(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년10월04일
(11) 등록번호 10-2712656
(24) 등록일자 2024년09월26일

(51) 국제특허분류(Int. Cl.)
C12N 9/22 (2006.01) *C12N 15/10* (2017.01)
C12N 15/113 (2010.01) *C12N 5/071* (2010.01)
(52) CPC특허분류
C12N 9/22 (2013.01)
C12N 15/102 (2013.01)
(21) 출원번호 10-2023-7044705(분할)
(22) 출원일자(국제) 2018년01월19일
심사청구일자 2023년12월22일
(85) 번역문제출일자 2023년12월22일
(65) 공개번호 10-2024-0001333
(43) 공개일자 2024년01월03일
(62) 원출원 특허 10-2019-7023931
원출원일자(국제) 2018년01월19일
심사청구일자 2020년12월29일
(86) 국제출원번호 PCT/US2018/014454
(87) 국제공개번호 WO 2018/136758
국제공개일자 2018년07월26일
(30) 우선권주장
62/449,335 2017년01월23일 미국(US)
(뒷면에 계속)
(56) 선행기술조사문헌
US06147066 A
US07820380 B
CN104698108 A
CN103520724 B

(73) 특허권자
리제너론 파마슈티칼스 인코포레이티드
미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리버 로드 777
(72) 발명자
아불-후슨 노라 에스.
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
가츠맨 옴리
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
(뒷면에 계속)
(74) 대리인
특허법인 광장리앤고

전체 청구항 수 : 총 40 항 심사관 : 문동현

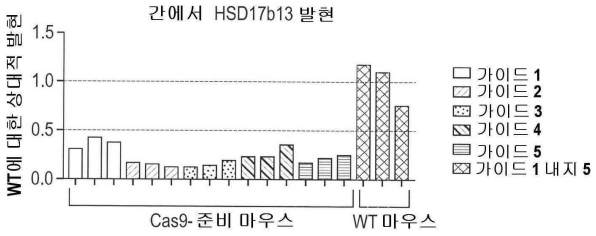
(54) 발명의 명칭 **HSD17B13 변종 및 이것의 용도**

(57) 요 약

*HSD17B13*의 변종과 관련된 분리된 핵산 및 단백질을 포함하여, *HSD17B13* 변종과 관련된 조성물, 및 이런 핵산 및 단백질을 포함하는 세포가 제공된다. 또한 *HSD17B13*의 변종과 관련된 방법이 제공된다. 이러한 방법은 재조합 *HSD17B13* 유전자 또는 HSD17B13 단백질을 암호화하는 핵산을 발현시키기 위하여 뉴클레아제 제제, 외인성 도너

(뒷면에 계속)

대 표 도



도 17A

서열, 전사 활성화제, 전사 억제자, 및 발현 벡터의 임의의 조합을 사용을 통해 세포를 변형시키는 방법을 포함한다. 또한 만성 간 질환을 가졌거나 발병의 위험이 있는 대상체를 치료하기 위한 치료 및 예방 방법이 제공된다.

(52) CPC특허분류
C12N 15/113 (2013.01)
C12N 5/067 (2013.01)
C12N 2310/20 (2017.05)

(72) 발명자
리 알렉산더
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
챙 시핑
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
신 유롱
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
페파니스 에반겔로스
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
하트포드 수잔느
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
그로마다 제스퍼
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
듀이 프레데릭 이.
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
바라스 아리스
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
쉴디너 알란
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내

(30) 우선권주장
62/472,972 2017년03월17일 미국(US)
62/581,918 2017년11월06일 미국(US)

명 세 서

청구범위

청구항 1

세포에서 *HSD17B13* 유전자를 변형시키는 방법에 사용하기 위해, Cas 단백질 또는 Cas 단백질을 암호화하는 핵산, 및 가이드 RNA 또는 가이드 RNA를 암호화하는 DNA를 포함하는 조성물로서,

가이드 RNA는 CRISPR RNA (crRNA) 부분 및 트랜스-활성화 CRISPR RNA (tracrRNA) 부분을 포함하고, *HSD17B13* 유전자 내의 가이드 RNA 표적 서열을 표적화하고, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치를 포함하거나 이에 근접하고,

방법은 Cas 단백질 또는 Cas 단백질을 암호화하는 핵산, 및 가이드 RNA 또는 가이드 RNA를 암호화하는 DNA를 세포에 도입하는 단계를 포함하며, 여기서 가이드 RNA는 Cas 단백질과 복합체를 형성하여 가이드 RNA 표적 서열을 표적화하는 것이고, Cas 단백질은 가이드 RNA 표적 서열을 절단하여 *HSD17B13* 유전자에서 표적화된 유전자 변형을 생성하는 것인,

조성물.

청구항 2

제1항에 있어서,

HSD17B13 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때, *HSD17B13* 유전자가 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 갖지 않는, 조성물.

청구항 3

제1항 또는 제2항에 있어서,

(a) 가이드 RNA 표적 서열이 SEQ ID NO: 226 내지 239, 및 264 내지 268 중 어느 하나를 포함하거나;

(b) 가이드 RNA가 SEQ ID NO: 1629 내지 1642, 및 1648 내지 1652 중 어느 하나를 포함하는 DNA-표적화 분절을 포함하거나;

(c) 가이드 RNA가 SEQ ID NO: 706 내지 719, 936 내지 949, 1166 내지 1179, 1396 내지 1409, 725 내지 729, 955 내지 959, 1185 내지 1189, 및 1415 내지 1419 중 어느 하나를 포함하는, 조성물.

청구항 4

제1항 또는 제2항에 있어서,

HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때, 가이드 RNA 표적 서열이 SEQ ID NO: 2의 엑손 6 또는 인트론 6에 상응하는 영역 내에 있거나, SEQ ID NO: 2의 인트론 6 또는 엑손 7에 상응하는 영역 내에 있는, 조성물.

청구항 5

제1항 또는 제2항에 있어서,

HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때, 가이드 RNA 표적 서열이 SEQ ID NO: 2의 위치 12666에 상응하는 위치에 대해 1000, 500, 400, 300, 200, 100, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5 뉴클레오타이드 내에 있는, 조성물.

청구항 6

제1항 또는 제2항에 있어서,

HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때, 가이드 RNA 표적 서열이 SEQ ID NO: 2의 위치 12666에

상응하는 위치를 포함하는, 조성물.

청구항 7

제1항 또는 제2항에 있어서,

조성물이 하나 이상의 추가 가이드 RNA 또는 하나 이상의 추가 가이드 RNA를 암호화하는 하나 이상의 DNA를 추가로 포함하고, 하나 이상의 추가 가이드 RNA는 *HSD17B13* 유전자 내의 하나 이상의 추가 가이드 RNA 표적 서열을 표적화하며,

방법이 하나 이상의 추가 가이드 RNA 또는 하나 이상의 추가 가이드 RNA를 암호화하는 하나 이상의 DNA를 세포에 도입하는 단계를 추가로 포함하고, 하나 이상의 추가 가이드 RNA는 Cas 단백질과 하나 이상의 복합체를 형성하여 하나 이상의 추가 가이드 RNA 표적 서열을 표적화하며,

Cas 단백질은 하나 이상의 추가 가이드 RNA 표적 서열을 절단하는 것인,

조성물.

청구항 8

제1항 또는 제2항에 있어서,

방법이 *HSD17B13* 유전자의 인트론 6의 도너 스플라이스 부위의 파괴를 초래하는, 조성물.

청구항 9

제1항 또는 제2항에 있어서,

방법이 *HSD17B13* 유전자의 기능의 상실을 초래하는, 조성물.

청구항 10

제1항 또는 제2항에 있어서,

표적화된 유전자 변형이 비상동성 단부 연합에 의한 절단된 가이드 RNA 표적 서열의 복구에 의해 생성되는, 조성물.

청구항 11

제1항 또는 제2항에 있어서,

조성물이 외인성 도너 서열을 추가로 포함하고,

방법이 외인성 도너 서열을 세포 내로 도입하는 단계를 추가로 포함하며, 외인성 도너 서열은 *HSD17B13* 유전자의 표적 게놈 유전자좌와 재조합하여 표적화된 유전자 변형을 생성하는 것인, 조성물.

청구항 12

제11항에 있어서,

외인성 도너 서열에 의한 *HSD17B13* 유전자의 복구가 비상동성 단부 연합-매개된 삽입을 통해 일어나는, 조성물.

청구항 13

제11항에 있어서,

외인성 도너 서열에 의한 *HSD17B13* 유전자의 복구가 상동성 지시된 복구를 통해 일어나는, 조성물.

청구항 14

제13항에 있어서,

외인성 도너 서열이 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 5'측 표적 서열에 혼성화하는 5' 상동성 팔 및 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 3'측 표적 서열에 혼성화하는 3' 상동성 팔을 포함하고, 외인

성 도너 서열이 *HSD17B13* 유전자와 재조합되는, 조성물.

청구항 15

제14항에 있어서,

외인성 도너 서열이 5' 상동성 팔 및 3' 상동성 팔에 의해 플랭킹된 핵산 삽입물을 추가로 포함하는, 조성물.

청구항 16

제15항에 있어서,

핵산 삽입물이 티민을 포함하고, 외인성 도너 서열이 *HSD17B13* 유전자와 재조합 시, *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때, 티민이 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입되는, 조성물.

청구항 17

제11항에 있어서,

외인성 도너 서열의 길이가 50개 뉴클레오타이드 내지 1 kb 사이인, 조성물.

청구항 18

제17항에 있어서,

외인성 도너 서열의 길이가 80개 뉴클레오타이드 내지 200개 뉴클레오타이드 사이인, 조성물.

청구항 19

제11항에 있어서,

외인성 도너 서열이 단일 가닥 올리고데옥시뉴클레오타이드인, 조성물.

청구항 20

제1항 또는 제2항에 있어서,

Cas 단백질이 Cas9 단백질인, 조성물.

청구항 21

제1항 또는 제2항에 있어서,

Cas 단백질을 암호화하는 핵산을 포함하는 조성물.

청구항 22

제21항에 있어서,

Cas 단백질을 암호화하는 핵산이 DNA를 포함하는, 조성물.

청구항 23

제21항에 있어서,

Cas 단백질을 암호화하는 핵산이 RNA를 포함하는, 조성물.

청구항 24

제1항 또는 제2항에 있어서,

가이드 RNA를 RNA의 형태로 포함하는 조성물.

청구항 25

제1항 또는 제2항에 있어서,

가이드 RNA를 암호화하는 DNA를 포함하는 조성물.

청구항 26

제1항 또는 제2항에 있어서,

Cas 단백질 또는 Cas 단백질을 암호화하는 핵산, 또는 가이드 RNA 또는 가이드 RNA를 암호화하는 DNA가 지질 나노입자에 있는, 조성물.

청구항 27

제1항 또는 제2항에 있어서,

Cas 단백질을 암호화하는 핵산 또는 가이드 RNA를 암호화하는 DNA가 아데노-관련 바이러스 벡터에 있는, 조성물.

청구항 28

제1항 또는 제2항에 있어서,

가이드 RNA가 crRNA 부분이 tracrRNA 부분에 연결되어 있는 단일 분자 가이드 RNA인, 조성물.

청구항 29

제28항에 있어서,

가이드 RNA가 SEQ ID NO: 1420, 256, 257, 또는 258에 제시된 서열을 포함하는, 조성물.

청구항 30

제1항 또는 제2항에 있어서,

crRNA 부분 및 tracrRNA 부분이 별도의 RNA 분자인, 조성물.

청구항 31

제30항에 있어서,

crRNA 부분이 SEQ ID NO: 1421에 제시된 서열을 포함하거나, tracrRNA 부분이 SEQ ID NO: 1422에 제시된 서열을 포함하는, 조성물.

청구항 32

제1항 또는 제2항에 있어서,

가이드 RNA가 변형되거나 조절된 안정성을 제공하는 변형을 포함하는, 조성물.

청구항 33

제1항 또는 제2항에 있어서,

세포가 생체외 또는 생체내에 있는, 조성물.

청구항 34

제1항 또는 제2항에 있어서,

세포가 마우스 세포, 래트 세포 또는 인간 세포인, 조성물.

청구항 35

제1항 또는 제2항에 있어서,

세포가 인간 간 세포, 마우스 간 세포, 마우스 만능 세포 또는 래트 만능 세포인, 조성물.

청구항 36

제35항에 있어서,

세포가 인간 간 세포인, 조성물.

청구항 37

제36항에 있어서,

세포가 생체내에 있는, 조성물.

청구항 38

제37항에 있어서,

Cas 단백질 또는 Cas 단백질을 암호화하는 핵산, 및 가이드 RNA 또는 가이드 RNA를 암호화하는 DNA가 생체내에서 간에 도입되는, 조성물.

청구항 39

제1항 또는 제2항에 있어서,

세포는 인간 세포이고,

(a) 가이드 RNA 표적 서열이 SEQ ID NO: 226 내지 239 중 어느 하나를 포함하거나;

(b) DNA-표적화 분절이 SEQ ID NO: 1629 내지 1642 중 어느 하나를 포함하거나;

(c) 가이드 RNA가 SEQ ID NO: 706 내지 719, 936 내지 949, 1166 내지 1179, 및 1396 내지 1409 중 어느 하나를 포함하는,

조성물.

청구항 40

제1항 또는 제2항에 있어서,

세포는 마우스 세포이고,

(a) 가이드 RNA 표적 서열이 SEQ ID NO: 264 내지 268 중 어느 하나를 포함하거나;

(b) DNA-표적화 분절이 SEQ ID NO: 1648 내지 1652 중 어느 하나를 포함하거나;

(c) 가이드 RNA가 SEQ ID NO: 725 내지 729, 955 내지 959, 1185 내지 1189, 및 1415 내지 1419 중 어느 하나를 포함하는,

조성물.

발명의 설명

기 술 분 야

[0001] **관련 출원에 대한 상호 참조**

[0002] 본 출원은 2017년 1월 23일에 출원된 미국 출원 번호 62/449,335, 2017년 3월 17 출원된 미국 출원 번호 62/472,972, 2017년 11 6에 출원된 미국 출원 번호 62/581,918의 이익을 주장하며, 상기 출원들은 각각 본원에 그 전문이 참조로 포함된다.

[0003] **EFS 웹을 통해 텍스트 파일로서 제출된 서열 목록에 대한 참조**

[0004] 파일 507242SEQLIST.txt로 필기된 서열 목록은 507 킬로바이트이며, 2018년 1월 19일에 생성되었고, 본원에 참

조로 포함된다.

[0005] 기술분야

[0006] 본 발명은 HSD17B13 변종 및 이것의 용도에 관한 것이다.

배 경 기 술

[0007] 만성 간 질환 및 간경변은 미국에서 이병률 및 사망률의 선두 요인이며, 2014년에 38,170건 (총 사망건수의 1.5%)의 사망을 차지한다 (Kochanek et al. (2016) *Natl Vital Stat Rep* 65:1-122, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 미국에서 간경변의 가장 공통적인 병인은 알코올성 간 질환, 만성 C형 간염, 및 비알코올성 지방간 질환 (NAFLD)이며, 이것들은 함께 2004년과 2013년 사이에 간 이식을 기다리고 있는 환자들의 약 80%를 차지한다 (Wong et al. (2015) *Gastroenterology* 148:547-555, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 미국에서 NAFLD의 추정된 유병률은 19 내지 46 퍼센트이고 (Browning et al. (2004) *Hepatology* 40:1387-1395; Lazo et al. (2013) *Am J Epidemiol* 178:38-45; 및 Williams et al. (2011) *Gastroenterology* 140:124-131, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨) 시간이 지남에 따라 상승하고 있는데 (Younossi et al. (2011) *Clin Gastroenterol Hepatol* 9:524-530 e1; quiz e60 (2011), 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 그것의 주요 위험 인자인 비만의 증가 속도와 유사하다 (Cohen et al. (2011) *Science* 332:1519-1523, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). C형 간염의 치료에서 상당한 진보가 이루어진 한편 (Morgan et al. (2013) *Ann Intern Med* 158:329-337 및 van der Meer et al. (2012) *JAMA* 308:2584-2593, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 현재 알코올성 또는 비알코올성 간 질환 및 간경변에 대한 증거-기반 치료는 없다.

[0008] 이전의 전 게놈 관련 연구 (genome wide association studies, GWAS)로 만성 간 질환과 관련된 제한된 수의 유전자 및 변종이 확인되었다. 오늘날까지 가장 강력하게 입증된 유전자 연관성 연구는 초기에 비알코올성 지방간 질환 (NAFLD)의 증가된 위험과 관련이 있는 것으로 발견되고 (Romeo et al. (2008) *Nat. Genet.* 40:1461-1465 및 Speliotes et al. (2011) *PLoS Genet.* 7:e1001324, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 계속해서 질환의 중증도 (Rotman et al. (2010) *Hepatology* 52:894-903 및 Sookoian et al. (2009) *J. Lipid Res.* 50:2111-2116, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨) 및 진행 (Trepo et al. (2016) *J. Hepatol.* doi:10.1016/j.jhep.2016.03.011, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨)과 관련이 있는 것으로 밝혀진, 3개의 유전자를 함유한 파타틴-유사 포스포리파아제 도메인의 공통 미스센스 변종에 대한 것이다. 경막 6 수퍼패밀리 구성원 2 (*TM6SF2*) 유전자의 변이는 또한 NAFLD에 대한 증가된 위험을 부여하는 것으로 나타났다 (Kozlitina et al. (2014) *Nat. Genet.* 46:352-356; Liu et al. (2014) *Nat. Commun.* 5:4309; 및 Sookoian et al. (2015) *Hepatology* 61:515-525, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 이 두 단백질의 정상 기능은 잘 알려져 있지 않지만, 둘 다 간세포 지질 대사에 관여하는 것으로 제안되어 왔다. *PNPLA3* 및 *TM6SF2*에서의 변종이 어떻게 간 질환의 증가된 위험에 기여하는지는 아직 설명되지 않았다. GWAS는 또한 임상적으로 빈번하게 측정되는 간세포 손상 및 간 지방 축적의 정량적 마커인 혈청 알라닌 아미노전이효소 (ALT) 및 아스파테이트 아미노전이효소 (AST) (Chambers et al. (2011) *Nat. Genet.* 43:131-1138 및 Yuan et al. (2008) *Am. J. Hum. Genet.* 83:520-528, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨)와 관련이 있는 몇몇 유전적 인자를 확인하였다. 오늘날까지, 만성 간 질환에 대한 보호성 유전자 변종은 기술된 것이 없다. 다른 환경에서의 보호성 유전자 변종의 발견, 예컨대 심혈관 질환의 위험을 감소시키는 PCSK9에서의 기능-상실 변종의 발견은 새로운 부류의 치료제의 개발에 대한 촉매가 되었다.

