비바이러스성) 간 질환	212 (20)	-	-	-
비알코올성 간경변	100 (19)	-	-	-
간세포 암종	44 (9)	-	-	-
간질환 없음	-	4,279 (100)	-	244 (100)

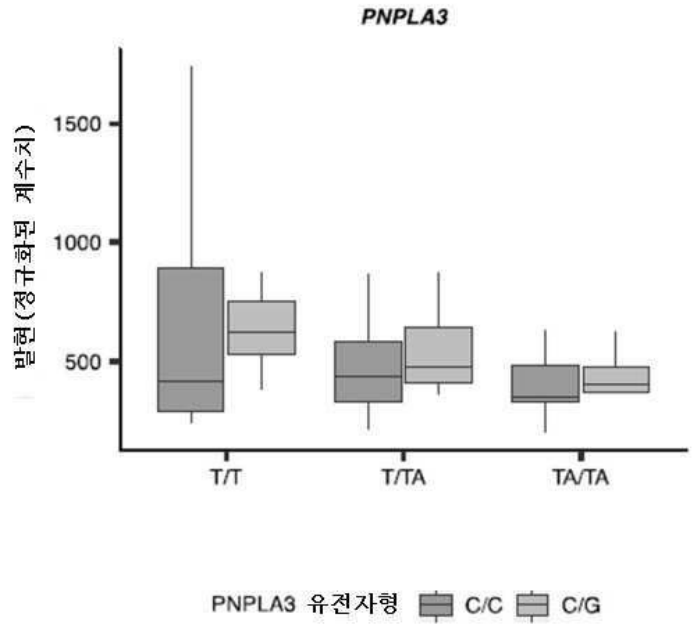
도면6


A.
r²
0.8
0.6
0.4
0.2
-log10(p-값)
재조합률 (cM/Mb)
KLHL8
RP11-529H2.1
MIR5705
RP11_529H2.2
RN7SL681P NUDT9
GAPDHP60
HSD17813
HSD17811
Y_RNA
RP11-710E1.1
chr4 상의 위치 (Mb)
B.
r²
0.8
0.6
0.4
0.2
-log10(p-값)
재조합률 (cM/Mb)
KLHL8
RP11-529H2.1
MIR5705
RP11_529H2.2
RN7SL681P NUDT9
GAPDHP60
HSD17813
HSD17811
Y_RNA
RP11-710E1.1
chr4 상의 위치 (Mb)

도면7


PNPLA3
발현 (정규화된 계수치)
1500
1000
500
T/T
T/TA
TA/TA
HSD17B13 rs72613567 유전자형

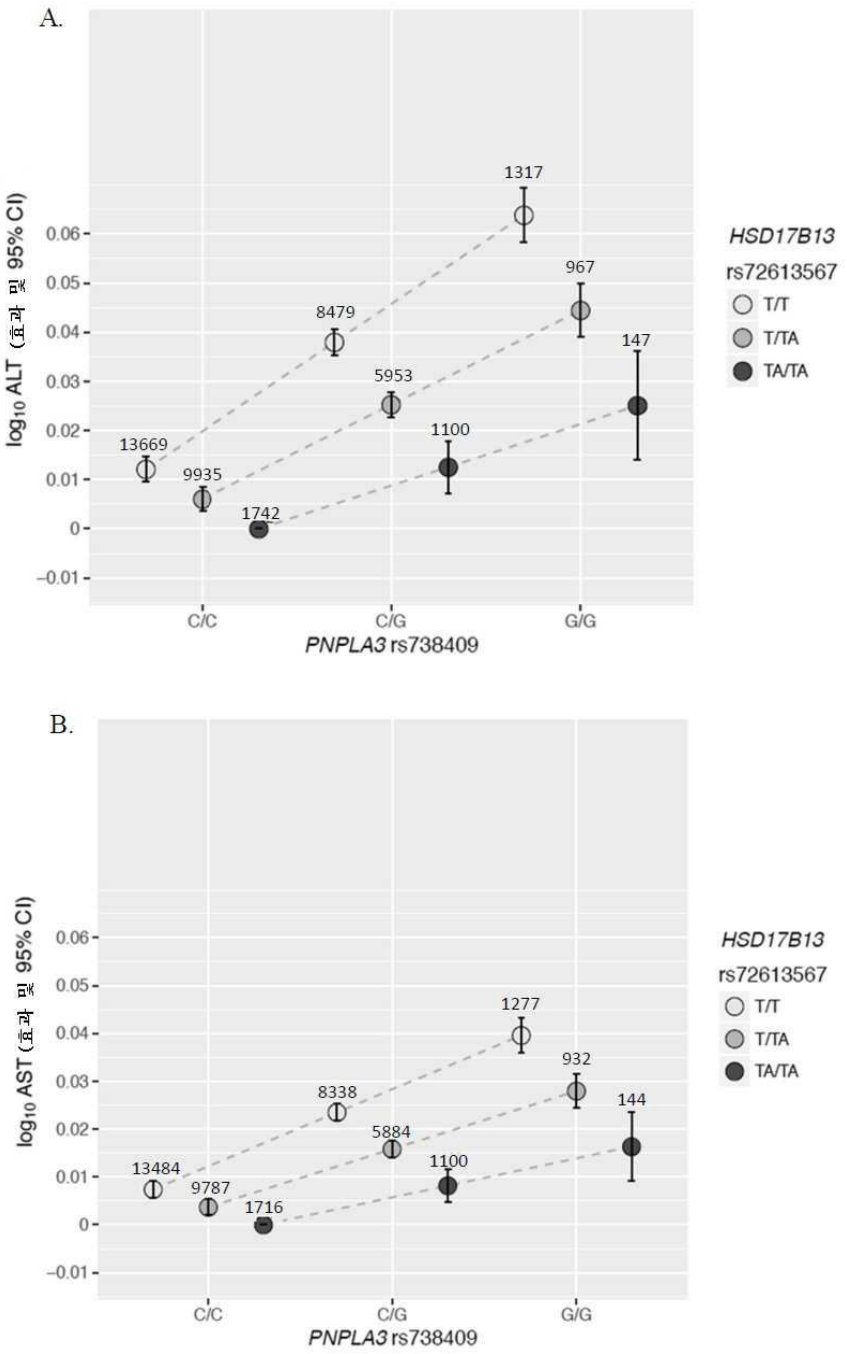
도면8


PNPLA3
발현 (정규화된 계수치)
1500
1000
500
T/T
T/TA
TA/TA
PNPLA3 유전자형
C/C
C/G

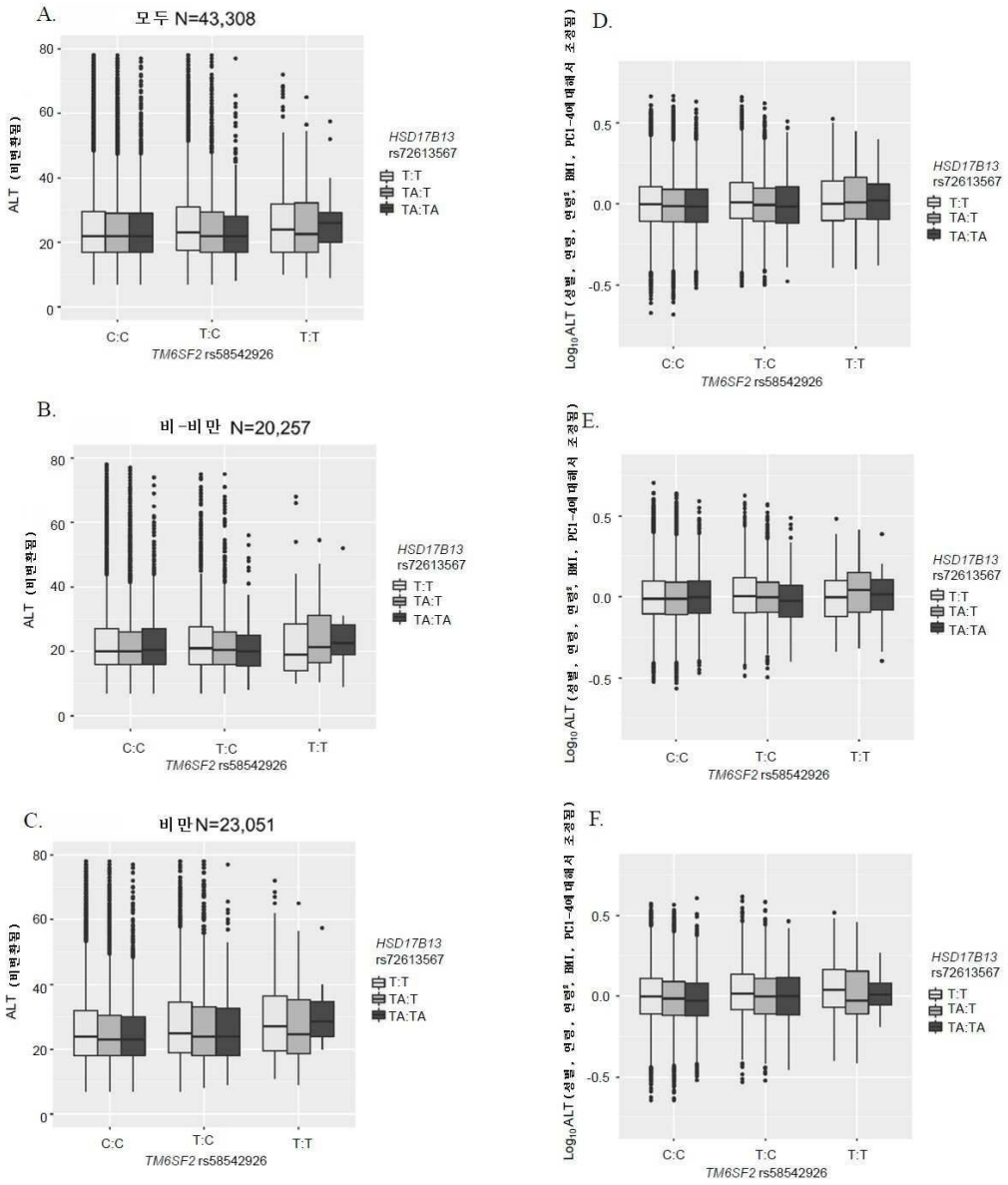
도면9

표현형/하위세트	N		*HSD17B13 rs72613567 :* [상호작용 효과]		*HSD17B13 rs72613567* [주요 효과]			*PNPLA3 rs738409* [주요 효과]		
log_{10} (ALT)										
하위세트	N		효과 (95% CI)	P	AAF	효과 (95% CI)	P	AAF	효과 (95% CI)	P
모두	43,309		-0.007 (-0.011,-0.002)	**1.80E-03**	26.36%	-0.006 (-0.009,-0.003)	**1.88E-04**	23.54%	0.026 (0.022,0.029)	**1.97E-50**
비만	23,051		-0.01 (-0.015,-0.004)	**1.01E-03**	26.48%	-0.009 (-0.013,-0.005)	**6.53E-05**	23.57%	0.037 (0.032,0.042)	**3.13E-55**
비-비만	20,258		-0.004 (-0.01,0.002)	1.49E-01	26.22%	-0.002 (-0.007,0.002)	3.53E-01	23.51%	0.013 (0.008,0.018)	**1.56E-07**
log_{10} (AST)										
하위세트	N		효과 (95% CI)	P	AAF	효과 (95% CI)	P	AAF	효과 (95% CI)	P
모두	42,662		-0.004 (-0.007,-0.001)	**4.53E-03**	26.40%	-0.004 (-0.006,-0.002)	**4.78E-04**	23.47%	0.016 (0.014,0.018)	**9.69E-46**
비만	22,719		-0.006 (-0.01,-0.003)	**1.04E-03**	26.51%	-0.006 (-0.009,-0.003)	**1.42E-04**	23.53%	0.023 (0.02,0.027)	**7.65E-49**
비-비만	19,943		-0.001 (-0.005,0.003)	4.97E-01	26.26%	-0.001 (-0.004,0.002)	3.42E-01	23.41%	0.008 (0.005,0.011)	**1.56E-06**
비알코올성 간 질환										
하위세트	N 대조군	N 사례	OR (95% CI)	P	AAF	OR (95% CI)	P	AAF	OR (95% CI)	P
모두	29,928	1,857	0.919 (0.812,1.039)	1.78E-01	26.43%	0.88 (0.787,0.983)	**2.40E-02**	23.51%	1.764 (1.606,1.938)	**1.76E-32**
비만	14,243	1,445	0.906 (0.786,1.044)	1.74E-01	26.36%	0.894 (0.788,1.012)	**7.81E-02**	23.65%	1.714 (1.537,1.91)	**2.06E-22**
비-비만	15,685	412	0.964 (0.75,1.235)	7.71E-01	26.50%	0.845 (0.662,1.069)	1.67E-01	23.38%	1.887 (1.566,2.269)	**1.96E-11**
알코올성 간 질환										
하위세트	N 대조군	N 사례	OR (95% CI)	P	AAF	OR (95% CI)	P	AAF	OR (95% CI)	P
모두	29,928	190	1.112 (0.749,1.637)	5.94E-01	26.57%	0.578 (0.391,0.834)	**4.56E-03**	22.99%	1.689 (1.298,2.185)	**7.80E-05**
비만	14,243	97	1.295 (0.741,2.224)	3.56E-01	26.59%	0.501 (0.283,0.839)	**1.22E-02**	22.83%	1.533 (1.052,2.201)	**2.30E-02**
비-비만	15,685	93	0.956 (0.544,1.65)	8.72E-01	26.55%	0.666 (0.382,1.105)	1.33E-01	23.13%	1.853 (1.275,2.664)	**1.00E-03**

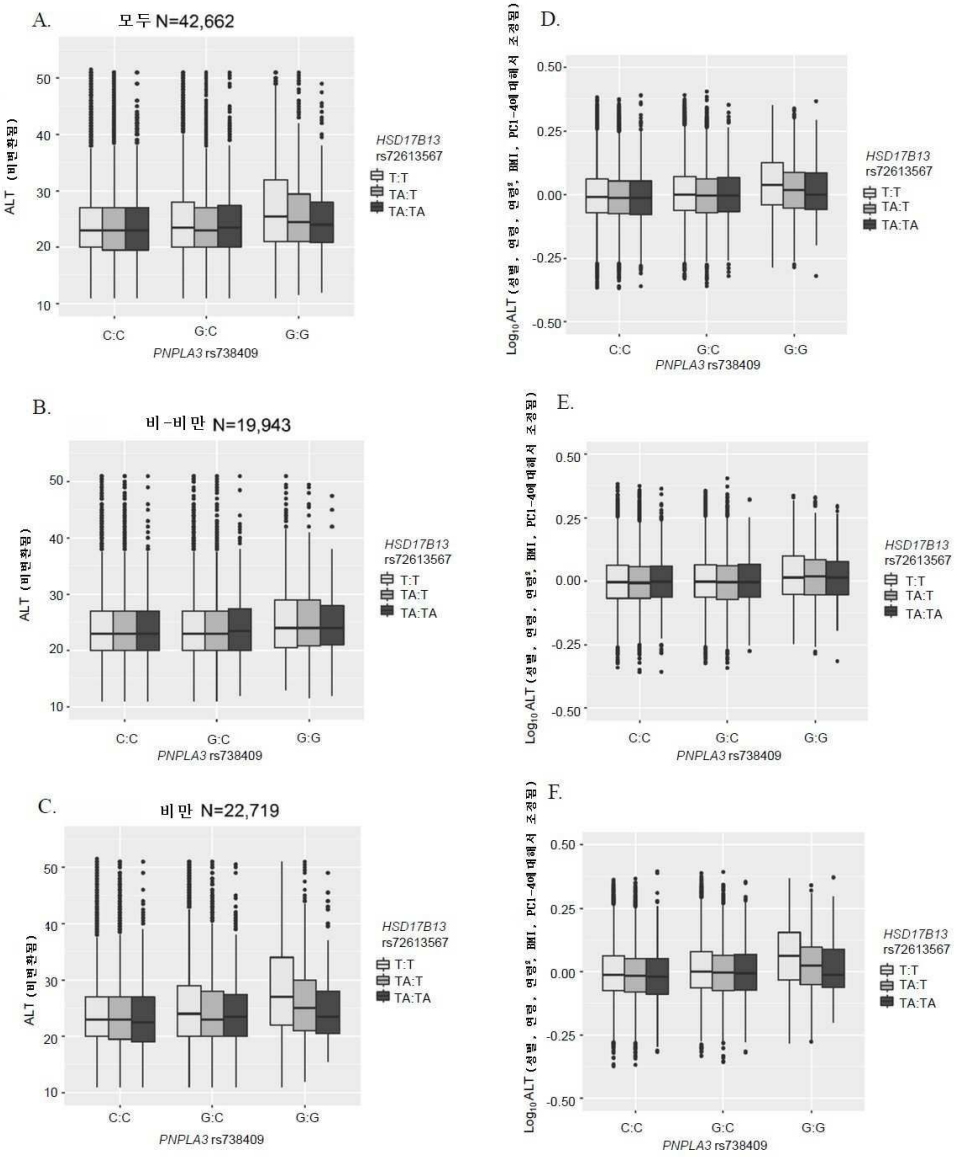
도면10


A.
log10 ALT (효과 및 95% CI)
0.06
0.05
0.04
0.03
0.02
0.01
0
−0.01
1317
967
8479
147
5953
1100
13669
9935
1742
C/C
C/G
G/G
PNPLA3 rs738409
HSD17B13
rs72613567
T/T
T/TA
TA/TA
B.
log10 AST (효과 및 95% CI)
0.06
0.05
0.04
0.03
0.02
0.01
0
−0.01
1277
932
8338
144
5884
1100
13484
9787
1716
C/C
C/G
G/G
PNPLA3 rs738409
HSD17B13
rs72613567
T/T
T/TA
TA/TA

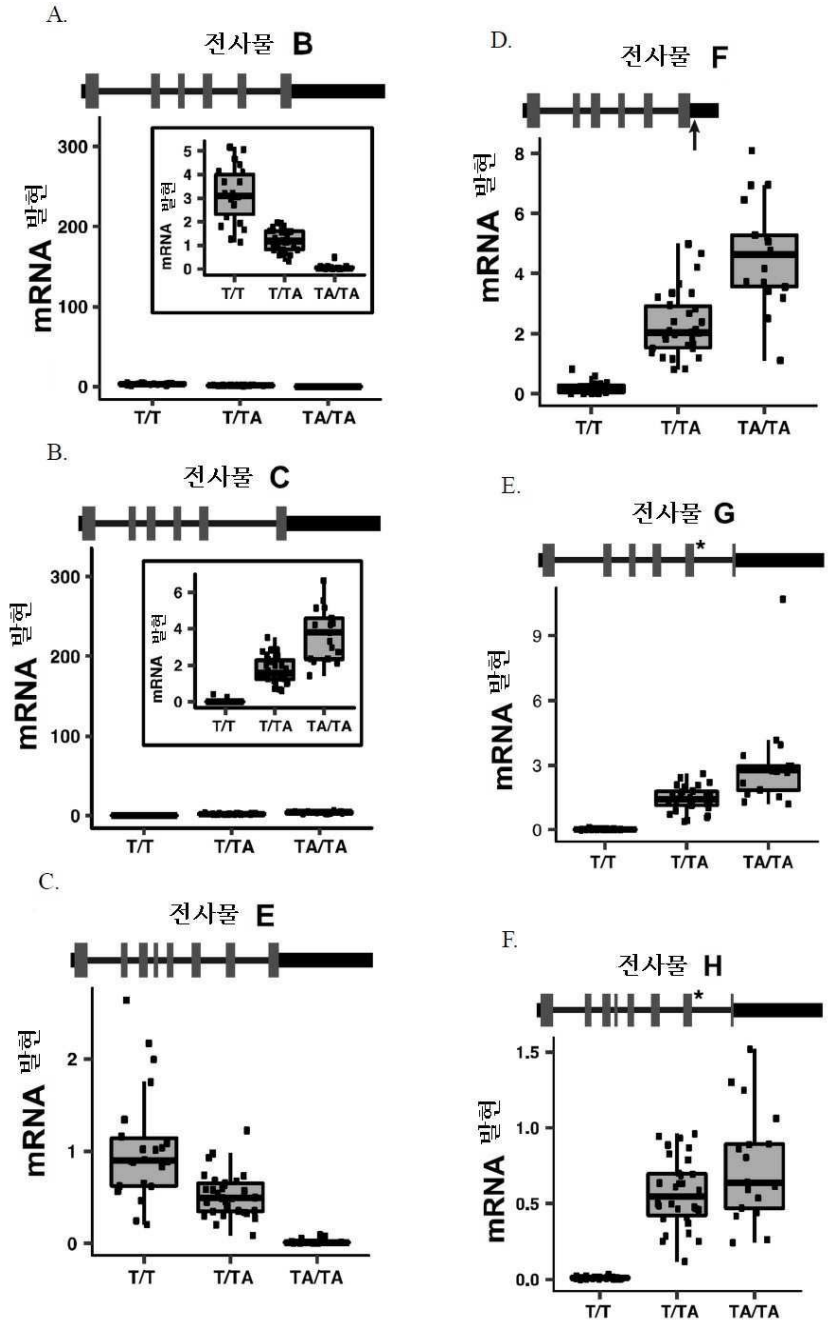
도면11


A.
모두 N=43,308
B.
비-비만 N=20,257
C.
비만N=23,051
D.
E.
F.
ALT (비변환됨)
Log10ALT (성별, 연령, 연령², BMI, PCI-4에 대해서 조정됨)
HSD17B13
rs72613567
T:T
TA:T
TA:TA
C:C
T:C
T:T
TM6SF2 rs58542926

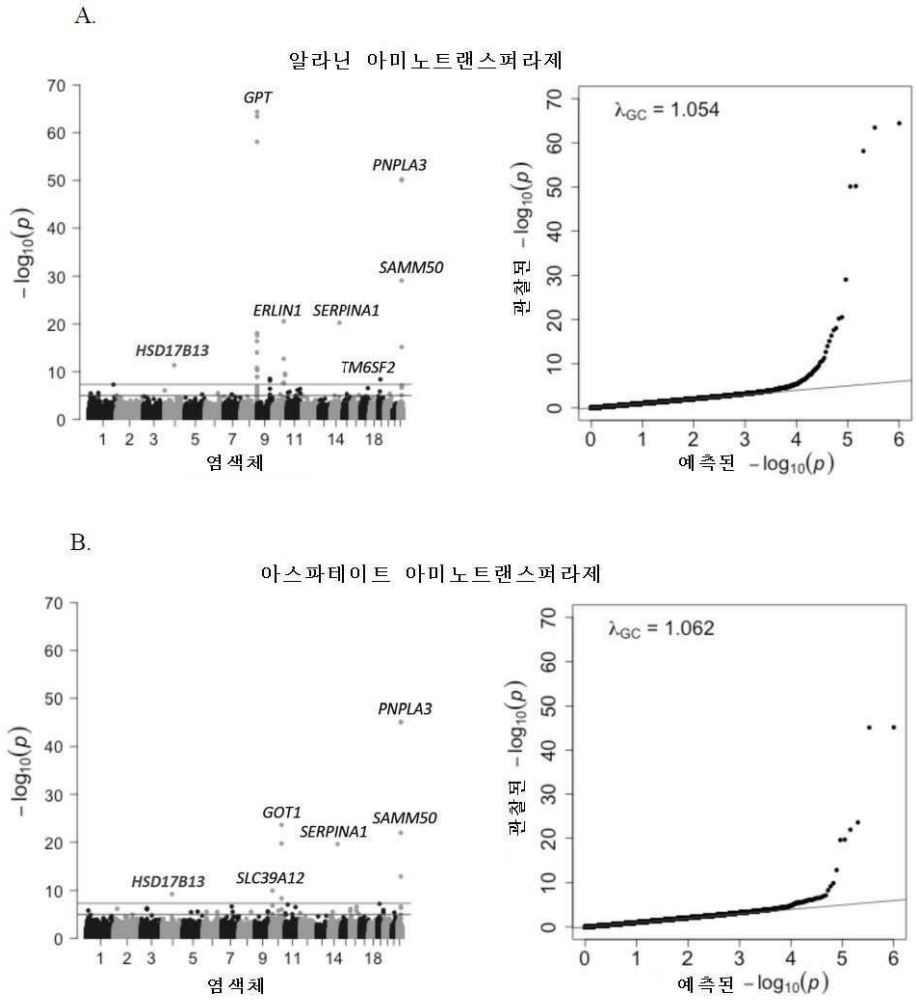
도면12


A.
모두 N=42,662
ALT (비변환됨)
HSD17B13
rs72613567
T:T
TA:T
TA:TA
C:C
G:C
G:G
PNPLA3 rs738409
D.
Log10ALT (성별, 연령, 연령², BMI, PC1-4에대해서 조정됨)
B.
비-비만 N=19,943
E.
C.
비만 N=22,719
F.

도면13

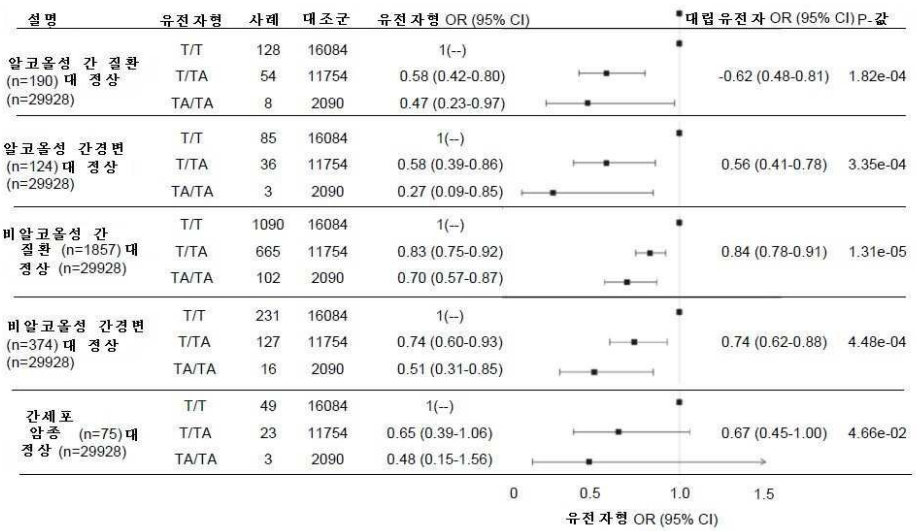


도면14

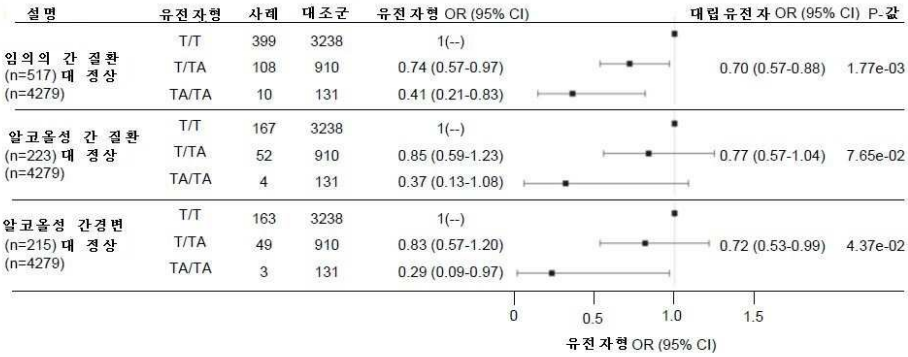

A.
알라닌 아미노트랜스퍼라제
GPT
PNPLA3
SAMM50
ERLIN1
SERPINA1
HSD17B13
TM6SF2
염색체
λGC = 1.054
관찰된 −log10(p)
예측된 −log10(p)
B.
아스파테이트 아미노트랜스퍼라제
PNPLA3
GOT1
SAMM50
SERPINA1
HSD17B13
SLC39A12
염색체
λGC = 1.062
관찰된 −log10(p)
예측된 −log10(p)

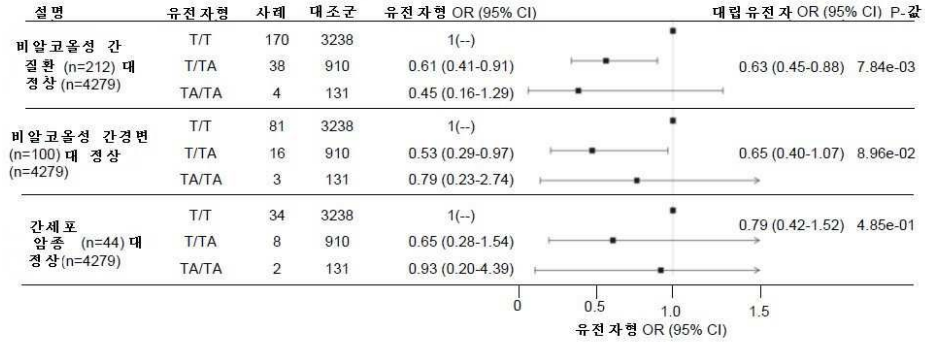
도면15

A.

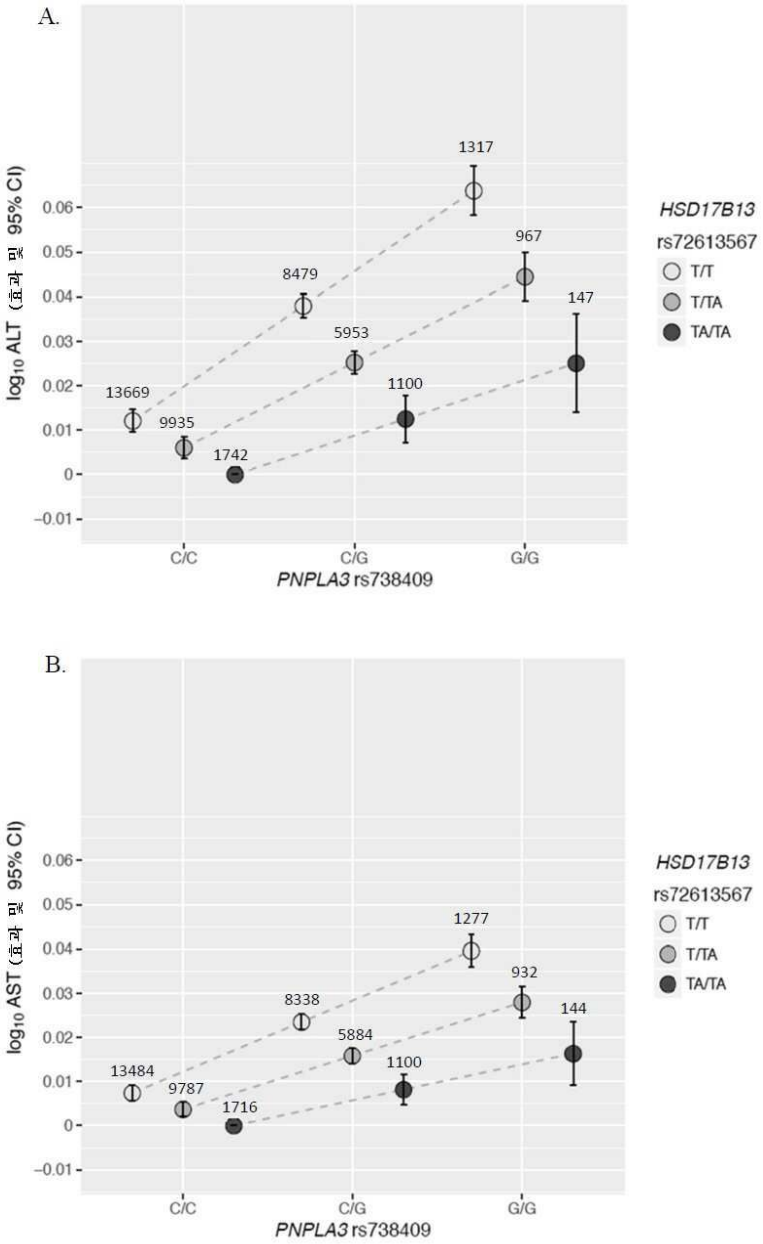

설명 유전자형 사례 대조군 유전자형 OR (95% CI) 대립유전자 OR (95% CI) P-값
알코올성 간 질환 (n=190) 대 정상 (n=29928)
T/T 128 16084 1(--)
T/TA 54 11754 0.58 (0.42-0.80) -0.62 (0.48-0.81) 1.82e-04
TA/TA 8 2090 0.47 (0.23-0.97)
알코올성 간경변 (n=124) 대 정상 (n=29928)
T/T 85 16084 1(--)
T/TA 36 11754 0.58 (0.39-0.86) 0.56 (0.41-0.78) 3.35e-04
TA/TA 3 2090 0.27 (0.09-0.85)
비알코올성 간 질환 (n=1857) 대 정상 (n=29928)
T/T 1090 16084 1(--)
T/TA 665 11754 0.83 (0.75-0.92) 0.84 (0.78-0.91) 1.31e-05
TA/TA 102 2090 0.70 (0.57-0.87)
비알코올성 간경변 (n=374) 대 정상 (n=29928)
T/T 231 16084 1(--)
T/TA 127 11754 0.74 (0.60-0.93) 0.74 (0.62-0.88) 4.48e-04
TA/TA 16 2090 0.51 (0.31-0.85)
간세포 암종 (n=75) 대 정상 (n=29928)
T/T 49 16084 1(--)
T/TA 23 11754 0.65 (0.39-1.06) 0.67 (0.45-1.00) 4.66e-02
TA/TA 3 2090 0.48 (0.15-1.56)
0 0.5 1.0 1.5
유전자형 OR (95% CI)

B.


설명 유전자형 사례 대조군 유전자형 OR (95% CI) 대립유전자 OR (95% CI) P-값
임의의 간 질환 (n=517) 대 정상 (n=4279)
T/T 399 3238 1(--)
T/TA 108 910 0.74 (0.57-0.97) 0.70 (0.57-0.88) 1.77e-03
TA/TA 10 131 0.41 (0.21-0.83)
알코올성 간 질환 (n=223) 대 정상 (n=4279)
T/T 167 3238 1(--)
T/TA 52 910 0.85 (0.59-1.23) 0.77 (0.57-1.04) 7.65e-02
TA/TA 4 131 0.37 (0.13-1.08)
알코올성 간경변 (n=215) 대 정상 (n=4279)
T/T 163 3238 1(--)
T/TA 49 910 0.83 (0.57-1.20) 0.72 (0.53-0.99) 4.37e-02
TA/TA 3 131 0.29 (0.09-0.97)
0 0.5 1.0 1.5
유전자형 OR (95% CI)

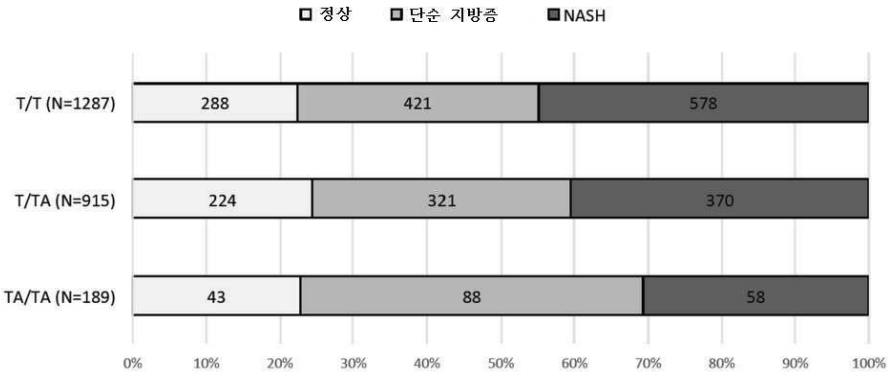

설명 유전자형 사례 대조군 유전자형 OR (95% CI) 대립유전자 OR (95% CI) P-값
비알코올성 간 질환 (n=212) 대 정상 (n=4279)
T/T 170 3238 1(--)
T/TA 38 910 0.61 (0.41-0.91) 0.63 (0.45-0.88) 7.84e-03
TA/TA 4 131 0.45 (0.16-1.29)
비알코올성 간경변 (n=100) 대 정상 (n=4279)
T/T 81 3238 1(--)
T/TA 16 910 0.53 (0.29-0.97) 0.65 (0.40-1.07) 8.96e-02
TA/TA 3 131 0.79 (0.23-2.74)
간세포 암종 (n=44) 대 정상 (n=4279)
T/T 34 3238 1(--)
T/TA 8 910 0.65 (0.28-1.54) 0.79 (0.42-1.52) 4.85e-01
TA/TA 2 131 0.93 (0.20-4.39)
0 0.5 1.0 1.5
유전자형 OR (95% CI)

도면16

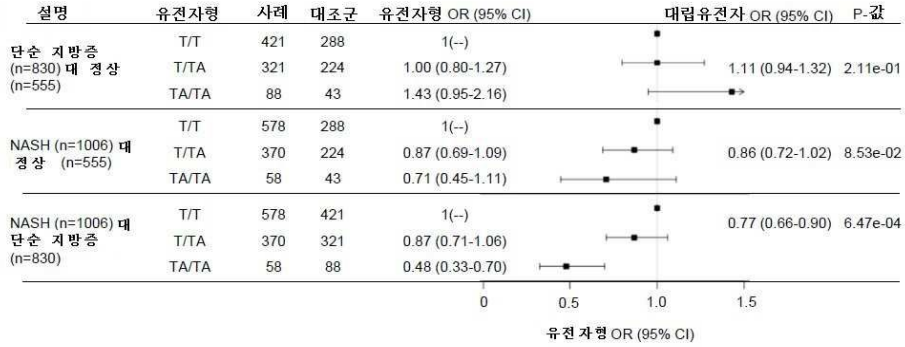

A.
log10 ALT (효과 및 95% CI)
0.06
0.05
0.04
0.03
0.02
0.01
0
−0.01
1317
967
8479
147
5953
13669
1100
9935
1742
C/C
C/G
G/G
PNPLA3 rs738409
HSD17B13
rs72613567
T/T
T/TA
TA/TA
B.
log10 AST (효과 및 95% CI)
0.06
0.05
0.04
0.03
0.02
0.01
0
−0.01
1277
932
8338
144
5884
13484
1100
9787
1716
C/C
C/G
G/G
PNPLA3 rs738409
HSD17B13
rs72613567
T/T
T/TA
TA/TA

도면17

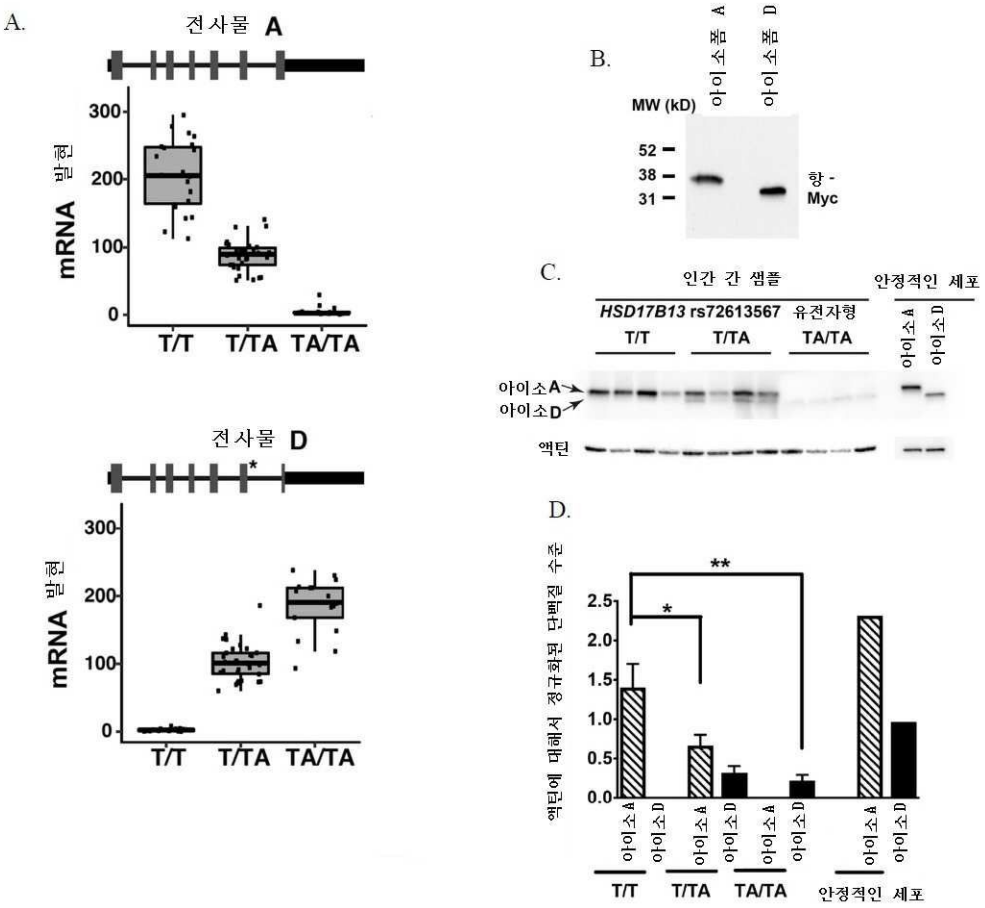
A.