발명의 내용

해결하려는 과제

[0009] 만성 간 질환의 발달 및 진행의 기저에 있는 유전적 인자들에 대한 지식은 위험 계층화를 개선시킬 수 있고 신규한 치료적 전략에 대한 기초를 제공할 수 있을 것이다. 위험 계층화를 개선하고 간 질환에 대한 신규한 치료법을 생성하기 위해 기저에 있는 유전적 인자들에 대한 더 나은 이해가 필요하다.

과제의 해결 수단

[0010] *HSD17B13* rs72613567 변종 유전자, 변종 *HSD17B13* 전사물, 및 변종 HSD17B13 단백질 동형(isoform)에 관련된

방법 및 조성물이 제공된다.

[0011] 한 측면으로, *HSD17B13* rs72613567 변종 유전자로부터의 돌연변이 잔기를 포함하는 분리된 핵산이 제공된다. 이러한 분리된 핵산은 *HSD17B13* 유전자의 적어도 15개의 연속 뉴클레오타이드를 포함하며 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665와 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 갖는다. 선택적으로, 연속 뉴클레오타이드는 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666을 포함한 SEQ ID NO: 2에서의 상응하는 서열에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일하다. 선택적으로, *HSD17B13* 유전자는 인간 *HSD17B13* 유전자이다. 선택적으로, 분리된 핵산은 SEQ ID NO: 2의 적어도 15, 적어도 20, 적어도 25, 적어도 30, 적어도 35, 적어도 40, 적어도 45, 적어도 50, 적어도 60, 적어도 70, 적어도 80, 적어도 90, 적어도 100, 적어도 200, 적어도 300, 적어도 400, 적어도 500, 적어도 600, 적어도 700, 적어도 800, 적어도 900, 적어도 1000, 적어도 2000, 적어도 3000, 적어도 4000, 적어도 5000, 적어도 6000, 적어도 7000, 적어도 8000, 적어도 9000, 적어도 10000, 적어도 11000, 적어도 12000, 적어도 13000, 적어도 14000, 적어도 15000, 적어도 16000, 적어도 17000, 적어도 18000, 또는 적어도 19000개의 연속 뉴클레오타이드를 포함한다.

[0012] 일부 이러한 분리된 핵산은 유전자의 하나 이상의 비필수 분절이 상응하는 야생형 *HSD17B13* 유전자와 관련하여 삭제되어 있는 *HSD17B13* 미니유전자를 포함한다. 선택적으로, 삭제된 분절은 하나 이상의 인트론성 서열을 포함한다. 선택적으로, 분리된 핵산은 추가로 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 인트론 6에 상응하는 인트론을 포함한다. 선택적으로, 인트론은 SEQ ID NO: 2의 인트론이다.

[0013] 다른 측면으로, 상이한 *HSD17B13* mRNA 전사물 또는 cDNA와 상응하는 분리된 핵산이 제공된다. 일부 이러한 분리된 핵산은 HSD17B13 단백질의 전부 또는 일부를 암호화하는 적어도 15개의 연속 뉴클레오타이드를 포함하며, 여기서 연속하는 핵산은 SEQ ID NO: 4에 존재하지 않는 (*HSD17B13* 전사물 A) SEQ ID NO: 7 (*HSD17B13* 전사물 D), SEQ ID NO: 10 (*HSD17B13* 전사물 G), 및 SEQ ID NO: 11 (*HSD17B13* 전사물 H)에 존재하는 분절에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절을 포함한다. 선택적으로, 연속 뉴클레오타이드는 SEQ ID NO: 11 (*HSD17B13* 전사물 H)에 존재하지 않는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 존재하는 분절에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절을 추가로 포함하며, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 10 (*HSD17B13* 전사물 G)에 존재하지 않는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 존재하는 분절에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절을 추가로 포함한다. 선택적으로, 연속 뉴클레오타이드는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 존재하지 않는 SEQ ID NO: 11 (*HSD17B13* 전사물 H)에 존재하는 분절에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절을 추가로 포함한다. 선택적으로, 연속 뉴클레오타이드는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 존재하지 않는 SEQ ID NO: 10 (*HSD17B13* 전사물 G)에 존재하는 분절에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절을 추가로 포함한다.

[0014] 일부 이러한 분리된 핵산은 HSD17B13 단백질의 전부 또는 일부를 암호화하는 적어도 15개의 연속 뉴클레오타이드를 포함하며, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하지 않는 SEQ ID NO: 8 (*HSD17B13* 전사물 E)에 존재하는 분절에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절을 포함한다. 선택적으로, 연속 뉴클레오타이드는 SEQ ID NO: 11 (*HSD17B13* 전사물 H)에 존재하지 않는 SEQ ID NO: 8 (*HSD17B13* 전사물 E)에 존재하는 분절에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절을 포함한다.

[0015] 일부 이러한 분리된 핵산은 HSD17B13 단백질의 전부 또는 일부를 암호화하는 적어도 15개의 연속 뉴클레오타이드를 포함하며, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하지 않는 SEQ ID NO: 9 (*HSD17B13* 전사물 F)에 존재하는 분절에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절을 포함한다.

[0016] 일부 이러한 분리된 핵산은 HSD17B13 단백질의 전부 또는 일부를 암호화하는 적어도 15개의 연속 뉴클레오타이드를 포함하며, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하지 않는 SEQ ID NO: 6 (*HSD17B13* 전사물 C)에 존재하는 분절에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절을 포함한다.

[0017] 선택적으로, HSD17B13 단백질은 인간 HSD17B13 단백질이다. 선택적으로, 분리된 핵산은 HSD17B13 단백질의 전부 또는 일부를 암호화하는 적어도 20, 적어도 25, 적어도 30, 적어도 35, 적어도 40, 적어도 45, 적어도 50, 적어

도 60, 적어도 70, 적어도 80, 적어도 90, 적어도 100, 적어도 200, 적어도 300, 적어도 400, 적어도 500, 적어도 600, 적어도 700, 적어도 800, 적어도 900, 적어도 1000, 또는 적어도 2000개의 연속 뉴클레오타이드를 포함한다.

[0018] 일부 이러한 분리된 핵산은 각각 SEQ ID NO: 6, 7, 8, 9, 10, 또는 11 (*HSD17B13* 전사물 C, D, E, F, G, 또는 H)에 제시된 서열에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하고 SEQ ID NO: 14, 15, 16, 17, 18, 또는 19 (HSD17B13 동형 C, D, E, F, G, 또는 H)에서 제시된 서열을 포함하는 HSD17B13 단백질을 암호화하는 서열을 포함한다.

[0019] 상기 핵산들 중 임의의 핵산에서, 연속 뉴클레오타이드는 개재된 인트론 없이 *HSD17B13* 유전자의 적어도 2개의 상이한 엑손으로부터의 서열을 선택적으로 포함할 수 있다.

[0020] 또 다른 측면으로, 상기 분리된 핵산들 중 임의의 핵산에 의해 암호화된 단백질이 제공된다.

[0021] 또 다른 측면으로, *HSD17B13* rs72613567 변종 유전자로부터의 돌연변이 잔기와 혼성화하는 또는 그 잔기와 유사한 분리된 핵산이 제공된다. 이러한 분리된 핵산은 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 1000, 500, 400, 300, 200, 100, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5개의 뉴클레오타이드를 포함하는 또는 그 안에 있는 분절에서 *HSD17B13* 유전자와 혼성화하는 적어도 15개의 연속 뉴클레오타이드를 포함할 수 있다. 선택적으로, 분절은 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2에서의 상응하는 서열에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일하다. 선택적으로, 분절은 SEQ ID NO: 2의 적어도 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 또는 2000개의 연속 뉴클레오타이드를 포함한다. 선택적으로, 분절은 SEQ ID NO: 2의 위치 12666 또는 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치를 포함한다. 선택적으로, *HSD17B13* 유전자는 인간 *HSD17B13* 유전자이다. 선택적으로, 분리된 핵산은 약 30, 40, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개까지의 뉴클레오타이드 길이를 가진다. 선택적으로, 분리된 핵산은 이종성 핵산에 연결되거나 이종성 표지를 포함한다. 선택적으로, 이종성 표지는 형광 표지이다.

[0022] 또 다른 측면으로, 상이한 *HSD17B13* mRNA 전사물 또는 cDNA에 혼성화하는 분리된 핵산이 제공된다. 일부 이러한 분리된 핵산은 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하며, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하지 않는 SEQ ID NO: 7 (*HSD17B13* 전사물 D), SEQ ID NO: 10 (*HSD17B13* 전사물 G), 및 SEQ ID NO: 11 (*HSD17B13* 전사물 H)에 존재하는 분절에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절을 포함한다.

[0023] 일부 이러한 분리된 핵산은 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하며, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하지 않는 SEQ ID NO: 8 (*HSD17B13* 전사물 E) 및 SEQ ID NO: 11 (*HSD17B13* 전사물 H)에 존재하는 분절에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 분절을 포함한다.

[0024] 일부 이러한 분리된 핵산은 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하며, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하지 않는 SEQ ID NO: 9 (*HSD17B13* 전사물 F)에 존재하는 분절에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 분절을 포함한다.

[0025] 일부 이러한 분리된 핵산은 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하며, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하지 않는 SEQ ID NO: 6 (*HSD17B13* 전사물 C)에 존재하는 분절에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 분절을 포함한다.

[0026] 선택적으로, HSD17B13 단백질은 인간 HSD17B13 단백질이다. 선택적으로, 분리된 핵산은 약 30, 40, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개까지의 뉴클레오타이드 길이를 가진다. 선택적으로, 분리된 핵산은 이종성 핵산에 연결되거나 이종성 표지를 포함한다. 선택적으로, 이종성 표지는 형광 표지이다.

[0027] 선택적으로, 상기 분리된 핵산들 중 임의의 핵산은 DNA를 포함한다. 선택적으로, 상기 분리된 핵산들 중 임의의 핵산은 RNA를 포함한다. 선택적으로, 상기 분리된 핵산들 중 임의의 핵산은 안티센스 RNA, 짧은 헤어핀 RNA, 또는 작은-간섭 RNA이다. 선택적으로, 상기 분리된 핵산들 중 임의의 핵산은 비-천연 뉴클레오타이드를 포함할 수

있다.

[0028] 또 다른 측면으로, 상기 분리된 핵산들 중 임의의 핵산 및 이종성 핵산 서열을 포함하는 벡터 및 외인성 도너 서열이 제공된다.

[0029] 또 다른 측면으로, 대상체에서 *HSD17B13* rs72613567 변종을 검출하는 방법, 대상체에서 *HSD17B13* 전사물 C, D, E, F, G, 또는 H의 존재를 검출하는 방법, 만성 간 질환이 발달하는 데 대한 대상체의 민감성을 측정하는 방법, 지방간 질환을 가진 대상체를 진단하는 방법, 또는 세포에서 *HSD17B13* 유전자를 변형시키는 방법, 세포에서 *HSD17B13* 유전자의 발현을 변경시키는 방법에서 상기 분리된 핵산들 중 임의의 핵산, 벡터, 또는 외인성 도너 서열의 용도가 제공된다.

[0030] 또 다른 측면으로, *HSD17B13* 유전자를 표적으로 하는 가이드 RNA가 제공된다. 이러한 가이드 RNA는 Cas 효소가 *HSD17B13* 유전자에 결합하거나 *HSD17B13* 유전자를 절단하도록 지시하는데 효과적일 수 있고, 이때 가이드 RNA는 *HSD17B13* 유전자 내의 가이드 RNA 인식 서열에 혼성화하는 DNA-표적화 분절을 포함한다. 즉, 이러한 가이드 RNA는 Cas 효소가 *HSD17B13* 유전자에 결합하거나 *HSD17B13* 유전자를 절단하도록 지시하는데 효과적일 수 있고, 이때 가이드 RNA는 *HSD17B13* 유전자 내의 가이드 RNA 표적 서열을 표적으로 하는 DNA-표적화 분절을 포함한다. 이러한 가이드 RNA는 Cas 효소가 *HSD17B13* 유전자에 결합하거나 *HSD17B13* 유전자를 절단하도록 지시하는데 효과적일 수 있고, 이때 가이드 RNA는 *HSD17B13* 유전자가 SEQ ID NO:2와 최적으로 정열될 때 SEQ ID NO:2의 위치 12666에 상응하는 위치를 포함하거나 근접한 *HSD17B13* 유전자 내의 가이드 RNA 표적 서열을 표적으로 하는 DNA-표적화 분절을 포함한다. 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 226 내지 239 및 264 내지 268 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 임의의 하나로 구성된다. 선택적으로, DNA-표적화 분절은 SEQ ID NO: 1629 내지 1642 및 1648 내지 1652 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 임의의 하나로 구성된다. 선택적으로, 가이드 RNA는 SEQ ID NO: 706 내지 719; 936 내지 949; 1166 내지 1179, 1396 내지 1409, 725 내지 729, 955 내지 959, 1185 내지 1189, 및 1415 내지 1419 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 임의의 하나로 구성된다. 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 226 내지 239 또는 SEQ ID NO: 230 및 231로부터 선택된다. 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 226 내지 230 및 264 내지 268로부터 선택된다. 선택적으로, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 엑손 6 및/또는 인트론 6에 상응하는 영역 내에 있다. 선택적으로, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 엑손 6 및/또는 인트론 6 및/또는 엑손 7에 상응하는 영역 내에 있다. 선택적으로, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 약 1000, 500, 400, 300, 200, 100, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5개의 뉴클레오타이드 내에 있다. 선택적으로, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치를 포함한다.

[0031] 이러한 가이드 RNA는 Cas 효소가 *HSD17B13* 유전자에 결합하거나 *HSD17B13* 유전자를 절단하도록 지시하는데 효과적일 수 있고, 이때 가이드 RNA는 *HSD17B13* 유전자의 개시 코돈을 포함하는 또는 근접한 *HSD17B13* 유전자 내에 있는 가이드 RNA 표적 서열을 표적으로 하는 DNA-표적화 분절을 포함한다. 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 20 내지 81 및 259 내지 263 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 임의의 하나로 구성된다. 선택적으로, DNA-표적화 분절은 SEQ ID NO: 1423 내지 1484 및 1643 내지 1647 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 임의의 하나로 구성된다. 선택적으로, 가이드 RNA는 SEQ ID NO: 500 내지 561, 730 내지 791, 960 내지 1021, 1190 내지 1251, 720 내지 724, 950 내지 954, 1180 내지 1184, 및 1410 내지 1414 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 임의의 하나로 구성된다. 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 20 내지 81 및 259 내지 263으로부터 선택된다. 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 21 내지 23, 33, 및 35로부터 선택된다. 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 33 및 35로부터 선택된다. 선택적으로, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 엑손 1에 상응하는 영역 내에 있다. 선택적으로, 가이드 RNA 표적 서열은 개시 코돈의 약 1000, 500, 400, 300, 200, 100, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5개 뉴클레오타이드 내에 있다.

[0032] 이러한 가이드 RNA는 Cas 효소가 *HSD17B13* 유전자에 결합하거나 *HSD17B13* 유전자를 절단하도록 지시하는데 효과적일 수 있고, 이때 가이드 RNA는 *HSD17B13* 유전자의 종결 코돈을 포함하는 또는 근접한 *HSD17B13* 유전자 내의 가이드 RNA 표적 서열을 표적으로 하는 DNA-표적화 분절을 포함한다. 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 82 내지 225 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 임의의 하나로 구성

된다. 선택적으로, DNA-표적화 분절은 SEQ ID NO: 1485 내지 1628 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 임의의 하나로 구성된다. 선택적으로, 가이드 RNA는 SEQ ID NO: 562 내지 705, 792 내지 935, 1022 내지 1165, 및 1252 내지 1395 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 임의의 하나로 구성된다. 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 82 내지 225로부터 선택된다. 선택적으로, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 엑손 7에 상응하는 영역 내에 있다. 선택적으로, 가이드 RNA 표적 서열은 종결 코돈의 약 1000, 500, 400, 300, 200, 100, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5개 뉴클레오타이드 내에 있다.

[0033] 선택적으로, *HSD17B13* 유전자는 인간 *HSD17B13* 유전자이다. 선택적으로, *HSD17B13* 유전자는 SEQ ID NO: 2를 포함한다.

[0034] 일부 이런 가이드 RNA는 DNA-표적화 분절을 포함하는 규칙적인 간격을 갖는 짧은 회문구조 반복 서열 (CRISPR) RNA (crRNA) 및 트랜스-활성화 CRISPR RNA (tracrRNA)를 포함한다. 선택적으로, 가이드 RNA는 crRNA 및 tracrRNA가 서로에게 혼성화하는 별도의 분자인 모듈식 가이드 RNA이다. 선택적으로, crRNA는 ID NO: 1421에 제시된 서열을 포함하거나, 본질적으로 그것으로 구성되거나, 또는 그것으로 구성되고, tracrRNA는 SEQ ID NO: 1422에 제시된 서열을 포함하거나, 본질적으로 그것으로 구성되거나, 또는 그것으로 구성된다. 선택적으로, 가이드 RNA는 crRNA가 링커를 통해 tracrRNA에 융합된 단일-가이드 RNA이다. 선택적으로, 단일-가이드 RNA는 SEQ ID NO: 1420 및 256 내지 258 중 임의의 하나에 제시된 서열을 포함하거나, 본질적으로 그것으로 구성되거나, 그것으로 구성된다.

[0035] 또 다른 측면으로, 본원에 개시된 *HSD17B13* 전사물 내의 서열에 혼성화하는 안티센스 RNA, siRNA, 또는 shRNA가 제공된다. 이러한 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A) 내의 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 세포에서 *HSD1B13* 전사물 A의 발현을 감소시킬 수 있다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 존재하지 않는, SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하는 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)의 엑손 7 내의 서열 또는 엑손 6-엑손 7 경계에 걸쳐 있는 서열에 혼성화한다. 일부 이러한 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 7 (*HSD17B13* 전사물 D) 내의 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 세포에서 *HSD17B13* 전사물 D의 발현을 감소시킬 수 있다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하지 않는, SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 존재하는 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)의 엑손 7 내의 서열 또는 엑손 6-엑손 7 경계에 걸쳐 있는 서열에 혼성화한다.

[0036] 또 다른 측면으로, 상기 가이드 RNA, 안티센스 RNA, siRNA, 또는 shRNA 중 임의의 것을 암호화하는 DNA가 제공된다. 또 다른 측면으로, 상기 가이드 RNA, 안티센스 RNA, siRNA, 또는 shRNA 중 임의의 것을 암호화하는 DNA 및 이종성 핵산을 포함하는 벡터가 제공된다. 또 다른 측면으로, 세포에서 *HSD17B13* 유전자를 변형시키는 방법 또는 세포에서 *HSD17B13* 유전자의 발현을 변경시키는 방법에서의, 상기 가이드 RNA, 안티센스 RNA, siRNA, 또는 shRNA, 가이드 RNA, 안티센스 RNA, siRNA, 또는 shRNA를 암호화하는 DNA 또는 가이드 RNA, 안티센스 RNA, siRNA, 또는 shRNA를 암호화하는 DNA를 포함하는 벡터 중 임의의 것의 용도가 제공된다.

[0037] 또 다른 측면으로, 상기 분리된 핵산들 중 임의의 것, 상기 가이드 RNA들 중 임의의 것, 상기 분리된 폴리펩타이드들 중 임의의 것, 상기 안티센스 RNA, siRNA, 또는 shRNA 들 중 임의의 것, 상기 벡터들 중 임의의 것, 또는 상기 외인성 도너 서열들 중 임의의 것을 포함하는 조성물이 제공된다. 선택적으로, 조성물은 상기 가이드 RNA들 중 임의의 것 및 Cas 단백질, 예컨대 Cas9 단백질을 포함한다. 선택적으로, 이러한 조성물은 분리된 폴리펩타이드, 가이드 RNA, 안티센스 RNA, siRNA, shRNA, 분리된 핵산, 벡터, 또는 외인성 도너 서열의 안정성을 증가시키는 담체를 포함한다. 선택적으로, 담체는 폴리(락트산) (PLA) 마이크로스피어, 폴리(D,L-락틱-코글리콜산) (PLGA) 마이크로스피어, 리포좀, 미셀, 역미셀, 지질 코클레이트(lipid cochleate), 또는 지질 미세관(lipid microtubule)을 포함한다.

[0038] 또한 상기 분리된 핵산들 중 임의의 것, 상기 가이드 RNA들 중 임의의 것, 상기 안티센스 RNA, siRNA, 또는 shRNA 들 중 임의의 것, 상기 분리된 폴리펩타이드들 중 임의의 것, 또는 상기 벡터들 중 임의의 것을 포함하는 세포가 제공된다. 선택적으로, 세포는 인간 세포, 설치류 세포, 마우스 세포, 또는 래트 세포이다. 선택적으로, 상기 세포들 중 임의의 것은 간 세포 또는 만능 세포이다.

[0039] 또한 세포에서 *HSD17B13* 유전자를 변형시키는 방법 또는 세포에서 *HSD17B13* 유전자의 발현을 변형시키는 방법에

서의 상기 가이드 RNA들 중 임의의 것의 용도가 제공된다. 또한 세포에서 *HSD17B13* 유전자의 발현을 변경시키는 방법에서의 상기 안티센스 RNA, siRNA, 또는 shRNA 중 임의의 것의 용도가 제공된다.

[0040] 또한 세포를 변형시키는 방법, *HSD17B13* 유전자를 변형시키는 방법, 또는 *HSD17B13* 유전자의 발현을 변경시키는 방법이 제공된다. 일부 이러한 방법은 세포에서 *HSD17B13* 유전자를 변형시키기 위한 것으로, 방법은 세포의 게놈을 (a) Cas 단백질; 및 (b) Cas 단백질과 복합체를 형성하고 *HSD17B13* 유전자 내의 가이드 RNA 표적 서열을 표적으로 하는 가이드 RNA와 접촉시키는 단계를 포함하며, 여기서 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치를 포함하거나 그 위치에 근접하고, Cas 단백질은 *HSD17B13* 유전자를 절단한다. 선택적으로, Cas 단백질은 Cas9 단백질이다. 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 226 내지 239 및 264 내지 268 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 또는 임의의 하나로 구성된다. 선택적으로, DNA-표적화 분절은 SEQ ID NO: 1629 내지 1642 및 1648 내지 1652 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 또는 임의의 하나로 구성된다. 선택적으로, 가이드 RNA는 SEQ ID NO: 706 내지 719; 936 내지 949; 1166 내지 1179, 1396 내지 1409, 725 내지 729, 955 내지 959, 1185 내지 1189, 및 1415 내지 1419 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 또는 임의의 하나로 구성된다. 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 226 내지 239로부터 선택되거나, 또는 가이드 RNA 표적 서열은 SEQ ID NO: 230 및 231로부터 선택된다. 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 226 내지 239 및 264 내지 268로부터 선택되거나 또는 SEQ ID NO: 264 내지 268로부터 선택된다. 선택적으로, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO:2와 최적으로 정렬될 때 SEQ ID NO:2의 엑손 6 및/또는 인트론 6에 상응하는 영역 내에 있다. 선택적으로, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO:2와 최적으로 정렬될 때 SEQ ID NO:2의 엑손 6 및/또는 인트론 6 및/또는 엑손 7에 상응하는 영역 내에 있다. 선택적으로, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO:2와 최적으로 정렬될 때 SEQ ID NO:2의 위치 12666에 상응하는 위치의 약 1000, 500, 400, 300, 200, 100, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5개 뉴클레오타이드 내에 있다. 선택적으로, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO:2와 최적으로 정렬될 때 SEQ ID NO:2의 위치 12666에 상응하는 위치를 포함한다.

[0041] 일부 이러한 방법은 추가로 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 표적 서열 5'에 혼성화하는 5' 상동성 팔 및 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 표적 서열 3'에 혼성화하는 3' 상동성 팔을 포함하는 외인성 도너 서열과 게놈을 접촉시키는 단계를 포함하고, 여기서 외인성 도너 서열은 *HSD17B13* 유전자와 재조합한다. 선택적으로, 외인성 도너 서열은 추가로 5' 상동성 팔 및 3' 상동성 팔에 의해 플랭킹된 핵산 삽입물을 포함한다. 선택적으로, 핵산 삽입물은 티민을 포함하고, 외인성 도너 서열이 *HSD17B13* 유전자와 재조합될 때, 티민은 *HSD17B13* 유전자가 SEQ ID NO:1과 최적으로 정렬될 때 SEQ ID NO:1의 위치 12665와 12666에 상응하는 뉴클레오타이드 사이에 삽입된다. 선택적으로, 외인성 도너 서열은 길이가 약 50개 뉴클레오타이드 내지 약 1 kb이거나 또는 길이가 약 80개 뉴클레오타이드 내지 약 200개 뉴클레오타이드이다. 선택적으로, 외인성 도너 서열은 단일-가닥 올리고데옥시뉴클레오타이드이다.

[0042] 일부 이러한 방법은 세포에서 *HSD17B13* 유전자를 변형시키기 위한 것으로, 방법은 세포의 게놈을 (a) Cas 단백질; 및 (b) Cas 단백질과 복합체를 형성하고 *HSD17B13* 유전자 내의 제1 가이드 RNA 표적 서열을 표적으로 하는 제1 가이드 RNA와 접촉시키는 단계를 포함하며, 여기서 제1 가이드 RNA 표적 서열은 *HSD17B13* 유전자에 대한 개시 코돈을 포함하거나 또는 개시 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있거나 또는 SEQ ID NO: 20 내지 81 및 259 내지 263으로부터 선택되고, Cas 단백질은 *HSD17B13* 유전자를 절단하거나 발현을 변경시킨다. 선택적으로, 제1 가이드 RNA 표적 서열은 SEQ ID NO: 20 내지 81 및 259 내지 263 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 또는 임의의 하나로 구성된다. 선택적으로, 제1 가이드 RNA 표적 서열은 SEQ ID NO: 20 내지 41 중 임의의 하나, SEQ ID NO: 21 내지 23, 33, 및 35 중 임의의 하나, 또는 SEQ ID NO: 33 및 35 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 또는 임의의 하나로 구성된다. 선택적으로, 제1 가이드 RNA는 SEQ ID NO: 1423 내지 1484 및 1643 내지 1647 중 임의의 하나를 포함하는 DNA-표적화 분절을 포함하거나, 본질적으로 그것으로 구성되거나, 또는 그것으로 구성된다. 선택적으로, 제1 가이드 RNA는 SEQ ID NO: 1447 내지 1468 중 임의의 하나, SEQ ID NO: 1448 내지 1450, 1460, 및 1462 중 임의의 하나; 또는 SEQ ID NO: 1460 및 1462 중 임의의 하나를 포함하는 DNA-표적화 분절을 포함하거나, 본질적으로 그것으로 구성되거나, 또는 그것으로 구성된다. 선택적으로, 제1 가이드 RNA는 SEQ ID NO: 500 내지 561, 730 내지 791, 960 내지 1021, 1190 내지 1251, 720 내지 724, 950 내지 954, 1180 내지 1184, 및 1410 내지 1414 중 임의의 하나를 포함하거나, 본질적으로 그것으로 구성되거나, 그것으로 구성된다. 선택적으로, 제1 가이드 RNA는 SEQ ID NO: 524 내지 545, 754 내지 775, 984 내지 1005, 및 1214 내지 1235 중 임의의 하나, 또는 SEQ ID NO: 295 내지 297, 525 내지 527, 755 내지 757,

985 내지 987, 1215 내지 1217, 307, 309, 537, 539, 767, 769, 997, 999, 1227, 및 1229 중 임의의 하나, 또는 SEQ ID NO: 307, 309, 537, 539, 767, 769, 997, 999, 1227, 및 1229 중 임의의 하나를 포함하거나, 본질적으로 그것으로 구성되거나, 그것으로 구성된다. 선택적으로, 제1 가이드 RNA 표적 서열은 SEQ ID NO: 20 내지 41로부터 선택되거나, SEQ ID NO: 21 내지 23, 33, 및 35로부터 선택되거나, 또는 SEQ ID NO: 33 및 35로부터 선택된다. 선택적으로, Cas 단백질은 Cas9 단백질이다. 선택적으로, Cas 단백질은 뉴클레아제-활성 Cas 단백질이다. 선택적으로, Cas 단백질은 전사 활성화제 도메인에 융합된 뉴클레아제-비활성 Cas 단백질 또는 전사 억제자 도메인에 융합된 뉴클레아제-비활성 Cas 단백질이다.

[0043] 일부 이러한 방법은 추가로 세포의 게놈을, Cas 단백질과 복합체를 형성하고 *HSD17B13* 유전자 내의 제2 가이드 RNA 표적 서열을 표적으로 하는 제2 가이드 RNA와 접촉시키는 단계를 포함하며, 여기서 제2 가이드 RNA 표적 서열은 *HSD17B13* 유전자에 대한 종결 코돈을 포함하거나 종결 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있거나 또는 SEQ ID NO: 82 내지 225로부터 선택되며, 세포는 제1 가이드 RNA 표적 서열과 제2 가이드 RNA 표적 서열 사이에서 삭제를 포함하도록 변형된다. 선택적으로, 제2 가이드 RNA 표적 서열은 SEQ ID NO: 82 내지 225 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 임의의 하나로 구성된다. 선택적으로, 제2 가이드 RNA는 SEQ ID NO: 1485 내지 1628 중 임의의 하나를 포함하는 DNA-표적화 분절을 포함하거나, 본질적으로 그것으로 구성되거나, 또는 그것으로 구성된다. 선택적으로, 제2 가이드 RNA는 SEQ ID NO: 562 내지 705, 792 내지 935, 1022 내지 1165, 및 1252 내지 1395 중 임의의 하나를 포함하거나, 본질적으로 임의의 하나로 구성되거나, 임의의 하나로 구성된다.

[0044] 일부 이러한 방법은 세포에서 *HSD17B13* 유전자의 발현을 감소시키거나 세포에서 특정 *HSD17B13* 전사물 (예컨대, 전사물 A 또는 전사물 D)의 발현을 감소시키기 위한 것이다. 일부 이러한 방법은 세포에서 *HSD17B13* 유전자의 발현을 감소시키기 위한 것으로, 세포의 게놈을 SEQ ID NO: 4 (*HSD17B13* 전사물 A)의 엑손 7 내의 서열에 혼성화하고 *HSD17B13* 전사물 A의 발현을 감소시키는 안티센스 RNA, siRNA, 또는 shRNA와 접촉시키는 단계를 포함한다. 일부 이러한 방법은 세포에서 *HSD17B13* 유전자의 발현을 감소시키기 위한 것으로, 세포의 게놈을 본원에 개시된 *HSD17B13* 전사물 내의 서열에 혼성화하는 안티센스 RNA, siRNA, 또는 shRNA와 접촉시키는 단계를 포함한다. 일부 이러한 방법에서, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A) 내의 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 세포에서 *HSD17B13* 전사물 A의 발현을 감소시킬 수 있다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 존재하지 않는, SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하는 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)의 엑손 7 내의 서열 또는 엑손 6-엑손 7 경계에 걸쳐 있는 서열에 혼성화한다. 일부 이러한 방법에서, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 7 (*HSD17B13* 전사물 D) 내의 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 세포에서 *HSD17B13* 전사물 D의 발현을 감소시킬 수 있다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하지 않는, SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 존재하는 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)의 엑손 7 내의 서열 또는 엑손 6-엑손 7 경계에 걸쳐 있는 서열에 혼성화한다.

[0045] *HSD17B13* 유전자를 변형시키기 위한 또는 *HSD17B13* 유전자의 발현을 변경시키기 위한 상기 방법들 중 임의의 방법에서, 방법은 추가로 세포 안으로 발현 벡터를 도입하는 단계를 포함할 수 있고, 여기서 발현 벡터는 재조합 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는 재조합 *HSD17B13* 유전자를 포함한다. 선택적으로, 재조합 *HSD17B13* 유전자는 인간 유전자이다. 선택적으로, 재조합 *HSD17B13* 유전자는 유전자의 하나 이상의 비필수 분절이 상응하는 야생형 *HSD17B13* 유전자와 관련하여 삭제된 *HSD17B13* 미니유전자이다. 선택적으로, 삭제된 분절은 하나 이상의 인트론성 서열을 포함한다. 선택적으로, *HSD17B13* 미니유전자는 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 인트론 6에 상응하는 인트론을 포함한다.

[0046] *HSD17B13* 유전자를 변형시키기 위한 또는 *HSD17B13* 유전자의 발현을 변경시키기 위한 상기 방법들 중 임의의 방법에서, 방법은 추가로 발현 벡터를 세포 안으로 도입하는 방법을 포함할 수 있고, 여기서 발현 벡터는 SEQ ID NO: 15 (HSD17B13 동형 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하는 핵산을 포함한다. 선택적으로, HSD17B13 단백질을 암호화하는 핵산은 SEQ ID NO: 7과 최적으로 정렬될 때 SEQ ID NO: 7 (*HSD17B13* 전사물 D)과 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다.

[0047] *HSD17B13* 유전자를 변형시키기 위한 또는 *HSD17B13* 유전자의 발현을 변경시키기 위한 상기 방법들 중 임의의 방

법에서, 방법은 추가로 세포 안으로 HSD17B13 단백질 또는 이것의 단편을 도입시키는 단계를 포함할 수 있다. 선택적으로, HSD17B13 단백질 또는 이것의 단편은 SEQ ID NO: 15 (HSD17B13 동형 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다.

[0048] 일부 이러한 방법은 세포를 변형시키기 위한 것으로, 세포 안으로 발현 벡터를 도입시키는 단계를 포함하며, 여기서 발현 벡터는 재조합 HSD17B13 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는 재조합 *HSD17B13* 유전자를 포함한다. 선택적으로, 재조합 *HSD17B13* 유전자는 인간 유전자이다. 선택적으로, 재조합 *HSD17B13* 유전자는 유전자의 하나 이상의 비필수 분절이 상응하는 야생형 *HSD17B13* 유전자와 관련하여 삭제되어 있는 *HSD17B13* 미니유전자이다. 선택적으로, 삭제된 분절은 하나 이상의 인트론성 서열을 포함한다. 선택적으로, *HSD17B13* 미니유전자는 SEQ ID NO: 2와 최적으로 배열될 때 SEQ ID NO: 2의 인트론 6에 상응하는 인트론을 포함한다.