정상
단순 지방증
NASH
T/T (N=1287)
288
421
578
T/TA (N=915)
224
321
370
TA/TA (N=189)
43
88
58
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%

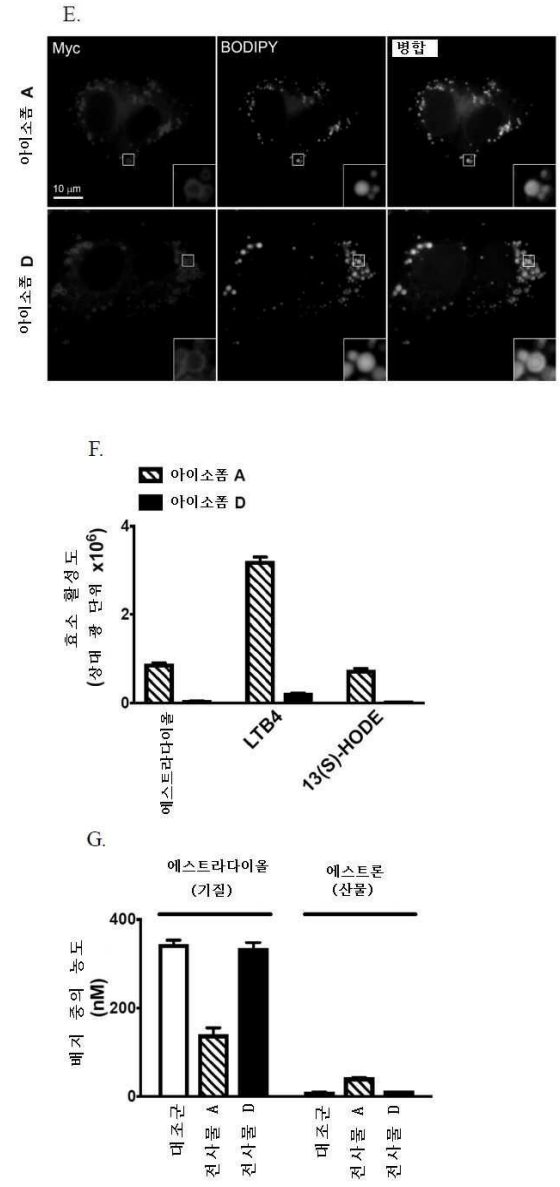
B.


설명
유전자형
사례
대조군
유전자형 OR (95% CI)
대립유전자 OR (95% CI)
P-값
단순 지방증 (n=830) 대 정상 (n=555)
T/T 421 288 1(--)
T/TA 321 224 1.00 (0.80-1.27)
TA/TA 88 43 1.43 (0.95-2.16)
1.11 (0.94-1.32) 2.11e-01
NASH (n=1006) 대 정상 (n=555)
T/T 578 288 1(--)
T/TA 370 224 0.87 (0.69-1.09)
TA/TA 58 43 0.71 (0.45-1.11)
0.86 (0.72-1.02) 8.53e-02
NASH (n=1006) 대 단순 지방증 (n=830)
T/T 578 421 1(--)
T/TA 370 321 0.87 (0.71-1.06)
TA/TA 58 88 0.48 (0.33-0.70)
0.77 (0.66-0.90) 6.47e-04
0
0.5
1.0
1.5
유전자형 OR (95% CI)

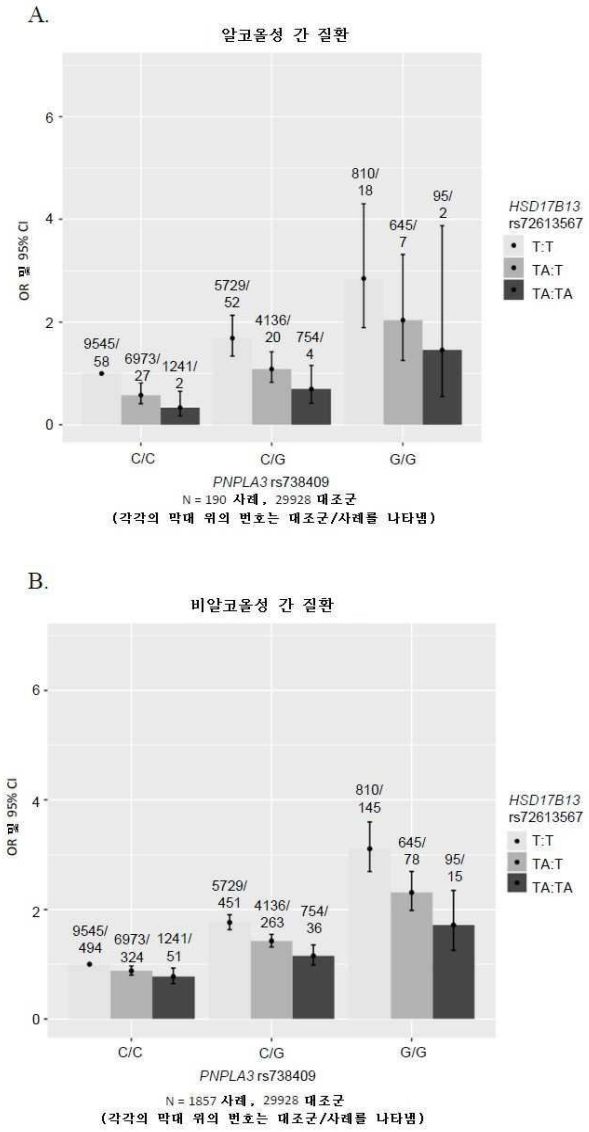
도면18ad


A.
전사물 A
mRNA 발현
300
200
100
0
T/T
T/TA
TA/TA
전사물 D
mRNA 발현
300
200
100
0
T/T
T/TA
TA/TA
B.
아이소폼 A
아이소폼 D
MW (kD)
52
38
31
항-Myc
C.
인간 간 샘플
안정적인 세포
HSD17B13 rs72613567 유전자형
T/T
T/TA
TA/TA
아이소 A
아이소 D
아이소A
아이소D
액틴
D.
액틴에 대해서 정규화된 단백질 수준
**
*
2.5
2.0
1.5
1.0
0.5
0.0
아이소 A
아이소 D
T/T
T/TA
TA/TA
안정적인 세포

도면18eg


E.
Myc
BODIPY
병합
아이소폼 A
아이소폼 D
10 μm
F.
아이소폼 A
아이소폼 D
효소 활성도 (상대 광 단위 x10^6)
4
2
0
에스트라디올
LTB4
13(S)-HODE
G.
에스트라디올 (기질)
에스트론 (산물)
베지 중의 농도 (nM)
400
200
0
대조군
전사물 A
전사물 D
대조군
전사물 A
전사물 D

도면19



서 열 목 록

SEQUENCE LISTING

<110> Regeneron Pharmaceuticals, Inc.

<120> Inhibition Of HSD17B13 In The Treatment Of Liver Disease In Patients Expressing The PNPLA3 I148M Variation

<130> 189238.01002 (10390WO01)

<150> 62/570,985

<151> 2017-10-11

<160> 66

<170> PatentIn version 3.5

<210> 1

<211> 19118

<212> DNA

<213> Homo sapien

<400> 1

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60

aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120

tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180

ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240

gacagagcat attggttctg tgggatatta ataaggtaat gtatacatct tccaactttt 300

taaagtcaca gagtaagata tgtattttaa gaattatttg acttaccatc tacttatctt 360

tgtatttttg tttttcaaag tttgataaat tccctggtcc cttagtctgt atatgtgtca 420

ggttagttag atgaagggaa tgtaattaag aactaagcag cgatttttat gacatggtgt 480

gcaggttgat agaaagactc aggagccagt ctccttccaa gctgctaaat gaggcaagtc 540

acatattatc tctcagcctg ttttcttggc tctgaagtgg ggataataac ttaggggatg 600

ggcaagaacg ggatctgaaa attacagcta caaacaaaag tcaaacgaag aacttgcaac 660

agaaaccttt agtgcctccc ctcatgcaca agcaacacag ttctaaaata tttactgtct 720

gacccctttac agaaaatgtt tgccagtccg tagtcaaaag gattaaataa gtaatatttt 780

cagcacttag catatgataa acgatacgtg gcacatgata aacaataact gtgttaaata 840

aaatatgtgc gcagtgagtc aggcttttcc ttggacatta gtatttttcc tgtgttctta 900

cttgtaaaca ctacattaac aaccccaaat aaaactgaag gaactgaaat cttgtatcat 960

tttctctaaa cttgtaaatt ctggtaaggc catgaaaata tatgcagaga agtgtttaca 1020

ggattttagg attggaaaaa ttgtgaagta ctccttgaga atcacatttt ctgcaaatta 1080

cagtggtttt aattaccatt atattattac tttctcatgt tctttgctgt catgtttagt 1140

tgaaacctaa aatgtctctt acacttagag aactaattct tttctgtttt ttttctgaat 1200

agtgaagaat actatacaaa aaagctacta cattttatt taacagatat gagcatttat 1260

ataatagagg agttgatgta tataaaaatg atttgccatc tttttggtct ttgaagaaat 1320

tcgaatgaac tttctggaag atagcaagaa tttacaaata gagaaaattg ttgcctgctg 1380

ttctcaggca tttgtccaaa aatataaata agtataaatc tatgaaaagg gcttgatgaa 1440

atctaacctt caaatctctt tccagatgtg tatttttggg gaaagggcta tatttattaa 1500

gtttttttta aattttaaaa tttccagaga caagagaaaa gtaaattaga aggaagtcgt 1560

attaaaaatg acttaagggc gggtgcagtg gctcacacct gtaatcccag cactttggga 1620

gacggaggtg ggcagattgc tggagcccag gagttcaaga ccagcctggg cagcacagca 1680

aaaccccaa ctctacaaaa aatacaaaaa ttagctgggt gcgggggtgc acacccgtag 1740

tcccagctac tcgggaggct gaggtgggag gatcgtttca gttcaggaag ccaaggctgc 1800

aatgagctat gatggcatca ttgcactcca agctgggcaa tagagccagg ctctgtctca 1860

aaaaaaataa aaaaagactt aagaaaaata ggtaacccaa cctcaaaaat tctctttgaa 1920

tcattaaatt tcatggttaa acatttaagc tactgaatga ttcactctaa ggctgtaatg 1980

taactcagat ctcctttagg cgaggaagat gctggctgag ttttcatcat aactggctcc 2040

ttttgccctg tgagatgaga gacacagtag cagtttggct cttatgcaat ctaaactgtt 2100

gcgttgggaa tacggttcaa aaaacacatt ggagtttaag ctaaagcaag tgttttgcta 2160

acaaaaagac aaggcatcac attttgcaat tgtctagctc agttataaaa cagaagaata 2220

ggccggacgc ggtggctcac gcctgtaatc ccagcacttt gggaggccga gacgggcgga 2280

tcacgaggtc aggagatcga gaccatcctg gataacacag tgaaaccccg tctctactaa 2340

aaatacaaaa aaattagcca ggcgtagtgg cgggcgcctg tagtcccagc tactcgggag 2400

gctgaggcag gagaatggtg tgaacccggg aggcggagct tgcagtgagc cgagatgacg 2460

ccactgcact ccagcctggg cgacagagcg agactccgtc tcaaaaaaaa aaaaaaaaaa 2520

aaaactgaag aataattaat tcttcaatca aaacatctga tgaatgctct ggtaacttat 2580

gctctctact gacctagaaa caaatgagag agtatggtgt ggtttgtgca atctggcagt 2640

gagcaagcta ccaactaaat cagtgaaaga ctctcctatt ctttttttac tcttctgcaa 2700

tcccacaaaa ggctatttga ggggatactg actttgagac tgggtcctaa catccatgtt 2760

tggggagttc aggctgctgc tccagggttt agcctacagt agcgaaatac aaaggaccca 2820

gagaccactc attcaaggtt tgccctaaat agcagcaaca ccactgtcat ctcaatacac 2880

gaagaatagg gcttttcagg tatccttgcc tctttgtcac agagaagagt ttacagattg 2940

tgagacggaa aagtataatt tttaaaacct tataatattt tctataaaag tcacctgagg 3000

tgaaaacttg aaaagaatta taattttcca gaatgtgagt caagaaacat tagagcaatt 3060

ttatcttagg aaagaggtct ttgaatttag gctgaaagta aattgctctg tctccatgtc 3120

ctatggttat gggcaagttt ggtacataaa tgagaaatcc atccagtggc cttgcccatc 3180

tcactcccaa acacctgaag aatgtaatgt tatatctcct agagtagcag catggtctcc 3240

ctatgaaagt ccttcttctt taaggagact tctttccctt ccctcctagg aggatgagtc 3300

agaatcatca agaaaaatat gatgggcaga ggcatacagt ttaccattac cactagttta 3360

gaattactac ttagcacttt actgcctatt acatagttgg tgctcaacaa atgtatgata 3420

aattaatggt tgagtttttc tttcttctcc atattcatct tccatgacac cacgaagagc 3480

aatgtttttc aagaatgttc ttcaaggttt gaaagtagcc tgctttagag aaactgccta 3540
ctgtacagcc tccaaccaag aggaaaagct gaaaaaagca tgaagggatt ttgttttgtt 3600
ttgtttgttt tggttttaat atgagcattc cctggcagaa aagccagggg taatctcatt 3660
gcaactaggc aatcactctc aagaaatttt ctaacaaata aggaggccaa tttttatttt 3720
attttgagac gaagtcccac tctgtcaccc aggttggagt gcaatggaat gatttcagct 3780
cactgcaacc tccgcctccc gggttcaagt gattctcctg tctaaacttc ccgagtagct 3840
gggattacag gctcccacca ccacgcccag ctaatttttt gtatttttag tagagatggg 3900

gtttcaccat tttggccaga ctggtctcaa actcctgacc tcaagtgatc caccctcctc 3960
ggcctcctaa agtgctggga ttacaggcgt gagccaccac acctgaccca ggaggccaat 4020
ttttaaaagg ttaactaatc ttcatgtcca aaatgaatgt taattgttca ttttggacat 4080
gaatgttaat ttttttttt ttttttttg agacagagtc tcactctgtt gcccaggctg 4140
gagtccagtg gcactatctc cactcactgc aacttcctcc tcccaggttc aagcaattat 4200
cctgcctcag cctcccaagt agctgggatt acaggcccac accatcaggc ctggctaatt 4260
tttgtatttt tagtagagac ggggtttcac catgttggcc aggctggtct tgaactcctg 4320

acctcgtgat ccgccctcct cggccaacca aagtgctggg attacaggcg tgagccaccg 4380
cgcctagccg aatgttaatt gtctaaaaat ttttcttctc caatgtcttc tcctccactt 4440
ttttcggaat ttgtttcttc ctaattacag cgcggtgtgg aggaaactgc agctgagtgc 4500
cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 4560
tatcgctctc taaatcaggt gagactgcag gttcacaaat ttcttcagat tattttgttt 4620
cctaggacgc tgacgtggaa aatgagaaag gtctttatga ctgcctgatt taaattggat 4680
tttagctgct aactgaagta gttatgtcac caaggaagga tatatacttt ttttcttgta 4740

tgtaatccac tcagctctgc ccattattat tgttcatatt attaatcaat ttcattctga 4800
tcagaagtgt gagcagtggc acagagtgac tgacaaaaga tttatcatca gggaatatgg 4860
atcacttcct agttttgttt tagtcctatt aactttgcag taattccagc ttctctttaa 4920
ttatttccct tgtgagattt tattttggtg ttaatgtagt cttctgtaga aaatgtaata 4980
ttaataatta ttatcacaat tattttaaaa gagtaaatac caaataatca caatgaacta 5040
agcactctaa caaactttac attttttaat tcaatcccta caataactct gtaaacttca 5100
ttttacagat aagcaaatta tgactcagag aggttaagcc agacccaggt catgtagtta 5160

ttaggttatg aaaccaggat ttctcaacca gcactttaga ccaggtgcgg tggttcacac 5220

atgtaatccc agcactttgt gaggccaagg tggaaggatc acatgagacc aagagttcaa 5280

gaccagccca ggcaacatag tgagacccta tctctaaaaa aaaaaaaaaa aaaaaaaaaa 5340

aaagtttaaa gaaaaacaca tttttaaaaa atgaacactt taaaaatatt tggtcagaat 5400

ttatatagga atttatcaac ataaatgtta atttcacttt actgataaac ttgcaaaaca 5460

tgatgtgctg ggtactgaaa tttagatgtt aaaagaacag tttatcccac ctttatgaca 5520

gtgttccctt ggcctccacg atttgagctc aacagtctgt cttgcctgaa ctctgagaga 5580

cctcatacaa tagaagaaag actctcatct ttggattata ttggtcccaa aactttgagt 5640

ttgaataata cacccagtga aagtgttctt tcaatttcaa aaggtgaaga aagaagtggg 5700

tgatgtaaca atcgtggtga ataatgctgg gacagtatat ccagccgatc ttctcagcac 5760

caaggatgaa gagattacca agacatttga ggtcaacatc ctaggacatt tttgggtgag 5820

tgtgagtcag aaacatttct gatttgtgca ccttctctta agatacatga aacttataac 5880

ggagttcaca tacttctgga caggaaactg gccagatctt tgccttaatc aagaatcatt 5940

aaatttgttt gagtagaaga gccacagagt ctctgacaca aggacacaga attcaagtgg 6000

acacaacaca ccagaatgta agctacttgg tctgtcttgt ccaccagtat ctgacacaaa 6060

gcttggcatg taccaggagc tcaacaaatg tttgtggagg tttgttaagg gttgtcagtg 6120

tacatctttt caatgctgtc acttgtgact tcatttttt ccctccacac catgattttg 6180

taatgtgtcc tcattttgtg gaattttaga atggaaagga catcagaagt aattacttgg 6240

atgtatatag gatcgaggac acttttggac gagactctga ggcaagtgtt ctagatccat 6300

ggggtgctgg aactgagaaa tgcagctata cagacctcat ataattggtt agttttgtgg 6360

gagatggaaa tatcaacttc aactgccttt gtatagaaat ttttatgatt aatcttccag 6420

tgcctcaata ttagtgtaga atctagggca gatctggatt ctagaagaaa gaagaaaaaa 6480

aagagatgtg tccccccttac ctttaccagc tcttcacata tgtgaattgg ctcccatgcc 6540

caccaaacta cacggagacc tcatacatta gctacctata gctgcataac aaattataca 6600

aaacttagtg gtttaaagca acaatgtatg ttcactatcc tctcacagtt tctatgggtt 6660

gggaatttgg aggtagcttg ggttgggagt tctagttcta tgaatttgca taggatttat 6720

taaattctta taaaatttta ttgatgtttc tcacaaaaga ggtttttgga aaaaaagaaa 6780

gacttgtttt ctgtaacatc aacatataat atacaatatt acaaataggg agatagtgaa 6840

ttcaatcatg attcattagt gtggtgtaga actctcagct tacactactc aactgtctta 6900

atacagttac acaagatttc actcttttaa ttagaatgat aaagccccaa accaaaaaat 6960

tatatgacac caaattatca taaggaataa ttttagttct gaaaactctg aatttttccc 7020

ttaatattgt ttagatgaca tatccaaaaa aggatctatt tgattccttc tgaagggaag 7080

gagggggagt actgagatta gtgttggcat ggggcttacc ataccaataa atttgtatct 7140

ttatttctat catttgtaaa gaattaatca tggaatgctt ggaagtattt tatttcattg 7200

tataagttct ctcaaatgcc tttctgtctt aacaaaaata aaactacctg atttggaaac 7260

ctaacgtcta tgtcattgtc tttcttcttt ctgcaatgat ccttaagatc acaaaagcac 7320

ttcttccatc gatgatggag agaaatcatg gccacatcgt cacagtggct tcagtgtgcg 7380

gccacgaagg gattccttac ctcatcccat attggtaagt atcacatgcc agccatgtta 7440

tatattttta tactttgaag ggagcattac acttcaaatt gttaccactg gagagtcctg 7500

gttcttggca tcttgaacaa agaattggac aaaactcacc aacaaagcca ggaaagaatg 7560

aagcaacaaa agcagagatt tattgaaaat gaaagtacgc tttacagggt gggagtgggc 7620

ccaagcacag gggctcaaga gccaattaca gaattttctg gggtttaaat acccccctaga 7680

ggtttccact ggttacttgg tgtacgccct atgtaaatga agaggatgaa ttaaagttac 7740

agagtcgttt actcagtgta caccatatgt aaatggagag gatatttcct gtcatagctg 7800

gagtgtttcc atttgattta gttctaggaa gtcagcatga atcggcctta tgttccctgc 7860

ctccagaccc tgttctcctg cctcaagatt acaatgctga gagcagagtg atttggattt 7920

acagaattta aatttatagt agtttagaat gatttttaa atgacttttt ctaaaacaat 7980

gaaaccaggt tgtaattata tttaagatat ttttagattt ctgcaaactc ctctgtagaa 8040

caatgagaga aaacagtaat gccaagcatg tttccattgt ttcctggaat aagaaacaga 8100

aaccccacag actgagaagc aaaacctaca gaagctaaaa tgaacacatg tctatgtcat 8160

ggccttggtg cccaagataa gacaatcaga gtggtccctg gatcaaaaca ttttacagtg 8220

tgcttgtgcc atgaaagtgt gtgtgtgtgt gtgtgtgtgt gtgtgagaga gagagagaga 8280

gagaaaacga ctctacctga ctaaaagttg cagataccac actccatgca ccaccaaaga 8340

cataaaggga aggaggtgag aggcgttaag gatgtactgc tgtatttgcc aaatatcctt 8400

tcctgtaaac tcttctccag atcctcataa taaaattaag aggccaaagt ggcaaccatt 8460

gtcaagagaa aaactatcaa ccattgtcaa gagaataact cagttattga gagagagagg 8520

agaaatgagc agagtcctac agaagtctgt caacacagat accagttttg tagaatttct 8580

aaatgtattt ttcctgattc atatttttca aaataaaagc agcaataaaa actgattaga 8640

aaacagtttg aagattcaat ggaaaaacct tacatgtagg atggaaaact gaacattaag 8700

ccaatcaata gagttatttt tgttcttttg ttatcattgt tgtttaagaa atgagatacg 8760

ttcacaattc tgcttaatca tgtaagaaaa tgaaaatgaa ttgccattta tactctcaga 8820

aaaatcacaa gtggctgatt tttggcttcc acttgttctt aagccaaatg ataccgcctt 8880

ctcacagaaa gctgaggatt ggtttcactc tcccttagct aacaatgctt aataattctc 8940

ttacagttcc agcaaatttg ccgctgttgg ctttcacaga ggtctgacat cagaacttca 9000
ggccttggga aaaactggta tcaaaacctc atgtctctgc ccagtttttg tgaatactgg 9060
gttcaccaaa aatccaagca caaggtaagg tcaaaatcaa gttagaatgg gtatgtggta 9120
tgataaattg atatgaaaac taatgagaaa tgtttaggca ggccaactaa tagaagaaaa 9180
tgaagaagga aaaataattt ttcttattat tattattatc ttgaaattaa aggaataaag 9240
ggggaaaaca cattagggac tagcaggaat gatcagccac cgatgaactg ggatatttat 9300
ttgtgtccgg gagaaagcac atacatttga tcaccgttac caccctgtct ttaaaatgca 9360

aatgttccaa ggaccagcaa ataaattgag tatctagctc cttagtcaag gtgaatttct 9420
gcaagaactc ttgtctctgg tgagacagga tttgagacca caagagaaga aaaattagtc 9480
ctgaaaggag aagaaaaaag caggaaggtg tggataagaa cccgaaaatt aagccatctg 9540
cttaacaaat ttttctaatc ctagtatata ttctgctgca ggttaacaaa atatactaag 9600
cttaatgatt cgaaaccaat tttttactgg aagggaatta atcctaaata tattcattca 9660
aaagaactaa acaattctct gttgagtgcc gcctcatttg aggatactga ctcttacagc 9720
ctgagttagc tatgtggtct ctgcagctgg aatcactccc tgccactgga gtccttcatg 9780

gtgttagacc ataggtactg ttgactaaag aaaaaaaaaa gtttttgttt ttatttttgt 9840
ttttttgag acagagtctc actctgtcac ccaggctgga gtacagtggc gcgatctcag 9900
ctcaccgcaa cctccgcctt tctgggttca agcaattctc cttcctcagc ctcctgagta 9960
tttggattac aggcgcccac caccacgcct ggctaatttt tgtattttta gtagagacgg 10020
ggtttcacca tgttggccag gctggtctca aactcctgac ctcaggtgtc ctacctgcct 10080
tggcctccta aaatgctggg attacaggag tgagccacca tgcccggcca aaaaaataag 10140
tttttaaaga attaaaggtc atcctggcta acacagtgaa accccgtctc tactaaaaaa 10200

cacaaaaaaa ttagccgggc gtggtggcgg gcgcctgtag tcccagctgc gcgggaggct 10260
gaggcaggag aatggcgtga acccgggagg cggagcttgc agtgagccga gatcgcgcca 10320
ctgcactcca gcctgggcga cagagcgaga ctccgtctca aaaaaaaaaa aaaaaaaaaa 10380
aaaaaaaaga attaaaggtg ttaattttat ttagaagcct tactgaagac tacagtcgga 10440
ggcctatagc ctgagagcag ccctttagag aggttcagtt gaactgttct gatagtgggg 10500
gccatgtgct ctatcctgta ttgtcttcaa agcatctttc cagagagctg cacattgtca 10560
cagagtcagg gactttgtga aattatgctg acaaccagaa gtgagtaaac gtggcttctt 10620

acatttgcta cgttgtctca cagtacttaa taagtatgca atatgtaagt aaatactata 10680

gtactattgc aactcctgat tgttttctta gacaaggaat tgggcccaat aaaaaccctc 10740

ttggtaggca ttcaggcttc gtgtaccatg agctttccta agggtatcct gccactcttg 10800

gggaaggcat gatagatgag gggagtaagg ataatggaac tctgggtaca gggttcctgg 10860

gggctaactt agaggtagac acaggcaatg ctaaatattt gggattgatt ttatagaggt 10920

tgctagattg tgaatttcct tagtaagggc taaggcattg atatgtaatg tcacacttgg 10980

ctccgaggct gggttgttgg atccatgtag atgaaatcag ggagagaaag ggcagaacgg 11040

agtaatttag aaatgtattg atttgtatta ctctctgttg gcttgctatt caaggcagtg 11100

gagaactcaa tcacataata atctgcagca aaccacagat catcccaggg aatgaagttt 11160

taacattcgc tggctcccta actcctcacc cagcctttac attcactggc tgttcagtcc 11220

atgcctggac atcttaattt gaatacaaca ttttaaatcc atttttctgt catcatcttg 11280

cactaacaga caattctaca ctaagcctat gtttatgaat atttctcaag agtacatgta 11340

cacagccttc agtataagga aaactggaag tatgacatac ctccagttgt catactcctt 11400

gggcccctct taaattctca ttaaactgca ggataggcaa gtcagaggtg aatctcaaat 11460

acgaaattct taccggaaag gggttccaat ccagacccca agagagggtt cttagatttc 11520

tcgcaagaaa taattcgggg caaggccaca gtgcaaagca aaagcaagtt tattaggaaa 11580

gtaaaggagt agagaacagc tactccatgg agaagaatgg cttgagctgc tccaccaagg 11640

gtatttagag ttacttcttg attatatgct aaacaagggg tggattattc atgagttttc 11700

cgggaaaagg gtgagcaatt cccagaactg agatttcctc ccctttttag gccatatagg 11760

gtaacttcct gccattgcca tggtatttgt aaactgtcat agtgctggtg gaagtgtctc 11820

ttagcttgct aatgtattat agttagctta taatgagcag tgaggacaac cagaggtcac 11880

tttcatcacc atcttggttt tggtgggttt tggccggctt ctttactgca ccctatttta 11940

tcaacaaggt ctttatgacc tgaatcttgt gccaacctcc tatctcatcc tgtgacaaag 12000

aatgccttaa cttcctggga atgcagccca gtaggtgtca gccttatttt acccagaccc 12060

tattcaagat ggagttgctc tgatttaaac gcctctgaca aaatgacgac ctcaaaacaa 12120

tccagcttta tggaatacct ccacaagaaa gaaagtatac ttagctatag aattttctcc 12180

ttgcatccaa caggctttga gatgtcagat gtttccttcc tgtcccatga ttaatcctag 12240

ccattcctct ttcttgtctg gctccactac tccttaccat ctaatgcctc gccaccattt 12300

tgatattttg actaagtgag ctatgaaaca cacctactgg atatgaaagt ataagtttct 12360

gataacaaaa catcaacatg ggatgtggag gaagtgggta gggtggcatt aatgcagcaa 12420

atcctggaat attttaaatc ttcattctaa atttagtaaa aatataggat aattttcctg 12480

ccatcattta cttataaaat taaaatttta gaaaataaaa ataatatttt cctcttttta 12540

atcacagatt atggcctgta ttggagacag atgaagtcgt aagaagtctg atagatggaa 12600

tacttaccaa taagaaaatg atttttgttc catcgtatat caatatcttt ctgagactac 12660

agaagtaagt acagcacaga acacccaaat actaaaacac caatagagct ttttttttg 12720

ctttttttt ttttagacag agtctcactc tgtcaccctg gctggattgc ggtggttgca 12780

gtggcatgat cttggctcac tgcaacctcc gcctcctggg ttcaagcaat tctcatgcct 12840

cagaccccca agtaactggg attataggtg tgtgctgcca cactacaccc agctaatttt 12900

tgtatttttt gatagagaca ggtttcccca tgttggccag gctggactcg aactcctgac 12960

ctcaagttat cctcctgtct cggcctccca aagtgctggg attacagtca tgagccacca 13020

tgcctggccc aatagagcta ttattatgga gcatctttca gttgtgaaaa ttggcatgga 13080

aactctccat ccctggggag aacagttatt tcctctgtta ttttcctacc cagtctataa 13140

aaagagagtg attcattttc tctaccaaat ctactgtctc tgcccaaact ttgctgaaga 13200

ctattctaac taaaggaaac acagtttaaa aagaatgcaa tatagtgaag tagttaataa 13260

taaagactcc atttttaaaa gtctgctgga agtttggttg ggattgcact gaatctatag 13320

agcaattggg gagtattgac atatcaacaa tattgagttt tctaatccaa gaacataata 13380

tctattttta aaatcttctt caaaatcttt aaatctttaa attgtatttt gtagtttttg 13440

gtgtttaagt cttgcacata ttttgtcaga tttattccaa agtatttcac gggttctttt 13500

ttttttttt ttttttttt ttgagacaga gtttcaccct tgttgcccag gctggagtgc 13560

agtggcgtga tcttggctca ctgcagcttc tgcctcctgg cttcaagtga ttctcctgcc 13620

tcagcctccc aagtagctgg gattacaggc acctgccccc tcgcccaact aacttttgt 13680

gtttgtagta gagacagggt ttcaccatgt tggccaggct ggtctcgaac tcctgacctc 13740

atgtgatcca cctgcctcag cctcccaaag tgctgggatt acaggcatga gccatcatgc 13800

ccagccctat ttgacggttt ttgacgctaa tgcaagtggc attttaaaaa attttatatt 13860

tcccattgtt tgttgtcagt atatattgga tttttgtaat ttgatctcat attttgcagt 13920

cttgctaaat tgctaaacct ctttttgcta aactcgataa gctttttttt ttttggtaga 13980

ttcctgggcc tctaattttc tttatgggaa agtttttaat tacaaattta atttctttaa 14040

tagctacatg gctattcaat ttacttatta attcttggta atgtgtgtct ttcaaggaat 14100

ttgtccattt catctaagtt gtagaatttc tttggcataa atttgtacat aacattccct 14160

tattatcctt ttaatgtctt tagaatgtct tatttattta tttatttatt tttattatat 14220

ttttttgaga cagagtctcg ctctgttgcc caggctggag tgcagtggca caatcttggc 14280

tcactgcaag ctccgccttc tgggttcatg ccattctcct gcctcagcct ccctagttgc 14340

tgggactaca ggcgcctgca accatgccca gcttattttt ttttttttt ttttttttt 14400

ttttttttt ttttttttt tagtagagac ggggtttcac cctgttagcc aggatggtct 14460
cgatctcctg acctggtgat ccgcccgcct cagcctccca aagtgctggg attacaggcg 14520
tgagccacca agcccagcct atttatttat ttagtagaga cagtctcact ttgctgccca 14580
ggcaacaaag gttttgaatg cctggcctca agcagtcctc ctgccttggc ctcccaaagt 14640
gctgggatta caggcatgag ccactgcacc tggccaaatg aatatgctga taatatcttc 14700
tttataagga tgacataaga ataaaataat gtaatacaaa caaagcccct gtcactgaaa 14760
atgtatagac ttcaaatgtt aaagtcttag agaacagaat ttatatgaaa tagcaacagc 14820

aacaatttcc cagaggaaat actctctcag ctttcttctg aggagcagtt tctaaattga 14880
aattgtatca gtgagaagat aactatacta acttcataag ccttgggcct ttttgaaaca 14940
aatccatata aactatgaac aaacttgaaa tagaacaatt tgagaacagg gtacaaactg 15000
cattggtgta tcaatttcag tatttggttt tagcttaaat agactgactt gagataacat 15060
aaggagaacc ttgaccccca agcaacatca tctcgcgagt tgactaggcc gggtgtggtg 15120
tctcacgcct gtaattccag cactttggga ggccacagca ggcagatcac ttgaggtcag 15180
gcattcgaga ccagcctggc caacatggtg aaacctcagc tctactaaag atacgaaaat 15240

tagcaggcat agtggcctgc acctgtaata ccaggcactc gcaggagaat cccttgaacc 15300
cggaaggcgg agattgcagt aaaccatgat tgtgccactg cactccagcc tgggcaacag 15360
gagactctgt ctcggaaaaa taaatttttt aaaaaaatga aaaaaaataa aagttgacta 15420
aattagtgtc ttggtactaa gcactgtagg aagtgagttt catggaaccc caactctctt 15480
ggggcccaaa gcaagtcata ttaatattga aaattacatg catatacatg catatgacca 15540
aggtgataaa aacaattatt ctgcctgagt tggagaatag tatcccagta aaataaacaa 15600
gagtctcaaa gtcttttgta tcctttgaag ctgtcatggt ggtttgtaac taggcaacag 15660

gtatatattg ttaatcttct ttgcatttaa ttccttttat agagagacac aattttacga 15720
gcagatgcaa ttactagcat gaaggtttct ttgtgagggt agttaaaagg cccacatgag 15780
ctctcttctt atccttgtcc ttctttcagc cagatcttcc ctgccccttt gctcattcca 15840
tctttcaccc acctaccccc aaaacaagga agtaaatctt gcattagtca acaataccaa 15900
agtgattttc aatatgactt tctctgcaga atgttattat ttctgcctct ttacattcac 15960
atactgtctt cctttttttt ttttttttt ttttttttt tagattgggt ctcactctgt 16020
tgcccaggct ggagtgcagt ggcttgatct cagctcactg taacctccac ctcctgagtt 16080

caagcaattc tcctgcctca gcctcctgag tagctgggat tacaggcatg tgccaccaca 16140

cctggctagt ttttttgtat ttttagtaga gacagggttt caccatgttg gtcaagctgg 16200

tctcgaactc ctgacctcat gatctgacca cctgtgcctc tcaaagtgct gggattacag 16260

gcgtgagcca ccgggccagc cactctcttc ctttcagttg cctactcatc tcttatgcat 16320

tcctggacat cagttgtcct tttgaagctt tcctccacta tcccagccca tgtgaatcct 16380

ccttccagtt atagccctta attctagatg gctgatattt ttcaataatt gttttaagat 16440

gaccatttta gcctatcagc taaacaatat caaagacaat agctattttt caagtacttt 16500

agtttacctt attatagagt gcataataga tattcagtaa atagtaaagg agaggtgaag 16560

gcttgcatag aatggattct ggtggtgtct cttggtgagc ttttagcatc aagattaatc 16620

agcagtttca gcaatgagct cagaccttca gttttagatc tttactcata tcagataaga 16680

gagtgagaag agtggtatgt atcagtgctt tatttatatt tgcatccaat ttgaactatg 16740

aatattacaa aggtgcacac ataggttcag acagattgat ttaaaatgac caaagatgac 16800

ctgtcgtaag caacctgggt atcttaagat gcactccttg gagagggaat gttcctaaaa 16860

acattttcag agggacgaac tgtatgaaat tcagtaaaac ataaatcatg aggaaaactg 16920

attactctct ttttgacatg aaatgagagt tttaatgcat ggttacgatt attaacgtac 16980

tccgctgcaa gacgttaata aagttactgt tttgcaggct agaatgtctt gatgctgtaa 17040

tcagaacaca ctttttcccc tttcttccag cttcaaatgc agattcataa ttgggctgac 17100

ttctaataac tgcaatgttt tctgccttgg gcttgcagca gaagcctgac aaaatagtgt 17160

ttgtttaggc aataatttat ttatttattt attgagatgg agtttcattc ttgtcgccca 17220

ggctggagtg caatggcgtg atctcggctc actgcaacct ctgtgttcag gcaataattt 17280

agactttacc ttacttgtga ttactatagc aattactata gccacaaggc ataattttac 17340

tgtctcattt caattttatg aatttgaatg tttttacact tttcctaatg aagtccacta 17400

tgaagttatg tcaaaaaaaa aaaagaaaaa gaaagatgca cacgtaaaag agaggtggtt 17460

gcaagagaag aaaagaacgg aggaaagtta aacgcaaacc agataactct cagcgtattc 17520

taaatgacca aaaacagaac tctgttgtca aagattttaa atggaaaatt tttcaatttt 17580

tttttctttt ttgtacaggt ttcttcctga acgcgcctca gcgattttaa atcgtatgca 17640

gaatattcaa tttgaagcag tggttggcca caaaatcaaa atgaaatgaa taaataagct 17700

ccagccagag atgtatgcat gataatgata tgaatagttt cgaatcaatg ctgcaaagct 17760

ttatttcaca ttttttcagt cctgataata ttaaaaacat tggtttggca ctagcagcag 17820

tcaaacgaac aagattaatt acctgtcttc ctgtttctca agaatattta cgtagttttt 17880

cataggtctg tttttccttt catgcctctt aaaaacttct gtgcttacat aaacatactt 17940

aaaaggtttt ctttaagata ttttatttt ccatttaaag gtggacaaaa gctacctccc 18000

taaaagtaaa tacaaagaga acttatttac acagggaagg tttaagactg ttcaagtagc 18060

attccaatct gtagccatgc cacagaatat caacaagaac acagaatgag tgcacagcta 18120

agagatcaag tttcagcagg cagctttatc tcaacctgga catattttaa gattcagcat 18180

ttgaaagatt tccctagcct cttccttttt cattagccca aaacggtgca actctattct 18240

ggactttatt acttgattct gtcttctgta taactctgaa gtccaccaaa agtggaccct 18300

ctatatttcc tccctttta tagtcttata agatacatta tgaaaggtga ccgactctat 18360

tttaaatctc agaattttaa gttctagccc catgataacc tttttctttg taatttatgc 18420

tttcatatat ccttggtccc agagatgttt agacaatttt aggctcaaaa attaaagcta 18480

acacaggaaa aggaactgta ctggctatta cataagaaac aatggaccca agagaagaaa 18540

aggaagaaag aaaggttttt tggtttttgt tttgttttgt tttgtttttt gttttttga 18600

gatggagtct cactctttcg cccaggctgg agtgcagtgg tatgatctca gctcactgca 18660

agctccacct cccgggttca cgccattctc ctgcctcagc ctcctgagta gctgggacta 18720

caggcgcccg ccaccacacc cggctaattt tttgtatttt ttgtagagac ggggtttcac 18780

catgttagcc aagatggtct cgatctcctg acctcgtgat ccacctgcct cggcctccca 18840

aagtgctggg attacgggtg tgagccaccg tgcccagcct ttttttttt aatagaaaaa 18900

ataatccgac tcccactaca tcaagactaa tcttgttttg tgtgtttttc acatgtatta 18960

tagaatgctt ttgcatggac tatcctcttg tttttattaa aaacaaatga ttttttaaa 19020

agtcacaaaa acaattcact aaaaataaat atgtcattgt gctttaaaaa aataacctct 19080

tgtagttata aaataaaacg tttgacttct aaactctg 19118

<210> 2

<211> 19119

<212> DNA

<213> Homo sapien

<400> 2

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60

aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120

tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180

ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240

gacagagcat attggttctg tgggatatta ataaggtaat gtatacatct tccaactttt 300

taaagtcaca gagtaagata tgtattttaa gaattatttg acttaccatc tacttatctt 360

tgtatttttg tttttcaaag tttgataaat tccctggtcc cttagtctgt atatgtgtca 420

ggttagttag atgaagggaa tgtaattaag aactaagcag cgatttttat gacatggtgt 480

gcaggttgat agaaagactc aggagccagt ctccttccaa gctgctaaat gaggcaagtc 540

acatattatc tctcagcctg ttttcttggc tctgaagtgg ggataataac ttaggggatg 600

ggcaagaacg ggatctgaaa attacagcta caaacaaaag tcaaacgaag aacttgcaac 660

agaaaccttt agtgcctccc ctcatgcaca agcaacacag ttctaaaata tttactgtct 720

gacccтttac agaaaatgtt tgccagtccg tagtcaaaag gattaaataa gtaatatttt 780

cagcacttag catatgataa acgatacgtg gcacatgata aacaataact gtgttaaata 840

aaatatgtgc gcagtgagtc aggcttttcc ttggacatta gtatttttcc tgtgttctta 900

cttgtaaaca ctacattaac aaccccaaat aaaactgaag gaactgaaat cttgtatcat 960

tttctctaaa cttgtaaatt ctggtaaggc catgaaaata tatgcagaga agtgtttaca 1020

ggattttagg attggaaaaa ttgtgaagta ctccttgaga atcacatttt ctgcaaatta 1080

cagtggtttt aattaccatt atattattac tttctcatgt tctttgctgt catgtttagt 1140

tgaaacctaa aatgtctctt acacttagag aactaattct tttctgtttt ttttctgaat 1200

agtgaagaat actatacaaa aaagctacta cattttтatt taacagatat gagcatttat 1260

ataatagagg agttgatgta tataaaaatg atttgccatc tttttggtct ttgaagaaat 1320

tcgaatgaac tttctggaag atagcaagaa tttacaaata gagaaaattg ttgcctgctg 1380

ttctcaggca tttgtccaaa aatataaata agtataaatc tatgaaaagg gcttgatgaa 1440

atctaacctt caaatctctt tccagatgtg tatttttggg gaaagggcta tatttattaa 1500

gtttttttta aattttaaaa tttccagaga caagagaaaa gtaaattaga aggaagtcgt 1560

attaaaaatg acttaagggc gggtgcagtg gctcacacct gtaatcccag cactttggga 1620

gacggaggtg ggcagattgc tggagcccag gagttcaaga ccagcctggg cagcacagca 1680

aaacccccaa ctctacaaaa aatacaaaaa ttagctgggt gcgggggtgc acacccgtag 1740

tcccagctac tcgggaggct gaggtgggag gatcgtttca gttcaggaag ccaaggctgc 1800

aatgagctat gatggcatca ttgcactcca agctgggcaa tagagccagg ctctgtctca 1860

aaaaaaataa aaaaagactt aagaaaaata ggtaacccaa cctcaaaaat tctctttgaa 1920

tcattaaatt tcatggttaa acatttaagc tactgaatga ttcactctaa ggctgtaatg 1980

taactcagat ctcctttagg cgaggaagat gctggctgag ttttcatcat aactggctcc 2040

ttttgccctg tgagatgaga gacacagtag cagtttggct cttatgcaat ctaaactgtt 2100

gcgttgggaa tacggttcaa aaaacacatt ggagtttaag ctaaagcaag tgttttgcta 2160

acaaaaagac aaggcatcac attttgcaat tgtctagctc agttataaaa cagaagaata 2220

ggccggacgc ggtggctcac gcctgtaatc ccagcacttt gggaggccga gacgggcgga 2280

tcacgaggtc aggagatcga gaccatcctg gataacacag tgaaaccccg tctctactaa 2340

aaatacaaaa aaattagcca ggcgtagtgg cgggcgcctg tagtcccagc tactcgggag 2400

gctgaggcag gagaatggtg tgaacccggg aggcggagct tgcagtgagc cgagatgacg 2460

ccactgcact ccagcctggg cgacagagcg agactccgtc tcaaaaaaaa aaaaaaaaaa 2520

aaaactgaag aataattaat tcttcaatca aaacatctga tgaatgctct ggtaacttat 2580

gctctctact gacctagaaa caaatgagag agtatggtgt ggtttgtgca atctggcagt 2640

gagcaagcta ccaactaaat cagtgaaaga ctctcctatt ctttttttac tcttctgcaa 2700

tcccacaaaa ggctatttga ggggatactg actttgagac tgggtcctaa catccatgtt 2760

tggggagttc aggctgctgc tccagggttt agcctacagt agcgaaatac aaaggaccca 2820

gagaccactc attcaaggtt tgccctaaat agcagcaaca ccactgtcat ctcaatacac 2880

gaagaatagg gcttttcagg tatccttgcc tctttgtcac agagaagagt ttacagattg 2940

tgagacggaa aagtataatt tttaaaacct tataatattt tctataaaag tcacctgagg 3000

tgaaaacttg aaaagaatta taattttcca gaatgtgagt caagaaacat tagagcaatt 3060

ttatcttagg aaagaggtct ttgaatttag gctgaaagta aattgctctg tctccatgtc 3120

ctatggttat gggcaagttt ggtacataaa tgagaaatcc atccagtggc cttgcccatc 3180

tcactcccaa acacctgaag aatgtaatgt tatatctcct agagtagcag catggtctcc 3240

ctatgaaagt ccttcttctt taaggagact tctttccctt ccctcctagg aggatgagtc 3300

agaatcatca agaaaaatat gatgggcaga ggcatacagt ttaccattac cactagttta 3360

gaattactac ttagcacttt actgcctatt acatagttgg tgctcaacaa atgtatgata 3420

aattaatggt tgagtttttc tttcttctcc atattcatct tccatgacac cacgaagagc 3480

aatgtttttc aagaatgttc ttcaaggttt gaaagtagcc tgctttagag aaactgccta 3540

ctgtacagcc tccaaccaag aggaaaagct gaaaaaagca tgaagggatt ttgttttgtt 3600

ttgtttgttt tggttttaat atgagcattc cctggcagaa aagccagggg taatctcatt 3660

gcaactaggc aatcactctc aagaaatttt ctaacaaata aggaggccaa tttttatttt 3720

attttgagac gaagtcccac tctgtcaccc aggttggagt gcaatggaat gatttcagct 3780

cactgcaacc tccgcctccc gggttcaagt gattctcctg tctaaacttc ccgagtagct 3840

gggattacag gctcccacca ccacgcccag ctaatttttt gtatttttag tagagatggg 3900

gtttcaccat tttggccaga ctggtctcaa actcctgacc tcaagtgatc caccctcctc 3960

ggcctcctaa agtgctggga ttacaggcgt gagccaccac acctgaccca ggaggccaat 4020

ttttaaaagg ttaactaatc ttcatgtcca aaatgaatgt taattgttca ttttggacat 4080
gaatgttaat ttttttttt ttttttttg agacagagtc tcactctgtt gcccaggctg 4140
gagtccagtg gcactatctc cactcactgc aacttcctcc tcccaggttc aagcaattat 4200
cctgcctcag cctcccaagt agctgggatt acaggcccac accatcaggc ctggctaatt 4260
tttgtatttt tagtagagac ggggtttcac catgttggcc aggctggtct tgaactcctg 4320
acctcgtgat ccgccctcct cggccaacca aagtgctggg attacaggcg tgagccaccg 4380
cgcctagccg aatgttaatt gtctaaaaat ttttcttctc caatgtcttc tcctccactt 4440