[0049] 일부 이러한 방법은 세포를 변형시키기 위한 것으로, 세포 안으로 발현 벡터를 도입시키는 단계를 포함하며, 여기서 발현 벡터는 SEQ ID NO: 15 (HSD17B13 동형 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하는 핵산을 포함한다. 선택적으로, HSD17B13 단백질을 암호화하는 핵산은 SEQ ID NO: 7과 최적으로 정렬될 때 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다.

[0050] 일부 이러한 방법은 세포를 변형시키기 위한 것으로, 세포 안으로 HSD17B13 단백질 또는 이것의 단편을 도입시키는 단계를 포함한다. 선택적으로, HSD17B13 단백질 또는 이것의 단편은 SEQ ID NO: 15 (HSD17B13 동형 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다.

[0051] 세포를 변형시키는, *HSD17B13* 유전자를 변형시키는, 또는 *HSD17B13* 유전자의 발현을 변경시키는 상기 방법들 중 임의의 방법에서, 세포는 인간 세포, 설치류 세포, 마우스 세포, 또는 래트 세포일 수 있다. 임의의 세포는 만능 세포 또는 분화된 세포일 수 있다. 임의의 세포는 간 세포일 수 있다. 세포를 변형시키는, *HSD17B13* 유전자를 변형시키는, 또는 *HSD17B13* 유전자의 발현을 변경시키는 상기 방법들 중 임의의 방법에서, 방법 또는 세포는 생체외에서 또는 생체내에서 있을 수 있다. 상기 방법들 중 임의의 방법에서 사용된 가이드 RNA는 서로에게 혼성화하는 별도의 crRNA 및 tracrRNA 분자를 포함하는 모듈식 가이드 RNA 또는 crRNA 부분이 tracrRNA 부분에 융합되어 있는 (예컨대, 링커에 의해) 단일-가이드 RNA일 수 있다.

[0052] 또 다른 측면으로, 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법이 제공된다. 또 다른 측면으로, 알코올성 또는 비알코올성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법이 제공된다. 그러한 대상체는, 예를 들어, *HSD17B13* rs72613567 변종의 보균자가 아닌 대상체 또는 *HSD17B13* rs72613567 변종의 동형접합성 보균자가 아닌 대상체일 수 있다. 일부 이러한 방법은 *HSD17B13* rs72613567 변종의 보균자가 아닌 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법을 포함하며, 방법은 대상체 안으로 (a) Cas 단백질 또는 Cas 단백질을 암호화하는 핵산; (b) 가이드 RNA 또는 가이드 RNA를 암호화하는 핵산, 여기서 가이드 RNA는 Cas 단백질과 복합체를 형성하며 *HSD17B13* 유전자 내의 가이드 RNA 표적 서열을 표적으로 하고, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치를 포함하거나 근접하며; 및 (c) SEQ ID NO: 2의 위치 12666에 상응하는 위치의 표적 서열 5'에 혼성화하는 5' 상동성 팔, SEQ ID NO: 2의 위치 12666에 상응하는 위치의 표적 서열 3'에 혼성화하는 3' 상동성 팔, 및 5' 상동성 팔 및 3' 상동성 팔에 의해 플랭킹된 티민을 포함하는 핵산 삽입물을 포함하는 외인성 도너 서열을 도입하는 단계를 포함하고, 여기서 Cas 단백질은 대상체의 간 세포에서 *HSD17B13* 유전자를 절단하며 외인성 도너 서열은 간 세포에서 *HSD17B13* 유전자와 재조합하고, 외인성 도너 서열과 *HSD17B13* 유전자와의 재조합시에, 티민은 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된다.

[0053] 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 226 내지 239로부터 선택되거나, 또는 가이드 RNA 표적 서열은 SEQ ID NO: 230 및 231로부터 선택된다. 선택적으로, 가이드 RNA 표적 서열은 SEQ ID NO: 226 내지 239 및 264 내지 268로부터 선택된다. 선택적으로, 가이드 RNA 표적 서열은 HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 엑손 6 및/또는 인트론 6에 상응하는 영역 내에 있다. 선택적으로, 가이드 RNA 표적 서열은 HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 엑손 6 및/또는 인트론 6 및/또는 엑손 7에 상응하는 영역 내에 있다. 선택적으로, 가이드 RNA 표적 서열은 HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 약 1000, 500, 400, 300, 200, 100, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5개 뉴클레오타이드 내에 있다. 선택적으로, 가이드 RNA 표적 서열은

HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치를 포함한다.

[0054] 선택적으로, 외인성 도너 서열은 길이가 약 50개 뉴클레오타이드 내지 약 1 kb이다. 선택적으로, 외인성 도너 서열은 길이가 약 80개 뉴클레오타이드 내지 약 200개 뉴클레오타이드이다. 선택적으로, 외인성 도너 서열은 단일-가닥 올리고데옥시뉴클레오타이드이다.

[0055] 일부 이러한 방법은 *HSD17B13* rs72613567 변종의 보균자가 아닌 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법을 포함하며, 방법은 대상체 안으로 (a) Cas 단백질 또는 Cas 단백질을 암호화하는 핵산; (b) 제1 가이드 RNA 또는 제1 가이드 RNA를 암호화하는 핵산, 여기서 제1 가이드 RNA는 Cas 단백질과 복합체를 형성하며 *HSD17B13* 유전자 내의 제1 가이드 RNA 표적 서열을 표적으로 하고, 제1 가이드 RNA 표적 서열은 *HSD17B13* 유전자에 대한 개시 코돈을 포함하거나 또는 개시 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있거나 또는 SEQ ID NO: 20 내지 81 및 259 내지 263으로부터 선택되며; 및 (c) 재조합 HSD17B13 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는 재조합 *HSD17B13* 유전자를 포함하는 발현 벡터를 도입하는 단계를 포함하고, 여기서 Cas 단백질은 대상체의 간세포에서 *HSD17B13* 유전자를 절단하거나 *HSD17B13* 유전자의 발현을 변경시키고 발현 벡터는 대상체의 간세포에서 재조합 *HSD17B13* 유전자를 발현한다. 일부 이러한 방법은 *HSD17B13* rs72613567 변종의 보균자가 아닌 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법을 포함하며, 방법은 대상체 안으로 (a) Cas 단백질 또는 Cas 단백질을 암호화하는 핵산; (b) 제1 가이드 RNA 또는 제1 가이드 RNA를 암호화하는 핵산, 여기서 제1 가이드 RNA는 Cas 단백질과 복합체를 형성하며 *HSD17B13* 유전자 내의 제1 가이드 RNA 표적 서열을 표적으로 하고, 제1 가이드 RNA 표적 서열은 *HSD17B13* 유전자에 대한 개시 코돈을 포함하거나 또는 개시 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있거나 또는 SEQ ID NO: 20 내지 81 및 259 내지 263으로부터 선택되며; 및 선택적으로 (c) 재조합 HSD17B13 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는 재조합 *HSD17B13* 유전자를 포함하는 발현 벡터를 도입하는 단계를 포함하고, 여기서 Cas 단백질은 대상체의 간세포에서 *HSD17B13* 유전자를 절단하거나 *HSD17B13* 유전자의 발현을 변경시키고 발현 벡터는 대상체의 간세포에서 재조합 *HSD17B13* 유전자를 발현한다.

[0056] 선택적으로, 제1 가이드 RNA 표적 서열은 SEQ ID NO: 20 내지 41로부터 선택되거나, SEQ ID NO: 21 내지 23, 33, 및 35로부터 선택되거나, 또는 SEQ ID NO: 33 및 35로부터 선택된다. 선택적으로, Cas 단백질은 뉴클레아제-활성 Cas 단백질이다. 선택적으로, Cas 단백질은 전사 억제자 도메인에 융합된 뉴클레아제-비활성 Cas 단백질이다.

[0057] 이러한 방버은 추가로 제2 가이드 RNA를 대상체 안으로 도입시키는 단계를 포함할 수 있고, 여기서 제2 가이드 RNA는 Cas 단백질과 복합체를 형성하며 *HSD17B13* 유전자 내의 제2 가이드 RNA 표적 서열을 표적으로 하고, 제2 가이드 RNA 표적 서열은 *HSD17B13* 유전자에 대한 종결 코돈을 포함하거나 또는 종결 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있거나 또는 SEQ ID NO: 82 내지 225로부터 선택되며, Cas 단백질은 간세포에서 제1 가이드 RNA 표적 서열 및 제2 가이드 RNA 표적 서열 둘 다의 내부에 있는 *HSD17B13* 유전자를 절단하고, 간세포는 제1 가이드 RNA 표적 서열 및 제2 가이드 RNA 표적 서열 사이의 삭제를 포함하도록 변형된다.

[0058] 선택적으로, 재조합 *HSD17B13* 유전자는 유전자의 하나 이상의 비필수 분절이 상응하는 야생형 *HSD17B13* 유전자와 관련하여 삭제되어 있는 *HSD17B13* 미니유전자이다. 선택적으로, 삭제된 분절은 하나 이상의 인트론성 서열을 포함한다. 선택적으로, *HSD17B13* 미니유전자는 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 인트론 6에 상응하는 인트론을 포함한다.

[0059] 상기 치료적 또는 예방적 방법들 중 임의의 방법에서, Cas 단백질은 Cas9 단백질일 수 있다. 상기 치료적 또는 예방적 방법들 중 임의의 방법에서, 대상체는 인간일 수 있다. 상기 치료적 또는 예방적 방법들 중 임의의 방법에서, 만성 간 질환은 지방간 질환, 비알코올성 지방간 질환 (NAFLD), 알코올성 간 지방간 질환, 간경변, 또는 간세포 암종일 수 있다. 유사하게, 상기 방법들 중 임의의 방법에서, 치료 또는 예방 방법은 알코올성 간 질환 또는 비알코올성 간 질환인 간 질환에 대한 것일 수 있다.

[0060] 일부 이러한 방법은 *HSD17B13* rs72613567 변종의 보균자가 아닌 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법을 포함하며, 방법은 SEQ ID NO: 4 (*HSD17B13* 전사물 A) 내의 엑손 7 내의 서열 또는 엑손

6-엑손 7 경계에 걸쳐 있는 서열에 혼성화하고 대상체의 간세포에서 *HSD17B13* 전사물 A의 발현을 감소시키는 안티센스 RNA, siRNA, 또는 shRNA를 대상체 안으로 도입하는 단계를 포함한다. 일부 이러한 방법은 *HSD17B13* rs72613567 변종의 보균자가 아닌 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법을 포함하며, 방법은 본원에 개시된 *HSD17B13* 전사물 내의 서열에 혼성화하는 안티센스 RNA, siRNA, 또는 shRNA를 대상체 안으로 도입하는 단계를 포함한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A) 내의 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 세포에서 *HSD17B13* 전사물 A의 발현을 감소시킬 수 있다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 존재하지 않는, SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하는 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A) 내의 엑손 7 내의 서열 또는 엑손 6-엑손 7 경계에 걸쳐 있는 서열에 혼성화한다.

[0061] 선택적으로, 이러한 방법은 추가로 발현 벡터를 대상체 안으로 도입시키는 단계를 포함하며, 여기서 발현 벡터는 재조합 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는 재조합 *HSD17B13* 유전자를 포함하고, 발현 벡터는 대상체의 간세포에서 재조합 *HSD17B13* 유전자를 발현한다.

[0062] 선택적으로, 이러한 방법은 추가로 발현 벡터를 대상체 안으로 도입시키는 단계를 포함하며, 여기서 발현 벡터는 SEQ ID NO: 15 (HSD17B13 동형 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하는 핵산을 포함하고, 발현 벡터는 대상체의 간세포에서 HSD17B13 단백질을 암호화하는 핵산을 발현한다. 선택적으로, HSD17B13 단백질을 암호화하는 핵산은 SEQ ID NO: 7과 최적으로 정렬될 때 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다.

[0063] 선택적으로, 이러한 방법은 추가로 메신저 RNA를 대상체 안으로 도입시키는 단계를 포함하며, 여기서 메신저 RNA는 SEQ ID NO: 15 (HSD17B13 동형 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하고, mRNA는 대상체의 간세포에서 HSD17B13 단백질을 발현시킨다. 선택적으로, 메신저 RNA로부터 역전사된 상보하는 DNA는 SEQ ID NO: 7과 최적으로 정렬될 때 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다.

[0064] 선택적으로, 이러한 방법은 추가로 HSD17B13 단백질 또는 이것의 단편을 대상체 안으로 도입시키는 단계를 포함한다. 선택적으로, HSD17B13 단백질 또는 이것의 단편은 SEQ ID NO: 15 (HSD17B13 동형 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다.

[0065] 일부 이러한 방법은 발현 벡터를 대상체 안으로 도입시키는 단계를 포함하는, *HSD17B13* rs72613567 변종의 보균자가 아닌 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법을 포함하며, 여기서 발현 벡터는 재조합 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는 재조합 *HSD17B13* 유전자를 포함하고, 발현 벡터는 대상체의 간세포에서 재조합 *HSD17B13* 유전자를 발현한다.

[0066] 상기 방법들 중 임의의 방법에서, 재조합 *HSD17B13* 유전자는 인간 유전자일 수 있다. 상기 방법들 중 임의의 방법에서, 재조합 *HSD17B13* 유전자는 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일할 수 있다. 상기 방법들 중 임의의 방법에서, 재조합 *HSD17B13* 유전자는 유전자의 하나 이상의 비필수 분절이 상응하는 야생형 *HSD17B13* 유전자와 관련하여 삭제되어 있는 *HSD17B13* 미니유전자일 수 있다. 선택적으로, 삭제된 분절은 하나 이상의 인트론성 서열을 포함한다. 선택적으로, *HSD17B13* 미니유전자는 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 인트론 6에 상응하는 인트론을 포함한다.

[0067] 일부 이러한 방법은 발현 벡터를 대상체 안으로 도입시키는 단계를 포함하는, *HSD17B13* rs72613567 변종의 보균자가 아닌 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법을 포함하며, 여기서 발현 벡터는 SEQ ID NO: 15 (HSD17B13 동형 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하는 핵산을 포함하고, 발현 벡터는 대상체의 간세포에서 HSD17B13 단백질을 암호화하는 핵산을 발현시킨다. 선택적으로, HSD17B13 단백질을 암호화하는 핵산은 SEQ ID NO: 7과 최적으로 정렬될 때 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다.

[0068] 일부 이러한 방법은 메신저 RNA를 대상체 안으로 도입시키는 단계를 포함하는, *HSD17B13* rs72613567 변종의 보균자가 아닌 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법을 포함하며, 여기서 메신저 RNA는 SEQ ID NO: 15 (HSD17B13 동형 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하고, 메신저 RNA는 대상체의 간세포에서 HSD17B13 단백질을 발현시킨다. 선택적으로, 메신저 RNA로부터 역전사된 상보하는 DNA는 SEQ ID NO: 7과 최적으로 정렬될 때 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다.

[0069] 일부 이러한 방법은 HSD17B13 단백질 또는 이것의 단편을 대상체 안으로 도입시키는 단계를 포함하는, *HSD17B13* rs72613567 변종의 보균자가 아닌 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법을 포함한다. 선택적으로, HSD17B13 단백질 또는 이것의 단편은 SEQ ID NO: 15 (HSD17B13 동형 D)에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다.

[0070] 상기 방법들 중 임의의 방법에서, 대상체는 인간일 수 있다. 상기 방법들 중 임의의 방법에서, 만성 간 질환은 비알코올성 지방간 질환 (NAFLD), 알코올성 간 지방간 질환, 간경변, 또는 간세포 암종일 수 있다. 유사하게, 상기 방법들 중 임의의 방법에서, 치료 또는 예방 방법은 알코올성 간 질환 또는 비알코올성 간 질환인 간 질환에 대한 것일 수 있다. 상기 방법들 중 임의의 방법에서, 대상체 안으로의 도입은 유체역학적 전달, 바이러스-매개 전달, 지질-나노입자-매개 전달, 또는 정맥내 주입을 포함할 수 있다.

도면의 간단한 설명

[0071] **도 1A** 및 **1B**는 GHS 개발 집단에서 중앙값의 알라닌 아미노전이효소 (ALT; **도 1A**) 및 아스파테이트 아미노전이효소 (AST; 도 **1B**) 수준과 관련하여 단일 뉴클레오타이드 변종의 맨하탄 도표 (좌측) 및 분위수 대 분위수 도표 (우측)를 도시한다. **도 1A**는 16개 유전자 중 $P < 1.0 \times 10^{-7}$에서 ALT 수준 (N=41,414)과 유의하게 관련된 31개의 변종이 있었음을 보여준다. **도 1B**는 10개 유전자 중 $P < 1.0 \times 10^{-7}$에서 AST 수준 (N=40,753)과 유의하게 관련된 12개의 변종이 있었음을 보여준다. 모든 유의한 관련은 **표 2**에 나타낸다. 3개의 별도의 유럽-혈통 집단의 반복 메타-분석에서 ALT 또는 AST와 유의하게 관련된 채로 유지된, *HSD17B13*을 포함한, 9개의 유전자 (유전자 명칭에 의해 표시됨) 중 13개의 변종이 있었다 (**표 3**). 연관성 테스트는 엑솜-전체 분위수 대 분위수 도표 및 게놈 대조군 람다 값에 의해 나타난 바와 같이, 잘 보정되었다 (**도 1A** 및 **도 1B**).

도 2A 및 **2B**는 *HSD17B13* rs72613567:TA가 개발 집단에서 알코올성 및 비알코올성 간 질환 표현형의 위험의 감소와 관련되며 (**도 2A**), 베리아트릭 수술 집단에서 단순 지방증으로부터 지방간염 및 섬유증으로의 진행의 위험의 감소과 관련된 (**도 2B**) 것을 보여준다. 오즈비(odds ratio)는 로지스틱 회귀를 사용하여 계산되었고, 이때 연령, 연령^2, 성별, BMI, 및 혈통의 주요 구성요소들에 대해 조정되었다. 이형접합성 (Het OR) 및 동형접합성 (Hom OR) 보균자에 대한 유전자형 오즈비가 또한 도시된다. **도 2A**의 GHS 개발 집단에서, 변종 *HSD17B13*은 대립유전자 용량-의존 방식으로 비알코올성 및 알코올성 간 질환, 간경변, 및 간세포 암종의 유의하게 감소된 위험과 관련되었다. **도 2B**의 GHS 베리아트릭 수술 집단에서, *HSD17B13* rs72613567은 이형접합성 및 동형접합성 TA 보균자에서, 각각, 비알코올성 지방간염 (NASH)의 13% 및 52% 더 낮은 오즈 및 섬유증의 13% 및 61% 더 낮은 오즈와 관련이 있었다.

도 3A 내지 3D는 *HSD17B13* rs72613567 스플라이스 변종의 동형접합성 기준 (T/T), 이형접합성 (T/TA), 및 동형접합성 대체물 (TA/TA) 보균자에서 4개의 *HSD17B13* 전사물 (A 내지 D)의 발현을 도시한다. 각 전사물은 상응하는 유전자 모델로 예시된다. 유전자 모델에서 코딩 영역들은 줄무늬 박스로 표시되고 미번역 영역은 검은색 박스로 표시된다. **도 3A**는 전사물 A의 대표도 및 전사물 A에 대한 발현 데이터를 나타낸다. **도 3B**는 전사물 B의 대표도 및 전사물 B에 대한 발현 데이터를 나타낸다. 전사물 B에서, 엑손 2는 빠져 있다. **도 3C**는 전사물 C의 대표도 및 전사물 C에 대한 발현 데이터를 나타낸다. 전사물 C에서, 엑손 6은 빠져 있다. **도 3D**는 전사물 D의 대표도 및 전사물 D에 대한 발현 데이터를 나타낸다. 전사물 D에서 별표는 단백질의 미성숙 절단으로 이어지는, 엑손 6의 3' 단부에서의 rs72613567로부터의 G의 삽입을 나타낸다. 전사물 D는 *HSD17B13* 스플라이스 변종의 동형접합성 보균자에서 우세한 전사물이 된다. 유전자 발현은 FPKM 단위 (백만개의 지도화된 판독 당 전사물의 킬로염기당 단편)로 표현된다. **도 3B** 및 **도 3C**의 삽도는 확대된 관점을 보여준다.

도 4는 인간 간의 RNA-Seq 연구가 6개의 신규한 *HSD17B13* 전사물 (전사물 C 내지 H)을 포함한, 8개의 *HSD17B13* 전사물을 드러냈음을 도시한다. 전사물의 발현은 FPKM 단위 (백만개의 지도화된 판독 당 전사물의 킬로염기당

단편)로 표현된다. 전사물의 구조는 도면의 우측에 제공된다.

도 5A 및 **5B**는 ALT 및 AST에 대한 GHS 개발 집단에서, 각각, *HSD17B13*의 유전자좌-줌 도표 (*HSD17B13* 주변 영역의 위치적 연관성 도표)를 도시한다. 영역 전체에서 유의미한 재조합은 관찰되지 않았다. 다이아몬드형은 스플라이스 변종 rs72613567을 나타낸다. 각각의 원형은 단일 뉴클레오타이드 변종을 나타내고 원형의 색은 변종과 rs72613567 사이의 연쇄 불균형 (DiscovEHR 집단에서 계산된 r^2)을 나타낸다. 선은 HapMap에서 추정된 재조합 속도를 나타낸다. 하부 패널은 유전자좌의 각 유전자의 상대적인 위치 및 전사된 가닥을 보여준다. ALT 또는 AST와 인접한 유전자 *HSD17B11*의 코딩 또는 스플라이스 영역 변종들 사이에 유의한 연관성은 없었다 (ALT 및 AST에 대한 가장 유의한 P-값은 각각 $1.4x10^{-1}$ 및 $4.3x10^{-2}$임).

도 6A 내지 6D는 *HSD17B13* 스플라이스 변종의 동형접합성 기준 (T/T), 이형접합성 (T/TA), 및 동형접합성 대체물 (TA/TA) 보균자에서 4개의 추가의 신규한 *HSD17B13* 전사물 (E 내지 H)에 대한 mRNA 발현을 도시한다. 각 전사물은 상응하는 유전자 모델로 예시된다. 유전자 모델에서 코딩 영역들은 줄무늬 박스로 표시되고 미번역 영역은 검은색 박스로 표시된다. **도 6A** 및 **6D**는 전사물 E 및 H가 엑손 3과 4 사이에 추가적인 엑손을 함유하는 것을 보여준다. **도 6B**는 전사물 F가 엑손 5으로부터 인트론 6에 이르기까지 철저하게 판독한 것을 보여준다. **도 6C**는 전사물 G에서, 엑손 2가 빠져 있는 것을 보여준다. 전사물 G 및 H (각각 **도 6C** 및 **6D**)에서 별표는 단백질의 미성숙 절단으로 이어지는, 엑손 6의 3' 단부에서의 rs72613567로부터의 G의 삽입을 나타낸다. 전사물은 박스 도표에서 나타나는 바, *HSD17B13* 유전자형에 따라 차등적으로 발현된다. mRNA 발현은 FPKM 단위 (백만개의 지도화된 판독 당 전사물의 킬로염기당 단편)로 표현된다.

도 7A 및 **7B**는 HSD17B13 단백질 동형 A 내지 H의 단백질 서열 정렬을 도시한다.

도 8은 *HSD17B13* rs72613567:TA가 알코올성 및 비알코올성 간 질환 표현형의 위험의 감소과 관련이 있음을 도시한다. 구체적으로, **도 8**은 댈러스 간 연구에서, *HSD17B13* rs72613567이 대립유전자 용량-의존 방식으로 임의의 간 질환의 하부 오즈와 관련되었음을 보여준다. 유사한 대립유전자 용량-의존 효과가 간 질환 아형 전체에서 관찰되었다. 오즈비는 로지스틱 회귀를 사용하여 계산되었고, 연령, 연령^2, 성별, BMI, 및 자체-보고된 민족성에 대해 조정되었다.

도 9는 *HSD17B13* rs72613567이 단순 지방증으로부터 지방간염 및 섬유증으로의 진행의 위험의 감소과 관련된 것을 보여준다. 구체적으로, 도 9는 조직병리학적으로 특성화된 간 질환의 유병률이 GHS 베리아트릭 수술 집단으로부터의 간 생검을 얻은 2,391명의 개인에서 *HSD17B13* rs72613567 유전자형에 따르는 것을 보여준다. 정상 간의 유병률은 유전자형에 의해 상이한 것으로 나타나지 않았지만 (P = 0.5, 비율로 트렌드에 대한 카이 제곱 테스트에 의해), NASH의 유병률은 감소되었고 (P = $1.6x10^{-4}$) 단순 지방증의 유병률은 각 TA 대립유전자로 증가되었다 (P = $1.1x10^{-3}$).

도 10A 내지 10E는 신규한 *HSD17B13* 전사물의 발현, 세포 이하 위치, 및 효소적 활성을 도시한다. **도 10A**는 *HSD17B13* 전사물 A 및 D를 과다발현하는 HepG2 세포로부터의 웨스턴 블롯을 보여주고 *HSD17B13* 전사물 D가 *HSD17B13* 전사물 A와 비교하여 저분자량을 가진 절단된 단백질로 번역되었음을 보여준다. **도 10B**는 신선한 냉동 인간 간 및 HEK293 세포 샘플로부터의 HSD17B13 웨스턴 블롯을 도시한다. 인간 간 샘플은 *HSD17B13* rs72613567 스플라이스 변종의 동형접합성 기준 (T/T), 이형접합성 (T/TA), 및 동형접합성 대체물 (TA/TA) 보균자로부터 유래된다. 세포 샘플은 비-태그된 *HSD17B13* 전사물 A 및 D를 과다발현하는 HEK293 세포로부터 유래된다. *HSD17B13* 전사물 D는 HSD17B13 IsoA보다 저 분자량의 절단된 단백질 IsoD로 번역되었다. **도 10C**는 HSD17B13 IsoD 단백질 수준이 인간 간 (좌측) 및 세포 (우측) 샘플로부터의 IsoA 단백질 수준보다 더 낮았음을 보여준다. 액틴으로 표준화된 단백질 수준은 막대 칼럼에 나타낸다; ** P<0.001, *P<0.05. **도 10D**는 17-베타 에스트라디이올 (에스트라디이올), 류코트라이엔 B4 (LTB4), 및 13-하이드록시옥타데카다이엔산 (13(S)-HODE)에 대한 HSD17B13 동형 A 및 D의 효소적 활성을 보여준다. HSD17B13 동형 D는 동형 A에 대한 상응하는 값의 10% 미만의 효소적 활성을 나타낸다. **도 10E**는 HEK 293 세포에서 과다발현되었을 때 배양 배지에서 측정되는 바 에스트라디이올 (기질)의 에스트론 (생성물)으로의 많은 전환을 나타내지 못한 한편, 과다발현된 HSD17B13 동형 A는 왕성한 전환을 나타낸 것을 보여준다.

도 11A 내지 11C는 HSD17B13 동형 D 단백질이 저분자량을 가지며 HEK 293 세포에서 과다발현될 때 불안정한 것을 도시한다. **도 11A**는 *HSD17B13* 전사물 A (IsoA) 및 D (IsoD)를 과다발현하는 HEK 293 세포로부터의 *HSD17B13*의 RT-PCR을 도시하며, *HSD17B13* IsoD RNA 수준이 IsoA RNA 수준보다 더 높았음을 나타낸다. **도 11B**는 동일한

세포주로부터의 웨스턴 블롯을 도시하며, *HSD17B13* 전사물 D가 *HSD17B13* 전사물 A와 비교하여 저분자량을 가지는 절단된 단백질로 번역되었음을 나타낸다. **도 11C**는 비록 RNA 수준이 더 높긴 할지라도 HSD17B13 IsoD 단백질 수준이 IsoA 단백질 수준보다 낮았음을 보여준다. HSD17B13 단백질 수준은 액틴에 대해 표준화되었다; *P<0.05.

도 12는 HepG2 안정적인 세포주로부터 유래된 분리된 지질 방울 (lipid droplet, LD)에 대한 HSD17B13 동형 A 및 동형 D의 유사한 국지화 패턴을 도시한다. ADRP 및 TIP47이 지질 방울 마커로서 사용되었다. AMP1, 칼레티쿨린(calreticulin), 및 COX IV가 각각 리소좀, 소포체, 및 미토콘드리아 구획에 대한 마커로서 사용되었다. GAPDH는 시토졸 마커로서 포함되었고, 액틴은 세포골격의 마커로서 사용되었다. 이 실험은 HrpG2 세포에서 2회 반복되었고, 이때 상기는 두 번의 작동의 대표적인 것이다. PNS = 핵 뒤의 분획; TM = 총 막.

도 13A 내지 13D는 *HSD17B13* 전사물 A 또는 D를 과다발현하는 HepG2 세포에서 올레산이 트라이글리세리드 함량을 증가시켰음을 도시한다. **도 13A**는 증가하는 농도의 올레산으로의 처리가 대조군 (GFP 과다발현 세포) 및 *HSD17B13* 전사물 A 및 D 세포주에서 트라이글리세리드 (TG) 함량을 유사한 정도로 증가시켰음을 보여준다. **도 13B**는 세포주에서 유사한 *HSD17B13* 전사물 A 및 D RNA 수준을 보여준다. RNA 수준은 백만개의 지도화된 판독당 전사물의 킬로염기당 판독 (RPKM)으로 나타낸다. **도 13C**는 *HSD17B13* 전사물 A 및 D를 과다발현하는 HepG2 세포로부터의 웨스턴 블롯을 보여준다. *HSD17B13* 전사물 D는 *HSD17B13* 전사물 A와 비교하여 저분자량을 가진 절단된 단백질로 번역되었다. **도 13D**는 HSD17B13 IsoD 단백질 수준이 IsoA 단백질 수준보다 낮은 것을 보여준다. 단백질 수준은 액틴에 대해 표준화되었다; **P<0.01.

도 14는 정제된 재조합 HSD17B13 단백질을 사용하는 에스트라다이올에 대한 K_m 및 V_{max} 값을 도시한다. K_m 및 V_{max} 측정에 대해, 0.2 μM 내지 200 μM의 17β-에스트라다이올의 용량 범위로 및 5분 내지 180분의 시점에서, 500 μM NAD^+ 및 228 nM HSD17B13으로 수행되었다. 그런 후 V_{max} 및 K_m은 미하엘리스-멘텐 모델(Michaelis-Menten model) 및 Prism 소프트웨어 (GraphPad Software, USA)를 사용하여 측정되었다.

도 15는 하이브리드 야생형 마우스 (75% C57BL/6NTac 25% 129S6/SvEvTac)로부터 분리된 일차 간세포에서 차세대 서열분석 (NGS)에 의해 측정된 바 마우스 *Hsd17b13* 유전자좌에서의 퍼센트 게놈 편집 (용해된 세포의 풀로부터 PCR 반응에서의 서열 판독의 총 수를 능가하는 Cas9-유도된 DNA 끊김(break)의 어느 한 쪽에서의 윈도우 20 염기쌍 내에서 관찰된 삽입 또는 삭제의 총 수)을 도시한다. 테스트된 샘플은 Cas9 및 마우스 *Hsd17b13* 유전자좌를 표적화하도록 설계된 가이드 RNA를 함유한 리보핵단백질 복합체로 처리된 간세포를 포함하였다.

도 16은 마우스 *Hsd17b13*을 표적화하도록 설계된 sgRNA 발현 카세트를 함유한 AAV8의 Cas9-준비된 마우스 안으로의 주입 후 3주 후에 마우스 간으로부터 분리된 샘플에서 차세대 서열분석 (NGS)에 의해 측정된 바 마우스 *Hsd17b13* 유전자좌에서의 퍼센트 게놈 편집 (용해된 세포의 풀로부터 PCR 반응에서의 서열 판독의 총 수를 능가하여 관찰된 삽입 또는 삭제의 총 수)을 도시한다. sgRNA 발현 카세트 전부를 함유하는 AAV8이 주입된, 어떠한 Cas9도 발현하지 않는 야생형 마우스가 음성 대조군으로서 사용되었다.

도 17A 및 **17B**는 마우스 *Hsd17b13*을 표적화하도록 설계된 가이드 RNA 발현 카세트를 운반하는 AAV8로 처리된 Cas9-준비된 마우스로부터의 간 샘플에서 RT-PCR에 의해 측정된 바, 마우스 *Hsd17b13* 및 비-표적 HSD 패밀리 구성원에 대한 각각의 상대적인 mRNA 발현을 도시한다. 모든 가이드 RNA에 대한 가이드 RNA 발현 카세트를 함유하는 AAV8이 주입된, 어떠한 Cas9도 발현하지 않는 야생형 마우스가 음성 대조군으로서 사용되었다.

발명을 실시하기 위한 구체적인 내용

[0072] **정의**

[0073] 본원에서 상호교환적으로 사용되는 용어 "단백질", "폴리펩타이드", 및 "펩타이드"는 암호화된 및 암호화되지 않은 아미노산 및 화학적으로 또는 생화학적으로 변형된 또는 유도체화된 아미노산을 포함하는, 임의의 길이의 아미노산들의 폴리머 형태를 포함한다. 용어는 또한 변형된 폴리머, 예컨대 변형된 펩타이드 골격을 가지는 폴리펩타이드를 포함한다.

[0074] 단백질은 "N-말단" 및 "C-말단"을 가진다고 말할 수 있다. 용어 "N-말단"은 단백질 또는 폴리펩타이드의 시작과 관련되며, 유리 아민기 (-NH2)를 가진 아미노산에 의해 종결된다. 용어 "C-말단"은 아미노산 사슬 (단백질 또는 폴리펩타이드)의 단부와 관련되며, 유리 카르복실기 (-COOH)에 의해 종결된다.

[0075] 본원에서 상호교환적으로 사용되는 용어 "핵산" 및 "폴리뉴클레오타이드"는 임의의 길이의 뉴클레오타이드의 폴리머 형태를 포함하며, 리보뉴클레오타이드, 데옥시리보뉴클레오타이드, 또는 이것들의 유사체 또는 변형된 버

전을 포함한다. 이것들은 단일-, 이중-, 및 다중-가닥의 DNA 또는 RNA, 게놈 DNA, cDNA, DNA-RNA 하이브리드, 및 퓨린 염기, 피리미딘 염기, 또는 다른 천연의, 화학적으로 변형된, 생화학적으로 변형된, 비-천연의, 또는 유도체화된 뉴클레오타이드 염기를 포함하는 폴리머를 포함한다.