ttttcggaat ttgtttcttc ctaattacag cgcggtgtgg aggaaactgc agctgagtgc 4500
cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 4560
tatcgctctc taaatcaggt gagactgcag gttcacaaat ttcttcagat tattttgttt 4620
cctaggacgc tgacgtggaa aatgagaaag gtctttatga ctgcctgatt taaattggat 4680
tttagctgct aactgaagta gttatgtcac caaggaagga tatatacttt ttttcttgta 4740
tgtaatccac tcagctctgc ccattattat tgttcatatt attaatcaat ttcattctga 4800
tcagaagtgt gagcagtggc acagagtgac tgacaaaaga tttatcatca gggaatatgg 4860

atcacttcct agttttgttt tagtcctatt aactttgcag taattccagc ttctctttaa 4920
ttatttccct tgtgagattt tattttggtg ttaatgtagt cttctgtaga aaatgtaata 4980
ttaataatta ttatcacaat tattttaaaa gagtaaatac caaataatca caatgaacta 5040
agcactctaa caaactttac atttttttaat tcaatcccta caataactct gtaaacttca 5100
ttttacagat aagcaaatta tgactcagag aggttaagcc agacccaggt catgtagtta 5160
ttaggttatg aaaccaggat ttctcaacca gcactttaga ccaggtgcgg tggttcacac 5220
atgtaatccc agcactttgt gaggccaagg tggaaggatc acatgagacc aagagttcaa 5280

gaccagccca ggcaacatag tgagacccta tctctaaaaa aaaaaaaaaa aaaaaaaaaa 5340
aaagtttaaa gaaaaacaca tttttaaaaa atgaacactt taaaaatatt tggtcagaat 5400
ttatatagga atttatcaac ataaatgtta atttcacttt actgataaac ttgcaaaaca 5460
tgatgtgctg ggtactgaaa tttagatgtt aaaagaacag tttatcccac ctttatgaca 5520
gtgttcccct ggcctccacg atttgagctc aacagtctgt cttgcctgaa ctctgagaga 5580
cctcatacaa tagaagaaag actctcatct ttggattata ttggtcccaa aactttgagt 5640
ttgaataata cacccagtga aagtgttctt tcaatttcaa aaggtgaaga aagaagtggg 5700

tgatgtaaca atcgtggtga ataatgctgg gacagtatat ccagccgatc ttctcagcac 5760

caaggatgaa gagattacca agacatttga ggtcaacatc ctaggacatt tttgggtgag 5820

tgtgagtcag aaacatttct gatttgtgca ccttctctta agatacatga aacttataac 5880

ggagttcaca tacttctgga caggaaactg gccagatctt tgccttaatc aagaatcatt 5940

aaatttgttt gagtagaaga gccacagagt ctctgacaca aggacacaga attcaagtgg 6000

acacaacaca ccagaatgta agctacttgg tctgtcttgt ccaccagtat ctgacacaaa 6060

gcttggcatg taccaggagc tcaacaaatg tttgtggagg tttgttaagg gttgtcagtg 6120

tacatctttt caatgctgtc acttgtgact tcatttttt ccctccacac catgattttg 6180

taatgtgtcc tcattttgtg gaattttaga atggaaagga catcagaagt aattacttgg 6240

atgtatatag gatcgaggac acttttggac gagactctga ggcaagtgtt ctagatccat 6300

ggggtgctgg aactgagaaa tgcagctata cagacctcat ataattggtt agttttgtgg 6360

gagatggaaa tatcaacttc aactgccttt gtatagaaat ttttatgatt aatcttccag 6420

tgcctcaata ttagtgtaga atctagggca gatctggatt ctagaagaaa gaagaaaaaa 6480

aagagatgtg tccccccttac ctttaccagc tcttcacata tgtgaattgg ctcccatgcc 6540

caccaaacta cacggagacc tcatacatta gctacctata gctgcataac aaattataca 6600

aaacttagtg gtttaaagca acaatgtatg ttcactatcc tctcacagtt tctatgggtt 6660

gggaatttgg aggtagcttg ggttgggagt tctagttcta tgaatttgca taggatttat 6720

taaattctta taaaatttta ttgatgtttc tcacaaaaga ggtttttgga aaaaaagaaa 6780

gacttgtttt ctgtaacatc aacatataat atacaatatt acaaataggg agatagtgaa 6840

ttcaatcatg attcattagt gtggtgtaga actctcagct tacactactc aactgtctta 6900

atacagttac acaagatttc actcttttaa ttagaatgat aaagccccaa accaaaaaat 6960

tatatgacac caaattatca taaggaataa ttttagttct gaaaactctg aatttttccc 7020

ttaatattgt ttagatgaca tatccaaaaa aggatctatt tgattccttc tgaagggaag 7080

gaggggggagt actgagatta gtgttggcat ggggcttacc ataccaataa atttgtatct 7140

ttatttctat catttgtaaa gaattaatca tggaatgctt ggaagtattt tatttcattg 7200

tataagttct ctcaaatgcc tttctgtctt aacaaaaata aaactacctg atttggaaac 7260

ctaacgtcta tgtcattgtc tttcttcttt ctgcaatgat ccttaagatc acaaaagcac 7320

ttcttccatc gatgatggag agaaatcatg gccacatcgt cacagtggct tcagtgtgcg 7380

gccacgaagg gattccttac ctcatcccat attggtaagt atcacatgcc agccatgtta 7440

tatattttta tactttgaag ggagcattac acttcaaatt gttaccactg gagagtcctg 7500

gttcttggca tcttgaacaa agaattggac aaaactcacc aacaaagcca ggaaagaatg 7560

aagcaacaaa agcagagatt tattgaaaat gaaagtacgc tttacagggt gggagtgggc 7620

ccaagcacag gggctcaaga gccaattaca gaatttttctg gggtttaaat acccccctaga 7680

ggtttccact ggttacttgg tgtacgccct atgtaaatga agaggatgaa ttaaagttac 7740

agagtcgttt actcagtgta caccatatgt aaatggagag gatatttcct gtcatagctg 7800

gagtgtttcc atttgattta gttctaggaa gtcagcatga atcggcctta tgttccctgc 7860

ctccagaccc tgttctcctg cctcaagatt acaatgctga gagcagagtg atttggattt 7920

acagaattta aatttatagt agtttagaat gatttttttaa atgacttttt ctaaaacaat 7980

gaaaccaggt tgtaattata tttaagatat ttttagattt ctgcaaactc ctctgtagaa 8040

caatgagaga aaacagtaat gccaagcatg tttccattgt ttcctggaat aagaaacaga 8100

aaccccacag actgagaagc aaaacctaca gaagctaaaa tgaacacatg tctatgtcat 8160

ggccttggtg cccaagataa gacaatcaga gtggtccctg gatcaaaaca ttttacagtg 8220

tgcttgtgcc atgaaagtgt gtgtgtgtgt gtgtgtgtgt gtgtgagaga gagagagaga 8280

gagaaaacga ctctacctga ctaaaagttg cagataccac actccatgca ccaccaaaga 8340

cataaaggga aggaggtgag aggcgttaag gatgtactgc tgtatttgcc aaatatcctt 8400

tcctgtaaac tcttctccag atcctcataa taaaattaag aggccaaagt ggcaaccatt 8460

gtcaagagaa aaactatcaa ccattgtcaa gagaataact cagttattga gagagagagg 8520

agaaatgagc agagtcctac agaagtctgt caacacagat accagttttg tagaatttct 8580

aaatgtattt ttcctgattc atatttttca aaataaaagc agcaataaaa actgattaga 8640

aaacagtttg aagattcaat ggaaaaacct tacatgtagg atggaaaact gaacattaag 8700

ccaatcaata gagttatttt tgttcttttg ttatcattgt tgtttaagaa atgagatacg 8760

ttcacaattc tgcttaatca tgtaagaaaa tgaaaatgaa ttgccattta tactctcaga 8820

aaaatcacaa gtggctgatt tttggcttcc acttgttctt aagccaaatg ataccgcctt 8880

ctcacagaaa gctgaggatt ggtttcactc tcccttagct aacaatgctt aataattctc 8940

ttacagttcc agcaaatttg ccgctgttgg ctttcacaga ggtctgacat cagaacttca 9000

ggccttggga aaaactggta tcaaaacctc atgtctctgc ccagtttttg tgaatactgg 9060

gttcaccaaa aatccaagca caaggtaagg tcaaaatcaa gttagaatgg gtatgtggta 9120

tgataaattg atatgaaaac taatgagaaa tgtttaggca ggccaactaa tagaagaaaa 9180

tgaagaagga aaaataattt ttcttattat tattattatc ttgaaattaa aggaataaag 9240

ggggaaaaca cattagggac tagcaggaat gatcagccac cgatgaactg ggatatttat 9300

ttgtgtccgg gagaaagcac atacatttga tcaccgttac caccctgtct ttaaaatgca 9360

aatgttccaa ggaccagcaa ataaattgag tatctagctc cttagtcaag gtgaatttct 9420

gcaagaactc ttgtctctgg tgagacagga tttgagacca caagagaaga aaaattagtc 9480

ctgaaaggag aagaaaaaag caggaaggtg tggataagaa cccgaaaatt aagccatctg 9540
cttaacaaat ttttctaatc ctagtatata ttctgctgca ggttaacaaa atatactaag 9600
cttaatgatt cgaaaccaat tttttactgg aagggaatta atcctaaata tattcattca 9660
aaagaactaa acaattctct gttgagtgcc gcctcatttg aggatactga ctcttacagc 9720
ctgagttagc tatgtggtct ctgcagctgg aatcactccc tgccactgga gtccttcatg 9780
gtgttagacc ataggtactg ttgactaaag aaaaaaaaaa gtttttgttt ttatttttgt 9840
ttttttgag acagagtctc actctgtcac ccaggctgga gtacagtggc gcgatctcag 9900

ctcaccgcaa cctccgcctt tctgggttca agcaattctc cttcctcagc ctcctgagta 9960
tttggattac aggcgcccac caccacgcct ggctaatttt tgtattttta gtagagacgg 10020
ggtttcacca tgttggccag gctggtctca aactcctgac ctcaggtgtc ctacctgcct 10080
tggcctccta aaatgctggg attacaggag tgagccacca tgcccggcca aaaaaataag 10140
tttttaaaga attaaaggtc atcctggcta acacagtgaa accccgtctc tactaaaaaa 10200
cacaaaaaaa ttagccgggc gtggtggcgg gcgcctgtag tcccagctgc gcgggaggct 10260
gaggcaggag aatggcgtga acccgggagg cggagcttgc agtgagccga gatcgcgcca 10320

ctgcactcca gcctgggcga cagagcgaga ctccgtctca aaaaaaaaaa aaaaaaaaaa 10380
aaaaaaaaga attaaaggtg ttaattttat ttagaagcct tactgaagac tacagtcgga 10440
ggcctatagc ctgagagcag ccctttagag aggttcagtt gaactgttct gatagtgggg 10500
gccatgtgct ctatcctgta ttgtcttcaa agcatctttc cagagagctg cacattgtca 10560
cagagtcagg gactttgtga aattatgctg acaaccagaa gtgagtaaac gtggcttctt 10620
acatttgcta cgttgtctca cagtacttaa taagtatgca atatgtaagt aaatactata 10680
gtactattgc aactcctgat tgttttctta gacaaggaat tgggcccaat aaaaaccctc 10740

ttggtaggca ttcaggcttc gtgtaccatg agctttccta agggtatcct gccactcttg 10800
gggaaggcat gatagatgag gggagtaagg ataatggaac tctgggtaca gggttcctgg 10860
gggctaactt agaggtagac acaggcaatg ctaaatattt gggattgatt ttatagaggt 10920
tgctagattg tgaatttcct tagtaagggc taaggcattg atatgtaatg tcacacttgg 10980
ctccgaggct gggttgttgg atccatgtag atgaaatcag ggagagaaag ggcagaacgg 11040
agtaatttag aaatgtattg atttgtatta ctctctgttg gcttgctatt caaggcagtg 11100
gagaactcaa tcacataata atctgcagca aaccacagat catcccaggg aatgaagttt 11160

taacattcgc tggctcccta actcctcacc cagcctttac attcactggc tgttcagtcc 11220

atgcctggac atcttaattt gaatacaaca ttttaaatcc atttttctgt catcatcttg 11280

cactaacaga caattctaca ctaagcctat gtttatgaat atttctcaag agtacatgta 11340

cacagccttc agtataagga aaactggaag tatgacatac ctccagttgt catactcctt 11400

gggcccctct taaattctca ttaaactgca ggataggcaa gtcagaggtg aatctcaaat 11460

acgaaattct taccggaaag gggttccaat ccagacccca agagagggtt cttagatttc 11520

tcgcaagaaa taattcgggg caaggccaca gtgcaaagca aaagcaagtt tattaggaaa 11580

gtaaaggagt agagaacagc tactccatgg agaagaatgg cttgagctgc tccaccaagg 11640

gtatttagag ttacttcttg attatatgct aaacaagggg tggattattc atgagttttc 11700

cgggaaaagg gtgagcaatt cccagaactg agatttcctc ccctttttag gccatatagg 11760

gtaacttcct gccattgcca tggtatttgt aaactgtcat agtgctggtg gaagtgtctc 11820

ttagcttgct aatgtattat agttagctta taatgagcag tgaggacaac cagaggtcac 11880

tttcatcacc atcttggttt tggtgggttt tggccggctt ctttactgca ccctatttta 11940

tcaacaaggt ctttatgacc tgaatcttgt gccaacctcc tatctcatcc tgtgacaaag 12000

aatgccttaa cttcctggga atgcagccca gtaggtgtca gccttatttt acccagaccc 12060

tattcaagat ggagttgctc tgatttaaac gcctctgaca aaatgacgac ctcaaaacaa 12120

tccagcttta tggaatacct ccacaagaaa gaaagtatac ttagctatag aattttctcc 12180

ttgcatccaa caggctttga gatgtcagat gtttccttcc tgtcccatga ttaatcctag 12240

ccattcctct ttcttgtctg gctccactac tccttaccat ctaatgcctc gccaccattt 12300

tgatattttg actaagtgag ctatgaaaca cacctactgg atatgaaagt ataagtttct 12360

gataacaaaa catcaacatg ggatgtggag gaagtgggta gggtggcatt aatgcagcaa 12420

atcctggaat attttaaatc ttcattctaa atttagtaaa aatataggat aattttcctg 12480

ccatcattta cttataaaat taaaatttta gaaaataaaa ataatatttt cctcttttta 12540

atcacagatt atggcctgta ttggagacag atgaagtcgt aagaagtctg atagatggaa 12600

tacttaccaa taagaaaatg atttttgttc catcgtatat caatatcttt ctgagactac 12660

agaagttaag tacagcacag aacacccaaa tactaaaaca ccaatagagc ttttttttttt 12720

gctttttttt tttttagaca gagtctcact ctgtcaccct ggctggattg cggtggttgc 12780

agtggcatga tcttggctca ctgcaacctc cgcctcctgg gttcaagcaa ttctcatgcc 12840

tcagaccccc aagtaactgg gattataggt gtgtgctgcc acactacacc cagctaattt 12900

ttgtattttt tgatagagac aggtttcccc atgttggcca ggctggactc gaactcctga 12960

cctcaagtta tcctcctgtc tcggcctccc aaagtgctgg gattacagtc atgagccacc 13020

atgcctggcc caatagagct attattatgg agcatctttc agttgtgaaa attggcatgg 13080

aaactctcca tccctgggga gaacagttat ttcctctgtt attttcctac ccagtctata 13140

aaaagagagt gattcatttt ctctaccaaa tctactgtct ctgcccaaac tttgctgaag 13200

actattctaa ctaaaggaaa cacagtttaa aaagaatgca atatagtgaa gtagttaata 13260

ataaagactc catttttaaa agtctgctgg aagtttggtt gggattgcac tgaatctata 13320

gagcaattgg ggagtattga catatcaaca atattgagtt ttctaatcca agaacataat 13380

atctattttt aaaatcttct tcaaaatctt taaatcttta aattgtattt tgtagttttt 13440

ggtgtttaag tcttgcacat attttgtcag atttattcca aagtatttca cgggttcttt 13500

ttttttttt ttttttttt tttgagacag agtttcaccc ttgttgccca ggctggagtg 13560

cagtggcgtg atcttggctc actgcagctt ctgcctcctg gcttcaagtg attctcctgc 13620

ctcagcctcc caagtagctg ggattacagg cacctgcccc ctcgcccaac taacttttg 13680

tgtttgtagt agagacaggg tttcaccatg ttggccaggc tggtctcgaa ctcctgacct 13740

catgtgatcc acctgcctca gcctcccaaa gtgctgggat tacaggcatg agccatcatg 13800

cccagcccta tttgacggtt tttgacgcta atgcaagtgg cattttaaaa aattttatat 13860

ttcccattgt ttgttgtcag tatatattgg atttttgtaa tttgatctca tattttgcag 13920

tcttgctaaa ttgctaaacc tctttttgct aaactcgata agctttttt tttttggtag 13980

attcctgggc ctctaatttt ctttatggga aagtttttaa ttacaaattt aatttcttta 14040

atagctacat ggctattcaa tttacttatt aattcttggt aatgtgtgtc tttcaaggaa 14100

tttgtccatt tcatctaagt tgtagaattt ctttggcata aatttgtaca taacattccc 14160

ttattatcct tttaatgtct ttagaatgtc ttatttattt atttatttat ttttattata 14220

ttttttttgag acagagtctc gctctgttgc ccaggctgga gtgcagtggc acaatcttgg 14280

ctcactgcaa gctccgcctt ctgggttcat gccattctcc tgcctcagcc tccctagttg 14340

ctgggactac aggcgcctgc aaccatgccc agcttatttt ttttttttt ttttttttt 14400

ttttttttt ttttttttt ttagtagaga cggggtttca ccctgttagc caggatggtc 14460

tcgatctcct gacctggtga tccgcccgcc tcagcctccc aaagtgctgg gattacaggc 14520

gtgagccacc aagcccagcc tatttattta tttagtagag acagtctcac tttgctgccc 14580

aggcaacaaa ggttttgaat gcctggcctc aagcagtcct cctgccttgg cctcccaaag 14640

tgctgggatt acaggcatga gccactgcac ctggccaaat gaatatgctg ataatatctt 14700

ctttataagg atgacataag aataaaataa tgtaatacaa acaaagcccc tgtcactgaa 14760

aatgtataga cttcaaatgt taaagtctta gagaacagaa tttatatgaa atagcaacag 14820

caacaatttc ccagaggaaa tactctctca gctttcttct gaggagcagt ttctaaattg 14880

aaattgtatc agtgagaaga taactatact aacttcataa gccttgggcc tttttgaaac 14940

aaatccatat aaactatgaa caaacttgaa atagaacaat ttgagaacag ggtacaaact 15000
gcattggtgt atcaatttca gtatttggtt ttagcttaaa tagactgact tgagataaca 15060
taaggagaac cttgaccccc aagcaacatc atctcgcgag ttgactaggc cgggtgtggt 15120
gtctcacgcc tgtaattcca gcactttggg aggccacagc aggcagatca cttgaggtca 15180
ggcattcgag accagcctgg ccaacatggt gaaacctcag ctctactaaa gatacgaaaa 15240
ttagcaggca tagtggcctg cacctgtaat accaggcact cgcaggagaa tcccttgaac 15300
ccggaaggcg gagattgcag taaaccatga ttgtgccact gcactccagc ctgggcaaca 15360

ggagactctg tctcggaaaa ataaattttt taaaaaaatg aaaaaaaata aaagttgact 15420
aaattagtgt cttggtacta agcactgtag gaagtgagtt tcatggaacc ccaactctct 15480
tggggcccaa agcaagtcat attaatattg aaaattacat gcatatacat gcatatgacc 15540
aaggtgataa aaacaattat tctgcctgag ttggagaata gtatcccagt aaaataaaca 15600
agagtctcaa agtcttttgt atcctttgaa gctgtcatgg tggtttgtaa ctaggcaaca 15660
ggtatatatt gttaatcttc tttgcattta attccttta tagagagaca caattttacg 15720
agcagatgca attactagca tgaaggtttc tttgtgaggg tagttaaaag gcccacatga 15780

gctctcttct tatccttgtc cttctttcag ccagatcttc cctgcccctt tgctcattcc 15840
atctttcacc cacctacccc caaaacaagg aagtaaatct tgcattagtc aacaatacca 15900
aagtgatttt caatatgact ttctctgcag aatgttatta tttctgcctc tttacattca 15960
catactgtct tccttttttt ttttttttt ttttttttt ttagattggg tctcactctg 16020
ttgcccaggc tggagtgcag tggcttgatc tcagctcact gtaacctcca cctcctgagt 16080
tcaagcaatt ctcctgcctc agcctcctga gtagctggga ttacaggcat gtgccaccac 16140
acctggctag ttttttttgta tttttagtag agacagggtt tcaccatgtt ggtcaagctg 16200

gtctcgaact cctgacctca tgatctgacc acctgtgcct ctcaaagtgc tgggattaca 16260
ggcgtgagcc accgggccag ccactctctt cctttcagtt gcctactcat ctcttatgca 16320
ttcctggaca tcagttgtcc ttttgaagct ttcctccact atcccagccc atgtgaatcc 16380
tccttccagt tatagccctt aattctagat ggctgatatt tttcaataat tgttttaaga 16440
tgaccatttt agcctatcag ctaaacaata tcaaagacaa tagctatttt tcaagtactt 16500
tagtttacct tattatagag tgcataatag atattcagta aatagtaaag gagaggtgaa 16560
ggcttgcata gaatggattc tggtggtgtc tcttggtgag cttttagcat caagattaat 16620

cagcagtttc agcaatgagc tcagaccttc agttttagat ctttactcat atcagataag 16680

agagtgagaa gagtggtatg tatcagtgct ttatttatat ttgcatccaa tttgaactat 16740

gaatattaca aaggtgcaca cataggttca gacagattga tttaaaatga ccaaagatga 16800

cctgtcgtaa gcaacctggg tatcttaaga tgcactcctt ggagagggaa tgttcctaaa 16860

aacattttca gagggacgaa ctgtatgaaa ttcagtaaaa cataaatcat gaggaaaact 16920

gattactctc tttttgacat gaaatgagag ttttaatgca tggttacgat tattaacgta 16980

ctccgctgca agacgttaat aaagttactg ttttgcaggc tagaatgtct tgatgctgta 17040

atcagaacac acttttttccc ctttcttcca gcttcaaatg cagattcata attgggctga 17100

cttctaataa ctgcaatgtt ttctgccttg ggcttgcagc agaagcctga caaaatagtg 17160

tttgtttagg caataattta tttatttatt tattgagatg gagtttcatt cttgtcgccc 17220

aggctggagt gcaatggcgt gatctcggct cactgcaacc tctgtgttca ggcaataatt 17280

tagactttac cttacttgtg attactatag caattactat agccacaagg cataatttta 17340

ctgtctcatt tcaattttat gaatttgaat gtttttacac ttttcctaat gaagtccact 17400

atgaagttat gtcaaaaaaa aaaaagaaaa agaaagatgc acacgtaaaa gagaggtggt 17460

tgcaagagaa gaaaagaacg gaggaaagtt aaacgcaaac cagataactc tcagcgtatt 17520

ctaaatgacc aaaaacagaa ctctgttgtc aaagatttta aatggaaaat ttttcaattt 17580

ttttttcttt tttgtacagg tttcttcctg aacgcgcctc agcgatttta aatcgtatgc 17640

agaatattca atttgaagca gtggttggcc acaaaatcaa aatgaaatga ataaataagc 17700

tccagccaga gatgtatgca tgataatgat atgaatagtt tcgaatcaat gctgcaaagc 17760

tttatttcac atttttcag tcctgataat attaaaaaca ttggtttggc actagcagca 17820

gtcaaacgaa caagattaat tacctgtctt cctgtttctc aagaatattt acgtagtttt 17880

tcataggtct gtttttcctt tcatgcctct taaaaacttc tgtgcttaca taaacatact 17940

taaaaggttt tctttaagat attttatttt tccatttaaa ggtggacaaa agctacctcc 18000

ctaaaagtaa atacaaagag aacttattta cacagggaag gtttaagact gttcaagtag 18060

cattccaatc tgtagccatg ccacagaata tcaacaagaa cacagaatga gtgcacagct 18120

aagagatcaa gtttcagcag gcagctttat ctcaacctgg acatatttta agattcagca 18180

tttgaaagat ttccctagcc tcttccttt tcattagccc aaaacggtgc aactctattc 18240

tggactttat tacttgattc tgtcttctgt ataactctga agtccaccaa aagtggaccc 18300

tctatatttc ctccctttt atagtcttat aagatacatt atgaaaggtg accgactcta 18360

ttttaaatct cagaatttta agttctagcc ccatgataac ctttttcttt gtaatttatg 18420

ctttcatata tccttggtcc cagagatgtt tagacaattt taggctcaaa aattaaagct 18480

aacacaggaa aaggaactgt actggctatt acataagaaa caatggaccc aagagaagaa 18540

aaggaagaaa gaaaggtttt ttggtttttg ttttgttttg ttttgttttt tgttttttg 18600

agatggagtc tcactctttc gcccaggctg gagtgcagtg gtatgatctc agctcactgc 18660

aagctccacc tcccgggttc acgccattct cctgcctcag cctcctgagt agctgggact 18720

acaggcgccc gccaccacac ccggctaatt ttttgtattt tttgtagaga cggggtttca 18780

ccatgttagc caagatggtc tcgatctcct gacctcgtga tccacctgcc tcggcctccc 18840

aaagtgctgg gattacgggt gtgagccacc gtgcccagcc ttttttttt taatagaaaa 18900

aataatccga ctcccactac atcaagacta atcttgtttt gtgtgttttt cacatgtatt 18960

atagaatgct tttgcatgga ctatcctctt gtttttatta aaaacaaatg attttttaa 19020

aagtcacaaa aacaattcac taaaaataaa tatgtcattg tgctttaaaa aaataacctc 19080

ttgtagttat aaaataaaac gtttgacttc taaactctg 19119

<210> 3

<211> 900

<212> DNA

<213> Homo sapien

<400> 3

augaacauca uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60

ucguugguga aguuuuucau uccucagagg agaaaaucug uggcugggga gauuguucuc 120

auuacuggag cugggcaugg aauaggcagg cagacuacuu augaauuugc aaaacgacag 180

agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240

cgaaaacuag gcgucacugc gcaugcguau gugguagacu gcagcaacag agaagagauc 300

uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaaucgu ggugaauaau 360

gcugggacag uauauccagc cgaucuucuc agcaccaagg augaagagau uaccaagaca 420

uuugagguca acauccuagg acauuuuugg aucacaaaag cacuucuucc aucgaugaug 480

gagagaaauc auggccacau cgucacagug gcuucagugu gcggccacga agggauuccu 540

uaccucaucc cauauuguuc cagcaaauuu gccgcuguug gcuuucacag aggucugaca 600

ucagaacuuc aggccuuggg aaaaacuggu aucaaaaccu caugucucug cccaguuuuu 660

gugaauacug gguucaccaa aaauccaagc acaagauuau ggccuguauu ggagacagau 720

gaagucguaa gaagucugau agauggaaua cuuaccaaua agaaaaugau uuuuguucca 780

ucguauauca auaucuuucu gagacuacag aaguuucuuc cugaacgcgc cucagcgauu 840

uuaaaucgua ugcagaauau ucaauuugaa gcagugguug gccacaaaau caaaaugaaa 900

<210> 4

<211> 792

<212> DNA

<213> Homo sapien

<400> 4

augaacauca uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60

ucguugguga aguuuuucau uccucagagg agaaaaucug uggcugggga gauuguucuc 120

auuacuggag cugggcaugg aauaggcagg cagacuacuu augaauuugc aaaacgacag 180

agcauauugg uucuguggga uauuaauaag gugaagaaag aaguggguga uguaacaauc 240

guggugaaua augcugggac aguauaucca gccgaucuuc ucagcaccaa ggaugaagag 300

auuaccaaga cauuugaggu caacauccua ggacauuuuu ggaucacaaa agcacuucuu 360

ccaucgauga uggagagaaa ucauggccac aucgucacag uggcuucagu gugcggccac 420

gaagggauuc cuuaccucau cccauauugu uccagcaaau uugccgcugu uggcuuucac 480

agaggucuga caucagaacu ucaggccuug ggaaaaacug guaucaaaac cucaugucuc 540

ugcccaguuu uugugaauac uggguucacc aaaaauccaa gcacaagauu auggccugua 600

uuggagacag augaagucgu aagaagucug auagauggaa uacuuaccaa uaagaaaaug 660

auuuuuguuc caucguauau caauaucuuu cugagacuac agaaguuucu uccugaacgc 720

gccucagcga uuuuaaaucg uaugcagaau auucaauuug aagcaguggu uggccacaaa 780

aucaaaauga aa 792

<210> 5

<211> 783

<212> DNA

<213> Homo sapien

<400> 5

augaacauca uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60

ucguugguga aguuuuucau uccucagagg agaaaaucug uggcugggga gauuguucuc 120

auuacuggag cugggcaugg aauaggcagg cagacuacuu augaauuugc aaaacgacag 180

agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240

cgaaaacuag gcgucacugc gcaugcguau gugguagacu gcagcaacag agaagagauc 300

uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaaucgu ggugaauaau 360

gcugggacag uauauccagc cgaucuucuc agcaccaagg augaagagau uaccaagaca 420

uuugagguca acauccuagg acauuuuugg aucacaaaag cacuucuucc aucgaugaug 480

gagagaaauc auggccacau cgucacagug gcuucagugu gcggccacga agggauuccu 540

uaccucaucc cauauuguuc cagcaaauuu gccgcuguug gcuuucacag aggucugaca 600

ucagaacuuc aggccuuggg aaaaacuggu aucaaaaccu caugucucug cccaguuuuu 660

gugaauacug gguucaccaa aaauccaagc acaagguuuc uuccugaacg cgccucagcg 720

auuuuaaauc guaugcagaa uauucaauuu gaagcagugg uuggccacaa aaucaaaaug 780

aaa 783

<210> 6

<211> 822

<212> DNA

<213> Homo sapien

<400> 6

augaacauca uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60

ucguugguga aguuuuucau uccucagagg agaaaaucug uggcugggga gauuguucuc 120

auuacuggag cugggcaugg aauaggcagg cagacuacuu augaauuugc aaaacgacag 180

agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240

cgaaaacuag gcgucacugc gcaugcguau gugguagacu gcagcaacag agaagagauc 300

uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaaucgu ggugaauaau 360

gcugggacag uauauccagc cgaucuucuc agcaccaagg augaagagau uaccaagaca 420

uuugagguca acauccuagg acauuuuugg aucacaaaag cacuucuucc aucgaugaug 480

gagagaaauc auggccacau cgucacagug gcuucagugu gcggccacga agggauuccu 540

uaccucaucc cauauuguuc cagcaaauuu gccgcuguug gcuuucacag aggucugaca 600

ucagaacuuc aggccuuggg aaaaacuggu aucaaaaccu caugucucug cccaguuuuu 660

gugaauacug gguucaccaa aaauccaagc acaagauuau ggccuguauu ggagacagau 720

gaagucguaa gaagucugau agauggaaua cuuaccaaua agaaaaugau uuuuguucca 780

ucguauauca auaucuuucu gagacuacag aagguuucuu cc 822

<210> 7

<211> 972

<212> DNA

<213> Homo sapien

<400> 7

augaacauca uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60

ucguugguga aguuuuucau uccucagagg agaaaaucug uggcugggga gauuguucuc 120

auuacuggag cugggcaugg aauaggcagg cagacuacuu augaauuugc aaaacgacag 180

agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240

cgaaaacuag gcgucacugc gcaugcguau gugguagacu gcagcaacag agaagagauc 300

uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaaucgu ggugaauaau 360

gcugggacag uauauccagc cgaucuucuc agcaccaagg augaagagau uaccaagaca 420

uuugagguca acauccuagg acauuuuugg aauggaaagg acaucagaag uaauuacuug 480

gauguauaua ggaucgagga cacuuuugga cgagacucug agaucacaaa agcacuucuu 540

ccaucgauga uggagagaaa ucauggccac aucgucacag uggcuucagu gugcggccac 600

gaagggauuc cuuaccucau cccauauugu uccagcaaau uugccgcugu uggcuuucac 660

agaggucuga caucagaacu ucaggccuug ggaaaaacug guaucaaaac cucaugucuc 720

ugcccaguuu uugugaauac uggguucacc aaaaauccaa gcacaagauu auggccugua 780

uuggagacag augaagucgu aagaagucug auagauggaa uacuuaccaa uaagaaaaug 840

auuuuuguuc caucguauau caauaucuuu cugagacuac agaaguuucu uccugaacgc 900

gccucagcga uuuuaaaaucg uaugcagaau auucaauuug aagcaguggu uggccacaaa 960

aucaaaauga aa 972

<210> 8

<211> 852

<212> DNA

<213> Homo sapien

<400> 8

augaacauca uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60

ucguugguga aguuuuucau uccucagagg agaaaaucug uggcugggga gauuguucuc 120

auuacuggag cugggcaugg aauaggcagg cagacuacuu augaauuugc aaaacgacag 180

agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240

cgaaaacuag gcgucacugc gcaugcguau gugguagacu gcagcaacag agaagagauc 300

uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaaucgu ggugaauaau 360

gcugggacag uauauccagc cgaucuucuc agcaccaagg augaagagau uaccaagaca 420

uuugagguca acauccuagg acauuuuugg aucacaaaag cacuucuucc aucgaugaug 480

gagagaaauc auggccacau cgucacagug gcuucagugu gcggccacga agggauuccu 540

uaccucaucc cauauuguuc cagcaaauuu gccgcuguug gcuuucacag aggucugaca 600

ucagaacuuc aggccuuggg aaaaacuggu aucaaaaccu caugucucug cccaguuuuu 660

gugaauacug gguucaccaa aaauccaagc acaagauuau ggccuguauu ggagacagau 720

gaagucguaa gaagucugau agauggaaua cuuaccaaua agaaaaugau uuuuguucca 780

ucguauauca auaucuuucu gagacuacag aaguuaagua cagcacagaa cacccaaaua 840

cuaaaacacc aa 852

<210> 9

<211> 714

<212> DNA

<213> Homo sapien

<400> 9

augaacauca uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60

ucguugguga aguuuuucau uccucagagg agaaaaucug uggcugggga gauuguucuc 120

auuacuggag cugggcaugg aauaggcagg cagacuacuu augaauuugc aaaacgacag 180

agcauauugg uucuguggga uauuaauaag gugaagaaag aaguggguga uguaacaauc 240

guggugaaua augcugggac aguauaucca gccgaucuuc ucagcaccaa ggaugaagag 300

auuaccaaga cauuugaggu caacauccua ggacauuuuu ggaucacaaa agcacuucuu 360

ccaucgauga uggagagaaa ucauggccac aucgucacag uggcuucagu gugcggccac 420

gaagggauuc cuuaccucau cccauauugu uccagcaaau uugccgcugu uggcuuucac 480

agaggucuga caucagaacu ucaggccuug ggaaaaacug guaucaaaac cucaugucuc 540

ugcccaguuu uugugaauac uggguucacc aaaaauccaa gcacaagauu auggccugua 600

uuggagacag augaagucgu aagaagucug auagauggaa uacuuaccaa uaagaaaaug 660

auuuuuguuc caucguauau caauaucuuu cugagacuac agaagguuuc uucc 714

<210> 10

<211> 894

<212> DNA

<213> Homo sapien

<400> 10

augaacauca uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60

ucguugguga aguuuuucau uccucagagg agaaaaucug uggcugggga gauuguucuc 120

auuacuggag cugggcaugg aauaggcagg cagacuacuu augaauuugc aaaacgacag 180

agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240

cgaaaacuag gcgucacugc gcaugcguau gugguagacu gcagcaacag agaagagauc 300

uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaaucgu ggugaauaau 360

gcugggacag uauauccagc cgaucuucuc agcaccaagg augaagagau uaccaagaca 420

uuugagguca acauccuagg acauuuuugg aauggaaagg acaucagaag uaauuacuug 480

gauguauaua ggaucgagga cacuuuugga cgagacucug agaucacaaa agcacuucuu 540

ccaucgauga uggagagaaa ucauggccac aucgucacag uggcuucagu gugcggccac 600

gaagggauuc cuuaccucau cccauauugu uccagcaaau uugccgcugu uggcuuucac 660

agaggucuga caucagaacu ucaggccuug ggaaaaacug guaucaaaac cucaugucuc 720

ugcccaguuu uugugaauac uggguucacc aaaaauccaa gcacaagauu auggccugua 780

uuggagacag augaagucgu aagaagucug auagauggaa uacuuaccaa uaagaaaaug 840

auuuuuguuc caucguauau caauaucuuu cugagacuac agaagguuuc uucc 894

<210> 11

<211> 813

<212> DNA

<213> Homo sapien

<400> 11

augaacauca uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60

ucguugguga aguuuuucau uccucagagg agaaaaucug uggcugggga gauuguucuc 120

auuacuggag cugggcaugg aauaggcagg cagacuacuu augaauuugc aaaacgacag 180

agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240

cgaaaacuag gcgucacugc gcaugcguau gugguagacu gcagcaacag agaagagauc 300

uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaaucgu ggugaauaau 360

gcugggacag uauauccagc cgaucuucuc agcaccaagg augaagagau uaccaagaca 420

uuugagguca acauccuagg acauuuuugg aucacaaaag cacuucuucc aucgaugaug 480

gagagaaauc auggccacau cgucacagug gcuucagugu gcggccacga agggauuccu 540

uaccucaucc cauauuguuc cagcaaauuu gccgcuguug gcuuucacag aggucugaca 600

ucagaacuuc aggccuuggg aaaaacuggu aucaaaaccu caugucucug cccaguuuuu 660

gugaauacug gguucaccaa aaauccaagc acaagauuau ggccuguauu ggagacagau 720

gaagucguaa gaagucugau agauggaaua cuuaccaaua agaaaaugau uuuuguucca 780

ucguauauca auaucuuucu gagacuacag aag 813

<210> 12

<211> 900

<212> DNA

<213> Homo sapien

<400> 12

atgaacatca tcctagaaat ccttctgctt ctgatcacca tcatctactc ctacttggag 60

tcgttggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120

attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180

agcatattgg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240

cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 300

tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360

gctgggacag tatatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420

tttgaggtca acatcctagg acatttttgg atcacaaaag cacttcttcc atcgatgatg 480

gagagaaatc atggccacat cgtcacagtg gcttcagtgt gcggccacga agggattcct 540

tacctcatcc catattgttc cagcaaattt gccgctgttg gctttcacag aggtctgaca 600

tcagaacttc aggccttggg aaaaactggt atcaaaacct catgtctctg cccagttttt 660

gtgaatactg ggttcaccaa aaatccaagc acaagattat ggcctgtatt ggagacagat 720

gaagtcgtaa gaagtctgat agatggaata cttaccaata agaaaatgat ttttgttcca 780

tcgtatatca atatctttct gagactacag aagtttcttc ctgaacgcgc ctcagcgatt 840

ttaaatcgta tgcagaatat tcaatttgaa gcagtggttg gccacaaaat caaaatgaaa 900

<210> 13

<211> 792

<212> DNA

<213> Homo sapien

<400> 13

atgaacatca tcctagaaat ccttctgctt ctgatcacca tcatctactc ctacttggag 60

tcgttggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120

attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180

agcatattgg ttctgtggga tattaataag gtgaagaaag aagtgggtga tgtaacaatc 240

gtggtgaata atgctgggac agtatatcca gccgatcttc tcagcaccaa ggatgaagag 300

attaccaaga catttgaggt caacatccta ggacattttt ggatcacaaa agcacttctt 360

ccatcgatga tggagagaaa tcatggccac atcgtcacag tggcttcagt gtgcggccac 420

gaagggattc ttacctcat cccatattgt tccagcaaat ttgccgctgt tggctttcac 480

agaggtctga catcagaact tcaggccttg ggaaaaactg gtatcaaaac ctcatgtctc 540

tgcccagttt ttgtgaatac tgggttcacc aaaaatccaa gcacaagatt atggcctgta 600

ttggagacag atgaagtcgt aagaagtctg atagatggaa tacttaccaa taagaaaatg 660

atttttgttc catcgtatat caatatcttt ctgagactac agaagtttct tcctgaacgc 720

gcctcagcga ttttaaatcg tatgcagaat attcaatttg aagcagtggt tggccacaaa 780

atcaaaatga aa 792

<210> 14

<211> 783

<212> DNA

<213> Homo sapien

<400> 14

atgaacatca tcctagaaat ccttctgctt ctgatcacca tcatctactc ctacttggag 60

tcgttggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120

attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180

agcatattgg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240

cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 300

tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360

gctgggacag tatatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420

tttgaggtca acatcctagg acatttttgg atcacaaaag cacttcttcc atcgatgatg 480

gagagaaatc atggccacat cgtcacagtg gcttcagtgt gcggccacga agggattcct 540

tacctcatcc catattgttc cagcaaattt gccgctgttg gctttcacag aggtctgaca 600

tcagaacttc aggccttggg aaaaactggt atcaaaacct catgtctctg cccagttttt 660

gtgaatactg ggttcaccaa aaatccaagc acaaggtttc ttcctgaacg cgcctcagcg 720

attttaaatc gtatgcagaa tattcaattt gaagcagtgg ttggccacaa aatcaaaatg 780

aaa 783

<210> 15

<211> 822

<212> DNA

<213> Homo sapien

<400> 15

atgaacatca tcctagaaat ccttctgctt ctgatcacca tcatctactc ctacttggag 60

tcgttggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120

attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180

agcatattgg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240
cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 300
tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360
gctgggacag tatatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420
tttgaggtca acatcctagg acatttttgg atcacaaaag cacttcttcc atcgatgatg 480
gagagaaatc atggccacat cgtcacagtg gcttcagtgt gcggccacga agggattcct 540
tacctcatcc catattgttc cagcaaattt gccgctgttg gctttcacag aggtctgaca 600

tcagaacttc aggccttggg aaaaactggt atcaaaacct catgtctctg cccagttttt 660
gtgaatactg ggttcaccaa aaatccaagc acaagattat ggcctgtatt ggagacagat 720
gaagtcgtaa gaagtctgat agatggaata cttaccaata agaaaatgat ttttgttcca 780
tcgtatatca atatctttct gagactacag aaggtttctt cc 822

<210> 16
<211> 972
<212> DNA
<213> Homo sapien
<400> 16

atgaacatca tcctagaaat ccttctgctt ctgatcacca tcatctactc ctacttggag 60
tcgttggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120

attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180
agcatattgg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240
cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 300
tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360
gctgggacag tatatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420
tttgaggtca acatcctagg acatttttgg aatggaaagg acatcagaag taattacttg 480
gatgtatata ggatcgagga cacttttgga cgagactctg agatcacaaa agcacttctt 540

ccatcgatga tggagagaaa tcatggccac atcgtcacag tggcttcagt gtgcggccac 600
gaagggattc cttacctcat cccatattgt tccagcaaat ttgccgctgt tggctttcac 660
agaggtctga catcagaact tcaggccttg ggaaaaactg gtatcaaaac ctcatgtctc 720
tgcccagttt ttgtgaatac tgggttcacc aaaaatccaa gcacaagatt atggcctgta 780
ttggagacag atgaagtcgt aagaagtctg atagatggaa tacttaccaa taagaaaatg 840

attttttgttc catcgtatat caatatcttt ctgagactac agaagtttct tcctgaacgc 900

gcctcagcga ttttaaatcg tatgcagaat attcaatttg aagcagtggt tggccacaaa 960

atcaaaatga aa 972

<210> 17

<211> 852

<212> DNA

<213> Homo sapien

<400> 17

atgaacatca tcctagaaat ccttctgctt ctgatcacca tcatctactc ctacttggag 60

tcgttggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120

attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180

agcatattgg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240

cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 300

tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360

gctgggacag tatatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420

tttgaggtca acatcctagg acatttttgg atcacaaaag cacttcttcc atcgatgatg 480

gagagaaatc atggccacat cgtcacagtg gcttcagtgt gcggccacga agggattcct 540

tacctcatcc catattgttc cagcaaattt gccgctgttg gctttcacag aggtctgaca 600

tcagaacttc aggccttggg aaaaactggt atcaaaacct catgtctctg cccagttttt 660

gtgaatactg ggttcaccaa aaatccaagc acaagattat ggcctgtatt ggagacagat 720

gaagtcgtaa gaagtctgat agatggaata cttaccaata agaaaatgat ttttgttcca 780

tcgtatatca atatctttct gagactacag aagttaagta cagcacagaa cacccaaata 840

ctaaaacacc aa 852

<210> 18

<211> 714

<212> DNA

<213> Homo sapien

<400> 18

atgaacatca tcctagaaat ccttctgctt ctgatcacca tcatctactc ctacttggag 60

tcgttggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120

attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180

agcatattgg ttctgtggga tattaataag gtgaagaaag aagtgggtga tgtaacaatc 240

gtggtgaata atgctgggac agtatatcca gccgatcttc tcagcaccaa ggatgaagag 300

attaccaaga catttgaggt caacatccta ggacattttt ggatcacaaa agcacttctt 360

ccatcgatga tggagagaaa tcatggccac atcgtcacag tggcttcagt gtgcggccac 420

gaagggattc cttacctcat cccatattgt tccagcaaat ttgccgctgt tggctttcac 480

agaggtctga catcagaact tcaggccttg ggaaaaactg gtatcaaaac ctcatgtctc 540

tgcccagttt ttgtgaatac tgggttcacc aaaaatccaa gcacaagatt atggcctgta 600

ttggagacag atgaagtcgt aagaagtctg atagatggaa tacttaccaa taagaaaatg 660

atttttgttc catcgtatat caatatcttt ctgagactac agaaggtttc ttcc 714

<210> 19

<211> 894

<212> DNA

<213> Homo sapien

<400> 19

atgaacatca tcctagaaat ccttctgctt ctgatcacca tcatctactc ctacttggag 60

tcgttggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120

attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180

agcatattgg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240

cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 300

tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360

gctgggacag tatatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420

tttgaggtca acatcctagg acatttttgg aatggaaagg acatcagaag taattacttg 480

gatgtatata ggatcgagga cacttttgga cgagactctg agatcacaaa agcacttctt 540

ccatcgatga tggagagaaa tcatggccac atcgtcacag tggcttcagt gtgcggccac 600

gaagggattc cttacctcat cccatattgt tccagcaaat ttgccgctgt tggctttcac 660

agaggtctga catcagaact tcaggccttg ggaaaaactg gtatcaaaac ctcatgtctc 720

tgcccagttt ttgtgaatac tgggttcacc aaaaatccaa gcacaagatt atggcctgta 780

ttggagacag atgaagtcgt aagaagtctg atagatggaa tacttaccaa taagaaaatg 840

atttttgttc catcgtatat caatatcttt ctgagactac agaaggtttc ttcc 894

<210> 20

<211> 813

<212> DNA

<213> Homo sapien

<400> 20

atgaacatca tcctagaaat ccttctgctt ctgatcacca tcatctactc ctacttggag 60

tcgttggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120

attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180

agcatattgg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240

cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 300

tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360

gctgggacag tatatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420

tttgaggtca acatcctagg acatttttgg atcacaaaag cacttcttcc atcgatgatg 480

gagagaaatc atggccacat cgtcacagtg gcttcagtgt gcggccacga agggattcct 540

tacctcatcc catattgttc cagcaaattt gccgctgttg gctttcacag aggtctgaca 600

tcagaacttc aggccttggg aaaaactggt atcaaaacct catgtctctg cccagttttt 660

gtgaatactg ggttcaccaa aaatccaagc acaagattat ggcctgtatt ggagacagat 720

gaagtcgtaa gaagtctgat agatggaata cttaccaata agaaaatgat ttttgttcca 780

tcgtatatca atatctttct gagactacag aag 813

<210> 21

<211> 300

<212> PRT

<213> Homo sapien

<400> 21

Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Leu Ile Thr Ile Ile Tyr

1 5 10 15

Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys

20 25 30

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

35 40 45

Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val

50 55 60

Leu Trp Asp Ile Asn Lys Arg Gly Val Glu Glu Thr Ala Ala Glu Cys

65 70 75 80

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

130 135 140
Ile Leu Gly His Phe Trp Ile Thr Lys Ala Leu Leu Pro Ser Met Met
145 150 155 160
Glu Arg Asn His Gly His Ile Val Thr Val Ala Ser Val Cys Gly His
165 170 175
Glu Gly Ile Pro Tyr Leu Ile Pro Tyr Cys Ser Ser Lys Phe Ala Ala
180 185 190
Val Gly Phe His Arg Gly Leu Thr Ser Glu Leu Gln Ala Leu Gly Lys

195 200 205
Thr Gly Ile Lys Thr Ser Cys Leu Cys Pro Val Phe Val Asn Thr Gly
210 215 220
Phe Thr Lys Asn Pro Ser Thr Arg Leu Trp Pro Val Leu Glu Thr Asp
225 230 235 240
Glu Val Val Arg Ser Leu Ile Asp Gly Ile Leu Thr Asn Lys Lys Met
245 250 255
Ile Phe Val Pro Ser Tyr Ile Asn Ile Phe Leu Arg Leu Gln Lys Phe

260 265 270
Leu Pro Glu Arg Ala Ser Ala Ile Leu Asn Arg Met Gln Asn Ile Gln
275 280 285
Phe Glu Ala Val Val Gly His Lys Ile Lys Met Lys
290 295 300
<210> 22
<211> 264
<212> PRT
<213> Homo sapien

<400> 22

Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Leu Ile Thr Ile Ile Tyr
1 5 10 15
Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys

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

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

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

210 215 220
Ser Tyr Ile Asn Ile Phe Leu Arg Leu Gln Lys Phe Leu Pro Glu Arg
225 230 235 240

Ala Ser Ala Ile Leu Asn Arg Met Gln Asn Ile Gln Phe Glu Ala Val
245 250 255

Val Gly His Lys Ile Lys Met Lys
260

<210> 23

<211> 261

<212> PRT

<213> Homo sapien

<400> 23

Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Leu Ile Thr Ile Ile Tyr

1 5 10 15

Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys
20 25 30

Ser Val Ala Gly Glu Ile Val Leu Ile Thr Gly Ala Gly His Gly Ile
35 40 45

Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val
50 55 60

Leu Trp Asp Ile Asn Lys Arg Gly Val Glu Glu Thr Ala Ala Glu Cys

65 70 75 80

Arg Lys Leu Gly Val Thr Ala His Ala Tyr Val Val Asp Cys Ser Asn
85 90 95

Arg Glu Glu Ile Tyr Arg Ser Leu Asn Gln Val Lys Lys Glu Val Gly
100 105 110

Asp Val Thr Ile Val Val Asn Asn Ala Gly Thr Val Tyr Pro Ala Asp
115 120 125

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

130 135 140

Ile Leu Gly His Phe Trp Ile Thr Lys Ala Leu Leu Pro Ser Met Met
145 150 155 160

Glu Arg Asn His Gly His Ile Val Thr Val Ala Ser Val Cys Gly His
165 170 175

Glu Gly Ile Pro Tyr Leu Ile Pro Tyr Cys Ser Ser Lys Phe Ala Ala

180 185 190

Val Gly Phe His Arg Gly Leu Thr Ser Glu Leu Gln Ala Leu Gly Lys

195 200 205

Thr Gly Ile Lys Thr Ser Cys Leu Cys Pro Val Phe Val Asn Thr Gly

210 215 220

Phe Thr Lys Asn Pro Ser Thr Arg Phe Leu Pro Glu Arg Ala Ser Ala

225 230 235 240

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

245 250 255

Lys Ile Lys Met Lys

260

<210

> 24

<211> 274

<212> PRT

<213> Homo sapien

<400> 24

Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Leu Ile Thr Ile Ile Tyr

1 5 10 15

Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys

20 25 30

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

35 40 45

Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val

50 55 60

Leu Trp Asp Ile Asn Lys Arg Gly Val Glu Glu Thr Ala Ala Glu Cys

65 70 75 80

Arg Lys Leu Gly Val Thr Ala His Ala Tyr Val Val Asp Cys Ser Asn

85 90 95

Arg Glu Glu Ile Tyr Arg Ser Leu Asn Gln Val Lys Lys Glu Val Gly

100 105 110

Asp Val Thr Ile Val Val Asn Asn Ala Gly Thr Val Tyr Pro Ala Asp

115 120 125

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

130 135 140

Ile Leu Gly His Phe Trp Ile Thr Lys Ala Leu Leu Pro Ser Met Met

145 150 155 160

Glu Arg Asn His Gly His Ile Val Thr Val Ala Ser Val Cys Gly His

165 170 175

Glu Gly Ile Pro Tyr Leu Ile Pro Tyr Cys Ser Ser Lys Phe Ala Ala

180 185 190

Val Gly Phe His Arg Gly Leu Thr Ser Glu Leu Gln Ala Leu Gly Lys

195 200 205

Thr Gly Ile Lys Thr Ser Cys Leu Cys Pro Val Phe Val Asn Thr Gly

210 215 220

Phe Thr Lys Asn Pro Ser Thr Arg Leu Trp Pro Val Leu Glu Thr Asp

225 230 235 240

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

245 250 255

Ile Phe Val Pro Ser Tyr Ile Asn Ile Phe Leu Arg Leu Gln Lys Val

260 265 270

Ser Ser

<210> 25

<211> 324

<212> PRT

<213> Homo sapien

<400> 25

Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Leu Ile Thr Ile Ile Tyr

1 5 10 15

Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys

20 25 30

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

35 40 45

Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val
50 55 60
Leu Trp Asp Ile Asn Lys Arg Gly Val Glu Glu Thr Ala Ala Glu Cys
65 70 75 80
Arg Lys Leu Gly Val Thr Ala His Ala Tyr Val Val Asp Cys Ser Asn
85 90 95

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

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

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

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

290 295 300

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

305 310 315 320

Ile Lys Met Lys

<210> 26

<211> 284

<212> PRT

<213> Homo sapien

<400> 26

Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Leu Ile Thr Ile Ile Tyr

1 5 10 15

Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys

20 25 30

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

35 40 45

Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val

50 55 60

Leu Trp Asp Ile Asn Lys Arg Gly Val Glu Glu Thr Ala Ala Glu Cys

65 70 75 80

Arg Lys Leu Gly Val Thr Ala His Ala Tyr Val Val Asp Cys Ser Asn

85 90 95

Arg Glu Glu Ile Tyr Arg Ser Leu Asn Gln Val Lys Lys Glu Val Gly

100 105 110

Asp Val Thr Ile Val Val Asn Asn Ala Gly Thr Val Tyr Pro Ala Asp

115 120 125

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

130 135 140

Ile Leu Gly His Phe Trp Ile Thr Lys Ala Leu Leu Pro Ser Met Met

145 150 155 160

Glu Arg Asn His Gly His Ile Val Thr Val Ala Ser Val Cys Gly His

165 170 175

Glu Gly Ile Pro Tyr Leu Ile Pro Tyr Cys Ser Ser Lys Phe Ala Ala
180 185 190
Val Gly Phe His Arg Gly Leu Thr Ser Glu Leu Gln Ala Leu Gly Lys
195 200 205
Thr Gly Ile Lys Thr Ser Cys Leu Cys Pro Val Phe Val Asn Thr Gly

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

275 280

<210> 27
<211> 238
<212> PRT
<213> Homo sapien
<400> 27
Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Leu Ile Thr Ile Ile Tyr
1 5 10 15
Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys
20 25 30
Ser Val Ala Gly Glu Ile Val Leu Ile Thr Gly Ala Gly His Gly Ile
35 40 45

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

100 105 110

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

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

<210> 28
<211> 298
<212> PRT
<213> Homo sapien
<400> 28
Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Leu Ile Thr Ile Ile Tyr
1 5 10 15
Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys
20 25 30
Ser Val Ala Gly Glu Ile Val Leu Ile Thr Gly Ala Gly His Gly Ile
35 40 45
Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val

50 55 60
Leu Trp Asp Ile Asn Lys Arg Gly Val Glu Glu Thr Ala Ala Glu Cys

65 70 75 80

Arg Lys Leu Gly Val Thr Ala His Ala Tyr Val Val Asp Cys Ser Asn

85 90 95

Arg Glu Glu Ile Tyr Arg Ser Leu Asn Gln Val Lys Lys Glu Val Gly

100 105 110

Asp Val Thr Ile Val Val Asn Asn Ala Gly Thr Val Tyr Pro Ala Asp

115 120 125

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

130 135 140

Ile Leu Gly His Phe Trp Asn Gly Lys Asp Ile Arg Ser Asn Tyr Leu

145 150 155 160

Asp Val Tyr Arg Ile Glu Asp Thr Phe Gly Arg Asp Ser Glu Ile Thr

165 170 175

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

180 185 190

Thr Val Ala Ser Val Cys Gly His Glu Gly Ile Pro Tyr Leu Ile Pro

195 200 205

Tyr Cys Ser Ser Lys Phe Ala Ala Val Gly Phe His Arg Gly Leu Thr

210 215 220

Ser Glu Leu Gln Ala Leu Gly Lys Thr Gly Ile Lys Thr Ser Cys Leu

225 230 235 240

Cys Pro Val Phe Val Asn Thr Gly Phe Thr Lys Asn Pro Ser Thr Arg

245 250 255

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

260 265 270

Gly Ile Leu Thr Asn Lys Lys Met Ile Phe Val Pro Ser Tyr Ile Asn

275 280 285

Ile Phe Leu Arg Leu Gln Lys Val Ser Ser

290 295

<210> 29

<211> 271

<212> PRT

<213> Homo sapien

<400> 29

Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Leu Ile Thr Ile Ile Tyr

1 5 10 15

Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys

20 25 30

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

35 40 45

Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val

50 55 60

Leu Trp Asp Ile Asn Lys Arg Gly Val Glu Glu Thr Ala Ala Glu Cys

65 70 75 80

Arg Lys Leu Gly Val Thr Ala His Ala Tyr Val Val Asp Cys Ser Asn

85 90 95

Arg Glu Glu Ile Tyr Arg Ser Leu Asn Gln Val Lys Lys Glu Val Gly

100 105 110

Asp Val Thr Ile Val Val Asn Asn Ala Gly Thr Val Tyr Pro Ala Asp

115 120 125

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

130 135 140

Ile Leu Gly His Phe Trp Ile Thr Lys Ala Leu Leu Pro Ser Met Met

145 150 155 160

Glu Arg Asn His Gly His Ile Val Thr Val Ala Ser Val Cys Gly His

165 170 175

Glu Gly Ile Pro Tyr Leu Ile Pro Tyr Cys Ser Ser Lys Phe Ala Ala

180 185 190

Val Gly Phe His Arg Gly Leu Thr Ser Glu Leu Gln Ala Leu Gly Lys

195 200 205

Thr Gly Ile Lys Thr Ser Cys Leu Cys Pro Val Phe Val Asn Thr Gly

210 215 220

Phe Thr Lys Asn Pro Ser Thr Arg Leu Trp Pro Val Leu Glu Thr Asp

225 230 235 240

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

245 250 255

Ile Phe Val Pro Ser Tyr Ile Asn Ile Phe Leu Arg Leu Gln Lys

260 265 270

<210> 30

<211> 23830

<212> DNA

<213> Homo sapien

<400> 30

atggtccgag gggggcgggg ctgacgtcgc gctgggaatg ccctggccga gacactgagg 60

cagggtagag agcgcttgcg ggcgccgggc ggagctgctg cggatcagga cccgagccga 120

ttcccgatcc cgacccagat cctaacccgc gcccccgccc cgccgccgcc gccatgtacg 180

acgcagagcg cggctggagc ttgtccttcg cgggctgcgg cttcctgggc ttctaccacg 240

tcggggcgac ccgctgcctg agcgagcacg ccccgcacct cctccgcgac gcgcgcatgt 300

tgttcggcgc ttcggccggg gcgttgcact gcgtcggcgt cctctccggt atcccgctgg 360

gtgcgtctgg ggacgctgcc cgggctccac gtgcggagtg ggtgcccccc aggccgggga 420

gcgggggatc cccaggggtc gcggggccct ggaggagcgg gcatcggacg cggacacggc 480

ggggtgcatc ccgagggccc cctccgaggc agatgcttcc tgcgggggcg ctgttcctgg 540

gcccgggaag ggggcgttgg aaccccgagc ggtccgggcc gaagcctggg actctcgtgc 600

gtccccaccc ctacccccat caggcgcccg tgcatgaagg gagaccctca cctccggact 660

gagagtcgga gcgtctcgga gcgacgggga gtagggagcg ggacccgggg cggagggtag 720

tgctggcccc tgcggactcc gggtcccctg tgtcctctcg ggaggggctg gacgggctga 780

gctgccgagg ggccgatttg ccctgggccg gacaaagagt ggggctttgg ccggtccccc 840

acggtgggct ccttccctct ggggattgag ggactcaaga caccccgcgc ctgcgctttt 900

cttttctttt tttctttttt ttttttgag acggagtttc gctcagtcgc ccaggctgga 960

gtgcagtggc gtgatctcaa ctcactgcaa gctccacctc ccaggttcac gccattctcc 1020

tgcctcagcc tcccgagtag ctgggactac aggcgccagc caccaagccc ggctaatttt 1080

ttgtattttt tagtagagac ggggtttcac cgtgttagcc aggatggtct cgatctcctg 1140

acctcgtgat ctgcccacct cggcctccca gaatgctggg gttacaggcg tgagccactg 1200

ctccctgctg cctacgctct ctgggtcgca gcccagcctt ctgggggctg ggtagcctcc 1260

cagaagggca accctgggca tcctccaggg caggctaact ggagtctagt ggggaggggt 1320

accttgaaag aggaaagttg tttcctcctc ctcctcctcc tccagtgttt gggacccttc 1380

ctgggggctg gagtgcatcc ctggacaccc cccaatccca tcctcttctc tagtttccac 1440

tgacctaggc ccaccctccc ctctccggct cagtactcct ggaaatgaga ttccgtacat 1500

ttgaatcttg tcctaatgaa atatttgtcc atgtgggtac ctgtgtgtgt gtggtggggg 1560

tgcagacgga gggtttgttt ctcactagct ggaactactg gggtgtggta tgcttcctgg 1620

gaatttgtgt gccacagtcc tggaggcgag gaggggttg tgagccagta ggcaggggct 1680

ggggcaagta gcattgtgaa gctattgaca cccagacgtc cccaggcagg agattatgcc 1740

cccattagcc ccctttatc tgggcttcct taacaatgga ctctttgccc tgcctgccag 1800

agccagcagg gagtgactgt tcagtggtga ggaagcgggc agaggaagcc ctgccattgg 1860

gtaggagcag tgggcagccc ctgggctgac tgggaggtgg ggattaggga ttagacagtc 1920

ctggctgtct gccttcccct aagccagggg gagaggagca aagggcacga aatgtggcct 1980

ccaggaggat tagaccgcca catgatcatt tgcacaccct ggggtttagc aacaataaaa 2040

gtcagctttt ttgtatccca aggtggcctg tggacaccca catggacaaa tgtttacact 2100

gggacagaat tcaaatgcag aggtcccagg agcctaaagt acactcactc tggtatagaa 2160

aggattcctt actgggcaga ggacaggtgc agcctggggc tttcccaggc aggacacagg 2220

gaggctcagg aaccaccaag tccctggaag gtggatctgg aggcgttggc aggagccact 2280

ccctgggttc cagggctcca ggttcctgct ttaacccccc gtctcacaga gggctgtgca 2340

cttgggggct gctgagcatg tcccagaggc tgcatcctgg acacagcacc tcagtgcatc 2400

tgagctgagg ctaacttggc aggagggaca ggcagaacct gccagccacg tgcaattcca 2460

cccctctggc cactcaggga aggagagctg tgagtcaaga tcagatttgg gtcaggacag 2520

gctggggcct gcctgtccct gtgcatccca agatttatgg ctggccaggg gttgggctgg 2580

gaggggtggt cttgcatgcc aggagagtgc agatcagcct gagaggccag gccagtaagt 2640

gaggtcagat ctcctgcacc tgatagcatt aaggccatct acaccaaagc tctaatgctg 2700

atatgttcct ggcctctatg tggggcatgg aggtggggca tggaggtgag gcctgctcgc 2760

ctgggcttct ggaagtggga gactcattcc tgtggctgag gcctacagca gtgctgtgtg 2820

gtaggaatac actggaagcc atgatgtcat tgtgcatttt ctagaagcca cattgaataa 2880

agtaaaagac acaggtagaa ttaatttcat tgagcccaat atatccaaaa taatatcatt 2940

ttcacatcta ttcaatataa aaatttacta atgagatatt tcatactaag ccactgaaat 3000

ccagtttgta tcttacacat ctcagttttg acgagccaca tttcaagggc gtgatagcca 3060

catgtggctc ccatagtaga cagtactggt ctagagaaat gttggtggca tccttgctgt 3120

ctggtttctg gccttgccaa aagtattacc atcccagtgt ggtacattct ttcatgtatt 3180

tgtctcctgt ccccagagca gactctgcag gtcctctcag atcttgtgcg gaaggccagg 3240

agtcggaaca ttggcatctt ccatccatcc ttcaacttaa gcaagttcct ccgacagggt 3300

ctctgcaaat gcctcccggc caatgtccac cagctcatct ccggcaaaat aggcatctct 3360

cttaccagag tgtctgatgg ggaaaacgtt ctggtgtctg actttcggtc caaagacgaa 3420

gtcgtggatg taagcagttt gcttatctgg acgttgtcaa gttagaaaag ctgttttggg 3480

atgggtgtgg tggctcatgc ctgtcatccc ggcactttgg gaggccgaag cgggtgggtt 3540

gcttgagccc aggagctcga gaccaacatg atgaaaccca gtctctacaa aaattacaga 3600

aaaattagct aggcatggtg ttgtgggccc atagtcccag ctactaggga ggctgaggca 3660

ggagaattgc ttgagcctgg gaggtggagg ttgcagtaag tcatgatcat gccactgtac 3720

tccagcccgg gtgacagtga gatgctgtct ggaaaaaaaa aaaaaagaaa gactgttttg 3780

ttttggaagc aacacaggca gttgtaggcc ccctgtgcca gagtgacata aactctgtac 3840

acctccagtg atttggtcca tgtttgtaaa ccctgaatgt tccagggcag tttcttttct 3900

tcacttttta tctctttttt ttgggtgggg gggcggggta cagagtcttg ctctgtctcc 3960

caggctggag tgcagtggcg caatctcaac ctcccgagga gctgggacta caggcacagg 4020

ccatcacacc ttgctaatgt ttgtactttt tgtagagacg gggttttgcc ctgttgccca 4080

ggctggtccc aaactcctgc acccaagtaa tctgcccacc tctgcctggc agttacaatt 4140

tcaaataatt cctccctttc cttcaacact tggctcatga ccgtccagtc caaggaacct 4200

gtcctgcagg tgtgcctctc ccgagcttcc tctatgcatc ttccataatg aagatgcctt 4260

ctcactggaa accctacaag ggtgggaacg tgccttattt gcctgtatcc tcagggtcta 4320

gcagagagaa gataatttgt aataccaaaa caccattaaa ttcagctgat gctttcataa 4380

gcgctccttg gaggaaggac tccatttact tgacagatct gtgcaagaca gcagcctggc 4440

gcgtctaacc tgcagccagt tgcatcctct gtttaacctt gtttgcggaa gctttctcta 4500

aacagccagc acttgtctgt tcccacatgg gtccgttctc ccagtgaatc accgtggtgc 4560

ctgctgactg ctctgtagca cagtgcttcg caaagtgtga tcctgggacc agcagagcag 4620

cagctccttt gagcttattg gaatggcaga ccctcaggtc ccacctctga cctgctgcat 4680

gggaattctg gggagggacg cagaatctct ggttccacag gctctccggt gatgctaatg 4740

aataccggca tttgaacagc accgatctag ccccctttcag tccatgagcc aacaaccctt 4800

ggtcctgtct gtggtgaccc agtgtgactc tcatggggag caaggagagg aagttgaagt 4860

tcactgacag ggttgttaag gggattatgc aatagatgag acccatgggc ctgaagtccg 4920

agggtgtatg ttagttcccc gttcttttga cccatggatt aacctactct gtgcaaaggg 4980
cattttcaag tttgttgccc tgctcacttg gagaaagctt atgaaggatc aggaaaatta 5040
aaagggtgct ctcgcctata acttctctct cctttgcttt cacaggcctt ggtatgttcc 5100
tgcttcatcc ccttctacag tggccttatc cctccttcct tcagaggcgt ggtaagtcgg 5160
ctttctctgc tagcgctgag tcctgggggc ctctgaagtg tgctcacaca tctcctgcct 5220
gcagggcact ggtgtcgggc acctcagggt ctgtcccatg gtggagcccc atgcctcact 5280
gcctttcaga cagagtagcc acagctggcc ctatttccag gctacccggg cagcaaaact 5340

tactgcatgt gtaattaatt atttggctat ctgtaaggta aactggctgg ttcacttaat 5400
ctgcaccttа agcatcagat agcttctcag tgatctagtt aaactatatg atgttggcca 5460
ggcgcggtgg ctcatgtctg taatcccagc actttgggag cctgaagcag gcagatcact 5520
tgaggtcagg agttcgagac cagcctggcc aacagtgtga aactctgtct ctcctaaaaa 5580
tacaaaaatt agctgggcat ggtggtgtgc acctgtaatc ccagctgctc gggaggctga 5640
ggcaggagaa ttgcttgaac ttgggaggcg gaagttgcag tgagccaaga tcgcaccact 5700
gcactccatc ctgggtgaca gagcgagact ctatctcaaa aagaaaaaaa aaaaaaaggt 5760

aaataaagta tatgacactg aagaatctgt tacccctgga aggtggagct ttactcttag 5820
ggggaactat aacagtcata tatatatatt tttttctttt ctttttttt ttttttgaga 5880
tggagtctgg ctctgtcgcc caggctggag tgcagtggtg caatctcggc tcactgcaac 5940
ctccacctca caggttcagg caattctcct gcctcaacct cccgagtagc tgggattaca 6000
ggtgcctgcc gtcacgccaa gctaattttt gtatttttag tagagacagg gtttcatcat 6060
attggccagg ctggtctcca actcctgacc tcaggtgatc cgcccgcctt ggcctcccaa 6120
agtgctgaga ttacaggcgt gagccatggt gcccggccaa caatcacatg tgttgtaaac 6180

aacaacaaaa atctgtcagc ctggtctaac ctagatttgt gctttgtttt gttttgccac 6240
tttgtgatgc acaggaggaa gtttaggctg taaaatacta gccttttagg gtaatttttg 6300
aactcacaag agcagcagcg gaacctttga tgcaatcctg tatgtagcac cagcagagcc 6360
acgtggcaga gggactcgca ttaggagcct cccattacag actacgtgct cctgtgcgtt 6420
atcttatagg gtccccacaa ccaaggggag atgtgattat tcatcctgtg tggctgtggg 6480
gaacttgaga gtcatacttg cccaaagagc acggccagcg agcttgcacc caggtcactc 6540
tctgctcctc tgtcagaaca gggcatgtct tggttcactg cagggcggct cttctcattc 6600

tctgtagttt ggggtccagg atagtggtcc acggagccac tggagtgccc agctactgag 6660

tgaccaaagc atattttgga tttccgacat tgccacagca tggttgggca tcagcaggac 6720

cccaacccct tgttatgctg gtggctttat gtggttattt gatcttcccc agaactcagc 6780

aggagtgcac ccagcagcac cgtagtgatg ctctctggct ccccagtgca cggttctggc 6840

tttccttcct ggtcgagagt ttcaagccct ctgggtccta ctctgtcctt ttcagcccat 6900

agctttgttc aaaagctgct ggcagtgttc agatttggct gagttcagtg aatatgtgca 6960

ttggctgatt tctgagccat gccaggggga tggagaagcc gaagcaggag tgtttgttct 7020

gcaggctctg gagtaggcat tgggtctgtg ccggctcact tgctagtctt gcatccttcc 7080

ccaacccccct ctggggatgt ctggccacat cagaagacag tttgggttgt cagaactggg 7140

ggagtaccag gccgaggtgg gtggatcatg aggtcaggag atcgagacca tcctggctaa 7200

cacagtgaaa cctcatctct actaaacata cgaaaaaaat tagctgggcg tggtggcggg 7260

cgcctgtagt cccagctact cgggaggctg aggcaggaga atggtgtgaa cccgggggggc 7320

ggagcttgca gtgagctgag atcctgccac tgcactccag cctgggcaac aaagcgagac 7380

tccgtctcac aaacaaaaca aaacaaaaca aaacaaaatc tgggggagtg ccactggcat 7440

ctgatgtata gaggcccgag atgctgtgtc atcacccgtt gagtgcgctc ataggcatct 7500

tcctgacaat tagaacccat tattcttcaa attcaatgca agcaaattca aagcattact 7560

gtgtacatac cgcatgctaa tcaattgcac cactggagct cctaaattca aaacattact 7620

ataaaaaagt tcaaaatgca tggaaaagtt gtacatggca ggagaatatt tgggcttctg 7680

actacccctt gaatgaagat gatccaccag ccgccttcct ccttggtctt cactccagat 7740

tcctagcatt tcattctgtg tctctttatg cagtgaggtt tttgtttgtt ttttgagaca 7800

gagtctcact gtatcaccta ggcctggagt gcagtggcgc gatctcagct cactgcaacc 7860

ctcggctcct gggtttaagc gattctcctg cctcagcctc ccgagcagct gagattacaa 7920

gcacacatcc ccatgcccag ctaatttttg tatttttagc agagacaggg tttcaccatg 7980

ttgcccaggc tggtctcgaa ctcctggcct caagtgatcc atgtgcctca gccttccaaa 8040

gtgctgggat tacaggcgtg agccaccatg cccagctcct agtgaggttt ttgatgcctt 8100

gctacatctg ccctagaaat tgtgtgacta cgattttgga aatgttgctg tgtaaacttg 8160

tgatcatttc tggactccag gcaagaatct tgatggctaa ggtgtggctg aacatgtctg 8220

attctctcct ggacctgttt taggccaaac tctgctctga aattcctccg tgtggaaggg 8280

cgggctgggg agagcctccc agctggaatc ttttggatgc ctttctctgt gggtatctga 8340

tggctggctc tgatggctgg ctgtgatggc tgtggctgga aatcattgtt gacatgagtt 8400

tcacagatgc aggctctgtc caaattgtag caaaagctgc ctgccccagc cgagctatgg 8460

gcaataaggt ggtttaagga tatagatgaa ggaaaactca cccttagaat aatttatcca 8520

aaatgctgct gtgttgtggg ttagaggaca ttttctgagg tcccaggttc attgtttcat 8580

ttaagtctca aaagtccctc caggtgttgg ttctaattgt caaagcatgg ggggagatgg 8640

gctcatgggt taaaggtctt atcccagatt tctgtatcct ccttgcaagc agcaaagggg 8700

tctggatttg aatccatgac catgtttctc ctttgggttt ccatcacact ctgtccccgt 8760

gcactgagca ccctttagtt catatgaccc ccttaggcat gttacatggg cactcctata 8820

ggtgcccatc tggccctagg acttggccaa cacaacatgg actccagttt ccatctgcct 8880

ctttgccagg cacttttgtg cagtgcacac actgtacaac agtagacggc aaccctgaga 8940

gccagagtag agcctgtcct agcaccggaa tgctcggtaa ggatttgtcg caggagtgat 9000

tccaaagcca atgtcctccc tccatatcag cctgtttgtg gctctgagaa gctctgccca 9060

catgtgaaag cttgttaagc acttaagcac taacccagag cttcagacag tgccagtcct 9120

ttttcccctt ctttaaaagc gatatgtgga tggaggagtg agtgacaacg taccttcat 9180

tgatgccaaa acaaccatca ccgtgtcccc cttctatggg gagtacgaca tctgccctaa 9240

agtcaagtcc acgaactttc ttcatgtgga catcaccaag ctcagtctac gcctctgcac 9300

agggaacctc taccttctct cgagagcttt tgtccccccg gatctcaagg tgagttggtg 9360

gtgaggggggc aggtgttctg gggtgcagct cttctttgcc tccctgattg ccaggagcta 9420

ccagttactg tctgcacaat caaacagaaa tagacctgtc cttgatggtt aacggaaata 9480

aaaggcgctt gtcccagaag ctcaggtgag gcaccaccct gattatggga atcacctggg 9540

aacatatacc cagacctaaa actcagatcc acttcccagg ctgtggttat atagtcaggg 9600

gggtgcagta tgggtattag gatttttat tttttagtta taaagatttt tttttggttt 9660

gtttttgaga cagggtcttg ctctgccgct taggctggag tgcagtggtg caatcatagc 9720

tcactgaagc ctcagactcc tgggttcaag cagtcctccc acctcagcct cctaaggagc 9780

tgggacccac aggcatgcag caccacacct ggctaatttt taaaaatttt gtggagtgtt 9840

gcccaggctg gtctcacact cctggcctca agcgatcctc ccacccccagc ctcccaatgt 9900

gttgggatta caggcatgag ccattgtacc cagccactaa gatgattctt atttggaaac 9960

acggtcaaga acaactgcgt tcggtagttt aacctttttt gattgtggtg gttttagtat 10020

gccttaccac tctaccatag taagaaattt gcagaccatg tacaccaacc tttggtgctc 10080

ctggggagaa agaaagaagg ctatgcaatg caatgcatgc tcacagtcca agggagaggg 10140

aaagctgtct aacaggattg gttttcccgt gtgctttata agcagatgag tagaggagac 10200

agctcttatt gtcctagtgg caattgggat aggctgcaaa gtttgttagg gtggaggctt 10260

attccgggac caagggagcc caaagaaaca agctcctgcc aggcgcggtg gctcacgcct 10320

gtaatcccag cactttggga ggctgaggca ggtggatcac ctgaggtcag gagtttgaga 10380

ccagcctggc caacatggtg aaaccccgtc tccatgaaaa atacaaaaat tacccgggca 10440
tggtggcggg cacctgtaat cccagctact agggaggctg aggcaggaaa atggcttgaa 10500
cctcggaagt ggaggtggcc gttagccgag atcacgccac tgcactccag cctgggcaac 10560
agagcaagac tctgccttaa aaaaaaaaaa aaaaaaaaga aaagtaaaag gaaaaaaaag 10620
aggctctggc ctgctggggt gcctgcaaag tctccgtgga agggtgacat tcaagccgag 10680
acctccaggg aactgtctcc tgggagcaca gagcccttg ctcagccccc aggtggctca 10740
gtgccccag ccagcagact cagagcttgc atgattcttt ggtgctctct gcggtcttcc 10800