[0076] 핵산은 한 모노뉴클레오타이드 펜토오스 고리의 5' 포스페이트가 그것의 인접한 모노뉴클레오타이드의 3' 산소에 포스포다이에스테르 결합을 통해 한 방향으로 부착되는 그런 방식으로 모노뉴클레오타이드들이 반응하여 올리고뉴클레오타이드를 만들기 때문에 "5' 단부" 및 "3' 단부"를 가진다고 말할 수 있다. 올리고뉴클레오타이드의 단부는 만약 그것의 5' 포스페이트가 모노뉴클레오타이드 펜토오스 고리의 3' 산소에 연결되지 않는다면 "5' 단부"로서 언급된다. 올리고뉴클레오타이드의 단부는 만약 그것의 3' 산소가 또 다른 모노뉴클레오타이드 펜토오스 고리의 5' 포스페이트에 연결되지 않는다면 "3' 단부"로서 언급된다. 핵산 서열은, 더 큰 올리고뉴클레오타이드의 내부에 있어도, 또한 5' 및 3' 단부를 가진다고 말할 수 있다. 선형이든 원형이든 DNA 분자에서, 구별되는 요소들은 "하류" 또는 3' 요소의 "상류" 또는 5'인 것으로 언급된다.

[0077] 용어 "야생형"은 정상 (돌연변이, 병에 걸린, 변경된, 등등과 대비됨) 상태 또는 맥락에서 발견되는 대로의 구조 및/또는 활성을 가지는 실체를 포함한다. 야생형 유전자 및 폴리펩타이드는 종종 다수의 상이한 형태 (예컨대, 대립유전자)로 존재한다.

[0078] 단백질 및 핵산과 관련하여 용어 "분리된"은 제자리에서 정상적으로 존재할 수 있는 다른 박테리아, 바이러스, 또는 세포 구성요소들과 관련하여 상대적으로 정제된 단백질 및 핵산을 포함하고, 단백질 및 폴리뉴클레오타이드의 실질적으로 순수한 조제물까지 포함한다. 용어 "분리된"은 또한 본질적으로 발생하는 대응물을 갖지 않는, 화학적으로 합성된 및 그러므로 실질적으로는 다른 단백질 또는 핵산에 의해 오염되지 않은, 또는 그것들이 본질적으로 수반할 수 있는 대부분의 다른 세포 구성요소들 (예컨대, 다른 세포 단백질, 폴리뉴클레오타이드, 또는 세포 구성요소들)로부터 분리되었거나 정제된 단백질 및 핵산을 포함한다.

[0079] "외인성" 분자 또는 서열은 세포에 그 형태로 정상적으로 존재하지 않는 분자 또는 서열을 포함한다. 정상적인 서열은 세포의 특정 발달 단계 및 환경 조건과 관련된 존재를 포함한다. 외인성 분자 또는 서열은, 예를 들어, 세포 내의 상응하는 내인성 서열의 돌연변이된 버전을 포함할 수 있거나 세포내의 내인성 서열에 상응하지만 상이한 형태로 있는 (즉, 염색체 내에 있지 않는) 서열을 포함할 수 있다. 대조적으로, 내인성 분자 또는 서열은 특정 환경 조건 하에서 특정 발달 단계에서 특정 세포에서 그 형태로 정상적으로 존재하는 분자 또는 서열을 포함한다.

[0080] 용어 "이종성"은 핵산 또는 단백질의 맥락에서 사용될 때 핵산 또는 단백질이 본질적으로 함께 발생하지 않는 적어도 두 부분을 포함하는 것을 가리킨다. 마찬가지로, 용어 "이종성"은 단백질을 암호화하는 핵산에 작동 가능하게 연결된 프로모터의 맥락에서 사용될 때 프로모터 및 단백질을 암호화하는 핵산이 본질적으로 함께 발생하지 않는 것 (즉, 본질적으로 작동 가능하게 연결되지 않는 것)을 가리킨다. 예를 들어, 용어 "이종성"은, 핵산의 일부 또는 단백질의 일부와 관련하여 사용될 때, 핵산 또는 단백질이 본질적으로 서로에 대해 동일한 관계에서 발견되지 않는 (예컨대, 함께 연합되지 않는) 둘 이상의 하위서열을 포함하는 것을 가리킨다. 한 예로서, 핵산 벡터의 "이종성" 영역은 본질적으로 다른 분자와 회합된 상태로 발견되지 않는 또 다른 핵산 분자 내의 또는 그것에 부착된 핵산의 분절이다. 예를 들어, 핵산 벡터의 이종성 영역은 본질적으로 코딩 서열과 함께 발견되지 않은 서열들에 의해 플랭킹된 코딩 서열을 포함할 수 있을 것이다. 마찬가지로, 단백질의 "이종성" 영역은 본질적으로 다른 펩타이드 분자와 회합된 상태로 발견되지 않는 또 다른 펩타이드 분자 내의 또는 그것에 부착된 아미노산의 분절이다 (예컨대, 융합 단백질, 또는 태그가 달린 단백질). 유사하게, 핵산 또는 단백질은 이종성 표지 또는 이종성 분비 또는 국지화 서열을 포함할 수 있다.

[0081] 용어 "표지"는 표적 화합물에 부착되었을 때 직접 또는 간접적으로 (예컨대, 그것의 분광 특성, 형태, 또는 활성으로 인해) 검출 가능한 화학적 모이어티 또는 단백질을 나타낸다. 표지는 직접적으로 검출 가능하거나 (형광단) 또는 간접적으로 검출 가능할 수 있다 (합텐, 효소, 또는 형광단 퀜처). 이러한 표지는 분광학적, 광화학적, 생화학적, 면역화학적, 또는 화학적 수단에 의해 검출될 수 있다. 이러한 표지로는, 예를 들어, 방사선-카운팅 장치로 측정될 수 있는 방사성 표지; 분광계로 시각적으로 관찰되거나 측정될 수 있는 색소, 염료 또는 기타 발색원; 스핀 표지 분석기로 측정될 수 있는 스핀 표지; 및 형광 표지 (형광단)을 들 수 있고, 여기서 출력 신호는 적합한 분자 부가물의 여기에 의해 생성되고 염료에 의해 흡수된 광으로의 여기에 의해 시각화될 수 있거나 표준 형광측정기 또는 영상화 시스템으로 측정될 수 있다. 표지는 또한, 예를 들어, 화학발광 물질일 수 있거나, 이때 출력 신호는 신호 화합물의 화학적 변형에 의해 생성되며; 금속-함유 물질일 수 있거나; 또는 효소일 수 있고, 이때 신호의 효소-의존성 이차 생성, 예컨대 무색 기질로부터의 착색 생성물의 형성이 일어난

다. 용어 "표지"는 또한 콘쥬게이션된 분자가, 기질과 함께 후속적으로 첨가될 때, 검출 가능한 신호를 생성하기 위해 사용되도록 콘쥬게이션된 분자에 선택적으로 결합될 수 있는 "태그" 또는 합텐을 나타낼 수 있다. 예를 들어, 당업자는 태그로서 비오틴을 사용할 수 있고, 그런 후 태그에 결합시키기 위해 양고추냉이 과산화효소(HRP)의 아비딘 또는 스트렙트아비딘 콘쥬게이트를 사용할 수 있으며, 그런 후 HRP의 존재를 검출하기 위하여 비색 기질 (예컨대, 테트라메틸벤지딘 (TMB)) 또는 형광발생 기질을 사용할 수 있다. 용어 "표지"는 또한, 예를 들어, 정제를 용이하게 하기 위해 사용될 수 있는 태그를 나타낼 수 있다. 이러한 태그의 비제한적인 예로는, myc, HA, FLAG 또는 3XFLAG, 6XHis 또는 폴리히스트딘, 글루타티온-S-트랜스퍼라제 (GST), 말토오스 결합 단백질, 에피토프 태그, 또는 면역글로불린의 Fc 부분을 들 수 있다. 많은 표지가 알려져 있고, 예를 들어, 입자, 형광단, 합텐, 효소 및 이것들의 비색성, 형광발생성 및 화학발광성 기질 및 기타 표지를 들 수 있다.

[0082] "코돈 최적화"는 아미노산을 구체화하는 3-염기쌍 코돈 조합의 다중성에 의해 나타나는, 코돈의 축퇴성의 장점을 취하며, 일반적으로 천연 아미노산 서열을 유지하는 한편 천연 서열의 적어도 하나의 코돈을 숙주 세포의 유전자에서 보다 빈번하게 또는 가장 빈번하게 사용되는 코돈으로 교체함으로써 특히 숙주 세포에서의 증강된 발현을 위해 핵산 서열을 변형시키는 과정을 포함한다. 예를 들어, Cas9 단백질을 암호화하는 폴리뉴클레오타이드는 자연적으로 발생하는 핵산 서열에 비교하여, 박테리아 세포, 효모 세포, 인간 세포, 비인간 세포, 포유류 세포, 설치류 세포, 마우스 세포, 래트 세포, 햄스터 세포, 또는 기타 다른 숙주 세포를 포함하여, 주어진 원핵 또는 진핵 세포에서 고빈도로 사용되는 코돈을 치환하기 위해 변형될 수 있다. 코돈 용법 표들은 쉽게 입수 가능한데, 예를 들어, "코돈 용법 데이터베이스"에서 입수 가능하다. 이런 표들은 많은 방법으로 조정될 수 있다. Nakamura et al. (2000) *Nucleic Acids Research* 28:292 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)를 참조한다. 특정 숙주에서의 발현을 위해 특정 서열의 코돈 최적화를 위한 컴퓨터 알고리즘 또한 활용될 수 있다 (예컨대, Gene Forge 참조).

[0083] 용어 "유전자좌"는 유기체의 게놈의 염색체 상의 유전자 (또는 중요한 서열), DNA 서열, 폴리펩타이드-암호화 서열, 또는 위치의 특수한 위치를 나타낸다. 예를 들어, "*HSD17B13* 유전자좌"는 하기 서열이 있는 곳으로서 확인된 유기체의 게놈의 염색체 상의 *HSD17B13* 유전자, HSD17B13 DNA 서열, HSD17B13-암호화 서열, 또는 HSD17B13 위치의 특수한 위치를 나타낼 수 있다. "*HSD17B13* 유전자좌"는 *HSD17B13* 유전자의 조절 요소, 이를테면, 예를 들어, 인핸서, 프로모터, 5' 및/또는 3' UTR, 또는 이것들의 조합을 포함할 수 있다.

[0084] 용어 "유전자"는 생성물 (예컨대, RNA 생성물 및/또는 폴리펩타이드 생성물)을 암호화하고 하나 이상의 비코딩 인트론이 끼어 있는 코딩 영역 및 5' 및 3' 양 단부에서 그 코딩 영역에 인접하여 위치한 서열을 포함하여서 유전자가 전장 mRNA (5' 및 3' 미번역 서열을 포함함)에 상응하게 되는 염색체의 DNA 서열을 나타낸다. 용어 "유전자"는 또한 조절 서열 (예컨대, 프로모터, 인핸서, 및 전사 인자 결합 부위), 폴리아데닐화 신호, 내부 리보솜 유입 부위, 사일런서, 절연 서열(insulating sequence), 및 매트릭스 부착 영역을 포함하는 다른 비코딩 서열을 포함한다. 이런 서열들은 유전자의 코딩 영역에 가깝거나 (예컨대, 10 kb 이내) 또는 일정한 거리에 있을 수 있고, 유전자의 전산 및 번역의 수준 또는 속도에 영향을 준다. 용어 "유전자"는 또한 "미니유전자"를 포함한다.

[0085] 용어 "미니유전자"는 유전자의 하나 이상의 비필수 분절이 상응하는 자연적으로 발생하는 생식선 유전자와 관련하여 삭제되었지만, 적어도 하나의 인트론은 남아 있는 유전자를 나타낸다. 삭제된 분절은 인트론성 서열일 수 있다. 예를 들어, 삭제된 분절은 적어도 약 500 염기쌍 내지 수 킬로염기의 인트론성 서열일 수 있다. 전형적으로, 필수 조절 요소들을 포함하지 않는 인트론성 서열은 삭제될 수 있다. 미니유전자를 포함하는 유전자 분절은 전형적으로 생식선 유전자에 존재하는 것과 같이 동일한 선상의 순서로 배열되겠지만, 이것은 언제나 그런 것은 아니다. 일부 바람직한 조절 요소들 (예컨대, 인핸서, 사일런서)은 상대적으로 위치-둔감성이어서 조절 요소는 상응하는 생식선 유전자에서보더 미니유전자에서 상이한 곳에 위치하고 있을지라도 정확하게 기능할 것이다. 예를 들어, 인핸서는 프로모터로부터 상이한 거리에, 상이한 방향으로, 및/또는 상이한 선상의 순서로 위치할 수 있다. 예를 들어, 생식선 배치에서 프로모터에 대해 3'에 위치한 인핸서는 미니유전자에서는 프로모터에 대해 5'에 위치할 것이다. 유사하게, 일부 유전자는 RNA 수준에서 대안적으로 스플라이싱되는 엑손을 가질 수 있다. 그러므로, 미니유전자는 상응하는 생식선 유전자보다 더 적은 엑손 및/또는 상이한 선상 순서의 엑손을 가질 수 있고 여전히 기능성 유전자 생성물을 암호화할 것이다. 유전자 생성물을 암호화하는 cDNA가 또한 미니유전자를 구성하기 위해 사용될 수 있다 (예컨대, 하이브리드 cDNA-게놈 융합물).

[0086] 용어 "대립유전자"는 유전자의 변종 형태를 나타낸다. 일부 유전자는 동일한 위치에서 위치한 다양한 상이한 형태, 또는 유전자 상의 유전자좌를 염색체 상에 가진다. 이배체 유기체는 각 유전자 상의 유전자좌에 2개의 대립유전자를 가진다. 대립유전자의 각각의 쌍은 특이적 유전자 상의 유전자좌의 유전자형을 나타낸다. 유전자형은

두 동일한 대립유전자가 특정 유전자좌에 있다면 동형접합성으로서 및 두 대립유전자가 상이하다면 이형접합성으로서 기술된다.

[0087] 용어 "변종" 또는 "유전자 변종"은 집단에서 가장 우세한 서열과 상이한 (예컨대, 한 뉴클레오타이드 차이) 뉴클레오타이드 서열을 나타낸다. 예를 들어, 뉴클레오타이드 서열에서 일부 변이 또는 치환은 코돈을 변경시켜서 유전자 변종 폴리펩타이드를 초래하는 상이한 아미노산이 암호화된다. 용어 "변종"은 또한 암호화된 폴리펩타이드의 아미노산 서열을 변화시키지 않는 (즉 보존된 변화) 위치에서 집단에서 가장 우세한 서열과 서열이 상이한 유전자를 나타낼 수 있다. 유전자 변종은 위험과 관련이 있을 수 있거나, 보호화 관련이 있을 수 있거나, 또는 그 중간일 수 있다.

[0088] "프로모터"는 통상적으로 특정 폴리뉴클레오타이드 서열에 대한 적절한 전사 개시 부위에서 RNA 합성을 개시하기 위해 RNA 중합효소 II를 지시할 수 있는 TATA 박스를 포함하는 DNA의 조절 영역이다. 프로모터는 추가적으로 전사 개시 속도에 영향을 주는 다른 영역들을 포함할 수 있다. 본원에 개시된 프로모터 서열은 작동 가능하게 연결된 폴리뉴클레오타이드의 전사를 조절한다. 프로모터는 본원에 개시된 세포 유형 (예컨대, 원핵 세포, 비인간 포유류 세포, 인간 세포, 설치류 세포, 만능 세포, 분화된 세포, 또는 이것들의 조합 중 하나 이상의 유형에서 활성적일 수 있다. 프로모터는, 예를 들어, 구성적으로 활성인 프로모터, 조건부 프로모터, 유도성 프로모터, 일시적으로 제한된 프로모터 (예컨대, 발달상 조절된 프로모터), 또는 공간적으로 제한된 프로모터 (예컨대, 세포-특이적 또는 조직-특이적 프로모터)일 수 있다. 프로모터의 실례는, 예를 들어, WO 2013/176772 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)에서 찾아볼 수 있다.

[0089] 유도성 프로모터의 실례로는, 예를 들어, 화학적으로 조절된 프로모터 및 물리적으로 조절된 프로모터를 포함한다. 화학적으로 조절된 프로모터로는, 예를 들어, 알코올-조절된 프로모터 (예컨대, 알코올 탈수효소 (alcA) 유전자 프로모터), 테트라사이클린-조절된 프로모터 (예컨대, 테트라사이클린-반응성 프로모터, 테트라사이클린 오퍼레이터 서열 (tetO), tet-온 프로모터, 또는 tet-오프 프로모터), 스테로이드 조절된 프로모터 (예컨대, 래트 글루코코르티코이드 수용체, 에스트로겐 수용체의 프로모터, 또는 엑디손 수용체의 프로모터), 또는 금속-조절된 프로모터 (예컨대, 메탈로프로테인 프로모터)를 들 수 있다. 물리적으로 조절된 프로모터로는, 예를 들어, 온도-조절된 프로모터 (예컨대, 열충격 프로모터) 및 광-조절된 프로모터 (예컨대, 광-유도성 프로모터 또는 광-억제성 프로모터)를 들 수 있다.

[0090] 조직-특이적 프로모터는, 예를 들어, 뉴런-특이적 프로모터, 신경교-특이적 프로모터, 근육 세포-특이적 프로모터, 심장 세포-특이적 프로모터, 신장 세포-특이적 프로모터, 뼈 세포-특이적 프로모터, 내피 세포-특이적 프로모터, 또는 면역 세포-특이적 프로모터 (예컨대, B 세포 프로모터 또는 T 세포 프로모터)일 수 있다.

[0091] 발달 단계에서 조절된 프로모터로는, 예를 들어, 배아 발달 단계 중에만 활성인, 또는 성인 세포에서만 활성인 프로모터를 들 수 있다.