aatgatgctg aaataaatgg tgcttggtgt ctccctgctg tagtcccctt gcttgctttg 10860
ctcacaggtg ctgggagaga tatgccttcg aggatatttg gatgcattca ggttcttgga 10920
agagaagggt atgtatgggc tgggaggatc agccatgccc ttttgacaag catttactag 10980
cggtcttggt aaagacttga gatttgcctt agttctaaca cttagtgccc aacgccttcc 11040
ttgtgttgct caacctactc atgagcccag gagataggaa atctccgtcc cattgtacag 11100
atggggaaac agaattttgg aaaggagagc caagcagcac acacccctcc ctgaggggca 11160
gagccgagat ttgaactggg atgtcatgac tccagggccc tctccctccc cagggtcccc 11220

ttatctgaag gcggtttttc tttccagctc gacctcttgt gacccttagt ttaacaaggg 11280
ccgaagttaa agagtttctg cgcctggacc ccaaatgaag caatcagatt tctcatctcc 11340
agtcaggtgt gggtccaagc ccactagaca agtttgctct tcccagagca catttctgcc 11400
ttcaagtcat cctggcttgt cagggctggg ggagttctgc tgtagaaata ttagagtgga 11460
aggaaaaaga tgtgttggga gctatttttc tttaatacta aaagttggtt gatgaatttg 11520
tcgttggcca agaccaagga gactgcattt ttaaggacat atgtgtattt atctgctcag 11580
aaaatgttca ttgctgtgtg ctagggatac tgcagtgaac acagaggtgt gacccttgcc 11640

agccttgtga gagaagtgag cagataagta agcagaaggg tgatgctgtg tcgatgggaa 11700
agtacaggtg ccaatgagaa ggcacaggtg tcaaggagaa gacacaggat gctggaggct 11760
catgcaggat ggatctccaa ggcccagggg aagaagggcc tctcggagga cgtgaatcca 11820
cattaagact ttggggataa gtaggagcgc cttaggcatg gggacccatg gatgcgaggc 11880
ctgtaggaca cagacaggat ggcatgaagg cctgtgcaac tggaggggtg gggatgggga 11940
cactaagaga tggctggaag tgtgggggtg gggacactaa gagatgactg gagaagaggg 12000
ggtcaggagt ggtgaaaaat gggagaggag ggcaggctgg gccttttgga tacaggggga 12060

ttgcatcctg cagtggtagg gagccactga gggctgctgc agtaggagtg aggggatcag 12120

aggagagctt tggaagcccc ctggatgcgg gacaggaagc gagataccag tgtctaggag 12180

gccagtgagg cagccacagg ctccaccagg atcagggctg cgagggtcat gaggaggaaa 12240

ccaatttgaa ggagtccagg ggaataggac ttggaaatga ccgatgggac atttgggaag 12300

aggaagacag aagagcgcag tcccggcttc tggctttagc agttgggcaa ggggagatgg 12360

ggagatgtgc ccatgggttg agggttgagg acattaggag ggagccggta tggcaggaag 12420

agctggtgtg ccagagatgc tggaagcagc atctgcctga gaacagatac ctggcaatat 12480

tcctaaggga aagtgacatc tcggagggtg aggagggcat ctgatagggc ctggaaagag 12540

ccggggcaag catgaatgtg aggttatctt ggggggcaag gctcaggcgt tgaggagcag 12600

cccctggtct cttcagcctg aagttggaag ccagagttgg gccaggtgca gctgtggttg 12660

tctgaagtcc ccctccccca gcccagtgtg ccaatgctgt aagagcaagg gccgctcact 12720

ggtgctggtg gctgagtccc agcacccagg acagggcctg gcacatactg gtgcccaatc 12780

ctcccttctg ggtgcttctt ccaaggcctt gtgatggaag tgagtaccct cttcgacatc 12840

agacccagct tcaaatcctg gctctgctat gtattggctg cgtggcttta gacaagtctt 12900

ttaaccttgc tgtgcttctg atttctcagc tgaaaaatgg agatgatgat agtggtttct 12960

gtaaggcctt atggtgaagc acctagctca gggcctggaa ggcaggtgta accagtggtt 13020

cagttgttat aaaccaacac taaccctcgc ctttgcacct catgaaacca gatatgtaga 13080

tggagcccac aaagctagca ggagccaagc tcacgtgtgt cctgctttaa agccccatac 13140

ccctttctcc gggtgacaaa cacctgtgct cgttctcttc ccttcccctc ttccccttgc 13200

atttggctaa taacaggcca gctgcctgcc tccctgcagt ttggtagatg ggtgggtaac 13260

gaccaccact cccacgttcg cctgatgggc ttgttttccg tgcccttcac aggcatctgc 13320

aacaggcccc agccaggcct gaagtcatcc tcagaaggga tggatcctga ggtcgccatg 13380

cccagctggg caaacatgag tctggattct tccccggagt cggctgcctt ggctgtgagg 13440

ctggagggag atgagctgct agaccacctg cgtctcagca tcctgccctg ggatgagagc 13500

atcctggaca ccctctcgcc caggctcgct acaggtaccc actcctcggg gtgagcacgg 13560

gcagcacctt gttttctttc ttgtgcatta tggaggaaga tggtactgcc acatgggagc 13620

gatagggtga ggcaaccatg acaggtggtt gggaacatct ccttccatgt gtacagcctg 13680

ggctgctgcc atcactccca gcacagcccc caacccccca aatcctggaa ccttgccaag 13740

tctcccttcc catggggtca tgaccaggag gaaaacaaac tccagctgag ccccttgggg 13800

ttccccatat aggctcctgc ctgtggcagc tgggccctct gtaccccttt ccaactctgt 13860

ctccctaaca tggcacctga gctcctgcca tcctggattt catggacccc aaggatgggg 13920

gtcctgcatc tgggacttgg cctattactc ggagctcctt ttcagccgcc tccctccacc 13980

tgtccaccca cctcaaggct cctttcttga gacctctcct aatttctccc ttcccctaaa 14040

cccacaattt tgaacctcca tcgaatggtg ctgtatttta taatgtcatc aaatatcaaa 14100

tggagacagt gctatggtcc aaatgattgt gtacccccca gaatttgtct tttgaaatcc 14160

taacccccaa catgatggtc ttaggaggtg gggcctttgg gaggagatta ggtcatgagg 14220

aaagggctgt catgaatggg attggtgccc ttattaaaca gacccaagag aggtcccttg 14280

tcccttctac tgtgtgagga ctcagaaggt ggtgtctatg aagaagcagg ccctcaccag 14340

acaccaacat gtctgctgcc ccttgatctg ggaccttgca gcctctagaa ctctgaaaaa 14400

tcgatgtttg ttgttttata agccactcag ttggtggcat tttgttagag tagcctgaac 14460

acggactaag tcaaacagaa gaacccacaa accagctaca gagttgggca tttggagaaa 14520

ttcaaaaatg agtcagacat aactccttat tcttgaggtg ccctaagaga tgggacacag 14580

cagctgccca ggtgcattag tttgttctca cattgctata aagaaatacc tgagactggg 14640

taactcataa agaaagaggt tgaattggct cacagttgca caggctggac aggaagcatg 14700

gtgctggcat ctgctcagct tctggggagg cctcaggaaa cttacaatca tggcagaagg 14760

tgaacgggaa gcatgcacat cccatgactg gagcaggagt gagagagaga gggaaataga 14820

gggaaggtgc catacacttt taaacaacca gatctcatga gaacacattc actatcaaga 14880

gaacagcacc agtggggaaa tccgccccca tgatccaatc acctcccatc aggctccgcc 14940

tccaacactg ggaattacaa tttgacatga gatgtgggca gggacacaga tccaaaccat 15000

atgaccagat taatacgatt tgaggcatca cgaggtcatt aaagagaggg aataaaagac 15060

tggggctcca ggaagaaggc tctggaatcc agcagagggt caaggaccag cttgtaaagc 15120

tggtggtgcc tgagaagtac ctaggagaac atagatgctg tgacgtttga tgtagctgtt 15180

ttttgttttg tgttttggtt tttgagacag agtctcactc tgttgcccag gctggagtgt 15240

gcagtggcgt gatcttggct cactggagcc tccatctccc aggttcaaat gatcctcatg 15300

cctcagcctc ctgagttgct gggattacag gtgcacacca ccacgcctgg ctaatttttg 15360

tgttttcagt agagacaggg tttcaccatg ttggccaggc tggtcttgaa ctcctgacct 15420

caagtgatcc aacaacttca gcctcccaaa gtgctgggat gacaggcatg agccaccatg 15480

cccagcctga tgtagctgtt tctgtgcaca ttatttgctg tggggtatat tcagatttct 15540

taatacaaga tgattctttg cctcatgact tacacaccat tttctattta atttcagcta 15600

tgatattgga aatggacatg tcttttcaag gaaaataaaa gcaggctttc tggaatggcg 15660

acttccaaac atatttgtca atttaaagga gctgggagtg gggaccctat gctccgtaag 15720

cactctctta gctgttcttg gctgtgctcc ccgcttcagc ttcacactgc ccttgctgtg 15780

aagggagcag cctgggccgg gcgcggtggc ttacacctgt aatcctagca ctttgggagg 15840

ccgaggtggg tggatcacct gaggtcagga gttcaagacc agcctggcca acatggtgaa 15900
actccatctc tactaaaaat acaaaaaatt agctgggcat ggtggcaggt gcctgtaatc 15960
ccagctactt gggaggctga ggcagaagaa tcgcttgaac ccaggaggcg gaggttgcag 16020
tgagccgaga ttgcgccatt gcactccagc ctgggggcaa caagagcaaa actctgtctg 16080
gaaaaaaaag aaaggagcag cttggcaaac cccaccttgt cgcttttgtg agtgcctctg 16140
acccctttggc tgccaggacg ggcgtatttt atggaaatgc taagcaccaa cagagtaaag 16200
tggtttggtt tttcacagtg gtgggagata atagctccaa attgtctttt tcagcactga 16260

gtgaagaaat gaaagacaaa ggtggataca tgagcaagat ttgcaacttg ctacccatta 16320
ggataatgtc ttatgtaatg ctgccctgta ccctgcctgt ggaatctgcc attgcgattg 16380
tccagaggtg agcattttag gtggctccgt gtcttcctca cagggttgat atgaggatga 16440
aacaagatga tagatcatgg tggcatgtag tctgggacct ggattgtcgt gccacagatc 16500
acagctcaca gtctatgtgc aatgcccctg aatgttgccc acctgtcctc aagccacaca 16560
tgcacctgta actcagtgca agcccagaaa ctccccgtgg ggactcctag agctgtcagt 16620
ggcctcacat agcagctggt ccagtctctt gtgattgccc aaggaaactg aggcctggag 16680

agcttggggt cactgctctg aggccataga gatgcctagt agaagggcca ggcctagaag 16740
caggatcctt gctgcccctc tgagctgttt ccatttaaaa tcacatgaag gccggcgccg 16800
tggctcacgg ctgtaatccc agcattttgg gaggccaagg tgggtggatc atgtgaggtc 16860
aggagtttga gaccagcttg gccaacatgg tgaaatgcca tctgtactaa aaatacaaaa 16920
attagtggag catggtggca cgtgcctgta ctcccagcta cttggaaggc tggggcagaa 16980
gaatcgcttg agcctgggag gcagaggttg tagtgagcca agattgtacc actgcactcc 17040
agcctgggtg acaggagaga aaccctatct caaaataaaa tgaaaggtaa tgaaatgaat 17100

aaaataataa atcaagtcac ggccgggcac ggtggctcac acctgtaatc ccagcgcttt 17160
gggaggccga ggtgggtgga taatgaggtc aggagttcaa gaccagcctg gccaacatgg 17220
tgaaaccatg tctctactaa aaatacaaaa attagctggg catggtggtg catgcctgta 17280
atcccagcta ctccggaggc taaggcagga gaattgcttg aagcaggacc taggaggcgg 17340
aggttggttg cagtgagccg agatcatgcc actgcactct agcctgggct acagagcgaa 17400
actccgactc aaaaaaaaaa aaaaaaaaaa atcaaatcac atgaaagtag aacatаggga 17460
attccatctt tcgttctagg catagtttgt taatatgatt cagagccagc agttaggaga 17520

acacagtgtg actctcctag aacttcttga ttgggcttcc tctgattggg tttcctctga 17580

ttgggcttcc tctgaaagtg ggggggatgg ggggtgggga gcagaatggt cagagcttgg 17640

ctcagcagtc agactgctct tcttcaaatc ctggctgcat tgcttactac agctgtgtga 17700

ctccagatga ctgaatccac ctctctgtgc tgcagcttcc cgtctagaga gatcacctgg 17760

agcagagggt ggtcaggaga ctcaatctgg ttactgactc acagtgcagg agtactcatc 17820

ccatagtaag catccagcta gagatgttga tttctatttt caggtaataa tgatgatcgt 17880

aaaattagag acagataaaa ggtatgggca ttaggccagg gcactgcaat ttctaagctg 17940

tgtgacctca ggcaagttac tcgacttctc tgagcctcag cggtttcatc cgcaatatat 18000

ggataggaaa accgacctca gtgggttgtc tgacagtgga gggcacttga ttaaaaaaaa 18060

aaaaattacc ctggtctgaa tattaccctg gactgaaaga aaaatattga gctaatacag 18120

gcatcaggaa tggggctgca gggagtccag ggaagggaga acgaagagcc tgaaggtgtg 18180

aggaggtgcg agtgctgatc tgtctgctac aaagaggctg ctgagcctcc tgtggatgtg 18240

gccctggact tggcagttta atacctgagc tgttaaaata acctcagatg ctgtgttctt 18300

taaggggtag gattcagatt cctgctgaaa tgcttctgaa agggagggaa tgagccagcc 18360

catccccagt tgctttttaa gatcattggg aagttctggt cttgccattt gtccctggac 18420

cactcttagg tcctcctgcc ccacttccat ctgggtgtgt gccctgggct gtccaccaca 18480

cagctacatc ctgccatctt ccctcctgga gccactgtgc catgcatgga tctgtagctt 18540

cattttctt ggcttttccc tggtttttct ggagcagagt ctctagtaaa ctcccaagga 18600

agaaaacgtt tgactttatg tgtgttggga aacgtgcttt ttttctatta catctcagtg 18660

ataggttggc catgtctaga attgcaggtt gaaaatcatt tcctctcagt atattggtta 18720

gtgagaagcc tgggactgag acagtcacat tctcacttct ttgcaggtga gtgctcttag 18780

gactgtcttt ttatccctta tactctgaaa tgtcatatgt cttggtgtaa gtccttattt 18840

cagttattga gctggacaag tactggagac cccttcagtc aaagccttct gtcattctcc 18900

agctctagga aattatcttc tattgttatt tctgttattc cttcccttcc attttctttt 18960

ttcttttttt ttttttttt ttgagacagg gtcttactct ggtgcccagg ctggaatgca 19020

gtgacctgat catggtacac tgcagcctga acctcccaga ctcaagtgat cctcccacct 19080

caacctccta agtagctggg actgcaagca cacatcacca cacccaacaa atatttttta 19140

aaaattttgt aagatgggat cttactatgt tgcccagact ttttcttcct cttcctgggg 19200

ctcttattag gaagatgttt gacttcctgg gttggattcc tgtctccgtg tctgactttc 19260

tctctttgtc atatttttca tcactcgttg tctttttgcg tctgctctga cagatttcct 19320

caaattttgt cttctagtcc tatcctacag tttttacttt cagcaaatat aatttaatct 19380

ccaagagtac tctcttgttc ttttttctta gcattctgtt cttgttttat ggatgtaaca 19440

ttctcttgga atatttgctg tcctctagat catcccttct ccatttcttc ttgggctagt 19500

ttttctgttt cttcatcttt ctcttttatg ctacttattc tgggcgtgtt cttggtgggt 19560

tttttcccat atagcaacag aggacttgga gctcagggag aaaagggtag gtgcatcacc 19620

tggcagagct cccagacagt gacaggcagg ctgcgggaag gatgtctact tggcggtgct 19680

accgctttcc tagaaaccct ttccctggag ctggttgaac tgttgggttt tgccctggtg 19740

gtgaacgctg gctccccgtg ctctgcctgt ttcatcacca gccccctccc cttctgcctg 19800

gggtccagta atctgttgaa atatatatct tgctcattgg tgagctcctg ctccttcctc 19860

gttgctcttg cagatttatc acttctcgta aggctgcgct tgtacttcgg gattttctct 19920

gtgccacact gggaaacata gggtggttgc atgctgcagt cctgagcact tatttcactc 19980

acatctttac acgaagattt ggtgggtgtt tactttgttt ttagtaagtt agtctgtcat 20040

gtcctttgat cctttttttt tgtttttga gatggagtct ctctgtgtcc tccaggctgg 20100

agtgcaatgt cgcgatctca gctcactgca acctccacct cctgggctca agagattctc 20160

ctgcttcagt ctcctgagta gctgggatta caggcatgtg ccaccacacc tggctaattt 20220

ttgtattttt agtagaggtg gggtttggca tgttggccag cctggtctca aactcctgac 20280

ctcctgacct gcctgccttg gcctcccaaa gtgctgggat tacaggtgtg agccaccaca 20340

cctggccctg attaatcttt taatgcccag tctctccttc aaaagccggc tcctttctct 20400

ccctcgcctt cctagattcc ttctccactc cccaggatca gcctcctcct ccccacccca 20460

ccactgccgg ggggatgtct gtggtcaggc atttatcaga gaccctgagg tgggggtcct 20520

ttatgtgtct gggggatgga gagtctagag gaggtagcgt tcagacctct ccatggtgcc 20580

tctgctgggc tcacatgtga ccaagcacag caaaccatga ggcaggggat ggtcttgacc 20640

atgagagccc ttgcagcagc tgccatgggc ctcagctcct ctccaagctg ggaagagccc 20700

tgaaaagcca aggtgttttt ttttccctct ttatttcagt gtaagtccct tgagctttct 20760

tgaaccagaa gtgggctcat tttgctttag agatttcagg tgggcttgtc cttgtcctag 20820

catcccagat ccaccttctg ggaagtcatc agattggagg tgatgttggc agcttttgta 20880

aacaaagggt agtgttgtaa gctgttgtgt ctgcctatgt gtgtgtttgt gtacttggtc 20940

tcatctctgc agactggtga catggcttcc agatatgccc gacgatgtcc tgtggttgca 21000

gtgggtgacc tcacaggtgt tcactcgagt gctgatgtgt ctgctccccg cctccaggta 21060

aatactttgg ctgtgggtgt gtgggccgga cgggcacctc tctcatctga tgaggcctca 21120

cacgacattc tagaaacagc tggctgaaca ccaagcaagg agcttgccct tgggtgtggg 21180

gaccctgtct catgggaggc agctgagtca gtcagaggtc ctggcacacc tgctgagagc 21240

tgccacccag gccaacctga accggagcct gggaagactt cccgtcggat gagtctcttt 21300

gagtgcagca ttgatggtgg aagagcagag aggccccaga taagcaggga aaggtgcttc 21360
agacagagtg gctgggatga ggactgggga gtgtcagata gcgctggcgt gtctgagcga 21420
aggagctctg gcacccatgg cacaggaagg aggtgggacc ctggaggggc agggctagca 21480
gagctcctcg gagcgtgtgg ctaggtgcct ggtaatgcaa gccccctgtc ctccaccctc 21540
tgttgtactg agtcacagtc tccggggtga agcctagcag tctgcgttga caggccccag 21600
gggatgccgc tacttcctga attctgaatt ctggaaactg agccggagtt cagggcctgg 21660
ctcccattac cagggttggg cgttatcctg aaaatcatag gccttggttt cctcacttgg 21720

ctaacagggg tgatccccat cccctcaatg ggtttccgtg agctcctgag agcccgtagc 21780
atggtacttg gcacatgctg ggcatcagga ggtatggcct ctcttgctat tgttgttatt 21840
ggtagacaca gaaggattta aaagtagggg aatgcaaaga tccgatttgc tagggaagag 21900
ggcagtagtg gccaagtaga gggtggatcc tgggccctgg ctggcagcag gcagcaaggg 21960
gggctgccag ggcccaggca gggacgatct gtagaccgag aggcttccta aggctcttgg 22020
acaggaggag gtgtcggttc caagcctaag gagtggggca gccctggtga ctggtggtca 22080
gtggtgccag gcggtgggtg gtaggacacc ctggcaggca agtaggtttg tgtgggggaa 22140

actgataggc ccctccaggg attcgttggt ggacaacacc tgtgatgtcc agtgggaggt 22200
gtccaggtag ctgggagggc cacaggcttg gaagacctag gtggtgacat cagcccagca 22260
ctgagggcta gaagaagctg tgtctctggc tgtgacggca ccctagagtg tgtgtggtgc 22320
cctctactgg ccggcaatgt gggtccaccg tagctcagac tgcacactgc agcagcggga 22380
acggcctcta agccaacttc ctccatgtgt ttcaggtccc aaatgccagt gagcagccaa 22440
caggcctccc catgcacacc tgagcaggac tggccctgct ggactccctg ctcccccaag 22500
ggctgtccag cagagaccaa agcagaggcc accccgcggt ccatcctcag gtccagcctg 22560

aacttcttct tgggcaataa agtacctgct ggtgctgagg ggctctccac ctttcccagt 22620
ttttcactag agaagagtct gtgagtcact tgaggaggcg agtctagcag attctttcag 22680
aggtgctaaa gtttcccatc tttgtgcagc tacctccgca ttgctgtgta gtgacccctg 22740
cctgtgacgt ggaggatccc agcctctgag ctgagttggt tttatgaaaa gctaggaagc 22800
aacctttcgc ctgtgcagcg gtccagcact taactctaat acatcagcat gcgttaattc 22860
agctggttgg gaaatgacac caggaagccc agtgcagagg gtcccttact gactgtttcg 22920
tggccctatt aatggtcaga ctgttccagc atgaggttct tagaatgaca ggtgtttgga 22980

tgggtggggg ccttgtgatg gggggtaggc tggcccatgt gtgatcttgt ggggtggagg 23040

gaagagaata gcatgatccc acttccccat gctgtgggaa ggggtgcagt tcgtccccaa 23100

gaacgacact gcctgtcagg tggtctgcaa agatgataac cttgactact aaaaacgtct 23160

ccatggcggg ggtaacaaga tgataatcta cttaatttta gaacaccttt ttcacctaac 23220

taaaataatg tttaaagagt tttgtataaa aatgtaagga agcgttgtta cctgttgaat 23280

tttgtattat gtgaatcagt gagatgttag tagaataagc cttaaaaaaa aaaaaatcgg 23340

ttgggtgcag tggcacacgg ctgtaatccc agcactttgg gaggccaagg ttggcagatc 23400

acctgaggtc aggagttcaa gaccagtctg gccaacatag caaaaccctg tctctactaa 23460

aaatacaaaa attatctggg catggtggtg catgcctgta atcccagcta ttcggaaggc 23520

tgaggcagga gaatcacttg aacccaggag gcggaggttg cggtgagctg agattgcacc 23580

atttcattcc agcctgggca acatgagtga aagtctgact caaaaaaaaa aaatttaaaa 23640

aacaaaataa tctagtgtgc agggcattca cctcagcccc ccaggcagga gccaagcaca 23700

gcaggagctt ccgcctcctc tccactggag cacacaactt gaacctggct tattttctgc 23760

agggaccagc cccacatggt cagtgagttt ctccccatgt gtggcgatga gagagtgtag 23820

aaataaagac 23830

<210> 31

<211> 23830

<212> DNA

<213> Homo sapien

<400> 31

atggtccgag gggggcgggg ctgacgtcgc gctgggaatg ccctggccga gacactgagg 60

cagggtagag agcgcttgcg ggcgccgggc ggagctgctg cggatcagga cccgagccga 120

ttcccgatcc cgacccagat cctaacccgc gcccccgccc cgccgccgcc gccatgtacg 180

acgcagagcg cggctggagc ttgtccttcg cgggctgcgg cttcctgggc ttctaccacg 240

tcggggcgac ccgctgcctg agcgagcacg ccccgcacct cctccgcgac gcgcgcatgt 300

tgttcggcgc ttcggccggg gcgttgcact gcgtcggcgt cctctccggt atcccgctgg 360

gtgcgtctgg ggacgctgcc cgggctccac gtgcggagtg ggtgcccccc aggccgggga 420

gcgggggatc cccaggggtc gcggggccct ggaggagcgg gcatcggacg cggacacggc 480

ggggtgcatc ccgagggccc cctccgaggc agatgcttcc tgcgggggcg ctgttcctgg 540

gcccgggaag ggggcgttgg aaccccgagc ggtccgggcc gaagcctggg actctcgtgc 600

gtccccaccc ctacccccat caggcgcccg tgcatgaagg gagaccctca cctccggact 660

gagagtcgga gcgtctcgga gcgacgggga gtagggagcg ggacccgggg cggagggtag 720

tgctggcccc tgcggactcc gggtcccctg tgtcctctcg ggaggggctg gacgggctga 780

gctgccgagg ggccgatttg ccctgggccg gacaaagagt ggggctttgg ccggtccccc 840

acggtgggct ccttccctct ggggattgag ggactcaaga caccccgcgc ctgcgctttt 900

cttttctttt tttctttttt ttttttgag acggagtttc gctcagtcgc ccaggctgga 960

gtgcagtggc gtgatctcaa ctcactgcaa gctccacctc ccaggttcac gccattctcc 1020

tgcctcagcc tcccgagtag ctgggactac aggcgccagc caccaagccc ggctaatttt 1080

ttgtattttt tagtagagac ggggtttcac cgtgttagcc aggatggtct cgatctcctg 1140

acctcgtgat ctgcccacct cggcctccca gaatgctggg gttacaggcg tgagccactg 1200

ctccctgctg cctacgctct ctgggtcgca gcccagcctt ctgggggctg ggtagcctcc 1260

cagaagggca accctgggca tcctccaggg caggctaact ggagtctagt ggggaggggt 1320

accttgaaag aggaaagttg tttcctcctc ctcctcctcc tccagtgttt gggacccttc 1380

ctgggggctg gagtgcatcc ctggacaccc cccaatccca tcctcttctc tagtttccac 1440

tgacctaggc ccaccctccc ctctccggct cagtactcct ggaaatgaga ttccgtacat 1500

ttgaatcttg tcctaatgaa atatttgtcc atgtgggtac ctgtgtgtgt gtggtggggg 1560

tgcagacgga gggtttgttt ctcactagct ggaactactg gggtgtggta tgcttcctgg 1620

gaatttgtgt gccacagtcc tggaggcgag gaggggggttg tgagccagta ggcaggggct 1680

ggggcaagta gcattgtgaa gctattgaca cccagacgtc cccaggcagg agattatgcc 1740

cccattagcc ccctttttatc tgggcttcct taacaatgga ctctttgccc tgcctgccag 1800

agccagcagg gagtgactgt tcagtggtga ggaagcgggc agaggaagcc ctgccattgg 1860

gtaggagcag tgggcagccc ctgggctgac tgggaggtgg ggattaggga ttagacagtc 1920

ctggctgtct gccttcccct aagccagggg gagaggagca aagggcacga aatgtggcct 1980

ccaggaggat tagaccgcca catgatcatt tgcacaccct ggggtttagc aacaataaaa 2040

gtcagctttt ttgtatccca aggtggcctg tggacaccca catggacaaa tgtttacact 2100

gggacagaat tcaaatgcag aggtcccagg agcctaaagt acactcactc tggtatagaa 2160

aggattcctt actgggcaga ggacaggtgc agcctggggc tttcccaggc aggacacagg 2220

gaggctcagg aaccaccaag tccctggaag gtggatctgg aggcgttggc aggagccact 2280

ccctgggttc cagggctcca ggttcctgct ttaacccccc gtctcacaga gggctgtgca 2340

cttgggggct gctgagcatg tcccagaggc tgcatcctgg acacagcacc tcagtgcatc 2400

tgagctgagg ctaacttggc aggagggaca ggcagaacct gccagccacg tgcaattcca 2460

cccctctggc cactcaggga aggagagctg tgagtcaaga tcagatttgg gtcaggacag 2520

gctggggcct gcctgtccct gtgcatccca agatttatgg ctggccaggg gttgggctgg 2580

gaggggtggt cttgcatgcc aggagagtgc agatcagcct gagaggccag gccagtaagt 2640

gaggtcagat ctcctgcacc tgatagcatt aaggccatct acaccaaagc tctaatgctg 2700

atatgttcct ggcctctatg tggggcatgg aggtggggca tggaggtgag gcctgctcgc 2760

ctgggcttct ggaagtggga gactcattcc tgtggctgag gcctacagca gtgctgtgtg 2820

gtaggaatac actggaagcc atgatgtcat tgtgcatttt ctagaagcca cattgaataa 2880

agtaaaagac acaggtagaa ttaatttcat tgagcccaat atatccaaaa taatatcatt 2940

ttcacatcta ttcaatataa aaatttacta atgagatatt tcatactaag ccactgaaat 3000

ccagtttgta tcttacacat ctcagttttg acgagccaca tttcaagggc gtgatagcca 3060

catgtggctc ccatagtaga cagtactggt ctagagaaat gttggtggca tccttgctgt 3120

ctggtttctg gccttgccaa aagtattacc atcccagtgt ggtacattct ttcatgtatt 3180

tgtctcctgt ccccagagca gactctgcag gtcctctcag atcttgtgcg gaaggccagg 3240

agtcggaaca ttggcatctt ccatccatcc ttcaacttaa gcaagttcct ccgacagggt 3300

ctctgcaaat gcctcccggc caatgtccac cagctcatct ccggcaaaat aggcatctct 3360

cttaccagag tgtctgatgg ggaaaacgtt ctggtgtctg actttcggtc caaagacgaa 3420

gtcgtggatg taagcagttt gcttatctgg acgttgtcaa gttagaaaag ctgttttggg 3480

atgggtgtgg tggctcatgc ctgtcatccc ggcactttgg gaggccgaag cgggtgggtt 3540

gcttgagccc aggagctcga gaccaacatg atgaaaccca gtctctacaa aaattacaga 3600

aaaattagct aggcatggtg ttgtgggccc atagtcccag ctactaggga ggctgaggca 3660

ggagaattgc ttgagcctgg gaggtggagg ttgcagtaag tcatgatcat gccactgtac 3720

tccagcccgg gtgacagtga gatgctgtct ggaaaaaaaa aaaaaagaaa gactgttttg 3780

ttttggaagc aacacaggca gttgtaggcc ccctgtgcca gagtgacata aactctgtac 3840

acctccagtg atttggtcca tgtttgtaaa ccctgaatgt tccagggcag tttcttttct 3900

tcacttttta tctctttttt ttgggtgggg gggcggggta cagagtcttg ctctgtctcc 3960

caggctggag tgcagtggcg caatctcaac ctcccgagga gctgggacta caggcacagg 4020

ccatcacacc ttgctaatgt ttgtactttt tgtagagacg gggttttgcc ctgttgccca 4080

ggctggtccc aaactcctgc acccaagtaa tctgcccacc tctgcctggc agttacaatt 4140

tcaaataatt cctccctttc cttcaacact tggctcatga ccgtccagtc caaggaacct 4200

gtcctgcagg tgtgcctctc ccgagcttcc tctatgcatc ttccataatg aagatgcctt 4260

ctcactggaa accctacaag ggtgggaacg tgccttattt gcctgtatcc tcagggtcta 4320

gcagagagaa gataatttgt aataccaaaa caccattaaa ttcagctgat gctttcataa 4380

gcgctccttg gaggaaggac tccatttact tgacagatct gtgcaagaca gcagcctggc 4440

gcgtctaacc tgcagccagt tgcatcctct gtttaacctt gtttgcggaa gctttctcta 4500

aacagccagc acttgtctgt tcccacatgg gtccgttctc ccagtgaatc accgtggtgc 4560

ctgctgactg ctctgtagca cagtgcttcg caaagtgtga tcctgggacc agcagagcag 4620

cagctccttt gagcttattg gaatggcaga ccctcaggtc ccacctctga cctgctgcat 4680

gggaattctg gggagggacg cagaatctct ggttccacag gctctccggt gatgctaatg 4740

aataccggca tttgaacagc accgatctag ccccttttcag tccatgagcc aacaaccctt 4800

ggtcctgtct gtggtgaccc agtgtgactc tcatggggag caaggagagg aagttgaagt 4860

tcactgacag ggttgttaag gggattatgc aatagatgag acccatgggc ctgaagtccg 4920

agggtgtatg ttagttcccc gttcttttga cccatggatt aacctactct gtgcaaaggg 4980

cattttcaag tttgttgccc tgctcacttg gagaaagctt atgaaggatc aggaaaatta 5040

aaagggtgct ctcgcctata acttctctct cctttgcttt cacaggcctt ggtatgttcc 5100

tgcttcatgc ccttctacag tggccttatc cctccttcct tcagaggcgt ggtaagtcgg 5160

ctttctctgc tagcgctgag tcctgggggc ctctgaagtg tgctcacaca tctcctgcct 5220

gcagggcact ggtgtcgggc acctcagggt ctgtcccatg gtggagcccc atgcctcact 5280

gcctttcaga cagagtagcc acagctggcc ctatttccag gctacccggg cagcaaaact 5340

tactgcatgt gtaattaatt atttggctat ctgtaaggta aactggctgg ttcacttaat 5400

ctgcacctta agcatcagat agcttctcag tgatctagtt aaactatatg atgttggcca 5460

ggcgcggtgg ctcatgtctg taatcccagc actttgggag cctgaagcag gcagatcact 5520

tgaggtcagg agttcgagac cagcctggcc aacagtgtga aactctgtct ctcctaaaaa 5580

tacaaaaatt agctgggcat ggtggtgtgc acctgtaatc ccagctgctc gggaggctga 5640

ggcaggagaa ttgcttgaac ttgggaggcg gaagttgcag tgagccaaga tcgcaccact 5700

gcactccatc ctgggtgaca gagcgagact ctatctcaaa aagaaaaaaa aaaaaaaggt 5760

aaataaagta tatgacactg aagaatctgt tacccctgga aggtggagct ttactcttag 5820

ggggaactat aacagtcata tatatatatt tttttctttt ctttttttt ttttttgaga 5880

tggagtctgg ctctgtcgcc caggctggag tgcagtggtg caatctcggc tcactgcaac 5940

ctccacctca caggttcagg caattctcct gcctcaacct cccgagtagc tgggattaca 6000

ggtgcctgcc gtcacgccaa gctaattttt gtatttttag tagagacagg gtttcatcat 6060

attggccagg ctggtctcca actcctgacc tcaggtgatc cgcccgcctt ggcctcccaa 6120

agtgctgaga ttacaggcgt gagccatggt gcccggccaa caatcacatg tgttgtaaac 6180

aacaacaaaa atctgtcagc ctggtctaac ctagatttgt gctttgtttt gttttgccac 6240
tttgtgatgc acaggaggaa gtttaggctg taaaatacta gccttttagg gtaatttttg 6300
aactcacaag agcagcagcg gaacctttga tgcaatcctg tatgtagcac cagcagagcc 6360
acgtggcaga gggactcgca ttaggagcct cccattacag actacgtgct cctgtgcgtt 6420
atcttatagg gtccccacaa ccaaggggag atgtgattat tcatcctgtg tggctgtggg 6480
gaacttgaga gtcatacttg cccaaagagc acggccagcg agcttgcacc caggtcactc 6540
tctgctcctc tgtcagaaca gggcatgtct tggttcactg cagggcggct cttctcattc 6600

tctgtagttt ggggtccagg atagtggtcc acggagccac tggagtgccc agctactgag 6660
tgaccaaagc atattttgga tttccgacat tgccacagca tggttgggca tcagcaggac 6720
cccaacccct tgttatgctg gtggctttat gtggttattt gatcttcccc agaactcagc 6780
aggagtgcac ccagcagcac cgtagtgatg ctctctggct ccccagtgca cggttctggc 6840
tttccttcct ggtcgagagt ttcaagccct ctgggtccta ctctgtcctt ttcagcccat 6900
agctttgttc aaaagctgct ggcagtgttc agatttggct gagttcagtg aatatgtgca 6960
ttggctgatt tctgagccat gccaggggga tggagaagcc gaagcaggag tgtttgttct 7020

gcaggctctg gagtaggcat tgggtctgtg ccggctcact tgctagtctt gcatccttcc 7080
ccaacccccc ctggggatgt ctggccacat cagaagacag tttgggttgt cagaactggg 7140
ggagtaccag gccgaggtgg gtggatcatg aggtcaggag atcgagacca tcctggctaa 7200
cacagtgaaa cctcatctct actaaacata cgaaaaaaat tagctgggcg tggtggcggg 7260
cgcctgtagt cccagctact cgggaggctg aggcaggaga atggtgtgaa cccgggggggc 7320
ggagcttgca gtgagctgag atcctgccac tgcactccag cctgggcaac aaagcgagac 7380
tccgtctcac aaacaaaaca aaacaaaaca aaacaaaatc tgggggagtg ccactggcat 7440

ctgatgtata gaggcccgag atgctgtgtc atcacccgtt gagtgcgctc ataggcatct 7500
tcctgacaat tagaacccat tattcttcaa attcaatgca agcaaattca aagcattact 7560
gtgtacatac cgcatgctaa tcaattgcac cactggagct cctaaattca aaacattact 7620
ataaaaaagt tcaaaatgca tggaaaagtt gtacatggca ggagaatatt tgggcttctg 7680
actacccctt gaatgaagat gatccaccag ccgccttcct ccttggtctt cactccagat 7740
tcctagcatt tcattctgtg tctctttatg cagtgaggtt tttgtttgtt ttttgagaca 7800
gagtctcact gtatcaccta ggcctggagt gcagtggcgc gatctcagct cactgcaacc 7860

ctcggctcct gggtttaagc gattctcctg cctcagcctc ccgagcagct gagattacaa 7920

gcacacatcc ccatgcccag ctaatttttg tatttttagc agagacaggg tttcaccatg 7980

ttgcccaggc tggtctcgaa ctcctggcct caagtgatcc atgtgcctca gccttccaaa 8040

gtgctgggat tacaggcgtg agccaccatg cccagctcct agtgaggttt ttgatgcctt 8100

gctacatctg ccctagaaat tgtgtgacta cgattttgga aatgttgctg tgtaaacttg 8160

tgatcatttc tggactccag gcaagaatct tgatggctaa ggtgtggctg aacatgtctg 8220

attctctcct ggacctgttt taggccaaac tctgctctga aattcctccg tgtggaaggg 8280

cgggctgggg agagcctccc agctggaatc ttttggatgc ctttctctgt gggtatctga 8340

tggctggctc tgatggctgg ctgtgatggc tgtggctgga aatcattgtt gacatgagtt 8400

tcacagatgc aggctctgtc caaattgtag caaaagctgc ctgccccagc cgagctatgg 8460

gcaataaggt ggtttaagga tatagatgaa ggaaaactca cccttagaat aatttatcca 8520

aaatgctgct gtgttgtggg ttagaggaca ttttctgagg tcccaggttc attgtttcat 8580

ttaagtctca aaagtccctc caggtgttgg ttctaattgt caaagcatgg ggggagatgg 8640

gctcatgggt taaaggtctt atcccagatt tctgtatcct ccttgcaagc agcaaagggg 8700

tctggatttg aatccatgac catgtttctc ctttgggttt ccatcacact ctgtccccgt 8760

gcactgagca ccctttagtt catatgaccc ccttaggcat gttacatggg cactcctata 8820

ggtgcccatc tggccctagg acttggccaa cacaacatgg actccagttt ccatctgcct 8880

ctttgccagg cacttttgtg cagtgcacac actgtacaac agtagacggc aaccctgaga 8940

gccagagtag agcctgtcct agcaccggaa tgctcggtaa ggatttgtcg caggagtgat 9000

tccaaagcca atgtcctccc tccatatcag cctgtttgtg gctctgagaa gctctgccca 9060

catgtgaaag cttgttaagc acttaagcac taacccagag cttcagacag tgccagtcct 9120

ttttccccct ctttaaaagc gatatgtgga tggaggagtg agtgacaacg taccttcat 9180

tgatgccaaa acaaccatca ccgtgtcccc cttctatggg gagtacgaca tctgccctaa 9240

agtcaagtcc acgaactttc ttcatgtgga catcaccaag ctcagtctac gcctctgcac 9300

agggaacctc taccttctct cgagagcttt tgtcccccg gatctcaagg tgagttggtg 9360

gtgaggggc aggtgttctg gggtgcagct cttctttgcc tccctgattg ccaggagcta 9420

ccagttactg tctgcacaat caaacagaaa tagacctgtc cttgatggtt aacggaaata 9480

aaaggcgctt gtcccagaag ctcaggtgag gcaccaccct gattatggga atcacctggg 9540

aacatatacc cagacctaaa actcagatcc acttcccagg ctgtggttat atagtcaggg 9600

gggtgcagta tgggtattag gatttttat tttttagtta taaagatttt tttttggttt 9660

gtttttgaga cagggtcttg ctctgccgct taggctggag tgcagtggtg caatcatagc 9720

tcactgaagc ctcagactcc tgggttcaag cagtcctccc acctcagcct cctaaggagc 9780

tgggacccac aggcatgcag caccacacct ggctaatttt taaaaatttt gtggagtgtt 9840

gcccaggctg gtctcacact cctggcctca agcgatcctc ccaccccagc ctcccaatgt 9900

gttgggatta caggcatgag ccattgtacc cagccactaa gatgattctt atttggaaac 9960

acggtcaaga acaactgcgt tcggtagttt aacctttttt gattgtggtg gttttagtat 10020

gccttaccac tctaccatag taagaaattt gcagaccatg tacaccaacc tttggtgctc 10080

ctggggagaa agaaagaagg ctatgcaatg caatgcatgc tcacagtcca agggagaggg 10140

aaagctgtct aacaggattg gttttcccgt gtgctttata agcagatgag tagaggagac 10200

agctcttatt gtcctagtgg caattgggat aggctgcaaa gtttgttagg gtggaggctt 10260

attccgggac caagggagcc caaagaaaca agctcctgcc aggcgcggtg gctcacgcct 10320

gtaatcccag cactttggga ggctgaggca ggtggatcac ctgaggtcag gagtttgaga 10380

ccagcctggc caacatggtg aaaccccgtc tccatgaaaa atacaaaaat tacccgggca 10440

tggtggcggg cacctgtaat cccagctact agggaggctg aggcaggaaa atggcttgaa 10500

cctcggaagt ggaggtggcc gttagccgag atcacgccac tgcactccag cctgggcaac 10560

agagcaagac tctgccttaa aaaaaaaaaa aaaaaaaaga aaagtaaaag gaaaaaaaag 10620

aggctctggc ctgctggggt gcctgcaaag tctccgtgga agggtgacat tcaagccgag 10680

acctccaggg aactgtctcc tgggagcaca gagccctttg ctcagccccc aggtggctca 10740

gtgccccccag ccagcagact cagagcttgc atgattcttt ggtgctctct gcggtcttcc 10800

aatgatgctg aaataaatgg tgcttggtgt ctccctgctg tagtcccctt gcttgctttg 10860

ctcacaggtg ctgggagaga tatgccttcg aggatatttg gatgcattca ggttcttgga 10920

agagaagggt atgtatgggc tgggaggatc agccatgccc ttttgacaag catttactag 10980

cggtcttggt aaagacttga gatttgcctt agttctaaca cttagtgccc aacgccttcc 11040

ttgtgttgct caacctactc atgagcccag gagataggaa atctccgtcc cattgtacag 11100

atggggaaac agaattttgg aaaggagagc caagcagcac acacccctcc ctgaggggca 11160

gagccgagat ttgaactggg atgtcatgac tccagggccc tctccctccc cagggtcccc 11220

ttatctgaag gcggtttttc tttccagctc gacctcttgt gacccttagt ttaacaaggg 11280

ccgaagttaa agagtttctg cgcctggacc ccaaatgaag caatcagatt tctcatctcc 11340

agtcaggtgt gggtccaagc ccactagaca agtttgctct tcccagagca catttctgcc 11400

ttcaagtcat cctggcttgt cagggctggg ggagttctgc tgtagaaata ttagagtgga 11460

aggaaaaaga tgtgttggga gctatttttc tttaatacta aaagttggtt gatgaatttg 11520

tcgttggcca agaccaagga gactgcattt ttaaggacat atgtgtattt atctgctcag 11580

aaaatgttca ttgctgtgtg ctagggatac tgcagtgaac acagaggtgt gacccttgcc 11640

agccttgtga gagaagtgag cagataagta agcagaaggg tgatgctgtg tcgatgggaa 11700
agtacaggtg ccaatgagaa ggcacaggtg tcaaggagaa gacacaggat gctggaggct 11760
catgcaggat ggatctccaa ggcccagggg aagaagggcc tctcggagga cgtgaatcca 11820
cattaagact ttggggataa gtaggagcgc cttaggcatg gggacccatg gatgcgaggc 11880
ctgtaggaca cagacaggat ggcatgaagg cctgtgcaac tggaggggtg gggatgggga 11940
cactaagaga tggctggaag tgtgggggtg gggacactaa gagatgactg gagaagaggg 12000
ggtcaggagt ggtgaaaaat gggagaggag ggcaggctgg gccttttgga tacaggggga 12060

ttgcatcctg cagtggtagg gagccactga gggctgctgc agtaggagtg aggggatcag 12120
aggagagctt tggaagcccc ctggatgcgg gacaggaagc gagataccag tgtctaggag 12180
gccagtgagg cagccacagg ctccaccagg atcagggctg cgagggtcat gaggaggaaa 12240
ccaatttgaa ggagtccagg ggaataggac ttggaaatga ccgatgggac atttgggaag 12300
aggaagacag aagagcgcag tcccggcttc tggctttagc agttgggcaa ggggagatgg 12360
ggagatgtgc ccatgggttg agggttgagg acattaggag ggagccggta tggcaggaag 12420
agctggtgtg ccagagatgc tggaagcagc atctgcctga gaacagatac ctggcaatat 12480

tcctaaggga aagtgacatc tcggagggtg aggagggcat ctgatagggc ctggaaagag 12540
ccggggcaag catgaatgtg aggttatctt ggggggcaag gctcaggcgt tgaggagcag 12600
cccctggtct cttcagcctg aagttggaag ccagagttgg gccaggtgca gctgtggttg 12660
tctgaagtcc ccctccccca gcccagtgtg ccaatgctgt aagagcaagg gccgctcact 12720
ggtgctggtg gctgagtccc agcacccagg acagggcctg gcacatactg gtgcccaatc 12780
ctcccttctg ggtgcttctt ccaaggcctt gtgatggaag tgagtaccct cttcgacatc 12840
agacccagct tcaaatcctg gctctgctat gtattggctg cgtggcttta gacaagtctt 12900

ttaaccttgc tgtgcttctg atttctcagc tgaaaaatgg agatgatgat agtggtttct 12960
gtaaggcctt atggtgaagc acctagctca gggcctggaa ggcaggtgta accagtggtt 13020
cagttgttat aaaccaacac taaccctcgc ctttgcacct catgaaacca gatatgtaga 13080
tggagcccac aaagctagca ggagccaagc tcacgtgtgt cctgctttaa agccccatac 13140
ccctttctcc gggtgacaaa cacctgtgct cgttctcttc ccttcccctc ttccccttgc 13200
atttggctaa taacaggcca gctgcctgcc tccctgcagt ttggtagatg ggtgggtaac 13260
gaccaccact cccacgttcg cctgatgggc ttgttttccg tgcccttcac aggcatctgc 13320

aacaggcccc agccaggcct gaagtcatcc tcagaaggga tggatcctga ggtcgccatg 13380

cccagctggg caaacatgag tctggattct tccccggagt cggctgcctt ggctgtgagg 13440

ctggagggag atgagctgct agaccacctg cgtctcagca tcctgccctg ggatgagagc 13500

atcctggaca ccctctcgcc caggctcgct acaggtaccc actcctcggg gtgagcacgg 13560

gcagcacctt gttttctttc ttgtgcatta tggaggaaga tggtactgcc acatgggagc 13620

gatagggtga ggcaaccatg acaggtggtt gggaacatct ccttccatgt gtacagcctg 13680

ggctgctgcc atcactccca gcacagcccc caaccccccc aatcctggaa ccttgccaag 13740

tctcccttcc catggggtca tgaccaggag gaaaacaaac tccagctgag ccccttgggg 13800

ttccccatat aggctcctgc ctgtggcagc tgggccctct gtaccccttt ccaactctgt 13860

ctccctaaca tggcacctga gctcctgcca tcctggattt catggacccc aaggatgggg 13920

gtcctgcatc tgggacttgg cctattactc ggagctcctt ttcagccgcc tccctccacc 13980

tgtccaccca cctcaaggct cctttcttga gacctctcct aatttctccc ttcccctaaa 14040

cccacaattt tgaacctcca tcgaatggtg ctgtatttta taatgtcatc aaatatcaaa 14100

tggagacagt gctatggtcc aaatgattgt gtacccccca gaatttgtct tttgaaatcc 14160

taaccccaa catgatggtc ttaggaggtg gggcctttgg gaggagatta ggtcatgagg 14220

aaagggctgt catgaatggg attggtgccc ttattaaaca gacccaagag aggtcccttg 14280

tcccttctac tgtgtgagga ctcagaaggt ggtgtctatg aagaagcagg ccctcaccag 14340

acaccaacat gtctgctgcc ccttgatctg ggaccttgca gcctctagaa ctctgaaaaa 14400

tcgatgtttg ttgttttata agccactcag ttggtggcat tttgttagag tagcctgaac 14460

acggactaag tcaaacagaa gaacccacaa accagctaca gagttgggca tttggagaaa 14520

ttcaaaaatg agtcagacat aactccttat tcttgaggtg ccctaagaga tgggacacag 14580

cagctgccca ggtgcattag tttgttctca cattgctata aagaaatacc tgagactggg 14640

taactcataa agaaagaggt tgaattggct cacagttgca caggctggac aggaagcatg 14700

gtgctggcat ctgctcagct tctggggagg cctcaggaaa cttacaatca tggcagaagg 14760

tgaacgggaa gcatgcacat cccatgactg gagcaggagt gagagagaga gggaaataga 14820

gggaaggtgc catacacttt taaacaacca gatctcatga gaacacattc actatcaaga 14880

gaacagcacc agtggggaaa tccgccccca tgatccaatc acctcccatc aggctccgcc 14940

tccaacactg ggaattacaa tttgacatga gatgtgggca gggacacaga tccaaaccat 15000

atgaccagat taatacgatt tgaggcatca cgaggtcatt aaagagaggg aataaaagac 15060

tggggctcca ggaagaaggc tctggaatcc agcagagggt caaggaccag cttgtaaagc 15120

tggtggtgcc tgagaagtac ctaggagaac atagatgctg tgacgtttga tgtagctgtt 15180

ttttgttttg tgttttggtt tttgagacag agtctcactc tgttgcccag gctggagtgt 15240

gcagtggcgt gatcttggct cactggagcc tccatctccc aggttcaaat gatcctcatg 15300

cctcagcctc ctgagttgct gggattacag gtgcacacca ccacgcctgg ctaatttttg 15360

tgttttcagt agagacaggg tttcaccatg ttggccaggc tggtcttgaa ctcctgacct 15420

caagtgatcc aacaacttca gcctcccaaa gtgctgggat gacaggcatg agccaccatg 15480

cccagcctga tgtagctgtt tctgtgcaca ttatttgctg tggggtatat tcagatttct 15540

taatacaaga tgattctttg cctcatgact tacacaccat tttctattta atttcagcta 15600

tgatattgga aatggacatg tcttttcaag gaaaataaaa gcaggctttc tggaatggcg 15660

acttccaaac atatttgtca atttaaagga gctgggagtg gggaccctat gctccgtaag 15720

cactctctta gctgttcttg gctgtgctcc ccgcttcagc ttcacactgc ccttgctgtg 15780

aagggagcag cctgggccgg gcgcggtggc ttacacctgt aatcctagca ctttgggagg 15840

ccgaggtggg tggatcacct gaggtcagga gttcaagacc agcctggcca acatggtgaa 15900

actccatctc tactaaaaat acaaaaaatt agctgggcat ggtggcaggt gcctgtaatc 15960

ccagctactt gggaggctga ggcagaagaa tcgcttgaac ccaggaggcg gaggttgcag 16020

tgagccgaga ttgcgccatt gcactccagc ctgggggcaa caagagcaaa actctgtctg 16080

gaaaaaaaag aaaggagcag cttggcaaac cccaccttgt cgcttttgtg agtgcctctg 16140

acccctttggc tgccaggacg ggcgtatttt atggaaatgc taagcaccaa cagagtaaag 16200

tggtttggtt tttcacagtg gtgggagata atagctccaa attgtctttt tcagcactga 16260

gtgaagaaat gaaagacaaa ggtggataca tgagcaagat ttgcaacttg ctacccatta 16320

ggataatgtc ttatgtaatg ctgccctgta ccctgcctgt ggaatctgcc attgcgattg 16380

tccagaggtg agcattttag gtggctccgt gtcttcctca cagggttgat atgaggatga 16440

aacaagatga tagatcatgg tggcatgtag tctgggacct ggattgtcgt gccacagatc 16500

acagctcaca gtctatgtgc aatgcccctg aatgttgccc acctgtcctc aagccacaca 16560

tgcacctgta actcagtgca agcccagaaa ctccccgtgg ggactcctag agctgtcagt 16620

ggcctcacat agcagctggt ccagtctctt gtgattgccc aaggaaactg aggcctggag 16680

agcttggggt cactgctctg aggccataga gatgcctagt agaagggcca ggcctagaag 16740

caggatcctt gctgcccctc tgagctgttt ccatttaaaa tcacatgaag gccggcgccg 16800

tggctcacgg ctgtaatccc agcattttgg gaggccaagg tgggtggatc atgtgaggtc 16860

aggagtttga gaccagcttg gccaacatgg tgaaatgcca tctgtactaa aaatacaaaa 16920

attagtggag catggtggca cgtgcctgta ctcccagcta cttggaaggc tggggcagaa 16980

gaatcgcttg agcctgggag gcagaggttg tagtgagcca agattgtacc actgcactcc 17040

agcctgggtg acaggagaga aaccctatct caaaataaaa tgaaaggtaa tgaaatgaat 17100

aaaataataa atcaagtcac ggccgggcac ggtggctcac acctgtaatc ccagcgcttt 17160
gggaggccga ggtgggtgga taatgaggtc aggagttcaa gaccagcctg gccaacatgg 17220
tgaaaccatg tctctactaa aaatacaaaa attagctggg catggtggtg catgcctgta 17280
atcccagcta ctccggaggc taaggcagga gaattgcttg aagcaggacc taggaggcgg 17340
aggttggttg cagtgagccg agatcatgcc actgcactct agcctgggct acagagcgaa 17400
actccgactc aaaaaaaaaa aaaaaaaaaa atcaaatcac atgaaagtag aacataggga 17460
attccatctt tcgttctagg catagtttgt taatatgatt cagagccagc agttaggaga 17520

acacagtgtg actctcctag aacttcttga ttgggcttcc tctgattggg tttcctctga 17580
ttgggcttcc tctgaaagtg ggggggatgg ggggtgggga gcagaatggt cagagcttgg 17640
ctcagcagtc agactgctct tcttcaaatc ctggctgcat tgcttactac agctgtgtga 17700
ctccagatga ctgaatccac ctctctgtgc tgcagcttcc cgtctagaga gatcacctgg 17760
agcagagggt ggtcaggaga ctcaatctgg ttactgactc acagtgcagg agtactcatc 17820
ccatagtaag catccagcta gagatgttga tttctatttt caggtaataa tgatgatcgt 17880
aaaattagag acagataaaa ggtatgggca ttaggccagg gcactgcaat ttctaagctg 17940

tgtgacctca ggcaagttac tcgacttctc tgagcctcag cggtttcatc cgcaatatat 18000
ggataggaaa accgacctca gtgggttgtc tgacagtgga gggcacttga ttaaaaaaaa 18060
aaaaattacc ctggtctgaa tattaccctg gactgaaaga aaaatattga gctaatacag 18120
gcatcaggaa tggggctgca gggagtccag ggaagggaga acgaagagcc tgaaggtgtg 18180
aggaggtgcg agtgctgatc tgtctgctac aaagaggctg ctgagcctcc tgtggatgtg 18240
gccctggact tggcagttta atacctgagc tgttaaaata acctcagatg ctgtgttctt 18300
taaggggtag gattcagatt cctgctgaaa tgcttctgaa agggagggaa tgagccagcc 18360

catccccagt tgctttttaa gatcattggg aagttctggt cttgccattt gtccctggac 18420
cactcttagg tcctcctgcc ccacttccat ctgggtgtgt gccctgggct gtccaccaca 18480
cagctacatc ctgccatctt ccctcctgga gccactgtgc catgcatgga tctgtagctt 18540
catttttctt ggcttttccc tggtttttct ggagcagagt ctctagtaaa ctcccaagga 18600
agaaaacgtt tgactttatg tgtgttggga aacgtgcttt ttttctatta catctcagtg 18660
ataggttggc catgtctaga attgcaggtt gaaaatcatt tcctctcagt atattggtta 18720
gtgagaagcc tgggactgag acagtcacat tctcacttct ttgcaggtga gtgctcttag 18780

gactgtcttt ttatccctta tactctgaaa tgtcatatgt cttggtgtaa gtccttattt 18840

cagttattga gctggacaag tactggagac cccttcagtc aaagccttct gtcattctcc 18900

agctctagga aattatcttc tattgttatt tctgttattc cttcccttcc attttctttt 18960

ttctttttttt ttttttttttt ttgagacagg gtcttactct ggtgcccagg ctggaatgca 19020

gtgacctgat catggtacac tgcagcctga acctcccaga ctcaagtgat cctcccacct 19080

caacctccta agtagctggg actgcaagca cacatcacca cacccaacaa atatttttta 19140

aaaattttgt aagatgggat cttactatgt tgcccagact ttttcttcct cttcctgggg 19200

ctcttattag gaagatgttt gacttcctgg gttggattcc tgtctccgtg tctgactttc 19260

tctctttgtc atatttttca tcactcgttg tcttttgcg tctgctctga cagatttcct 19320

caaattttgt cttctagtcc tatcctacag tttttacttt cagcaaatat aatttaatct 19380

ccaagagtac tctcttgttc ttttttctta gcattctgtt cttgttttat ggatgtaaca 19440

ttctcttgga atatttgctg tcctctagat catcccttct ccatttcttc ttgggctagt 19500

ttttctgttt cttcatcttt ctcttttatg ctacttattc tgggcgtgtt cttggtgggt 19560

tttttcccat atagcaacag aggacttgga gctcagggag aaaagggtag gtgcatcacc 19620

tggcagagct cccagacagt gacaggcagg ctgcgggaag gatgtctact tggcggtgct 19680

accgctttcc tagaaaccct ttccctggag ctggttgaac tgttgggttt tgccctggtg 19740

gtgaacgctg gctccccgtg ctctgcctgt ttcatcacca gccccctccc cttctgcctg 19800

gggtccagta atctgttgaa atatatatct tgctcattgg tgagctcctg ctccttcctc 19860

gttgctcttg cagatttatc acttctcgta aggctgcgct tgtacttcgg gattttctct 19920

gtgccacact gggaaacata gggtggttgc atgctgcagt cctgagcact tatttcactc 19980

acatctttac acgaagattt ggtgggtgtt tactttgttt ttagtaagtt agtctgtcat 20040

gtcctttgat cctttttttt tgtttttga gatggagtct ctctgtgtcc tccaggctgg 20100

agtgcaatgt cgcgatctca gctcactgca acctccacct cctgggctca agagattctc 20160

ctgcttcagt ctcctgagta gctgggatta caggcatgtg ccaccacacc tggctaattt 20220

ttgtattttt agtagaggtg gggtttggca tgttggccag cctggtctca aactcctgac 20280

ctcctgacct gcctgccttg gcctcccaaa gtgctgggat tacaggtgtg agccaccaca 20340

cctggccctg attaatcttt taatgcccag tctctccttc aaaagccggc tcctttctct 20400

ccctcgcctt cctagattcc ttctccactc cccaggatca gcctcctcct ccccacccca 20460

ccactgccgg ggggatgtct gtggtcaggc atttatcaga gaccctgagg tgggggtcct 20520

ttatgtgtct gggggatgga gagtctagag gaggtagcgt tcagacctct ccatggtgcc 20580

tctgctgggc tcacatgtga ccaagcacag caaaccatga ggcaggggat ggtcttgacc 20640

atgagagccc ttgcagcagc tgccatgggc ctcagctcct ctccaagctg ggaagagccc 20700

tgaaaagcca aggtgttttt ttttccctct ttatttcagt gtaagtccct tgagctttct 20760

tgaaccagaa gtgggctcat tttgctttag agatttcagg tgggcttgtc cttgtcctag 20820

catcccagat ccaccttctg ggaagtcatc agattggagg tgatgttggc agcttttgta 20880

aacaaagggt agtgttgtaa gctgttgtgt ctgcctatgt gtgtgtttgt gtacttggtc 20940

tcatctctgc agactggtga catggcttcc agatatgccc gacgatgtcc tgtggttgca 21000

gtgggtgacc tcacaggtgt tcactcgagt gctgatgtgt ctgctccccg cctccaggta 21060

aatactttgg ctgtgggtgt gtgggccgga cgggcacctc tctcatctga tgaggcctca 21120

cacgacattc tagaaacagc tggctgaaca ccaagcaagg agcttgccct tgggtgtggg 21180

gaccctgtct catgggaggc agctgagtca gtcagaggtc ctggcacacc tgctgagagc 21240

tgccacccag gccaacctga accggagcct gggaagactt cccgtcggat gagtctcttt 21300

gagtgcagca ttgatggtgg aagagcagag aggccccaga taagcaggga aaggtgcttc 21360

agacagagtg gctgggatga ggactgggga gtgtcagata gcgctggcgt gtctgagcga 21420

aggagctctg gcacccatgg cacaggaagg aggtgggacc ctggaggggc agggctagca 21480

gagctcctcg gagcgtgtgg ctaggtgcct ggtaatgcaa gccccctgtc ctccaccctc 21540

tgttgtactg agtcacagtc tccggggtga agcctagcag tctgcgttga caggccccag 21600

gggatgccgc tacttcctga attctgaatt ctggaaactg agccggagtt cagggcctgg 21660

ctcccattac cagggttggg cgttatcctg aaaatcatag gccttggttt cctcacttgg 21720

ctaacagggg tgatccccat cccctcaatg ggtttccgtg agctcctgag agcccgtagc 21780

atggtacttg gcacatgctg ggcatcagga ggtatggcct ctcttgctat tgttgttatt 21840

ggtagacaca gaaggattta aaagtagggg aatgcaaaga tccgatttgc tagggaagag 21900

ggcagtagtg gccaagtaga gggtggatcc tgggccctgg ctggcagcag gcagcaaggg 21960

gggctgccag ggcccaggca gggacgatct gtagaccgag aggcttccta aggctcttgg 22020

acaggaggag gtgtcggttc caagcctaag gagtggggca gccctggtga ctggtggtca 22080

gtggtgccag gcggtgggtg gtaggacacc ctggcaggca agtaggtttg tgtgggggaa 22140

actgataggc ccctccaggg attcgttggt ggacaacacc tgtgatgtcc agtgggaggt 22200

gtccaggtag ctgggagggc cacaggcttg gaagacctag gtggtgacat cagcccagca 22260

ctgagggcta gaagaagctg tgtctctggc tgtgacggca ccctagagtg tgtgtggtgc 22320

cctctactgg ccggcaatgt gggtccaccg tagctcagac tgcacactgc agcagcggga 22380

acggcctcta agccaacttc ctccatgtgt ttcaggtccc aaatgccagt gagcagccaa 22440

caggcctccc catgcacacc tgagcaggac tggccctgct ggactccctg ctcccccaag 22500

ggctgtccag cagagaccaa agcagaggcc accccgcggt ccatcctcag gtccagcctg 22560

aacttcttct tgggcaataa agtacctgct ggtgctgagg ggctctccac ctttcccagt 22620
ttttcactag agaagagtct gtgagtcact tgaggaggcg agtctagcag attctttcag 22680
aggtgctaaa gtttcccatc tttgtgcagc tacctccgca ttgctgtgta gtgacccctg 22740
cctgtgacgt ggaggatccc agcctctgag ctgagttggt tttatgaaaa gctaggaagc 22800
aacctttcgc ctgtgcagcg gtccagcact taactctaat acatcagcat gcgttaattc 22860
agctggttgg gaaatgacac caggaagccc agtgcagagg gtcccttact gactgtttcg 22920
tggccctatt aatggtcaga ctgttccagc atgaggttct tagaatgaca ggtgtttgga 22980

tgggtggggg ccttgtgatg gggggtaggc tggcccatgt gtgatcttgt ggggtggagg 23040
gaagagaata gcatgatccc acttccccat gctgtgggaa ggggtgcagt tcgtccccaa 23100
gaacgacact gcctgtcagg tggtctgcaa agatgataac cttgactact aaaaacgtct 23160
ccatggcggg ggtaacaaga tgataatcta cttaatttta gaacaccttt ttcacctaac 23220
taaaataatg tttaaagagt tttgtataaa aatgtaagga agcgttgtta cctgttgaat 23280
tttgtattat gtgaatcagt gagatgttag tagaataagc cttaaaaaaa aaaaaatcgg 23340
ttgggtgcag tggcacacgg ctgtaatccc agcactttgg gaggccaagg ttggcagatc 23400

acctgaggtc aggagttcaa gaccagtctg gccaacatag caaaaccctg tctctactaa 23460
aaatacaaaa attatctggg catggtggtg catgcctgta atcccagcta ttcggaaggc 23520
tgaggcagga gaatcacttg aacccaggag gcggaggttg cggtgagctg agattgcacc 23580
atttcattcc agcctgggca acatgagtga aagtctgact caaaaaaaaa aaatttaaaa 23640
aacaaaataa tctagtgtgc agggcattca cctcagcccc ccaggcagga gccaagcaca 23700
gcaggagctt ccgcctcctc tccactggag cacacaactt gaacctggct tattttctgc 23760
agggaccagc cccacatggt cagtgagttt ctccccatgt gtggcgatga gagagtgtag 23820

aaataaagac 23830

<210> 32

<211> 1443

<212> DNA

<213> Homo sapien

<400> 32

auguacgacg cagagcgcgg cuggagcuug uccuucgcgg gcugcggcuu ccugggcuuc 60
uaccacgucg gggcgacccg cugccugagc gagcacgccc cgcaccuccu ccgcgacgcg 120
cgcauguugu ucggcgcuuc ggccggggcg uugcacugcg ucggcguccu cuccgguauc 180