[0092] "작동 가능한 결합" 또는 "작동 가능하게 연결된"은 두 구성요소가 정상적으로 기능하고 구성요소들 중 적어도 하나가 다른 구성요소들 중 적어도 하나에 대해 발휘되는 기능을 매개할 수 있는 가능성을 허용하도록 둘 이상의 구성요소 (예컨대, 프로모터와 또 다른 서열 요소)의 병렬을 포함한다. 예를 들어, 프로모터는 만약 프로모터가 하나 이상의 전사 조절 인자의 존재 또는 부재에 대한 반응으로 코딩 서열의 전사 수준을 제어한다면 코딩 서열에 작동 가능하게 연결될 수 있다. 작동 가능한 결합은 서로와 연속적인 또는 트랜스로 작용하는 이러한 서열들을 포함할 수 있다 (예컨대, 조절 서열은 코딩 서열의 전사를 제어하기 위해 일정 거리에서 작용할 수 있다).

[0093] 용어 "프라이머"는 폴리뉴클레오타이드에 상보하는 프라이머 연장 생성물의 합성이 촉매되는 조건 하에 놓이게 될 때 상보하는 가닥을 따라 폴리뉴클레오타이드 합성이 개시되는 지점으로서 작용할 수 있는 올리고뉴클레오타이드를 나타낸다. 이러한 조건은 적절한 완충액 (보조인자인, 또는 pH, 이온 강도, 등에 영향을 주는 치환기를 포함함)에서, 및 적합한 온도에서 4개의 상이한 뉴클레오타이드 트라이포스페이트 또는 뉴클레오사이드 유사체 및 중합반응을 위한 하나 이상의 제제, 예컨대 DNA 중합효소 및/또는 역전사효소의 존재를 포함한다. 서열 특이적 방식으로의 프라이머의 연장은, 예를 들어, PCR, DNA 서열분석, DNA 연장, DNA 중합반응, RNA 전사, 또는 역전사의 방법들을 포함할 수 있다. 프라이머는 중합효소를 위한 제제의 존재하에 연장 생성물의 합성을 프라이밍하기 위해 충분히 길어야 한다. 전형적인 프라이머는 길이가 표적 서열에 실질적으로 상보하는 서열의 적어도 약 5개 뉴클레오타이드이지만, 더 긴 프라이머가 바람직하다. 전형적으로, 프라이머는 길이가 약 15 내지 30개 뉴클레오타이드이지만, 더 긴 프라이머가 또한 사용될 수 있다. 프라미어 서열은 주형 또는 표적 서열에 정확하게 상보할 필요는 없지만 주형 또는 표적 서열과 혼성화하기 위해 충분히 상보적이어야 한다. 용어 "프라이머

쌍"은 증폭될 DNA 서열의 5' 단부에 혼성화하는 5' 상류 프라이머, 및 증폭될 서열의 3' 단부의 완료를 위해 혼성화하는 3' 하류 프라이머를 포함하는, 프라이머의 쌍을 의미한다. 프라이머 쌍은 표적 폴리뉴클레오타이드의 증폭에 사용될 수 있다 (예컨대, 중합효소 연쇄 반응 (PCR) 또는 다른 종래의 핵산 증폭 방법에 의해). "PCR" 또는 "중합효소 연쇄 반응"은 특이적 DNA 분절의 증폭에 사용되는 기법이다 (미국 특허 제 4,683,195호 및 4,800,159호 참조, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0094] 용어 "프로브"는 구조가 상이한 표적 분자들 사이를 검출 가능하게 구별할 수 있는 분자를 나타낸다. 검출은 사용된 프로브의 유형 및 표적 분자의 유형에 따라 좌우되는 다양한 상이한 방식으로 달성될 수 있다. 그러므로, 예를 들어, 검출은 표적 분자의 활성 수준의 차별을 기반으로 할 수 있지만, 바람직하게는 특이적 결합의 검출을 기반으로 한다. 이러한 특이적 결합의 실례로는 항체 결합 및 핵산 프로브 혼성화를 들 수 있다. 그러므로, 프로브는, 예를 들어, 효소 기질, 항체 및 항체 단편, 및 핵산 혼성화 프로브를 포함할 수 있다. 예를 들어, 프로브는 종래의 검출 가능한 표지 또는 리포터 분자, 예컨대 방사성 동위원소, 리간드, 화학발광제, 효소, 등에 부착된 분리된 폴리뉴클레오타이드일 수 있다. 이러한 프로브는 표적 폴리뉴클레오타이드, 예컨대 *HSD17B13* rs72613567 변종 또는 특이적 HSD17B13 mRNA 전사물을 포함하는 폴리뉴클레오타이드의 가닥에 상보한다. 데옥시리보핵산 프로브는 *HSD17B13*-mRNA/cDNA-특이적 프라이머 또는 *HSD17B13*-rs72613567-특이적 프라이머, 시험관내에서 합성된 올리고뉴클레오타이드 프로브, 또는 박테리아 인공 염색체, 포스미드, 또는 코스미드 라이브러리로부터 얻어진 DNA를 사용하는 PCR에 의해 생성된 것들을 포함할 수 있다. 프로브는 데옥시리보핵산 또는 리보핵산뿐만 아니라 폴리아미드 및 특이적으로 표적 DNA 서열의 존재를 검출할 수 있는 기타 프로브 물질을 포함한다. 핵산 프로브의 경우, 검출 시약으로는, 예를 들어, 방사성 표지된 프로브, 효소적으로 표지된 프로브 (예컨대, 양고추냉이 과산화효소 및 알칼리성 포스파타제), 친화도 표지된 프로브 (예컨대 비오틴, 아비딘, 및 스트렙트아비딘), 및 형광 표지된 프로브 (예컨대, 6-FAM, VIC, TAMRA, MGB, 플루오레세인, 로다민, 및 텍사스 레드)를 들 수 있다. 본원에 기술된 핵산 프로브는 잘 알려져 있는 확립된 키트 포맷 중 하나로 쉽게 혼입될 수 있다.

[0095] 용어 "안티센스 RNA"는 세포에서 전사된 메신저 RNA 가닥에 상보하는 단일-가닥 RNA을 나타낸다.

[0096] 용어 "작은 간섭 RNA (siRNA)"는 RNA 간섭 (RNAi) 경로를 유도하는 전형적으로 이중 가닥의 RNA 분자를 나타낸다. 이런 분자들은 길이가 다를 수 있고 (일반적으로 18 내지 30개 염기쌍) 안티센스 가닥에 그것들의 표적 mRNA에 대한 다양한 정도의 상보성을 함유할 수 있다. 전부는 아니지만, 일부 siRNA는 센스 가닥 및/또는 안티센스 가닥의 5' 또는 3' 단부에 쌍을 이루지 않은 돌출된 염기를 가진다. 용어 "siRNA"는 2개의 별도의 가닥의 듀플렉스와, 뿐만 아니라 듀플렉스 영역을 포함하는 헤어핀 구조를 형성할 수 있는 단일 가닥을 포함한다. 이중 가닥 구조는, 예를 들어, 길이가 20, 25, 30, 35, 40, 45, 또는 50개 뉴클레오타이드 미만일 수 있다. 예를 들어, 이중 가닥 구조는 길이가 약 21 내지 23개 뉴클레오타이드, 길이가 약 약 19 내지 25개 뉴클레오타이드, 또는 길이가 약 19 내지 23개 뉴클레오타이드일 수 있다.

[0097] 용어 "짧은 헤어핀 RNA (shRNA)"는 헤어핀 구조로 자체-혼성화하고 프로세싱시 RNA 간섭 (RNAi) 경로를 포함할 수 있는 RNA 염기들의 단일 가닥을 나타낸다. 이런 분자들은 길이가 다를 수 있다 (일반적으로 길이가 약 50 내지 90개 뉴클레오타이드이거나, 또는 일부 경우에, 예컨대 microRNA-적응된 shRNA의 경우, 길이가 250개 뉴클레오타이드보다 더 많을 수 있다). shRNA 분자는 siRNA를 형성하기 위하여 세포 내에서 프로세싱되고, 결국 유전자 발현을 녹다운시킬 수 있다. shRNA는 벡터 안으로 포함될 수 있다. 용어 "shRNA"는 또한 그것으로부터 짧은, 헤어핀 RNA 분자가 전사될 수 있는 DNA 분자를 나타낸다.

[0098] 핵산의 "상보성"은 핵산의 한 가닥의 뉴클레오타이드 서열이 그것의 핵염기 기의 방향으로 인해 반대쪽 핵산 가닥의 다른 서열과 수소 결합하는 것을 의미한다. DNA의 상보하는 염기는 전형적으로 A와 T 및 C와 G이다. RNA에서, 그것들은 전형적으로 C와 G 및 U와 A이다. 상보성은 완벽하거나 실질적이거나/충분할 수 있다. 두 핵산 사이의 완벽한 상보성은 두 핵산이 듀플렉스의 모든 염기가 왓슨-크릭 쌍형성에 의해 상보하는 가닥에 결합되는 듀플렉스를 형성할 수 있는 것을 의미한다. "실질적인" 또는 "충분한" 상보하는은 한 가닥의 서열이 반대쪽 가닥의 서열에 완전하게 및/또는 완벽하게 상보하는 것은 아니지만, 두 가닥 상의 염기들 사이에 충분한 결합이 발생하여 혼성화 조건의 세트 (예컨대, 염 농도 및 온도)에서 안정적인 하이브리드 복합체를 형성하는 것을 의미한다. 이러한 조건은 혼성화된 가닥의 Tm (용융 온도)를 예측하기 위한 서열 및 표준 수학적 계산을 사용함으로써, 또는 기본적인 방법을 사용하여 Tm의 실험적인 측정에 의해 예측될 수 있다. Tm은 두 핵산 가닥 사이에 형성된 혼성화 복합체들의 집단이 50% 변성되는 (즉, 이중 가닥 핵산 분자의 집단의 절반이 단일 가닥으로 분리되는) 온도를 포함한다. Tm보다 낮은 온도에서, 혼성화 복합체의 형성이 선호되는 반면, Tm을 초과하는 온도에서는, 혼성화 복합체의 가닥들의 용융 또는 분리가 우선된다. Tm은 수성 1 M NaCl 용액 중의 공지된 G + C 함량

을 가지는 핵산에 대해, 예컨대 Tm=81.5+0.41(% G+C)를 사용함으로써 추정될 수 있지만, 다른 공지의 Tm 계산은 핵산 구조 특징을 고려한다.

[0099] "혼성화 조건"은 한 핵산 가닥이 상보하는 가닥 상호작용 및 수소 결합에 의해 제2 핵산 가닥에 결합하여 혼성화 복합체를 생성하는 누적 환경을 포함한다. 이러한 조건은 핵산을 함유하는 수성 또는 유기 용액의 화학적 구성요소들 및 그것들의 농도 (예컨대, 염, 킬레이트화제, 포름아미드), 및 혼합물의 온도를 포함한다. 다른 인자들, 예컨대 인큐베이션 시간의 길이 또는 반응 챔버 크기가 환경에 기여할 수 있다. 예컨대, Sambrook et al., Molecular Cloning, A Laboratory Manual, 2.sup.nd ed., pp. 1.90-1.91, 9.47-9.51, 1 1.47-11.57 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989)을 참조함 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0100] 혼성화는 두 핵산이 상보하는 서열을 함유하는 것을 필요로 하지만, 염기들 사이의 미스매치도 가능하다. 두 핵산 사이의 혼성화에 적절한 조건은, 잘 알려져 있는 변수들인, 핵산의 길이 및 상보성 정도에 좌우된다. 두 뉴클레오타이드 서열 사이의 상보성의 정도가 클수록, 이런 서열을 가지는 핵산의 하이브리드에 대한 용융 온도(Tm)의 값이 더 커진다. 짧은 구간의 상보성 (예컨대 35개 이하, 30개 이하, 25개 이하, 22개 이하, 20개 이하, 또는 18개 이하의 뉴클레오타이드에 걸친 상보성)을 가진 핵산들 사이의 혼성화의 경우, 미스매치의 위치는 중요한 것이 된다 (Sambrook et al., 상기 문헌, 11.7-11.8). 전형적으로, 혼성화 가능한 핵산에 대한 길이는 적어도 약 10개 뉴클레오타이드이다. 혼성화 가능한 핵산에 대한 예시적인 최소 길이로는 적어도 약 15개 뉴클레오타이드, 적어도 약 20개 뉴클레오타이드, 적어도 약 22개 뉴클레오타이드, 적어도 약 25개 뉴클레오타이드, 및 적어도 약 30개 뉴클레오타이드를 들 수 있다. 나아가, 온도 및 세척액 염 농도가 상보성 영역의 길이 및 상보성 정도와 같은 인자들에 따라 필요에 따라 조정될 수 있다.

[0101] 폴리뉴클레오타이드의 서열은 특이적으로 혼성화할 수 있는 그것의 표적 핵산의 서열에 100% 상보할 필요는 없다. 더욱이, 폴리뉴클레오타이드는 하나 이상의 분절에 걸쳐 혼성화할 수 있어서 개재하는 또는 인접한 분절은 혼성화 사건 (예컨대, 루프 구조 또는 헤어핀 구조)에 관련되지 않는다. 폴리뉴클레오타이드 (예컨대, gRNA)는 그것이 표적화하는 표적 핵산 서열 내의 표적 영역에 대한 적어도 70%, 적어도 80%, 적어도 90%, 적어도 95%, 적어도 99%, 또는 100%의 서열 상보성을 포함할 수 있다. 예를 들어, 20개의 뉴클레오타이드 중 18개가 표적 영역에 상보하고, 따라서 특이적으로 혼성화할 수 있을 gRNA는 90% 상보성을 나타낼 것이다. 이 실례에서, 나머지 비-상보하는성 뉴클레오타이드는 상보하는 뉴클레오타이드와 클러스터를 형성하거나 그것들이 사이에 섞여있을 수 있다.

[0102] 핵산 내의 핵산 서열의 특정 구간 사이의 퍼센트 상보성은 기본적으로 BLAST 프로그램 (기본적인 국소 정렬 조사 도구) 및 프로그램 (Altschul et al. (1990) *J. Mol. Biol*. 215:403-410; Zhang 및 Madden (1997) *Genome Res*. 7:649-656)을 사용하여 또는 Gap 프로그램 (Wisconsin Sequence Analysis Package, 버전 8 for Unix, Genetics Computer Group, University Research Park, Madison Wis.)을 사용함으로써, Smith 및 Waterman의 알고리즘 (Adv. Appl. Math., 1981, 2, 482-489)을 사용하는 디폴트 세팅을 사용함으로써 측정될 수 있다.

[0103] 본원에 제공된 방법 및 조성물은 다양한 상이한 구성요소를 사용한다. 명세서 전체에서 일부 구성요소는 활성 변종 및 단편을 가질 수 있다. 이러한 구성요소로는, 예를 들어, Cas9 단백질, CRISPR RNA, tracrRNA, 및 가이드 RNA를 들 수 있다. 이 구성요소들 각각에 대한 생물학적 활성은 본원의 다른 곳에서 기술된다.

[0104] 두 폴리뉴클레오타이드 또는 폴리펩타이드 서열의 맥락에서 "서열 동일성" 또는 "동일성"은 명시된 비교 창(window)에 걸쳐 최대 상응성을 위해 정렬될 때 동일한 두 서열의 잔기들을 언급한다. 서열 동일성의 백분율이 단백질을 언급하여 사용될 때, 동일하지 않은 잔기 위치는 종종 보존성 아미노산 치환에 의해 상이하며, 여기서 아미노산 잔기는 유사한 화학적 특성 (예컨대, 전하 또는 소수성)을 가진 다른 아미노산 잔기에 대해 치환되며 따라서 분자의 기능적 특성은 변화하지 않는다. 서열이 보존성 치환으로 상이할 때, 퍼센트 서열 동일성은 치환의 보존성 성질에 대해 보정하기 위해 상방향으로 조정될 수 있다. 이러한 보존성 치환에 의해 상이한 서열들은 "서열 유사성" 또는 "유사성"을 가지는 것으로 말할 수 있다. 이런 조정을 만들기 위한 수단은 잘 알려져 있다. 전형적으로, 이것은 전체 미스매치보다는 부분적인 것으로서 보존성 치환을 채점하고, 그로써 백분율 서열 동일성이 증가한다. 그러므로, 예를 들어, 동일한 아미노산이 1의 점수로 주어지고 비보존성 치환이 0의 점수로 주어지는 경우, 보존성 치환은 0과 1 사이의 점수로 주어진다. 보존성 치환의 채점은, 예컨대 프로그램 PC/유전자 (Intelligenetics, Mountain View, California)에서 실행되는 것과 같이 계산된다.

[0105] "서열 동일성의 백분율"은 비교 창에 걸쳐 두개의 최적으로 정렬된 서열을 비교함으로써 측정된 값 (완벽하게 매치된 잔기들의 가장 큰 수)을 포함하고, 여기서 비교 창의 폴리뉴클레오타이드 서열의 부분은 두 서열의 최적

정렬을 위한 참조 서열 (첨가 또는 삭제를 포함하지 않음)과 비교하여 첨가 또는 삭제 (즉, 갭)을 포함할 수 있다. 백분율은 동일한 핵산 염기 또는 아미노산 잔기가 두 서열에서 발생하는 위치의 수를 측정하여 매치된 위치의 수를 산출하고, 매치된 위치의 수를 비교 창에서의 위치의 총 수로 나눈 후, 그 결과에 100을 곱하여 서열 동일성의 백분율을 산출함으로써 계산된다. 다르게 명시되지 않는 한 (예컨대, 더 짧은 서열은 연결된 이종성 서열을 포함함), 비교 창은 비교되는 두 창 중 더 짧은 서열의 전장이다.