ccgcuggagc agacucugca gguccucuca gaucuugugc ggaaggccag gagucggaac 240

auuggcaucu uccauccauc cuucaacuua agcaaguucc uccgacaggg ucucugcaaa 300

ugccucccgg ccaaugucca ccagcucauc uccggcaaaa uaggcaucuc ucuuaccaga 360

gugucugaug gggaaaacgu ucuggugucu gacuuucggu ccaaagacga agucguggau 420

gccuugguau guuccugcuu caucccccuuc uacaguggcc uuaucccucc uuccuucaga 480

ggcgugcgau auguggaugg aggagugagu gacaacguac ccuucauuga ugccaaaaca 540

accaucaccg ugucccccuu cuaugggggag uacgacaucu gcccuaaagu caaguccacg 600

aacuuucuuc auguggacau caccaagcuc agucuacgcc ucugcacagg gaaccucuac 660

cuucucucga gagcuuuugu ccccccggau cucaaggugc ugggagagau augccuucga 720

ggauauuugg augcauucag guucuuggaa gagaagggca ucugcaacag gccccagcca 780

ggccugaagu cauccucaga agggauggau ccugaggucg ccaugcccag cugggcaaac 840

augagucugg auucuucccc ggagucggcu gccuuggcug ugaggcugga gggagaugag 900

cugcuagacc accugcgucu cagcauccug cccugggaug agagcauccu ggacacccuc 960

ucgcccaggc ucgcuacagc acugagugaa gaaaugaaag acaaaggugg auacaugagc 1020

aagauuugca acuugcuacc cauuaggaua augucuuaug uaaugcugcc cuguacccug 1080

ccuguggaau cugccauugc gauuguccag agacugguga cauggcuucc agauaugccc 1140

gacgaugucc ugugguugca gugggugacc ucacaggugu ucacucgagu gcugaugugu 1200

cugcuccccg ccuccagguc ccaaaugcca gugagcagcc aacaggccuc cccaugcaca 1260

ccugagcagg acuggcccug cuggacuccc ugcucccccа agggcugucc agcagagacc 1320

aaagcagagg ccaccccgcg guccauccuc agguccagcc ugaacuucuu cuugggcaau 1380

aaaguaccug cuggugcuga ggggcucucc accuuuccca guuuuucacu agagaagagu 1440

cug 1443

<210> 33

<211> 1431

<212> DNA

<213> Homo sapien

<400> 33

auguacgacg cagagcgcgg cuggagcuug uccuucgcgg gcugcggcuu ccugggcuuc 60

uaccacgucg gggcgacccg cugccugagc gagcacgccc cgcaccuccu ccgcgacgcg 120

cgcauguugu ucggcgcuuc ggccggggcg uugcacugcg ucggcguccu cuccgagcag 180

acucugcagg uccucucaga ucuugugcgg aaggccagga gucggaacau uggcaucuuc 240

cauccauccu ucaacuuaag caaguuccuc cgacaggguc ucugcaaaug ccucccggcc 300

aauguccacc agcucaucuc cggcaaaaua ggcaucucuc uuaccagagu gucugauggg 360

gaaaacguuc uggugucuga cuuucggucc aaagacgaag ucguggaugc cuugguaugu 420

uccugcuuca uccccuucua caguggccuu aucccuccuu ccuucagagg cgugcgauau 480

guggauggag gagugaguga caacguaccc uucauugaug ccaaaacaac caucaccgug 540

ucccccuucu auggggagua cgacaucugc ccuaaaguca aguccacgaa cuuucuucau 600

guggacauca ccaagcucag ucuacgccuc ugcacaggga accucuaccu ucucucgaga 660

gcuuuugucc ccccggaucu caaggugcug ggagagauau gccuucgagg auauuuggau 720

gcauucaggu ucuuggaaga gaagggcauc ugcaacaggc cccagccagg ccugaaguca 780

uccucagaag ggauggaucc ugaggucgcc augcccagcu gggcaaacau gagucuggau 840

ucuuccccgg agucggcugc cuuggcugug aggcuggagg gagaugagcu gcuagaccac 900

cugcgucuca gcauccugcc cugggaugag agcauccugg acacccucuc gcccaggcuc 960

gcuacagcac ugagugaaga aaugaaagac aaagguggau acaugagcaa gauuugcaac 1020

uugcuaccca uuaggauaau gucuuaugua augcugcccu guacccugcc uguggaaucu 1080

gccauugcga uuguccagag acuggugaca uggcuuccag auaugcccga cgauguccug 1140

ugguugcagu gggugaccuc acagguguuc acucgagugc ugaugugucu gcuccccgcc 1200

uccaggaccc aaaugccagu gagcagccaa caggccuccc caugcacacc ugagcaggac 1260

uggcccugcu ggacucccug cucccccaag ggcuguccag cagagaccaa agcagaggcc 1320

accccgcggu ccauccucag guccagccug aacuucuucu ugggcaauaa aguaccugcu 1380

ggugcugagg ggcucuccac cuuucccagu uuuucacuag agaagagucu g 1431

<210> 34

<211> 1443

<212> DNA

<213> Homo sapien

<400> 34

auguacgacg cagagcgcgg cuggagcuug uccuucgcgg gcugcggcuu ccugggcuuc 60

uaccacgucg gggcgacccg cugccugagc gagcacgccc cgcaccuccu ccgcgacgcg 120

cgcauguugu ucggcgcuuc ggccggggcg uugcacugcg ucggcguccu cuccgguauc 180

ccgcuggagc agacucugca gguccucuca gaucuugugc ggaaggccag gagucggaac 240

auuggcaucu uccauccauc cuucaacuua agcaaguucc uccgacaggg ucucugcaaa 300

ugccucccgg ccaaugucca ccagcucauc uccggcaaaa uaggcaucuc ucuuaccaga 360

gugucugaug gggaaaacgu ucuggugucu gacuuucggu ccaaagacga agucguggau 420

gccuugguau guuccugcuu caugcccuuc uacaguggcc uuaucccucc uuccuucaga 480

ggcgugcgau auguggaugg aggagugagu gacaacguac ccuucauuga ugccaaaaca 540

accaucaccg ugucccccuu cuaugggggag uacgacaucu gcccuaaagu caaguccacg 600

aacuuucuuc auguggacau caccaagcuc agucuacgcc ucugcacagg gaaccucuac 660

cuucucucga gagcuuuugu ccccccggau cucaaggugc ugggagagau augccuucga 720

ggauauuugg augcauucag guucuuggaa gagaagggca ucugcaacag gccccagcca 780

ggccugaagu cauccucaga agggauggau ccugaggucg ccaugcccag cugggcaaac 840

augagucugg auucuucccc ggagucggcu gccuuggcug ugaggcugga gggagaugag 900

cugcuagacc accugcgucu cagcauccug cccugggaug agagcauccu ggacacccuc 960

ucgcccaggc ucgcuacagc acugagugaa gaaaugaaag acaaaggugg auacaugagc 1020

aagauuugca acuugcuacc cauuaggaua augucuuaug uaaugcugcc cuguacccug 1080

ccuguggaau cugccauugc gauuguccag agacugguga cauggcuucc agauaugccc 1140

gacgaugucc ugugguugca gugggugacc ucacaggugu ucacucgagu gcugaugugu 1200

cugcuccccg ccuccagguc ccaaaugcca gugagcagcc aacaggccuc cccaugcaca 1260

ccugagcagg acuggcccug cuggacuccc ugcuccccca agggcugucc agcagagacc 1320

aaagcagagg ccaccccgcg guccauccuc agguccagcc ugaacuucuu cuugggcaau 1380

aaaguaccug cuggugcuga ggggcucucc accuuuccca guuuuucacu agagaagagu 1440

cug 1443

<210> 35

<211> 1431

<212> DNA

<213> Homo sapien

<400> 35

auguacgacg cagagcgcgg cuggagcuug uccuucgcgg gcugcggcuu ccugggcuuc 60

uaccacgucg gggcgacccg cugccugagc gagcacgccc cgcaccuccu ccgcgacgcg 120

cgcauguugu ucggcgcuuc ggccggggcg uugcacugcg ucggcguccu cuccgagcag 180

acucugcagg uccucucaga ucuugugcgg aaggccagga gucggaacau uggcaucuuc 240

cauccauccu ucaacuuaag caaguuccuc cgacaggguc ucugcaaaug ccucccggcc 300

aauguccacc agcucaucuc cggcaaaaua ggcaucucuc uuaccagagu gucugauggg 360

gaaaacguuc uggugucuga cuuucggucc aaagacgaag ucguggaugc cuugguaugu 420

uccugcuuca ugcccuucua caguggccuu aucccuccuu ccuucagagg cgugcgauau 480

guggauggag gagugaguga caacguaccc uucauugaug ccaaaacaac caucaccgug 540

ucccccuucu auggggagua cgacaucugc ccuaaaguca aguccacgaa cuuucuucau 600

guggacauca ccaagcucag ucuacgccuc ugcacaggga accucuaccu ucucucgaga 660

gcuuuugucc ccccggaucu caaggugcug ggagagauau gccuucgagg auauuuggau 720

gcauucaggu ucuuggaaga gaagggcauc ugcaacaggc cccagccagg ccugaaguca 780

uccucagaag ggauggaucc ugaggucgcc augcccagcu gggcaaacau gagucuggau 840

ucuuccccgg agucggcugc cuuggcugug aggcuggagg gagaugagcu gcuagaccac 900

cugcgucuca gcauccugcc cugggaugag agcauccugg acacccucuc gcccaggcuc 960

gcuacagcac ugagugaaga aaugaaagac aaagguggau acaugagcaa gauuugcaac 1020

uugcuaccca uuaggauaau gucuuaugua augcugcccu guacccugcc uguggaaucu 1080

gccauugcga uuguccagag acuggugaca uggcuuccag auaugcccga cgauguccug 1140

ugguugcagu gggugaccuc acagguguuc acucgagugc ugaugugucu gcuccccgcc 1200

uccaggaccc aaaugccagu gagcagccaa caggccuccc caugcacacc ugagcaggac 1260

uggcccugcu ggacucccug cucccccaag ggcuguccag cagagaccaa agcagaggcc 1320

accccgcggu ccauccucag guccagccug aacuucuucu ugggcaauaa aguaccugcu 1380

ggugcugagg ggcucuccac cuuucccagu uuuucacuag agaagagucu g 1431

<210> 36

<211> 1443

<212> DNA

<213> Homo sapien

<400> 36

atgtacgacg cagagcgcgg ctggagcttg tccttcgcgg gctgcggctt cctgggcttc 60

taccacgtcg gggcgacccg ctgcctgagc gagcacgccc cgcacctcct ccgcgacgcg 120

cgcatgttgt tcggcgcttc ggccggggcg ttgcactgcg tcggcgtcct ctccggtatc 180

ccgctggagc agactctgca ggtcctctca gatcttgtgc ggaaggccag gagtcggaac 240

attggcatct tccatccatc cttcaactta agcaagttcc tccgacaggg tctctgcaaa 300

tgcctcccgg ccaatgtcca ccagctcatc tccggcaaaa taggcatctc tcttaccaga 360

gtgtctgatg gggaaaacgt tctggtgtct gactttcggt ccaaagacga agtcgtggat 420

gccttggtat gttcctgctt catccccttc tacagtggcc ttatccctcc ttccttcaga 480

ggcgtgcgat atgtggatgg aggagtgagt gacaacgtac ccttcattga tgccaaaaca 540

accatcaccg tgtccccctt ctatggggag tacgacatct gccctaaagt caagtccacg 600

aactttcttc atgtggacat caccaagctc agtctacgcc tctgcacagg gaacctctac 660

cttctctcga gagcttttgt ccccccggat ctcaaggtgc tgggagagat atgccttcga 720

ggatatttgg atgcattcag gttcttggaa gagaagggca tctgcaacag gccccagcca 780

ggcctgaagt catcctcaga agggatggat cctgaggtcg ccatgcccag ctgggcaaac 840

atgagtctgg attcttcccc ggagtcggct gccttggctg tgaggctgga gggagatgag 900

ctgctagacc acctgcgtct cagcatcctg ccctgggatg agagcatcct ggacaccctc 960

tcgcccaggc tcgctacagc actgagtgaa gaaatgaaag acaaaggtgg atacatgagc 1020

aagatttgca acttgctacc cattaggata atgtcttatg taatgctgcc ctgtaccctg 1080

cctgtggaat ctgccattgc gattgtccag agactggtga catggcttcc agatatgccc 1140

gacgatgtcc tgtggttgca gtgggtgacc tcacaggtgt tcactcgagt gctgatgtgt 1200

ctgctccccg cctccaggtc ccaaatgcca gtgagcagcc aacaggcctc cccatgcaca 1260

cctgagcagg actggccctg ctggactccc tgctccccca agggctgtcc agcagagacc 1320

aaagcagagg ccaccccgcg gtccatcctc aggtccagcc tgaacttctt cttgggcaat 1380

aaagtacctg ctggtgctga ggggctctcc acctttccca gtttttcact agagaagagt 1440

ctg 1443

<210> 37

<211> 1431

<212> DNA

<213> Homo sapien

<400> 37

atgtacgacg cagagcgcgg ctggagcttg tccttcgcgg gctgcggctt cctgggcttc 60

taccacgtcg gggcgacccg ctgcctgagc gagcacgccc cgcacctcct ccgcgacgcg 120

cgcatgttgt tcggcgcttc ggccggggcg ttgcactgcg tcggcgtcct ctccgagcag 180

actctgcagg tcctctcaga tcttgtgcgg aaggccagga gtcggaacat tggcatcttc 240

catccatcct tcaacttaag caagttcctc cgacagggtc tctgcaaatg cctcccggcc 300

aatgtccacc agctcatctc cggcaaaata ggcatctctc ttaccagagt gtctgatggg 360

gaaaacgttc tggtgtctga ctttcggtcc aaagacgaag tcgtggatgc cttggtatgt 420

tcctgcttca tccccttcta cagtggcctt atccctcctt ccttcagagg cgtgcgatat 480

gtggatggag gagtgagtga caacgtaccc ttcattgatg ccaaaacaac catcaccgtg 540

tccccttct atggggagta cgacatctgc cctaaagtca agtccacgaa ctttcttcat 600

gtggacatca ccaagctcag tctacgcctc tgcacaggga acctctacct tctctcgaga 660

gcttttgtcc ccccggatct caaggtgctg ggagagatat gccttcgagg atatttggat 720

gcattcaggt tcttggaaga gaagggcatc tgcaacaggc cccagccagg cctgaagtca 780

tcctcagaag ggatggatcc tgaggtcgcc atgcccagct gggcaaacat gagtctggat 840

tcttccccgg agtcggctgc cttggctgtg aggctggagg gagatgagct gctagaccac 900

ctgcgtctca gcatcctgcc ctgggatgag agcatcctgg acaccctctc gcccaggctc 960

gctacagcac tgagtgaaga aatgaaagac aaaggtggat acatgagcaa gatttgcaac 1020

ttgctaccca ttaggataat gtcttatgta atgctgccct gtaccctgcc tgtggaatct 1080

gccattgcga ttgtccagag actggtgaca tggcttccag atatgcccga cgatgtcctg 1140

tggttgcagt gggtgacctc acaggtgttc actcgagtgc tgatgtgtct gctccccgcc 1200

tccaggtccc aaatgccagt gagcagccaa caggcctccc catgcacacc tgagcaggac 1260

tggccctgct ggactccctg ctcccccaag ggctgtccag cagagaccaa agcagaggcc 1320

accccgcggt ccatcctcag gtccagcctg aacttcttct tgggcaataa agtacctgct 1380

ggtgctgagg ggctctccac ctttcccagt ttttcactag agaagagtct g 1431

<210> 38

<211> 1443

<212> DNA

<213> Homo sapien

<400> 38

atgtacgacg cagagcgcgg ctggagcttg tccttcgcgg gctgcggctt cctgggcttc 60

taccacgtcg gggcgacccg ctgcctgagc gagcacgccc cgcacctcct ccgcgacgcg 120

cgcatgttgt tcggcgcttc ggccggggcg ttgcactgcg tcggcgtcct ctccggtatc 180

ccgctggagc agactctgca ggtcctctca gatcttgtgc ggaaggccag gagtcggaac 240

attggcatct tccatccatc cttcaactta agcaagttcc tccgacaggg tctctgcaaa 300

tgcctcccgg ccaatgtcca ccagctcatc tccggcaaaa taggcatctc tcttaccaga 360

gtgtctgatg gggaaaacgt tctggtgtct gactttcggt ccaaagacga agtcgtggat 420

gccttggtat gttcctgctt catgcccttc tacagtggcc ttatccctcc ttccttcaga 480

ggcgtgcgat atgtggatgg aggagtgagt gacaacgtac ccttcattga tgccaaaaca 540

accatcaccg tgtccccctt ctatggggag tacgacatct gccctaaagt caagtccacg 600

aactttcttc atgtggacat caccaagctc agtctacgcc tctgcacagg gaacctctac 660

cttctctcga gagcttttgt ccccccggat ctcaaggtgc tgggagagat atgccttcga 720

ggatatttgg atgcattcag gttcttggaa gagaagggca tctgcaacag gccccagcca 780

ggcctgaagt catcctcaga agggatggat cctgaggtcg ccatgcccag ctgggcaaac 840

atgagtctgg attcttcccc ggagtcggct gccttggctg tgaggctgga gggagatgag 900

ctgctagacc acctgcgtct cagcatcctg ccctgggatg agagcatcct ggacaccctc 960

tcgcccaggc tcgctacagc actgagtgaa gaaatgaaag acaaaggtgg atacatgagc 1020

aagatttgca acttgctacc cattaggata atgtcttatg taatgctgcc ctgtaccctg 1080

cctgtggaat ctgccattgc gattgtccag agactggtga catggcttcc agatatgccc 1140

gacgatgtcc tgtggttgca gtgggtgacc tcacaggtgt tcactcgagt gctgatgtgt 1200

ctgctccccg cctccaggtc ccaaatgcca gtgagcagcc aacaggcctc cccatgcaca 1260

cctgagcagg actggccctg ctggactccc tgctccccca agggctgtcc agcagagacc 1320

aaagcagagg ccaccccgcg gtccatcctc aggtccagcc tgaacttctt cttgggcaat 1380

aaagtacctg ctggtgctga ggggctctcc acctttccca gtttttcact agagaagagt 1440

ctg 1443

<210> 39

<211> 1431

<212> DNA

<213> Homo sapien

<400> 39

atgtacgacg cagagcgcgg ctggagcttg tccttcgcgg gctgcggctt cctgggcttc 60

taccacgtcg gggcgacccg ctgcctgagc gagcacgccc cgcacctcct ccgcgacgcg 120

cgcatgttgt tcggcgcttc ggccggggcg ttgcactgcg tcggcgtcct ctccgagcag 180

actctgcagg tcctctcaga tcttgtgcgg aaggccagga gtcggaacat tggcatcttc 240

catccatcct tcaacttaag caagttcctc cgacagggtc tctgcaaatg cctcccggcc 300

aatgtccacc agctcatctc cggcaaaata ggcatctctc ttaccagagt gtctgatggg 360

gaaaacgttc tggtgtctga ctttcggtcc aaagacgaag tcgtggatgc cttggtatgt 420

tcctgcttca tgcccttcta cagtggcctt atccctcctt ccttcagagg cgtgcgatat 480

gtggatggag gagtgagtga caacgtaccc ttcattgatg ccaaaacaac catcaccgtg 540

tccccttct atggggagta cgacatctgc cctaaagtca agtccacgaa ctttcttcat 600

gtggacatca ccaagctcag tctacgcctc tgcacaggga acctctacct tctctcgaga 660

gcttttgtcc ccccggatct caaggtgctg ggagagatat gccttcgagg atatttggat 720

gcattcaggt tcttggaaga gaagggcatc tgcaacaggc cccagccagg cctgaagtca 780

tcctcagaag ggatggatcc tgaggtcgcc atgcccagct gggcaaacat gagtctggat 840

tcttccccgg agtcggctgc cttggctgtg aggctggagg gagatgagct gctagaccac 900

ctgcgtctca gcatcctgcc ctgggatgag agcatcctgg acaccctctc gcccaggctc 960

gctacagcac tgagtgaaga aatgaaagac aaaggtggat acatgagcaa gatttgcaac 1020

ttgctaccca ttaggataat gtcttatgta atgctgccct gtaccctgcc tgtggaatct 1080

gccattgcga ttgtccagag actggtgaca tggcttccag atatgcccga cgatgtcctg 1140

tggttgcagt gggtgacctc acaggtgttc actcgagtgc tgatgtgtct gctccccgcc 1200

tccaggtccc aaatgccagt gagcagccaa caggcctccc catgcacacc tgagcaggac 1260

tggccctgct ggactccctg ctcccccaag ggctgtccag cagagaccaa agcagaggcc 1320

accccgcggt ccatcctcag gtccagcctg aacttcttct tgggcaataa agtacctgct 1380

ggtgctgagg ggctctccac ctttcccagt ttttcactag agaagagtct g 1431

<210> 40

<211> 481

<212> PRT

<213> Homo sapien

<400> 40

Met Tyr Asp Ala Glu Arg Gly Trp Ser Leu Ser Phe Ala Gly Cys Gly
1 5 10 15

Phe Leu Gly Phe Tyr His Val Gly Ala Thr Arg Cys Leu Ser Glu His
20 25 30

Ala Pro His Leu Leu Arg Asp Ala Arg Met Leu Phe Gly Ala Ser Ala
35 40 45

Gly Ala Leu His Cys Val Gly Val Leu Ser Gly Ile Pro Leu Glu Gln
50 55 60

Thr Leu Gln Val Leu Ser Asp Leu Val Arg Lys Ala Arg Ser Arg Asn
65 70 75 80

Ile Gly Ile Phe His Pro Ser Phe Asn Leu Ser Lys Phe Leu Arg Gln
85 90 95

Gly Leu Cys Lys Cys Leu Pro Ala Asn Val His Gln Leu Ile Ser Gly
100 105 110

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

115 120 125

Val Ser Asp Phe Arg Ser Lys Asp Glu Val Val Asp Ala Leu Val Cys

130 135 140

Ser Cys Phe Ile Pro Phe Tyr Ser Gly Leu Ile Pro Pro Ser Phe Arg

145 150 155 160

Gly Val Arg Tyr Val Asp Gly Gly Val Ser Asp Asn Val Pro Phe Ile

165 170 175

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

180 185 190

Ile Cys Pro Lys Val Lys Ser Thr Asn Phe Leu His Val Asp Ile Thr

195 200 205

Lys Leu Ser Leu Arg Leu Cys Thr Gly Asn Leu Tyr Leu Leu Ser Arg

210 215 220

Ala Phe Val Pro Pro Asp Leu Lys Val Leu Gly Glu Ile Cys Leu Arg

225 230 235 240

Gly Tyr Leu Asp Ala Phe Arg Phe Leu Glu Glu Lys Gly Ile Cys Asn

245 250 255

Arg Pro Gln Pro Gly Leu Lys Ser Ser Ser Glu Gly Met Asp Pro Glu

260 265 270

Val Ala Met Pro Ser Trp Ala Asn Met Ser Leu Asp Ser Ser Pro Glu

275 280 285

Ser Ala Ala Leu Ala Val Arg Leu Glu Gly Asp Glu Leu Leu Asp His

290 295 300

Leu Arg Leu Ser Ile Leu Pro Trp Asp Glu Ser Ile Leu Asp Thr Leu

305 310 315 320

Ser Pro Arg Leu Ala Thr Ala Leu Ser Glu Glu Met Lys Asp Lys Gly

325 330 335

Gly Tyr Met Ser Lys Ile Cys Asn Leu Leu Pro Ile Arg Ile Met Ser

340 345 350

Tyr Val Met Leu Pro Cys Thr Leu Pro Val Glu Ser Ala Ile Ala Ile

355 360 365

Val Gln Arg Leu Val Thr Trp Leu Pro Asp Met Pro Asp Asp Val Leu
370 375 380
Trp Leu Gln Trp Val Thr Ser Gln Val Phe Thr Arg Val Leu Met Cys
385 390 395 400
Leu Leu Pro Ala Ser Arg Ser Gln Met Pro Val Ser Ser Gln Gln Ala
405 410 415
Ser Pro Cys Thr Pro Glu Gln Asp Trp Pro Cys Trp Thr Pro Cys Ser

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

<210> 41
<211> 477
<212> PRT

<213> Homo sapien
<400> 41
Met Tyr Asp Ala Glu Arg Gly Trp Ser Leu Ser Phe Ala Gly Cys Gly
1 5 10 15
Phe Leu Gly Phe Tyr His Val Gly Ala Thr Arg Cys Leu Ser Glu His
20 25 30
Ala Pro His Leu Leu Arg Asp Ala Arg Met Leu Phe Gly Ala Ser Ala
35 40 45
Gly Ala Leu His Cys Val Gly Val Leu Ser Glu Gln Thr Leu Gln Val
50 55 60

Leu Ser Asp Leu Val Arg Lys Ala Arg Ser Arg Asn Ile Gly Ile Phe
65 70 75 80
His Pro Ser Phe Asn Leu Ser Lys Phe Leu Arg Gln Gly Leu Cys Lys

85 90 95

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

100 105 110

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

115 120 125

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

130 135 140

Pro Phe Tyr Ser Gly Leu Ile Pro Pro Ser Phe Arg Gly Val Arg Tyr

145 150 155 160

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

165 170 175

Thr Ile Thr Val Ser Pro Phe Tyr Gly Glu Tyr Asp Ile Cys Pro Lys

180 185 190

Val Lys Ser Thr Asn Phe Leu His Val Asp Ile Thr Lys Leu Ser Leu

195 200 205

Arg Leu Cys Thr Gly Asn Leu Tyr Leu Leu Ser Arg Ala Phe Val Pro

210 215 220

Pro Asp Leu Lys Val Leu Gly Glu Ile Cys Leu Arg Gly Tyr Leu Asp

225 230 235 240

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

245 250 255

Gly Leu Lys Ser Ser Ser Glu Gly Met Asp Pro Glu Val Ala Met Pro

260 265 270

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

275 280 285

Ala Val Arg Leu Glu Gly Asp Glu Leu Leu Asp His Leu Arg Leu Ser

290 295 300

Ile Leu Pro Trp Asp Glu Ser Ile Leu Asp Thr Leu Ser Pro Arg Leu

305 310 315 320

Ala Thr Ala Leu Ser Glu Glu Met Lys Asp Lys Gly Gly Tyr Met Ser

325 330 335

Lys Ile Cys Asn Leu Leu Pro Ile Arg Ile Met Ser Tyr Val Met Leu
340 345 350

Pro Cys Thr Leu Pro Val Glu Ser Ala Ile Ala Ile Val Gln Arg Leu
355 360 365

Val Thr Trp Leu Pro Asp Met Pro Asp Asp Val Leu Trp Leu Gln Trp
370 375 380

Val Thr Ser Gln Val Phe Thr Arg Val Leu Met Cys Leu Leu Pro Ala
385 390 395 400

Ser Arg Ser Gln Met Pro Val Ser Ser Gln Gln Ala Ser Pro Cys Thr
405 410 415

Pro Glu Gln Asp Trp Pro Cys Trp Thr Pro Cys Ser Pro Lys Gly Cys
420 425 430

Pro Ala Glu Thr Lys Ala Glu Ala Thr Pro Arg Ser Ile Leu Arg Ser
435 440 445

Ser Leu Asn Phe Phe Leu Gly Asn Lys Val Pro Ala Gly Ala Glu Gly
450 455 460

Leu Ser Thr Phe Pro Ser Phe Ser Leu Glu Lys Ser Leu
465 470 475

<210> 42

<211> 481

<212> PRT

<213> Homo sapien

<400> 42

Met Tyr Asp Ala Glu Arg Gly Trp Ser Leu Ser Phe Ala Gly Cys Gly
1 5 10 15

Phe Leu Gly Phe Tyr His Val Gly Ala Thr Arg Cys Leu Ser Glu His
20 25 30

Ala Pro His Leu Leu Arg Asp Ala Arg Met Leu Phe Gly Ala Ser Ala
35 40 45

Gly Ala Leu His Cys Val Gly Val Leu Ser Gly Ile Pro Leu Glu Gln
50 55 60

Thr Leu Gln Val Leu Ser Asp Leu Val Arg Lys Ala Arg Ser Arg Asn

65 70 75 80

Ile Gly Ile Phe His Pro Ser Phe Asn Leu Ser Lys Phe Leu Arg Gln
85 90 95

Gly Leu Cys Lys Cys Leu Pro Ala Asn Val His Gln Leu Ile Ser Gly
100 105 110

Lys Ile Gly Ile Ser Leu Thr Arg Val Ser Asp Gly Glu Asn Val Leu
115 120 125

Val Ser Asp Phe Arg Ser Lys Asp Glu Val Val Asp Ala Leu Val Cys
130 135 140

Ser Cys Phe Met Pro Phe Tyr Ser Gly Leu Ile Pro Pro Ser Phe Arg
145 150 155 160

Gly Val Arg Tyr Val Asp Gly Gly Val Ser Asp Asn Val Pro Phe Ile
165 170 175

Asp Ala Lys Thr Thr Ile Thr Val Ser Pro Phe Tyr Gly Glu Tyr Asp
180 185 190

Ile Cys Pro Lys Val Lys Ser Thr Asn Phe Leu His Val Asp Ile Thr
195 200 205

Lys Leu Ser Leu Arg Leu Cys Thr Gly Asn Leu Tyr Leu Leu Ser Arg
210 215 220

Ala Phe Val Pro Pro Asp Leu Lys Val Leu Gly Glu Ile Cys Leu Arg
225 230 235 240

Gly Tyr Leu Asp Ala Phe Arg Phe Leu Glu Glu Lys Gly Ile Cys Asn
245 250 255

Arg Pro Gln Pro Gly Leu Lys Ser Ser Ser Glu Gly Met Asp Pro Glu
260 265 270

Val Ala Met Pro Ser Trp Ala Asn Met Ser Leu Asp Ser Ser Pro Glu
275 280 285

Ser Ala Ala Leu Ala Val Arg Leu Glu Gly Asp Glu Leu Leu Asp His
290 295 300

Leu Arg Leu Ser Ile Leu Pro Trp Asp Glu Ser Ile Leu Asp Thr Leu
305 310 315 320

Ser Pro Arg Leu Ala Thr Ala Leu Ser Glu Glu Met Lys Asp Lys Gly
325 330 335
Gly Tyr Met Ser Lys Ile Cys Asn Leu Leu Pro Ile Arg Ile Met Ser
340 345 350

Tyr Val Met Leu Pro Cys Thr Leu Pro Val Glu Ser Ala Ile Ala Ile
355 360 365
Val Gln Arg Leu Val Thr Trp Leu Pro Asp Met Pro Asp Asp Val Leu
370 375 380
Trp Leu Gln Trp Val Thr Ser Gln Val Phe Thr Arg Val Leu Met Cys
385 390 395 400
Leu Leu Pro Ala Ser Arg Ser Gln Met Pro Val Ser Ser Gln Gln Ala
405 410 415

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

Leu

<210> 43
<211> 477
<212> PRT
<213> Homo sapien
<400> 43
Met Tyr Asp Ala Glu Arg Gly Trp Ser Leu Ser Phe Ala Gly Cys Gly
1 5 10 15
Phe Leu Gly Phe Tyr His Val Gly Ala Thr Arg Cys Leu Ser Glu His
20 25 30
Ala Pro His Leu Leu Arg Asp Ala Arg Met Leu Phe Gly Ala Ser Ala

35 40 45

Gly Ala Leu His Cys Val Gly Val Leu Ser Glu Gln Thr Leu Gln Val

50 55 60

Leu Ser Asp Leu Val Arg Lys Ala Arg Ser Arg Asn Ile Gly Ile Phe

65 70 75 80

His Pro Ser Phe Asn Leu Ser Lys Phe Leu Arg Gln Gly Leu Cys Lys

85 90 95

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

100 105 110

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

115 120 125

Arg Ser Lys Asp Glu Val Val Asp Ala Leu Val Cys Ser Cys Phe Met

130 135 140

Pro Phe Tyr Ser Gly Leu Ile Pro Pro Ser Phe Arg Gly Val Arg Tyr

145 150 155 160

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

165 170 175

Thr Ile Thr Val Ser Pro Phe Tyr Gly Glu Tyr Asp Ile Cys Pro Lys

180 185 190

Val Lys Ser Thr Asn Phe Leu His Val Asp Ile Thr Lys Leu Ser Leu

195 200 205

Arg Leu Cys Thr Gly Asn Leu Tyr Leu Leu Ser Arg Ala Phe Val Pro

210 215 220

Pro Asp Leu Lys Val Leu Gly Glu Ile Cys Leu Arg Gly Tyr Leu Asp

225 230 235 240

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

245 250 255

Gly Leu Lys Ser Ser Ser Glu Gly Met Asp Pro Glu Val Ala Met Pro

260 265 270

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

275 280 285

Ala Val Arg Leu Glu Gly Asp Glu Leu Leu Asp His Leu Arg Leu Ser

290 295 300

Ile Leu Pro Trp Asp Glu Ser Ile Leu Asp Thr Leu Ser Pro Arg Leu

305 310 315 320

Ala Thr Ala Leu Ser Glu Glu Met Lys Asp Lys Gly Gly Tyr Met Ser

325 330 335

Lys Ile Cys Asn Leu Leu Pro Ile Arg Ile Met Ser Tyr Val Met Leu

340 345 350

Pro Cys Thr Leu Pro Val Glu Ser Ala Ile Ala Ile Val Gln Arg Leu

355 360 365

Val Thr Trp Leu Pro Asp Met Pro Asp Asp Val Leu Trp Leu Gln Trp

370 375 380

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

385 390 395 400

Ser Arg Ser Gln Met Pro Val Ser Ser Gln Gln Ala Ser Pro Cys Thr

405 410 415

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

420 425 430

Pro Ala Glu Thr Lys Ala Glu Ala Thr Pro Arg Ser Ile Leu Arg Ser

435 440 445

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

450 455 460

Leu Ser Thr Phe Pro Ser Phe Ser Leu Glu Lys Ser Leu

465 470 475

<210> 44

<211> 2397

<212> DNA

<213> Homo sapien

<400> 44

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60

aagccaugaa caucauccua gaaauccuuc ugcuucugau caccaucauc uacuccuacu 120

uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180

uucucauuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240

gacagagcau auugguucug ugggauauua auaagcgcgg uguggaggaa acugcagcug 300

agugccgaaa acuaggcguc acugcgcaug cguauguggu agacugcagc aacagagaag 360

agaucuaucg cucucuaaau caggugaaga aagaaguggg ugauguaaca aucgugguga 420

auaaugcugg gacaguauau ccagccgauc uucucagcac caaggaugaa gagauuacca 480

agacauuuga ggucaacauc cuaggacauu uuuggaucac aaaagcacuu cuuccaucga 540

ugauggagag aaaucauggc cacaucguca caguggcuuc agugugcggc cacgaaggga 600

uuccuuaccu caucccauau uguuccagca aauuugccgc uguuggcuuu cacagagguc 660

ugacaucaga acuucaggcc uugggaaaaa cugguaucaa aaccucaugu cucugcccag 720

uuuuugugaa uacuggguuc accaaaaauc caagcacaag auuauggccu guauuggaga 780

cagaugaagu cguaagaagu cugauagaug gaauacuuac caauaagaaa augauuuuug 840

uuccaucgua uaucaauauc uuucugagac uacagaaguu ucuuccugaa cgcgccucag 900

cgauuuuaaa ucguaugcag aauauucaau uugaagcagu gguuggccac aaaaucaaaa 960

ugaaaugaau aaauaagcuc cagccagaga uguaugcaug auaaugauau gaauaguuuc 1020

gaaucaaugc ugcaaagcuu uauuucacau uuuuucaguc cugauaauau uaaaaacauu 1080

gguuuggcac uagcagcagu caaacgaaca agauuaauua ccugucuucc uguuucucaa 1140

gaauauuuac guaguuuuuc auaggucugu uuuuccuuuc augccucuua aaaacuucug 1200

ugcuuacaua aacauacuua aaagguuuuc uuuaagauau uuuauuuuuc cauuuaaagg 1260

uggacaaaag cuaccucccu aaaaguaaau acaaagagaa cuuauuuaca cagggaaggu 1320

uuaagacugu ucaaguagca uuccaaucug uagccaugcc acagaauauc aacaagaaca 1380

cagaaugagu gcacagcuaa gagaucaagu uucagcaggc agcuuuaucu caaccuggac 1440

auauuuuaag auucagcauu ugaaagauuu cccuagccuc uuccuuuuuc auuagcccaa 1500

aacggugcaa cucuauucug gacuuuauua cuugauucug ucuucuguau aacucugaag 1560

uccaccaaaa guggacccuc uauauuuccu cccuuuuuau agucuuauaa gauacauuau 1620

gaaaggugac cgacucuauu uuaaaucuca gaauuuuaag uucuagcccc augauaaccu 1680

uuuucuuugu aauuuaugcu uucauauauc cuuggcccca gagauguuua gacaauuuua 1740

ggcucaaaaa uuaaagcuaa cacaggaaaa ggaacuguac uggcuauuac auaagaaaca 1800

auggacccaa gagaagaaaa ggaagaaaga aagguuuuuu gguuuuuguu uuguuuuguu 1860

uuguuuuuug uuuuuuugag auggagucuc acucuuucgc ccaggcugga gugcaguggu 1920

augaucucag cucacugcaa gcuccaccuc ccggguucac gccauucucc ugccucagcc 1980

uccugaguag cugggacuac aggcgcccgc caccacaccc ggcuaauuuu uuguauuuuu 2040

uguagagacg ggguuucacc auguuagcca agauggucuc gaucuccuga ccucgugauc 2100

caccugccuc ggccucccaa agugcuggga uuacgggugu gagccaccgu gcccagccuu 2160

uuuuuuuuua auagaaaaaa uaauccgacu cccacuacau caagacuaau cuuguuuugu 2220

guguuuuuca cauguauuau agaaugcuuu ugcauggacu auccucuugu uuuuauuaaa 2280

aacaaaugau uuuuuuaaaa gucacaaaaa caauucacua aaaauaaaua ugucauugug 2340

cuuuaaaaaa auaaccucuu guaguuauaa aauaaaacgu uugacuucua aacucug 2397

<210> 45

<211> 2289

<212> RNA

<213> Homo sapien

<400> 45

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60

aagccaugaa caucauccua gaaauccuuc ugcuucugau caccaucauc uacuccuacu 120

uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180

uucucauuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240

gacagagcau auugguucug ugggauauua auaaggugaa gaaagaagug ggugauguaa 300

caaucguggu gaauaaugcu gggacaguau auccagccga ucuucucagc accaaggaug 360

aagagauuac caagacauuu gaggucaaca uccuaggaca uuuuuggauc acaaaagcac 420

uucuuccauc gaugauggag agaaaucaug gccacaucgu cacaguggcu ucagugugcg 480

gccacgaagg gauuccuuac cucaucccau auuguuccag caaauuugcc gcuguuggcu 540

uucacagagg ucugacauca gaacuucagg ccuugggaaa aacugguauc aaaaccucau 600

gucucugccc aguuuuugug aauacugggu ucaccaaaaa uccaagcaca agauuauggc 660

cuguauugga gacagaugaa gucguaagaa gucugauaga uggaauacuu accaauaaga 720

aaaugauuuu uguuccaucg uauaucaaua ucuuucugag acuacagaag uuucuuccug 780

aacgcgccuc agcgauuuua aaucguaugc agaauauuca auuugaagca gugguuggcc 840

acaaaaucaa aaugaaauga auaaauaagc uccagccaga gauguaugca ugauaaugau 900

augaauaguu ucgaaucaau gcugcaaagc uuuauuucac auuuuuucag uccugauaau 960

auuaaaaaca uugguuuggc acuagcagca gucaaacgaa caagauuaau uaccugucuu 1020

ccuguuucuc aagaauauuu acguaguuuu ucauaggucu guuuuuccuu ucaugccucu 1080

uaaaaacuuc ugugcuuaca uaaacauacu uaaaagguuu ucuuuaagau auuuuauuuu 1140

uccauuuaaa gguggacaaa agcuaccucc cuaaaaguaa auacaaagag aacuuauuua 1200

cacagggaag guuuaagacu guucaaguag cauuccaauc uguagccaug ccacagaaua 1260

ucaacaagaa cacagaauga gugcacagcu aagagaucaa guuucagcag gcagcuuuau 1320

cucaaccugg acauauuuua agauucagca uuugaaagau uucccuagcc ucuuccuuuu 1380

ucauuagccc aaaacggugc aacucuauuc uggacuuuau uacuugauuc ugucuucugu 1440

auaacucuga aguccaccaa aaguggaccc ucuauauuuc cucccuuuuu auagucuuau 1500

aagauacauu augaaaggug accgacucua uuuuaaaucu cagaauuuua aguucuagcc 1560

ccaugauaac cuuuuucuuu guaauuuaug cuuucauaua uccuuggucc cagagauguu 1620

uagacaauuu uaggcucaaa aauuaaagcu aacacaggaa aaggaacugu acuggcuauu 1680

acauaagaaa caauggaccc aagagaagaa aaggaagaaa gaaagguuuu uugguuuuug 1740

uuuuguuuug uuuuguuuuu uguuuuuuug agauggaguc ucacucuuuc gcccaggcug 1800

gagugcagug guaugaucuc agcucacugc aagcuccacc ucccggguuc acgccauucu 1860

ccugccucag ccuccugagu agcugggacu acaggcgccc gccaccacac ccggcuaauu 1920

uuuuguauuu uuuguagaga cgggguuuca ccauguuagc caagaugguc ucgaucuccu 1980

gaccucguga uccaccugcc ucggccuccc aaagugcugg gauuacgggu gugagccacc 2040

gugcccagcc uuuuuuuuuu uaauagaaaa aauaauccga cucccacuac aucaagacua 2100

aucuuguuuu guguguuuuu cacauguauu auagaaugcu uuugcaugga cuauccucuu 2160

guuuuuauua aaaacaaaug auuuuuuuaa aagucacaaa aacaauucac uaaaaauaaa 2220

uaugucauug ugcuuuaaaa aaauaaccuc uuguaguuau aaaauaaaac guuugacuuc 2280

uaaacucug 2289

<210> 46

<211> 2280

<212> RNA

<213> Homo sapien

<400> 46

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60

aagccaugaa caucauccua gaaauccuuc ugcuucugau caccaucauc uacuccuacu 120

uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180

uucucauuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240

gacagagcau auugguucug ugggauauua auaagcgcgg uguggaggaa acugcagcug 300

agugccgaaa acuaggcguc acugcgcaug cguauguggu agacugcagc aacagagaag 360

agaucuaucg cucucuaaau caggugaaga aagaaguggg ugauguaaca aucgugguga 420

auaaugcugg gacaguauau ccagccgauc uucucagcac caaggaugaa gagauuacca 480

agacauuuga ggucaacauc cuaggacauu uuuggaucac aaaagcacuu cuuccaucga 540

ugauggagag aaaucauggc cacaucguca caguggcuuc agugugcggc cacgaaggga 600

uuccuuaccu caucccauau uguuccagca aauuugccgc uguuggcuuu cacagagguc 660

ugacaucaga acuucaggcc uugggaaaaa cugguaucaa aaccucaugu cucugcccag 720

uuuuugugaa uacuggguuc accaaaaauc caagcacaag guuucuuccu gaacgcgccu 780

cagcgauuuu aaaucguaug cagaauauuc aauuugaagc agugguuggc cacaaaauca 840

aaaugaaaug aauaaauaag cuccagccag agauguaugc augauaauga uaugaauagu 900

uucgaaucaa ugcugcaaag cuuuauuuca cauuuuuuca guccugauaa uauuaaaaac 960

auugguuugg cacuagcagc agucaaacga acaagauuaa uuaccugucu uccuguuucu 1020

caagaauauu uacguaguuu uucauagguc uguuuuuccu uucaugccuc uuaaaaacuu 1080

cugugcuuac auaaacauac uuaaaagguu uucuuuaaga uauuuuauuu uuccauuuaa 1140

agguggacaa aagcuaccuc ccuaaaagua aauacaaaga gaacuuauuu acacagggaa 1200

gguuuaagac uguucaagua gcauuccaau cuguagccau gccacagaau aucaacaaga 1260

acacagaaug agugcacagc uaagagauca aguuucagca ggcagcuuua ucucaaccug 1320

gacauauuuu aagauucagc auuugaaaga uuucccuagc cucuuccuuu uucauuagcc 1380

caaaacggug caacucuauu cuggacuuua uuacuugauu cugucuucug uauaacucug 1440

aaguccacca aaaguggacc cucuauauuu ccucccuuuu uauagucuua uaagauacau 1500

uaugaaaggu gaccgacucu auuuuaaauc ucagaauuuu aaguucuagc cccaugauaa 1560

ccuuuuucuu uguaauuuau gcuuucauau auccuugguc ccagagaugu uuagacaauu 1620

uuaggcucaa aaauuaaagc uaacacagga aaaggaacug uacuggcuau uacauaagaa 1680

acaauggacc caagagaaga aaaggaagaa agaaagguuu uuugguuuuu guuuuguuuu 1740

guuuuguuuu uuguuuuuuu gagauggagu cucacucuuu cgcccaggcu ggagugcagu 1800

gguaugaucu cagcucacug caagcuccac cucccggguu cacgccauuc uccugccuca 1860

gccuccugag uagcugggac uacaggcgcc cgccaccaca cccggcuaau uuuuuguauu 1920

uuuuguagag acgggguuuc accauguuag ccaagauggu cucgaucucc ugaccucgug 1980

auccaccugc cucggccucc caaagugcug ggauuacggg ugugagccac cgugcccagc 2040

cuuuuuuuuu uuaauagaaa aaauaauccg acucccacua caucaagacu aaucuuguuu 2100

uguguguuuu ucacauguau uauagaaugc uuuugcaugg acuauccucu uguuuuuauu 2160

aaaaacaaau gauuuuuuua aaagucacaa aaacaauuca cuaaaaauaa auaugucauu 2220

gugcuuuaaa aaaauaaccu cuuguaguua uaaaauaaaa cguuugacuu cuaaacucug 2280

<210> 47

<211> 2398

<212> RNA

<213> Homo sapien

<400> 47

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60

aagccaugaa caucauccua gaaauccuuc ugcuucugau caccaucauc uacuccuacu 120

uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180

uucucauuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240

gacagagcau auugguucug ugggauauua auaagcgcgg uguggaggaa acugcagcug 300

agugccgaaa acuaggcguc acugcgcaug cguauguggu agacugcagc aacagagaag 360

agaucuaucg cucucuaaau caggugaaga aagaaguggg ugauguaaca aucgugguga 420

auaaugcugg gacaguauau ccagccgauc uucucagcac caaggaugaa gagauuacca 480

agacauuuga ggucaacauc cuaggacauu uuuggaucac aaaagcacuu cuuccaucga 540

ugauggagag aaaucauggc cacaucguca caguggcuuc agugugcggc cacgaaggga 600

uuccuuaccu caucccauau uguuccagca aauuugccgc uguuggcuuu cacagagguc 660

ugacaucaga acuucaggcc uugggaaaaa cugguaucaa aaccucaugu cucugcccag 720

uuuuugugaa uacuggguuc accaaaaauc caagcacaag auuauggccu guauuggaga 780

cagaugaagu cguaagaagu cugauagaug gaauacuuac caauaagaaa augauuuuug 840

uuccaucgua uaucaauauc uuucugagac uacagaaggu uucuuccuga acgcgccuca 900

gcgauuuuaa aucguaugca gaauauucaa uuugaagcag ugguuggcca caaaaucaaa 960

augaaaugaa uaaauaagcu ccagccagag auguaugcau gauaaugaua ugaauaguuu 1020

cgaaucaaug cugcaaagcu uuauuucaca uuuuuucagu ccugauaaua uuaaaaacau 1080

ugguuuggca cuagcagcag ucaaacgaac aagauuaauu accugucuuc cuguuucuca 1140

agaauauuua cguaguuuuu cauaggucug uuuuuccuuu caugccucuu aaaaacuucu 1200

gugcuuacau aaacauacuu aaaaggguuu cuuuaagaua uuuuauuuuu ccauuuaaag 1260

guggacaaaa gcuaccuccc uaaaaguaaa uacaaagaga acuuauuuac acagggaagg 1320

uuuaagacug uucaaguagc auuccaaucu guagccaugc cacagaauau caacaagaac 1380

acagaaugag ugcacagcua agagaucaag uuucagcagg cagcuuuauc ucaaccugga 1440

cauauuuuaa gauucagcau uugaaagauu ucccuagccu cuuccuuuuu cauuagccca 1500

aaacggugca acucuauucu ggacuuuauu acuugauucu gucuucugua uaacucugaa 1560

guccaccaaa aguggacccu cuauauuucc ucccuuuuua uagucuuaua agauacauua 1620

ugaaagguga ccgacucuau uuuaaaucuc agaauuuuaa guucuagccc caugauaacc 1680

uuuuucuuug uaauuuaugc uuucauauau ccuuggcccc agagauguuu agacaauuuu 1740

aggcucaaaa auuaaagcua acacaggaaa aggaacugua cuggcuauua cauaagaaac 1800

aauggaccca agagaagaaa aggaagaaag aaagguuuuu ugguuuuugu uuuguuuugu 1860

uuuguuuuuu guuuuuuuga gauggagucu cacucuuucg cccaggcugg agugcagugg 1920

uaugaucuca gcucacugca agcuccaccu cccggguuca cgccauucuc cugccucagc 1980

cuccugagua gcugggacua caggcgcccg ccaccacacc cggcuaauuu uuuguauuuu 2040

uuguagagac gggguuucac cauguuagcc aagauggucu cgaucuccug accucgugau 2100

ccaccugccu cggccuccca aagugcuggg auuacgggug ugagccaccg ugcccagccu 2160

uuuuuuuuuu aauagaaaaa auaauccgac ucccacuaca ucaagacuaa ucuuguuuug 2220

uguguuuuuc acauguauua uagaaugcuu uugcauggac uauccucuug uuuuuauuaa 2280

aaacaaauga uuuuuuuaaa agucacaaaa acaauucacu aaaaauaaau augucauugu 2340

gcuuuaaaaa aauaaccucu uguaguuaua aaauaaaacg uuugacuucu aaacucug 2398

<210> 48

<211> 2469

<212> RNA

<213> Homo sapien

<400> 48

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60

aagccaugaa caucauccua gaaauccuuc ugcuucugau caccaucauc uacuccuacu 120

uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180

uucucauuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240

gacagagcau auugguucug ugggauauua auaagcgcgg uguggaggaa acugcagcug 300

agugccgaaa acuaggcguc acugcgcaug cguauguggu agacugcagc aacagagaag 360

agaucuaucg cucucuaaau caggugaaga aagaaguggg ugauguaaca aucgugguga 420

auaaugcugg gacaguauau ccagccgauc uucucagcac caaggaugaa gagauuacca 480

agacauuuga ggucaacauc cuaggacauu uuuggaaugg aaaggacauc agaaguaauu 540

acuuggaugu auauaggauc gaggacacuu uuggacgaga cucugagauc acaaaagcac 600

uucuuccauc gaugauggag agaaaucaug gccacaucgu cacaguggcu ucagugugcg 660

gccacgaagg gauuccuuac cucaucccau auuguuccag caaauuugcc gcuguuggcu 720

uucacagagg ucugacauca gaacuucagg ccuugggaaa aacugguauc aaaaccucau 780

gucucugccc aguuuuugug aauacugggu ucaccaaaaa uccaagcaca agauuauggc 840

cuguauugga gacagaugaa gucguaagaa gucugauaga uggaauacuu accaauaaga 900

aaaugauuuu uguuccaucg uauaucaaua ucuuucugag acuacagaag uuucuuccug 960

aacgcgccuc agcgauuuua aaucguaugc agaauauuca auuugaagca gugguuggcc 1020

acaaaaucaa aaugaaauga auaaauaagc uccagccaga gauguaugca ugauaaugau 1080

augaauaguu ucgaaucaau gcugcaaagc uuuauuucac auuuuuucag uccugauaau 1140

auuaaaaaca uugguuuggc acuagcagca gucaaacgaa caagauuaau uaccugucuu 1200

ccuguuucuc aagaauauuu acguaguuuu ucauaggucu guuuuuccuu ucaugccucu 1260

uaaaaacuuc ugugcuuaca uaaacauacu uaaaagguuu ucuuuaagau auuuuauuuu 1320

uccauuuaaa gguggacaaa agcuaccucc cuaaaaguaa auacaaagag aacuuauuua 1380

cacagggaag guuuaagacu guucaaguag cauuccaauc uguagccaug ccacagaaua 1440

ucaacaagaa cacagaauga gugcacagcu aagagaucaa guuucagcag gcagcuuuau 1500

cucaaccugg acauauuuua agauucagca uuugaaagau uucccuagcc ucuuccuuuu 1560

ucauuagccc aaaacggugc aacucuauuc uggacuuuau uacuugauuc ugucuucugu 1620

auaacucuga aguccaccaa aaguggaccc ucuauauuuc cucccuuuuu auagucuuau 1680

aagauacauu augaaaggug accgacucua uuuuaaaucu cagaauuuua aguucuagcc 1740

ccaugauaac cuuuuucuuu guaauuuaug cuuucauaua uccuuggucc cagagauguu 1800

uagacaauuu uaggcucaaa aauuaaagcu aacacaggaa aaggaacugu acuggcuauu 1860

acauaagaaa caauggaccc aagagaagaa aaggaagaaa gaaagguuuu uugguuuuug 1920

uuuuguuuug uuuuguuuuu uguuuuuuug agauggaguc ucacucuuuc gcccaggcug 1980

gagugcagug guaugaucuc agcucacugc aagcuccacc ucccggguuc acgccauucu 2040

ccugccucag ccuccugagu agcugggacu acaggcgccc gccaccacac ccggcuaauu 2100

uuuuguauuu uuuguagaga cgggguuuca ccauguuagc caagaugguc ucgaucuccu 2160

gaccucguga uccaccugcc ucggccuccc aaagugcugg gauuacgggu gugagccacc 2220

gugcccagcc uuuuuuuuuu uaauagaaaa aauaauccga cucccacuac aucaagacua 2280

aucuuguuuu guguguuuuu cacauguauu auagaaugcu uuugcaugga cuauccucuu 2340

guuuuuauua aaaacaaaug auuuuuuuaa aagucacaaa aacaauucac uaaaaauaaa 2400

uaugucauug ugcuuuaaaa aaauaaccuc uuguaguuau aaaauaaaac guuugacuuc 2460

uaaacucug 2469

<210> 49

<211> 1715

<212> RNA

<213> Homo sapien

<400> 49

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60

aagccaugaa caucauccua gaaauccuuc ugcuucugau caccaucauc uacuccuacu 120

uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180

uucucauuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240

gacagagcau auugguucug ugggauauua auaagcgcgg uguggaggaa acugcagcug 300

agugccgaaa acuaggcguc acugcgcaug cguauguggu agacugcagc aacagagaag 360

agaucuaucg cucucuaaau caggugaaga aagaaguggg ugauguaaca aucgugguga 420

auaaugcugg gacaguauau ccagccgauc uucucagcac caaggaugaa gagauuacca 480

agacauuuga ggucaacauc cuaggacauu uuuggaucac aaaagcacuu cuuccaucga 540

ugauggagag aaaucauggc cacaucguca caguggcuuc agugugcggc cacgaaggga 600

uuccuuaccu caucccauau uguuccagca aauuugccgc uguuggcuuu cacagagguc 660

ugacaucaga acuucaggcc uugggaaaaa cugguaucaa aaccucaugu cucugcccag 720

uuuuugugaa uacuggguuc accaaaaauc caagcacaag auuauggccu guauuggaga 780

cagaugaagu cguaagaagu cugauagaug gaauacuuac caauaagaaa augauuuuug 840

uuccaucgua uaucaauauc uuucugagac uacagaaguu aaguacagca cagaacaccc 900

aaauacuaaa acaccaauag agcuuuuuuu uuugcuuuuu uuuuuuuuag acagagucuc 960

acucugucac ccuggcugga uugcgguggu ugcaguggca ugaucuuggc ucacugcaac 1020

cuccgccucc uggguucaag caauucucau gccucagacc cccaaguaac ugggauuaua 1080

ggugugugcu gccacacuac acccagcuaa uuuuuguauu uuuugauaga gacagguuuc 1140

cccauguugg ccaggcugga cucgaacucc ugaccucaag uuauccuccu gucucggccu 1200

cccaaagugc ugggauuaca gucaugagcc accaugccug gcccaauaga gcuauuauua 1260

uggagcaucu uucaguugug aaaauuggca uggaaacucu ccaucccugg ggagaacagu 1320

uauuuccucu guuauuuucc uacccagucu auaaaaagag agugauucau uuucucuacc 1380

aaaucuacug ucucugccca aacuuugcug aagacuauuc uaacuaaagg aaacacaguu 1440

uaaaaagaau gcaauauagu gaaguaguua auaauaaaga cuccauuuuu aaaagucugc 1500

uggaaguuug guugggauug cacugaaucu auagagcaau uggggaguau ugacauauca 1560

acaauauuga guuuucuaau ccaagaacau aauaucuauu uuuaaaaucu ucuucaaaau 1620

cuuuaaaucu uuaaauugua uuuuguaguu uuugguguuu aagucuugca cauauuuugu 1680

cagauuuauu ccaaaguauu ucacggguuc uuuuu 1715

<210> 50

<211> 2290

<212> RNA

<213> Homo sapien

<400> 50

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60

aagccaugaa caucauccua gaaauccuuc ugcuucugau caccaucauc uacuccuacu 120

uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180

uucucauuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240

gacagagcau auugguucug ugggauauua auaaggugaa gaaagaagug ggugauguaa 300

caaucguggu gaauaaugcu gggacaguau auccagccga ucuucucagc accaaggaug 360

aagagauuac caagacauuu gaggucaaca uccuaggaca uuuuuggauc acaaaagcac 420

uucuuccauc gaugauggag agaaaucaug gccacaucgu cacaguggcu ucagugugcg 480

gccacgaagg gauuccuuac cucaucccau auuguuccag caaauuugcc gcuguuggcu 540

uucacagagg ucugacauca gaacuucagg ccuugggaaa aacugguauc aaaaccucau 600

gucucugccc aguuuuugug aauacugggu ucaccaaaaa uccaagcaca agauuauggc 660

cuguauugga gacagaugaa gucguaagaa gucugauaga uggaauacuu accaauaaga 720

aaaugauuuu uguuccaucg uauaucaaua ucuuucugag acuacagaag guuucuuccu 780

gaacgcgccu cagcgauuuu aaaucguaug cagaauauuc aauuugaagc agugguuggc 840

cacaaaauca aaaugaaaug aauaaauaag cuccagccag agauguaugc augauaauga 900

uaugaauagu uucgaaucaa ugcugcaaag cuuuauuuca cauuuuuuca guccugauaa 960

uauuaaaaac auugguuugg cacuagcagc agucaaacga acaagauuaa uuaccugucu 1020

uccuguuucu caagaauauu uacguaguuu uucauagguc uguuuuuccu uucaugccuc 1080

uuaaaaacuu cugugcuuac auaaacauac uuaaaagguu uucuuuaaga uauuuuauuu 1140

uuccauuuaa agguggacaa aagcuaccuc ccuaaaagua aauacaaaga gaacuuauuu 1200

acacagggaa gguuuaagac uguucaagua gcauuccaau cuguagccau gccacagaau 1260

aucaacaaga acacagaaug agugcacagc uaagagauca aguuucagca ggcagcuuua 1320

ucucaaccug gacauauuuu aagauucagc auuugaaaga uuucccuagc cucuuccuuu 1380

uucauuagcc caaaacggug caacucuauu cuggacuuua uuacuugauu cugucuucug 1440

uauaacucug aaguccacca aaaguggacc cucuauauuu ccucccuuuu uauagucuua 1500

uaagauacau uaugaaaggu gaccgacucu auuuuaaauc ucagaauuuu aaguucuagc 1560

cccaugauaa ccuuuuucuu uguaauuuau gcuuucauau auccuugguc ccagagaugu 1620

uuagacaauu uuaggcucaa aaauuaaagc uaacacagga aaaggaacug uacuggcuau 1680

uacauaagaa acaauggacc caagagaaga aaaggaagaa agaaagguuu uuugguuuuu 1740

guuuuguuuu guuuuguuuu uuguuuuuuu gagauggagu cucacucuuu cgcccaggcu 1800

ggagugcagu gguaugaucu cagcucacug caagcuccac cucccggguu cacgccauuc 1860

uccugccuca gccuccugag uagcugggac uacaggcgcc cgccaccaca cccggcuaau 1920

uuuuuguauu uuuuguagag acgggguuuc accauguuag ccaagauggu cucgaucucc 1980

ugaccucgug auccaccugc cucggccucc caaagugcug ggauuacggg ugugagccac 2040

cgugcccagc cuuuuuuuuu uuaauagaaa aaauaauccg acucccacua caucaagacu 2100

aaucuuguuu uguguguuuu ucacauguau uauagaaugc uuuugcaugg acuauccucu 2160

uguuuuuauu aaaaacaaau gauuuuuuua aaagucacaa aaacaauuca cuaaaaauaa 2220

auaugucauu gugcuuuaaa aaaauaaccu cuuguaguua uaaaauaaaa cguuugacuu 2280

cuaaacucug 2290

<210> 51

<211> 2470

<212> RNA

<213> Homo sapien

<400> 51

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60

aagccaugaa caucauccua gaaauccuuc ugcuucugau caccaucauc uacuccuacu 120

uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180

uucucauuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240

gacagagcau auugguucug ugggauauua auaagcgcgg uguggaggaa acugcagcug 300

agugccgaaa acuaggcguc acugcgcaug cguauguggu agacugcagc aacagagaag 360

agaucuaucg cucucuaaau caggugaaga aagaaguggg ugauguaaca aucgugguga 420

auaaugcugg gacaguauau ccagccgauc uucucagcac caaggaugaa gagauuacca 480

agacauuuga ggucaacauc cuaggacauu uuuggaaugg aaaggacauc agaaguaauu 540

acuuggaugu auauaggauc gaggacacuu uuggacgaga cucugagauc acaaaagcac 600

uucuuccauc gaugauggag agaaaucaug gccacaucgu cacaguggcu ucagugugcg 660

gccacgaagg gauuccuuac cucaucccau auuguuccag caaauuugcc gcuguuggcu 720

uucacagagg ucugacauca gaacuucagg ccuugggaaa aacugguauc aaaaccucau 780

gucucugccc aguuuuugug aauacugggu ucaccaaaaa uccaagcaca agauuauggc 840

cuguauugga gacagaugaa gucguaagaa gucugauaga uggaauacuu accaauaaga 900

aaaugauuuu uguuccaucg uauaucaaua ucuuucugag acuacagaag guuucuuccu 960

gaacgcgccu cagcgauuuu aaaucguaug cagaauauuc aauuugaagc agugguuggc 1020

cacaaaauca aaaugaaaug aauaaauaag cuccagccag agauguaugc augauaauga 1080

uaugaauagu uucgaaucaa ugcugcaaag cuuuauuuca cauuuuuuca guccugauaa 1140

uauuaaaaac auugguuugg cacuagcagc agucaaacga acaagauuaa uuaccugucu 1200

uccuguuucu caagaauauu uacguaguuu uucauagguc uguuuuuccu uucaugccuc 1260

uuaaaaacuu cugugcuuac auaaacauac uuaaaagguu uucuuuaaga uauuuuauuu 1320

uuccauuuaa agguggacaa aagcuaccuc ccuaaaagua aauacaaaga gaacuuauuu 1380

acacagggaa gguuuaagac uguucaagua gcauuccaau cuguagccau gccacagaau 1440

aucaacaaga acacagaaug agugcacagc uaagagauca aguuucagca ggcagcuuua 1500

ucucaaccug gacauauuuu aagauucagc auuugaaaga uuucccuagc cucuuccuuu 1560

uucauuagcc caaaacggug caacucuauu cuggacuuua uuacuugauu cugucuucug 1620

uauaacucug aaguccacca aaaguggacc cucuauauuu ccucccuuuu uauagucuua 1680

uaagauacau uaugaaaggu gaccgacucu auuuuaaauc ucagaauuuu aaguucuagc 1740

cccaugauaa ccuuuuucuu uguaauuuau gcuuucauau auccuugguc ccagagaugu 1800

uuagacaauu uuaggcucaa aaauuaaagc uaacacagga aaaggaacug uacuggcuau 1860

uacauaagaa acaauggacc caagagaaga aaaggaagaa agaaagguuu uuugguuuuu 1920

guuuuguuuu guuuuguuuu uuguuuuuuu gagauggagu cucacucuuu cgcccaggcu 1980

ggagugcagu gguaugaucu cagcucacug caagcuccac cucccggguu cacgccauuc 2040

uccugccuca gccuccugag uagcugggac uacaggcgcc cgccaccaca cccggcuaau 2100

uuuuuguauu uuuuguagag acgggguuuc accauguuag ccaagauggu cucgaucucc 2160

ugaccucgug auccaccugc cucggccucc caaagugcug ggauuacggg ugugagccac 2220

cgugcccagc cuuuuuuuuu uuaauagaaa aaauaauccg acucccacua caucaagacu 2280

aaucuuguuu uguguguuuu ucacauguau uauagaaugc uuuugcaugg acuauccucu 2340

uguuuuuauu aaaaacaaau gauuuuuuua aaagucacaa aaacaauuca cuaaaaauaa 2400

auaugucauu gugcuuuaaa aaaauaaccu cuuguaguua uaaaauaaaa cguuugacuu 2460

cuaaacucug 2470

<210> 52

<211> 1714

<212> RNA

<213> Homo sapien

<400> 52

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60

aagccaugaa caucauccua gaaauccuuc ugcuucugau caccaucauc uacuccuacu 120

uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180

uucucauuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240

gacagagcau auugguucug ugggauauua auaagcgcgg uguggaggaa acugcagcug 300

agugccgaaa acuaggcguc acugcgcaug cguauguggu agacugcagc aacagagaag 360

agaucuaucg cucucuaaau caggugaaga aagaaguggg ugauguaaca aucgugguga 420

auaaugcugg gacaguauau ccagccgauc uucucagcac caaggaugaa gagauuacca 480

agacauuuga ggucaacauc cuaggacauu uuuggaucac aaaagcacuu cuuccaucga 540

ugauggagag aaaucauggc cacaucguca caguggcuuc agugugcggc cacgaaggga 600

uuccuuaccu caucccauau uguuccagca aauuugccgc uguuggcuuu cacagagguc 660

ugacaucaga acuucaggcc uugggaaaaa cugguaucaa aaccucaugu cucugcccag 720

uuuuugugaa uacuggguuc accaaaaauc caagcacaag auuauggccu guauuggaga 780

cagaugaagu cguaagaagu cugauagaug gaauacuuac caauaagaaa augauuuuug 840

uuccaucgua uaucaauauc uuucugagac uacagaagua aguacagcac agaacaccca 900

aauacuaaaa caccaauaga gcuuuuuuuu uugcuuuuuu uuuuuuuaga cagagucuca 960

cucugucacc cuggcuggau ugcgguggguu gcaguggcau gaucuuggcu cacugcaacc 1020

uccgccuccu ggguucaagc aauucucaug ccucagaccc ccaaguaacu gggauuauag 1080

gugugugcug ccacacuaca cccagcuaau uuuuguauuu uuugauagag acagguuucc 1140

ccauguuggc caggcuggac ucgaacuccu gaccucaagu uauccuccug ucucggccuc 1200

ccaaagugcu gggauuacag ucaugagcca ccaugccugg cccaauagag cuauuauuau 1260

ggagcaucuu ucaguuguga aaauuggcau ggaaacucuc caucccuggg gagaacaguu 1320

auuuccucug uuauuuuccu acccagucua uaaaaagaga gugauucauu uucucuacca 1380

aaucuacugu cucugcccaa acuuugcuga agacuauucu aacuaaagga aacacaguuu 1440

aaaaagaaug caauauagug aaguaguuaa uaauaaagac uccauuuuua aaagucugcu 1500

ggaaguuugg uugggauugc acugaaucua uagagcaauu ggggaguauu gacauaucaa 1560

caauauugag uuuucuaauc caagaacaua auaucuauuu uuaaaaucuu cuucaaaauc 1620

uuuaaaucuu uaaauuguau uuuguaguuu uugguguuua agucuugcac auauuuuguc 1680

agauuuauuc caaaguauuu cacggguucu uuuu 1714

<210> 53

<211> 2397

<212> DNA

<213> Homo sapien

<400> 53

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60

aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120

tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180

ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240

gacagagcat attggttctg tgggatatta ataagcgcgg tgtggaggaa actgcagctg 300

agtgccgaaa actaggcgtc actgcgcatg cgtatgtggt agactgcagc aacagagaag 360

agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420

ataatgctgg gacagtatat ccagccgatc ttctcagcac caaggatgaa gagattacca 480

agacatttga ggtcaacatc ctaggacatt tttggatcac aaaagcactt cttccatcga 540

tgatggagag aaatcatggc cacatcgtca cagtggcttc agtgtgcggc cacgaaggga 600

ttccttacct catcccatat tgttccagca aatttgccgc tgttggcttt cacagaggtc 660

tgacatcaga acttcaggcc ttgggaaaaa ctggtatcaa aacctcatgt ctctgcccag 720

tttttgtgaa tactgggttc accaaaaatc caagcacaag attatggcct gtattggaga 780

cagatgaagt cgtaagaagt ctgatagatg gaatacttac caataagaaa atgatttttg 840

ttccatcgta tatcaatatc tttctgagac tacagaagtt tcttcctgaa cgcgcctcag 900

cgattttaaa tcgtatgcag aatattcaat ttgaagcagt ggttggccac aaaatcaaaa 960

tgaaatgaat aaataagctc cagccagaga tgtatgcatg ataatgatat gaatagtttc 1020

gaatcaatgc tgcaaagctt tatttcacat tttttcagtc ctgataatat taaaaacatt 1080

ggtttggcac tagcagcagt caaacgaaca agattaatta cctgtcttcc tgtttctcaa 1140

gaatatttac gtagtttttc ataggtctgt ttttcctttc atgcctctta aaaacttctg 1200

tgcttacata aacatactta aaaggttttc tttaagatat tttatttttc catttaaagg 1260

tggacaaaag ctacctccct aaaagtaaat acaaagagaa cttatttaca cagggaaggt 1320

ttaagactgt tcaagtagca ttccaatctg tagccatgcc acagaatatc aacaagaaca 1380

cagaatgagt gcacagctaa gagatcaagt ttcagcaggc agctttatct caacctggac 1440

atattttaag attcagcatt tgaaagattt ccctagcctc ttccttttc attagcccaa 1500

aacggtgcaa ctctattctg gactttatta cttgattctg tcttctgtat aactctgaag 1560

tccaccaaaa gtggaccctc tatatttcct ccctttttat agtcttataa gatacattat 1620

gaaaggtgac cgactctatt ttaaatctca gaattttaag ttctagcccc atgataacct 1680

ttttctttgt aatttatgct ttcatatatc cttggtccca gagatgttta gacaatttta 1740

ggctcaaaaa ttaaagctaa cacaggaaaa ggaactgtac tggctattac ataagaaaca 1800

atggacccaa gagaagaaaa ggaagaaaga aaggtttttt ggtttttgtt ttgttttgtt 1860

ttgttttttg ttttttgag atggagtctc actctttcgc ccaggctgga gtgcagtggt 1920

atgatctcag ctcactgcaa gctccacctc ccgggttcac gccattctcc tgcctcagcc 1980

tcctgagtag ctgggactac aggcgcccgc caccacaccc ggctaatttt ttgtattttt 2040

tgtagagacg gggtttcacc atgttagcca agatggtctc gatctcctga cctcgtgatc 2100

cacctgcctc ggcctcccaa agtgctggga ttacgggtgt gagccaccgt gcccagcctt 2160

ttttttttta atagaaaaaa taatccgact cccactacat caagactaat cttgttttgt 2220

gtgtttttca catgtattat agaatgcttt tgcatggact atcctcttgt ttttattaaa 2280

aacaaatgat ttttttaaaa gtcacaaaaa caattcacta aaaataaata tgtcattgtg 2340

ctttaaaaaa ataacctctt gtagttataa aataaaacgt ttgacttcta aactctg 2397

<210> 54

<211> 2289

<212> DNA

<213> Homo sapien

<400> 54

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60

aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120

tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180

ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240

gacagagcat attggttctg tgggatatta ataaggtgaa gaaagaagtg ggtgatgtaa 300

caatcgtggt gaataatgct gggacagtat atccagccga tcttctcagc accaaggatg 360

aagagattac caagacattt gaggtcaaca tcctaggaca tttttggatc acaaaagcac 420

ttcttccatc gatgatggag agaaatcatg gccacatcgt cacagtggct tcagtgtgcg 480

gccacgaagg gattccttac ctcatcccat attgttccag caaatttgcc gctgttggct 540

ttcacagagg tctgacatca gaacttcagg ccttgggaaa aactggtatc aaaacctcat 600

gtctctgccc agtttttgtg aatactgggt tcaccaaaaa tccaagcaca agattatggc 660

ctgtattgga gacagatgaa gtcgtaagaa gtctgataga tggaatactt accaataaga 720

aaatgatttt tgttccatcg tatatcaata tctttctgag actacagaag tttcttcctg 780

aacgcgcctc agcgatttta aatcgtatgc agaatattca atttgaagca gtggttggcc 840

acaaaatcaa aatgaaatga ataaataagc tccagccaga gatgtatgca tgataatgat 900

atgaatagtt tcgaatcaat gctgcaaagc tttatttcac atttttcag tcctgataat 960

attaaaaaca ttggtttggc actagcagca gtcaaacgaa caagattaat tacctgtctt 1020

cctgtttctc aagaatattt acgtagtttt tcataggtct gtttttcctt tcatgcctct 1080

taaaaacttc tgtgcttaca taaacatact taaaaggttt tctttaagat attttatttt 1140

tccatttaaa ggtggacaaa agctacctcc ctaaaagtaa atacaaagag aacttattta 1200

cacagggaag gtttaagact gttcaagtag cattccaatc tgtagccatg ccacagaata 1260

tcaacaagaa cacagaatga gtgcacagct aagagatcaa gtttcagcag gcagctttat 1320

ctcaacctgg acatatttta agattcagca tttgaaagat ttccctagcc tcttccttt 1380

tcattagccc aaaacggtgc aactctattc tggactttat tacttgattc tgtcttctgt 1440

ataactctga agtccaccaa aagtggaccc tctatatttc ctccctttt atagtcttat 1500

aagatacatt atgaaaggtg accgactcta ttttaaatct cagaatttta agttctagcc 1560

ccatgataac ctttttcttt gtaatttatg ctttcatata tccttggtcc cagagatgtt 1620

tagacaattt taggctcaaa aattaaagct aacacaggaa aaggaactgt actggctatt 1680

acataagaaa caatggaccc aagagaagaa aaggaagaaa gaaaggtttt ttggtttttg 1740

ttttgttttg ttttgttttt tgttttttg agatggagtc tcactctttc gcccaggctg 1800

gagtgcagtg gtatgatctc agctcactgc aagctccacc tcccgggttc acgccattct 1860

cctgcctcag cctcctgagt agctgggact acaggcgccc gccaccacac ccggctaatt 1920

ttttgtattt tttgtagaga cggggtttca ccatgttagc caagatggtc tcgatctcct 1980

gacctcgtga tccacctgcc tcggcctccc aaagtgctgg gattacgggt gtgagccacc 2040

gtgcccagcc ttttttttt taatagaaaa aataatccga ctcccactac atcaagacta 2100

atcttgtttt gtgtgttttt cacatgtatt atagaatgct tttgcatgga ctatcctctt 2160

gtttttatta aaaacaaatg atttttttaa aagtcacaaa aacaattcac taaaaataaa 2220

tatgtcattg tgctttaaaa aaataacctc ttgtagttat aaaataaaac gtttgacttc 2280

taaactctg 2289

<210> 55

<211> 2280

<212> DNA

<213> Homo sapien

<400> 55

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60

aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120

tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180

ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240

gacagagcat attggttctg tgggatatta ataagcgcgg tgtggaggaa actgcagctg 300

agtgccgaaa actaggcgtc actgcgcatg cgtatgtggt agactgcagc aacagagaag 360

agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420

ataatgctgg gacagtatat ccagccgatc ttctcagcac caaggatgaa gagattacca 480

agacatttga ggtcaacatc ctaggacatt tttggatcac aaaagcactt cttccatcga 540

tgatggagag aaatcatggc cacatcgtca cagtggcttc agtgtgcggc cacgaaggga 600

ttccttacct catcccatat tgttccagca aatttgccgc tgttggcttt cacagaggtc 660

tgacatcaga acttcaggcc ttgggaaaaa ctggtatcaa aacctcatgt ctctgcccag 720

tttttgtgaa tactgggttc accaaaaatc caagcacaag gtttcttcct gaacgcgcct 780

cagcgatttt aaatcgtatg cagaatattc aatttgaagc agtggttggc cacaaaatca 840

aaatgaaatg aataaataag ctccagccag agatgtatgc atgataatga tatgaatagt 900

ttcgaatcaa tgctgcaaag ctttatttca catttttca gtcctgataa tattaaaaac 960

attggtttgg cactagcagc agtcaaacga acaagattaa ttacctgtct tcctgtttct 1020

caagaatatt tacgtagttt ttcataggtc tgtttttcct ttcatgcctc ttaaaaactt 1080

ctgtgcttac ataaacatac ttaaaaggtt ttctttaaga tattttattt ttccatttaa 1140

aggtggacaa aagctacctc cctaaaagta aatacaaaga gaacttattt acacagggaa 1200

ggtttaagac tgttcaagta gcattccaat ctgtagccat gccacagaat atcaacaaga 1260

acacagaatg agtgcacagc taagagatca agtttcagca ggcagcttta tctcaacctg 1320

gacatatttt aagattcagc atttgaaaga tttccctagc ctcttccttt ttcattagcc 1380

caaaacggtg caactctatt ctggacttta ttacttgatt ctgtcttctg tataactctg 1440

aagtccacca aaagtggacc ctctatattt cctccctttt tatagtctta taagatacat 1500

tatgaaaggt gaccgactct attttaaatc tcagaatttt aagttctagc cccatgataa 1560

ccttttttctt tgtaatttat gctttcatat atccttggtc ccagagatgt ttagacaatt 1620

ttaggctcaa aaattaaagc taacacagga aaaggaactg tactggctat tacataagaa 1680

acaatggacc caagagaaga aaaggaagaa agaaaggttt tttggttttt gttttgtttt 1740

gttttgtttt ttgtttttttt gagatggagt ctcactcttt cgcccaggct ggagtgcagt 1800

ggtatgatct cagctcactg caagctccac ctcccgggtt cacgccattc tcctgcctca 1860

gcctcctgag tagctgggac tacaggcgcc cgccaccaca cccggctaat tttttgtatt 1920

ttttgtagag acggggtttc accatgttag ccaagatggt ctcgatctcc tgacctcgtg 1980

atccacctgc ctcggcctcc caaagtgctg ggattacggg tgtgagccac cgtgcccagc 2040

ctttttttttt ttaatagaaa aaataatccg actcccacta catcaagact aatcttgttt 2100

tgtgtgtttt tcacatgtat tatagaatgc ttttgcatgg actatcctct tgtttttatt 2160

aaaaacaaat gatttttta aaagtcacaa aaacaattca ctaaaaataa atatgtcatt 2220

gtgctttaaa aaaataacct cttgtagtta taaaataaaa cgtttgactt ctaaactctg 2280

<210> 56

<211> 2398

<212> DNA

<213> Homo sapien

<400> 56

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60

aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120

tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180

ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240

gacagagcat attggttctg tgggatatta ataagcgcgg tgtggaggaa actgcagctg 300

agtgccgaaa actaggcgtc actgcgcatg cgtatgtggt agactgcagc aacagagaag 360

agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420

ataatgctgg gacagtatat ccagccgatc ttctcagcac caaggatgaa gagattacca 480

agacatttga ggtcaacatc ctaggacatt tttggatcac aaaagcactt cttccatcga 540

tgatggagag aaatcatggc cacatcgtca cagtggcttc agtgtgcggc cacgaaggga 600

ttccttacct catcccatat tgttccagca aatttgccgc tgttggcttt cacagaggtc 660

tgacatcaga acttcaggcc ttgggaaaaa ctggtatcaa aacctcatgt ctctgcccag 720

tttttgtgaa tactgggttc accaaaaatc caagcacaag attatggcct gtattggaga 780

cagatgaagt cgtaagaagt ctgatagatg gaatacttac caataagaaa atgatttttg 840

ttccatcgta tatcaatatc tttctgagac tacagaaggt ttcttcctga acgcgcctca 900

gcgattttaa atcgtatgca gaatattcaa tttgaagcag tggttggcca caaaatcaaa 960

atgaaatgaa taaataagct ccagccagag atgtatgcat gataatgata tgaatagttt 1020

cgaatcaatg ctgcaaagct ttatttcaca ttttttcagt cctgataata ttaaaaacat 1080

tggtttggca ctagcagcag tcaaacgaac aagattaatt acctgtcttc ctgtttctca 1140

agaatattta cgtagttttt cataggtctg tttttccttt catgcctctt aaaaacttct 1200

gtgcttacat aaacatactt aaaaggtttt ctttaagata ttttatttt ccatttaaag 1260

gtggacaaaa gctacctccc taaaagtaaa tacaaagaga acttatttac acagggaagg 1320

tttaagactg ttcaagtagc attccaatct gtagccatgc cacagaatat caacaagaac 1380

acagaatgag tgcacagcta agagatcaag tttcagcagg cagctttatc tcaacctgga 1440

catattttaa gattcagcat ttgaaagatt tccctagcct cttccttttt cattagccca 1500

aaacggtgca actctattct ggactttatt acttgattct gtcttctgta taactctgaa 1560

gtccaccaaa agtggaccct ctatatttcc tccctttta tagtcttata agatacatta 1620

tgaaaggtga ccgactctat tttaaatctc agaattttaa gttctagccc catgataacc 1680

tttttctttg taatttatgc tttcatatat ccttggtccc agagatgttt agacaatttt 1740

aggctcaaaa attaaagcta acacaggaaa aggaactgta ctggctatta cataagaaac 1800

aatggaccca agagaagaaa aggaagaaag aaaggttttt tggtttttgt tttgttttgt 1860

tttgtttttt gttttttga gatggagtct cactctttcg cccaggctgg agtgcagtgg 1920

tatgatctca gctcactgca agctccacct cccgggttca cgccattctc ctgcctcagc 1980

ctcctgagta gctgggacta caggcgcccg ccaccacacc cggctaattt tttgtatttt 2040

ttgtagagac ggggtttcac catgttagcc aagatggtct cgatctcctg acctcgtgat 2100

ccacctgcct cggcctccca aagtgctggg attacgggtg tgagccaccg tgcccagcct 2160

ttttttttt aatagaaaaa ataatccgac tcccactaca tcaagactaa tcttgttttg 2220

tgtgtttttc acatgtatta tagaatgctt ttgcatggac tatcctcttg tttttattaa 2280

aaacaaatga ttttttaaa agtcacaaaa acaattcact aaaaataaat atgtcattgt 2340

gctttaaaaa aataacctct tgtagttata aaataaaacg tttgacttct aaactctg 2398

<210> 57

<211> 2469

<212> DNA

<213> Homo sapien

<400> 57

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60

aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120

tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180

ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240

gacagagcat attggttctg tgggatatta ataagcgcgg tgtggaggaa actgcagctg 300

agtgccgaaa actaggcgtc actgcgcatg cgtatgtggt agactgcagc aacagagaag 360

agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420

ataatgctgg gacagtatat ccagccgatc ttctcagcac caaggatgaa gagattacca 480

agacatttga ggtcaacatc ctaggacatt tttggaatgg aaaggacatc agaagtaatt 540

acttggatgt atataggatc gaggacactt ttggacgaga ctctgagatc acaaaagcac 600

ttcttccatc gatgatggag agaaatcatg gccacatcgt cacagtggct tcagtgtgcg 660

gccacgaagg gattccttac ctcatcccat attgttccag caaatttgcc gctgttggct 720

ttcacagagg tctgacatca gaacttcagg ccttgggaaa aactggtatc aaaacctcat 780

gtctctgccc agtttttgtg aatactgggt tcaccaaaaa tccaagcaca agattatggc 840

ctgtattgga gacagatgaa gtcgtaagaa gtctgataga tggaatactt accaataaga 900

aaatgatttt tgttccatcg tatatcaata tctttctgag actacagaag tttcttcctg 960

aacgcgcctc agcgatttta aatcgtatgc agaatattca atttgaagca gtggttggcc 1020

acaaaatcaa aatgaaatga ataaataagc tccagccaga gatgtatgca tgataatgat 1080

atgaatagtt tcgaatcaat gctgcaaagc tttatttcac atttttcag tcctgataat 1140

attaaaaaca ttggtttggc actagcagca gtcaaacgaa caagattaat tacctgtctt 1200

cctgtttctc aagaatattt acgtagtttt tcataggtct gtttttcctt tcatgcctct 1260

taaaaacttc tgtgcttaca taaacatact taaaaggttt tctttaagat attttatttt 1320

tccatttaaa ggtggacaaa agctacctcc ctaaaagtaa atacaaagag aacttattta 1380

cacagggaag gtttaagact gttcaagtag cattccaatc tgtagccatg ccacagaata 1440

tcaacaagaa cacagaatga gtgcacagct aagagatcaa gtttcagcag gcagctttat 1500

ctcaacctgg acatatttta agattcagca tttgaaagat ttccctagcc tcttcctttt 1560

tcattagccc aaaacggtgc aactctattc tggactttat tacttgattc tgtcttctgt 1620

ataactctga agtccaccaa aagtggaccc tctatatttc ctccctttt atagtcttat 1680

aagatacatt atgaaaggtg accgactcta ttttaaatct cagaatttta agttctagcc 1740

ccatgataac ctttttcttt gtaatttatg ctttcatata tccttggtcc cagagatgtt 1800

tagacaattt taggctcaaa aattaaagct aacacaggaa aaggaactgt actggctatt 1860

acataagaaa caatggaccc aagagaagaa aaggaagaaa gaaaggtttt ttggtttttg 1920

ttttgttttg ttttgttttt tgtttttttg agatggagtc tcactctttc gcccaggctg 1980

gagtgcagtg gtatgatctc agctcactgc aagctccacc tcccgggttc acgccattct 2040

cctgcctcag cctcctgagt agctgggact acaggcgccc gccaccacac ccggctaatt 2100

ttttgtattt tttgtagaga cggggtttca ccatgttagc caagatggtc tcgatctcct 2160

gacctcgtga tccacctgcc tcggcctccc aaagtgctgg gattacgggt gtgagccacc 2220

gtgcccagcc tttttttttt taatagaaaa aataatccga ctcccactac atcaagacta 2280

atcttgtttt gtgtgttttt cacatgtatt atagaatgct tttgcatgga ctatcctctt 2340

gtttttatta aaaacaaatg atttttttaa aagtcacaaa aacaattcac taaaaataaa 2400

tatgtcattg tgctttaaaa aaataacctc ttgtagttat aaaataaaac gtttgacttc 2460

taaactctg 2469

<210> 58

<211> 1715

<212> DNA

<213> Homo sapien

<400> 58

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60

aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120

tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180

ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240

gacagagcat attggttctg tgggatatta ataagcgcgg tgtggaggaa actgcagctg 300

agtgccgaaa actaggcgtc actgcgcatg cgtatgtggt agactgcagc aacagagaag 360

agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420

ataatgctgg gacagtatat ccagccgatc ttctcagcac caaggatgaa gagattacca 480

agacatttga ggtcaacatc ctaggacatt tttggatcac aaaagcactt cttccatcga 540

tgatggagag aaatcatggc cacatcgtca cagtggcttc agtgtgcggc cacgaaggga 600

ttccttacct catcccatat tgttccagca aatttgccgc tgttggcttt cacagaggtc 660

tgacatcaga acttcaggcc ttgggaaaaa ctggtatcaa aacctcatgt ctctgcccag 720

tttttgtgaa tactgggttc accaaaaatc caagcacaag attatggcct gtattggaga 780

cagatgaagt cgtaagaagt ctgatagatg gaatacttac caataagaaa atgatttttg 840

ttccatcgta tatcaatatc tttctgagac tacagaagtt aagtacagca cagaacaccc 900

aaatactaaa acaccaatag agctttttt tttgcttttt ttttttttag acagagtctc 960

actctgtcac cctggctgga ttgcggtggt tgcagtggca tgatcttggc tcactgcaac 1020

ctccgcctcc tgggttcaag caattctcat gcctcagacc cccaagtaac tgggattata 1080

ggtgtgtgct gccacactac acccagctaa tttttgtatt ttttgataga gacaggtttc 1140

cccatgttgg ccaggctgga ctcgaactcc tgacctcaag ttatcctcct gtctcggcct 1200

cccaaagtgc tgggattaca gtcatgagcc accatgcctg gcccaataga gctattatta 1260

tggagcatct ttcagttgtg aaaattggca tggaaactct ccatccctgg ggagaacagt 1320

tatttcctct gttattttcc tacccagtct ataaaaagag agtgattcat tttctctacc 1380

aaatctactg tctctgccca aactttgctg aagactattc taactaaagg aaacacagtt 1440

taaaaagaat gcaatatagt gaagtagtta ataataaaga ctccattttt aaaagtctgc 1500

tggaagtttg gttgggattg cactgaatct atagagcaat tggggagtat tgacatatca 1560

acaatattga gttttctaat ccaagaacat aatatctatt tttaaaatct tcttcaaaat 1620

ctttaaatct ttaaattgta ttttgtagtt tttggtgttt aagtcttgca catattttgt 1680

cagatttatt ccaaagtatt tcacgggttc ttttt 1715

<210> 59

<211> 2290

<212> DNA

<213> Homo sapien

<400> 59

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60

aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120

tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180

ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240

gacagagcat attggttctg tgggatatta ataaggtgaa gaaagaagtg ggtgatgtaa 300

caatcgtggt gaataatgct gggacagtat atccagccga tcttctcagc accaaggatg 360

aagagattac caagacattt gaggtcaaca tcctaggaca tttttggatc acaaaagcac 420

ttcttccatc gatgatggag agaaatcatg gccacatcgt cacagtggct tcagtgtgcg 480

gccacgaagg gattccttac ctcatcccat attgttccag caaatttgcc gctgttggct 540

ttcacagagg tctgacatca gaacttcagg ccttgggaaa aactggtatc aaaacctcat 600

gtctctgccc agtttttgtg aatactgggt tcaccaaaaa tccaagcaca agattatggc 660

ctgtattgga gacagatgaa gtcgtaagaa gtctgataga tggaatactt accaataaga 720

aaatgatttt tgttccatcg tatatcaata tctttctgag actacagaag gtttcttcct 780

gaacgcgcct cagcgatttt aaatcgtatg cagaatattc aatttgaagc agtggttggc 840

cacaaaatca aaatgaaatg aataaataag ctccagccag agatgtatgc atgataatga 900

tatgaatagt ttcgaatcaa tgctgcaaag ctttatttca catttttca gtcctgataa 960

tattaaaaac attggtttgg cactagcagc agtcaaacga acaagattaa ttacctgtct 1020

tcctgtttct caagaatatt tacgtagttt ttcataggtc tgtttttcct ttcatgcctc 1080

ttaaaaactt ctgtgcttac ataaacatac ttaaaaggtt ttctttaaga tattttattt 1140

ttccatttaa aggtggacaa aagctacctc cctaaaagta aatacaaaga gaacttattt 1200

acacagggaa ggtttaagac tgttcaagta gcattccaat ctgtagccat gccacagaat 1260

atcaacaaga acacagaatg agtgcacagc taagagatca agtttcagca ggcagcttta 1320

tctcaacctg gacatatttt aagattcagc atttgaaaga tttccctagc ctcttccttt 1380

ttcattagcc caaaacggtg caactctatt ctggacttta ttacttgatt ctgtcttctg 1440

tataactctg aagtccacca aaagtggacc ctctatattt cctccctttt tatagtctta 1500

taagatacat tatgaaaggt gaccgactct attttaaatc tcagaatttt aagttctagc 1560

cccatgataa ccttttttctt tgtaatttat gctttcatat atccttggtc ccagagatgt 1620

ttagacaatt ttaggctcaa aaattaaagc taacacagga aaaggaactg tactggctat 1680

tacataagaa acaatggacc caagagaaga aaaggaagaa agaaaggttt tttggttttt 1740

gttttgtttt gttttgtttt ttgtttttttt gagatggagt ctcactcttt cgcccaggct 1800

ggagtgcagt ggtatgatct cagctcactg caagctccac ctcccgggtt cacgccattc 1860

tcctgcctca gcctcctgag tagctgggac tacaggcgcc cgccaccaca cccggctaat 1920

tttttgtatt ttttgtagag acggggtttc accatgttag ccaagatggt ctcgatctcc 1980

tgacctcgtg atccacctgc ctcggcctcc caaagtgctg ggattacggg tgtgagccac 2040

cgtgcccagc cttttttttt ttaatagaaa aaataatccg actcccacta catcaagact 2100

aatcttgttt tgtgtgtttt tcacatgtat tatagaatgc ttttgcatgg actatcctct 2160

tgtttttatt aaaaacaaat gatttttta aaagtcacaa aaacaattca ctaaaaataa 2220

atatgtcatt gtgctttaaa aaaataacct cttgtagtta taaaataaaa cgtttgactt 2280

ctaaactctg 2290

<210> 60

<211> 2470

<212> DNA

<213> Homo sapien

<400> 60

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60

aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120

tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180

ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240

gacagagcat attggttctg tgggatatta ataagcgcgg tgtggaggaa actgcagctg 300

agtgccgaaa actaggcgtc actgcgcatg cgtatgtggt agactgcagc aacagagaag 360

agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420

ataatgctgg gacagtatat ccagccgatc ttctcagcac caaggatgaa gagattacca 480

agacatttga ggtcaacatc ctaggacatt tttggaatgg aaaggacatc agaagtaatt 540

acttggatgt atataggatc gaggacactt ttggacgaga ctctgagatc acaaaagcac 600

ttcttccatc gatgatggag agaaatcatg gccacatcgt cacagtggct tcagtgtgcg 660

gccacgaagg gattccttac ctcatcccat attgttccag caaatttgcc gctgttggct 720

ttcacagagg tctgacatca gaacttcagg ccttgggaaa aactggtatc aaaacctcat 780

gtctctgccc agtttttgtg aatactgggt tcaccaaaaa tccaagcaca agattatggc 840

ctgtattgga gacagatgaa gtcgtaagaa gtctgataga tggaatactt accaataaga 900

aaatgatttt tgttccatcg tatatcaata tctttctgag actacagaag gtttcttcct 960

gaacgcgcct cagcgatttt aaatcgtatg cagaatattc aatttgaagc agtggttggc 1020

cacaaaatca aaatgaaatg aataaataag ctccagccag agatgtatgc atgataatga 1080

tatgaatagt ttcgaatcaa tgctgcaaag ctttatttca catttttca gtcctgataa 1140

tattaaaaac attggtttgg cactagcagc agtcaaacga acaagattaa ttacctgtct 1200

tcctgtttct caagaatatt tacgtagttt ttcataggtc tgtttttcct ttcatgcctc 1260

ttaaaaactt ctgtgcttac ataaacatac ttaaaaggtt ttctttaaga tattttattt 1320

ttccatttaa aggtggacaa aagctacctc cctaaaagta aatacaaaga gaacttattt 1380

acacagggaa ggtttaagac tgttcaagta gcattccaat ctgtagccat gccacagaat 1440

atcaacaaga acacagaatg agtgcacagc taagagatca agtttcagca ggcagcttta 1500

tctcaacctg gacatatttt aagattcagc atttgaaaga tttccctagc ctcttccttt 1560

ttcattagcc caaaacggtg caactctatt ctggacttta ttacttgatt ctgtcttctg 1620

tataactctg aagtccacca aaagtggacc ctctatattt cctccctttt tatagtctta 1680

taagatacat tatgaaaggt gaccgactct attttaaatc tcagaatttt aagttctagc 1740

cccatgataa ccttttttctt tgtaatttat gctttcatat atccttggtc ccagagatgt 1800

ttagacaatt ttaggctcaa aaattaaagc taacacagga aaaggaactg tactggctat 1860

tacataagaa acaatggacc caagagaaga aaaggaagaa agaaaggttt tttggttttt 1920

gttttgtttt gttttgtttt ttgttttttt gagatggagt ctcactcttt cgcccaggct 1980

ggagtgcagt ggtatgatct cagctcactg caagctccac ctcccgggtt cacgccattc 2040

tcctgcctca gcctcctgag tagctgggac tacaggcgcc cgccaccaca cccggctaat 2100

tttttgtatt ttttgtagag acggggtttc accatgttag ccaagatggt ctcgatctcc 2160

tgacctcgtg atccacctgc ctcggcctcc caaagtgctg ggattacggg tgtgagccac 2220

cgtgcccagc cttttttttt ttaatagaaa aaataatccg actcccacta catcaagact 2280

aatcttgttt tgtgtgtttt tcacatgtat tatagaatgc ttttgcatgg actatcctct 2340

tgtttttatt aaaaacaaat gatttttta aaagtcacaa aaacaattca ctaaaaataa 2400

atatgtcatt gtgctttaaa aaaataacct cttgtagtta taaaataaaa cgtttgactt 2460

ctaaactctg 2470

<210> 61

<211> 1714

<212> DNA

<213> Homo sapien

<400> 61

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60

aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120

tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180

ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240

gacagagcat attggttctg tgggatatta ataagcgcgg tgtggaggaa actgcagctg 300

agtgccgaaa actaggcgtc actgcgcatg cgtatgtggt agactgcagc aacagagaag 360

agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420

ataatgctgg gacagtatat ccagccgatc ttctcagcac caaggatgaa gagattacca 480

agacatttga ggtcaacatc ctaggacatt tttggatcac aaaagcactt cttccatcga 540

tgatggagag aaatcatggc cacatcgtca cagtggcttc agtgtgcggc cacgaaggga 600

ttccttacct catcccatat tgttccagca aatttgccgc tgttggcttt cacagaggtc 660

tgacatcaga acttcaggcc ttgggaaaaa ctggtatcaa aacctcatgt ctctgcccag 720

tttttgtgaa tactgggttc accaaaaatc caagcacaag attatggcct gtattggaga 780

cagatgaagt cgtaagaagt ctgatagatg gaatacttac caataagaaa atgatttttg 840

ttccatcgta tatcaatatc tttctgagac tacagaagta agtacagcac agaacaccca 900

aatactaaaa caccaataga gctttttttt ttgctttttt ttttttaga cagagtctca 960

ctctgtcacc ctggctggat tgcggtggtt gcagtggcat gatcttggct cactgcaacc 1020

tccgcctcct gggttcaagc aattctcatg cctcagaccc ccaagtaact gggattatag 1080

gtgtgtgctg ccacactaca cccagctaat ttttgtattt tttgatagag acaggtttcc 1140

ccatgttggc caggctggac tcgaactcct gacctcaagt tatcctcctg tctcggcctc 1200

ccaaagtgct gggattacag tcatgagcca ccatgcctgg cccaatagag ctattattat 1260

ggagcatctt tcagttgtga aaattggcat ggaaactctc catccctggg gagaacagtt 1320

atttcctctg ttattttcct acccagtcta taaaaagaga gtgattcatt ttctctacca 1380

aatctactgt ctctgcccaa actttgctga agactattct aactaaagga aacacagttt 1440

aaaaagaatg caatatagtg aagtagttaa taataaagac tccattttta aaagtctgct 1500

ggaagtttgg ttgggattgc actgaatcta tagagcaatt ggggagtatt gacatatcaa 1560

caatattgag ttttctaatc caagaacata atatctattt ttaaaatctt cttcaaaatc 1620

tttaaatctt taaattgtat tttgtagttt ttggtgttta agtcttgcac atattttgtc 1680

agatttattc caaagtattt cacgggttct tttt 1714

<210> 62

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> forward primer

<400> 62

atgaacatca tcctagaaat ccttc 25

<210> 63

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> reverse primer

<400> 63

atcatgcata catctctggc tggag 25

<210> 64

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 64

atcagaactt caggccttgg 20

<210> 65

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> first exon

<400> 65

gcaaagccat gaacatcatc c 21

<210> 66

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> last exon

<400> 66

tcttgatgta gtgggagtcg gatt 24