[0106] 다르게 진술되지 않는 한, 서열 동일성/유사성 값은 다음의 매개변수를 사용하는 GAP 버전 10을 사용하여 얻어진 값을 포함한다: 50의 GAP 무게 및 3의 길이 무게, 및 nwsgapdna.cmp 채점표를 사용하는 뉴클레오타이드 서열에 대한 % 동일성 및 % 유사성; 8의 GAP 무게 및 2의 길이 무게, 및 BLOSUM62 채점표를 사용하는 아미노산 서열에 대한 % 동일성 및 % 유사성; 또는 이것들의 임의의 동등한 프로그램. "동등한 프로그램"은, 문제의 임의의 두 서열에 대해, GAP 버전 10에 의해 생성된 상응하는 정렬과 비교할 때 동일한 뉴클레오타이드 또는 아미노산 잔기 매치 및 동일한 퍼센트 서열을 가지는 정렬을 생성하는 임의의 서열 비교 프로그램을 포함한다.

[0107] 용어 "보존성 아미노산 치환"은 유사한 크기, 전하, 또는 극성의 상이한 아미노산을 가진 서열에 정상적으로 존재하는 아미노산의 치환을 나타낸다. 보존성 치환의 예로는 아이소류신, 발린, 또는 류신과 같은 비극성 (소수성) 잔기의 또 다른 비극성 잔기에 대한 치환을 들 수 있다. 마찬가지로, 보존성 치환의 예로는, 한 극성 (친수성) 잔기의 또 다른 극성 잔기에 대한, 예컨대 아르기닌과 리신 사이, 글루타민과 아스파라긴 사이, 또는 글리신과 세린 사이의 치환을 들 수 있다. 추가적으로, 리신, 아르기닌, 또는 히스티딘과 같은 염기성 잔기의, 또는 아스파르트산 또는 글루탐산과 같은 한 산성 잔기의 또 다른 산성 잔기에 대한 치환은 보존성 치환의 추가적인 예이다. 비보존성 치환의 예로는 비극성 (소수성) 아미노산 잔기, 예컨대 아이소류신, 발린, 류신, 알라닌, 또는 메티오닌의 극성 (친수성) 잔기, 예컨대 시스테인, 글루타민, 글루탐산 또는 리신에 대한 및/또는 극성 잔기의 비극성 잔기에 대한 치환을 들 수 있다. 전형적인 아미노산 범주화를 하기에 요약한다.

[0108]

알라닌	Ala	A	비극성	중성	1.8
아르기닌	Arg	R	극성	양성	-4.5
아스파라긴	Asn	N	극성	중성	-3.5
아스파르트산	Asp	D	극성	음성	-3.5
시스테인	Cys	C	비극성	중성	2.5
글루탐산	Glu	E	극성	음성	-3.5
글루타민	Gln	Q	극성	중성	-3.5
글리신	Gly	G	비극성	중성	-0.4
히스티딘	His	H	극성	양성	-3.2
아이소류신	Ile	I	비극성	중성	4.5
류신	Leu	L	비극성	중성	3.8
리신	Lys	K	극성	양성	-3.9
메티오닌	Met	M	비극성	중성	1.9
페닐알라닌	Phe	F	비극성	중성	2.8
프롤린	Pro	P	비극성	중성	-1.6
세린	Ser	S	극성	중성	-0.8
트레오닌	Thr	T	극성	중성	-0.7
트립토판	Trp	W	비극성	중성	-0.9
티로신	Tyr	Y	극성	중성	-1.3
발린	Val	V	비극성	중성	4.2

[0109] 프라이머 또는 가이드 RNA와 같은 대상 핵산은 위치의 약 1000, 500, 400, 300, 200, 100, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5개 뉴클레오타이드 내에 있을 때 참조 핵산에서 명시된 뉴클레오타이드 위치에 혼성화하거나 위치를 표적화하거나 또는 그 위치에 인접한 위치를 포함한다.

[0110] 용어 "생물학적 샘플"은 대상체 내에 있거나 대상체로부터 얻을 수 있는, 그것으로부터 핵산 또는 단백질을 회수할 수 있는 생물학적 물질의 샘플을 나타낸다. 용어 생물학적 샘플은 또한 샘플, 예컨대 세포 또는 그것의 자손을 가공함으로써 유래된 임의의 물질을 포함할 수 있다. 생물학적 샘플의 가공은 여과, 증류, 추출, 농축, 고정화, 간섭하는 성분들의 비활성화, 등 중 하나 이상을 포함할 수 있다. 일부 구체예에서, 생물학적 샘플은 핵산, 예컨대 게놈 DNA, cDNA, 또는 mRNA를 포함한다. 일부 구체예에서, 생물학적 샘플은 단백질을 포함한다. 대상체는 임의의 유기체, 이를테면, 예를 들어, 인간, 비인간 포유류, 설치류, 마우스, 또는 래트일 수 있다. 생물학적 샘플은 대상체로부터의 임의의 세포, 조직, 또는 생물학적 유체로부터 유래될 수 있다. 샘플은 임의의 임상적으로 관련된 조직, 예컨대 골수 샘플, 종양 생검, 미세 바늘 흡인, 또는 체액의 샘플, 예컨대 혈액,

혈장, 혈청, 림프, 복수, 낭액(cystic fluid), 또는 소변을 포함할 수 있다. 일부 경우에, 샘플은 구강 면봉을 포함한다. 본원에 개시된 방법에서 사용된 샘플은 검정 포맷, 검출 방법의 성질, 및 샘플로서 사용된 조직, 세포, 또는 추출물을 기반으로 다를 것이다.

[0111] 용어 "대조군 샘플"은 *HSD17B13* rs72613567 변종을 갖지 않은, 바람직하게는 *HSD17B13* 유전자의 야생형 대립유전자에 대해 동형접합성인 대상체로부터 얻어진 샘플을 나타낸다. 이러한 샘플은 생물학적 샘플로서 동시에 또는 상이한 기회에 얻어질 수 있다. 생물학적 샘플 및 대조군 샘플은 둘 다 동일한 조직 또는 체액으로부터 얻어질 수 있다.

[0112] "상동성" 서열 (예컨대, 핵산 서열)은 공지된 참조 서열에 대해 동일하거나 또는 실질적으로 유사한, 예를 들어, 공지된 참조 서열에 대해 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하게 되는 서열을 포함한다. 상동성 서열은, 예를 들어, 병렬상동성 서열(orthologous sequence) 및 직렬상동성 서열(paralogous sequences)을 포함할 수 있다. 상동성 유전자는, 예를 들어, 전형적으로 공통 조상 DNA 서열로부터, 종분화 사건 (병렬상동성 유전자) 또는 유전자 중복 사건 (직렬상동성 유전자) 어느 한 쪽을 통해 내려온다. "병렬상동성" 유전자는 종분화에 의해 공통 조상 유전자로부터 진화된 상이한 종에서의 유전자를 포함한다. 병렬상동체는 전형적으로 진화 과정에서 동일한 기능을 보유한다. "직렬상동성" 유전자는 게놈 내에서의 중복에 의해 관련된 유전자를 포함한다. 직렬상동체는 진화 과정에서 새로운 기능을 진화시킬 수 있다.

[0113] 용어 "시험관내에서"는 인공 환경을 포함하며 인공 환경 (예컨대, 시험관) 내에서 발생하는 가공 또는 반응을 포함한다. 용어 "생체내에서"는 천연 환경 (예컨대, 세포 또는 유기체 또는 신체, 예컨대 유기체 또는 신체 내의 세포)을 포함하며 천연 환경 내에서 발생하는 가공 또는 반응을 포함한다. 용어 "생체외에서"는 개체의 신체로부터 제거된 세포를 포함하며 이런 세포 내에서 발생하는 가공 또는 반응을 포함한다.

[0114] 하나 이상의 인용된 요소들을 "포함하는" 또는 "포함하여," 조성물 또는 방법은 구체적으로 인용되지 않은 다른 요소들을 포함할 수 있다. 예를 들어, 단백질을 "포함하는" 또는 "포함하여," 조성물은 단백질을 단독으로 또는 다른 성분들과 함께 함유할 수 있다. 이행 문구 "본질적으로 ~으로 구성되는"은 청구범위의 범주가 그 청구범위에서 인용되는 명시된 요소들 및 청구된 발명의 기본적인 및 신규한 특징(들)에 실질적으로 영향을 주지 않는 것들을 다 포함하는 것으로 해석될 것을 의미한다. 그러므로, 용어 "본질적으로 ~으로 구성되는"은 본 발명의 청구범위에서 사용될 때 "포함하는"과 동등한 것으로 해석되는 것으로 의도된 것이 아니다.

[0115] "선택적인" 또는 "선택적으로"는 후속적으로 기술된 사건 또는 상황이 일어날 수 있거나 일어나지 않을 수 있는 것을 의미하고 그런 기술은 그 사건 또는 상황이 일어난 경우 및 일어나지 않은 경우를 포함하는 것을 의미한다.

[0116] 값의 범위의 지정은 그 범위 내의 또는 그 범위를 한정하는 모든 정수, 및 그 범위 내의 정수들에 의해 한정된 모든 하위범위를 포함한다.

[0117] 맥락으로부터 다르게 표현되지 않는 한, 용어 "약"은 표시된 값의 측정의 표준 오차 범위 (SEM) 내의 값들을 포함한다.

[0118] 용어 "및/또는"은 관련된 열거된 항목들 중 하나 이상의 임의의 및 모든 가능한 조합, 뿐만 아니라 대안적으로 해석될 때 조합의 결핍 ("또는")을 나타내며 그러한 것들을 포함한다.

[0119] 용어 "또는"은 특정 목록의 임의의 한 구성원을 나타내며 또한 그 목록의 구성원들의 임의의 조합을 포함한다.

[0120] 물품의 단일 형태를 나타내는 용어들은 맥락이 분명하게 다른 것을 표시하지 않는한 복수의 대상물을 포함한다. 예를 들어, 용어 "Cas9 단백질" 또는 "적어도 하나의 Cas9 단백질"은 이것들의 혼합물을 포함하여, 다수의 Cas9 단백질을 포함할 수 있다.

[0121] 통계학적으로 유의한은 $p \leq 0.05$인 것을 의미한다.

[0122] **상세한 설명**

[0123] ***I. 개관***

[0124] 본원에는 감소된 알라닌 및 아스파테이트 아미노전이효소 수준; 비알코올성 및 알코올성 간 지방간 질환, 간경변, 및 간세포 암종을 포함하는 만성 간 질환의 위험의 감소; 및 단순 지방증으로부터 만성 간 질환의 보다 임상적으로 진전된 단계로의 진행의 감소와 관련된 것으로 발견된 *HSD17B13* 변종이 제공된다. 또한 본원에는 그

변종과 관련된 *HSD17B13* 유전자의, 이전에는 미확인되었던 전사물이 제공된다.

[0125] *HSD17B13*의 변종과 관련된 분리된 핵산 및 단백질, 및 그런 핵산 및 단백질을 포함하는 세포가 본원에 제공된다. 또한 재조합 *HSD17B13* 유전자 또는 HSD17B13 단백질을 암호화하는 핵산을 발현시키기 위하여 뉴클레아제 제제, 외인성 도너 서열, 전사 활성화제, 전사 억제제, 및 발현 벡터의 임의의 조합을 사용을 통해 세포를 변형시키는 방법이 제공된다. 또한 만성 간 질환을 가진 또는 만성 간 질환의 발병 위험이 있는 대상체를 치료하기 위한 치료적 및 예방적 방법이 제공된다.

[0126] ***II. HSD17B13 변종***

[0127] 본원에는 *HSD17B13*의 변종 (또한 하이드록시스테로이드 17-베타 탈수소효소 13, 17-베타-하이드록시스테로이드 탈수소효소 13, 17β-하이드록시스테로이드 탈수소효소-13, 17β-HSD13, 단쇄 탈수소효소/환원효소 9, *SCDR9*, *HMFN0376*, *NIL497*, 및 *SDR16C3*으로도 알려져 있음)에 관련된 분리된 핵산 및 단백질이 제공된다. 인간 *HSD17B13* 유전자는 길이가 대략 19 kb이고 게놈의 4q22.1에 위치한 7개의 엑손 및 6개의 인트론을 포함한다. 예시적인 인간 HSD17B13 단백질 서열은 UniProt 등록 번호 Q7Z5P4 (각각 SEQ ID NO: 240 및 241; Q7Z5P4-1 및 Q7Z5P4-2) 및 NCBI 참조 서열 번호 NP_835236 및 NP_001129702 (각각 SEQ ID NO: 242 및 243)이 부여된다. 예시적인 인간 HSD17B13 mRNA는 NCBI 참조 서열 번호 NM_178135 및 NM_001136230 (각각 SEQ ID NO: 244 및 245)가 부여된다.

[0128] 특히, 본원에는 인트론 6의 도너 스플라이스 부위에 인접하여 아데닌의 삽입을 가지는 *HSD17B13* (rs72613567)의 스플라이스 변종이 제공된다. 아데닌은 염색체의 전방 (플러스) 가닥에 삽입된 것으로, 염색체의 역방향 (마이너스)에서 삽입된 티민에 상응한다. 인간 *HSD17B13* 유전자는 역방향으로 전사되고, 이런 뉴클레오타이드 삽입은 SEQ ID NO: 1에서 제공된 예시적인 야생형 *HSD17B13* 유전자 서열에 비교하여 SEQ ID NO: 2에서 제공된 예시적인 *HSD17B13* rs72613567 변종 서열에서 삽입된 티민으로서 반영된다. 그러므로 삽입은 본원에서는 SEQ ID NO: 1의 위치 12665와 12666 사이에서 또는 SEQ ID NO: 2의 위치 12666에서 삽입된 티민으로서 언급될 것이다.

[0129] 2개의 mRNA 전사물 (각각 A 및 B; SEQ ID NO: 4 및 5)이 야생형 *HSD17B13* 유전자를 가진 대상체에서 발현되는 것으로 이전에 확인되었다. 전사물 A는 *HSD17B13* 유전자의 모두 7개의 엑손을 포함한 반면, 엑손 2는 전사물 B에서는 빠져 있다. 전사물 A는 야생형 대상체에서 우세한 전사물이다. 그러나, 본원에는, 이전에 확인된, 6개의 추가적인 발현되는 (각각 C 내지 H, SEQ ID NO: 6 내지 11) *HSD17B13* 전사물이 제공된다. 이들 전사물은 도 4에서 도시된다. 전사물 C에서, 엑손 6은 전사물 A와 비교하여 빠져 있다. 전사물 D에서, 엑손 6의 3'에 구아닌이 삽입되어 전사물 A와 비교하여 엑손 7에서의 프레임시프트 및 미성숙 절단을 초래한다. 전사물 E에서, 전사물 A와 비교하여 엑손 3과 4 사이에 추가적인 엑손이 있다. *HSD17B13* rs72613567 변종 보균자에서만 발현되는 전사물 F에서, 전사물 A와 비교하여 엑손 6으로부터 인트론 6으로까지 초과번역 (read-through)이 있다. 전사물 G에서, 엑손 2는 빠져 있고, 엑손 6의 3'에 구아닌이 삽입되어 전사물 A와 비교하여 엑손 7에서의 프레임시프트 및 미성숙 절단을 초래한다. 전사물 H에서, 엑손 3과 4 사이에 추가적인 엑손이 있고, 엑손 6의 3'에 구아닌이 삽입되어 전사물 A와 비교하여 엑손 7에서의 프레임시프트 및 미성숙 절단을 초래한다. 전사물 C, D, F, G, 및 H는 *HSD17B13* rs72613567 변종 보균자에서 우세하고, 이때 *HSD17B13* rs72613567 변종의 보균자에서 전사물 D가 가장 풍부한 전사물이다. 또한 본원에는 저수준으로 발현되는, 이전에 확인되지 않았던, 하나의 추가적인 *HSD17B13* 전사물 (F', SEQ ID NO: 246)이 제공된다. 전사물 F와 같이, 전사물 F'는 또한 전사물 A와 비교하여 엑손 6으로부터 인트론 6까지 초과번역을 포함하지만, 전사물 F와는 대조적으로, 초과번역은 *HSD17B13* rs72613567 변종 유전자에 존재하는 삽입된 티민을 포함하지 않는다. 각 전사물에 대한 *HSD17B13* 유전자 내의 엑손의 뉴클레오타이드 위치가 하기에 제공된다.

[0130] **야생형 *HSD17B13* 유전자에 대해 동형접합성인 대상체에서 보다 우세한 HSD17B13 전사물의 엑손에 대한 SEQ ID NO: 1에서의 뉴클레오타이드 위치**

	전사물 A	전사물 B	전사물 E	전사물 F'
엑손 1	1-275	1-275	1-275	1-275
엑손 2	4471-4578	건너뜀	4471-4578	4471-4578
엑손 3	5684-5815	5684-5815	5684-5815	5684-5815
엑손 3'	존재하지 않음	존재하지 않음	6210-6281	존재하지 않음
엑손 4	7308-7414	7308-7414	7308-7414	7308-7414
엑손 5	8947-9084	8947-9084	8947-9084	8947-9084
엑손 6	12548-12664	12548-12664	12548-12664	12548-13501*
엑손 7	17599-19118	17599-19118	17599-19118	건너뜀

[0131]

[0132] * 엑손 6으로부터 인트론 6까지의 초과번역을 포함함; 초과번역 = 위치 12665-13501

[0133] **rs72613567 *HSD17B13* 변종 유전자 (위치 12666에서 T의 삽입)에 대해 동형접합성인 대상체에서 보다 우세한 HSD17B13 전사물의 엑손에 대한 SEQ ID NO: 2에서의 뉴클레오타이드 위치**

	전사물 C	전사물 D	전사물 F	전사물 G	전사물 H
엑손 1	1-275	1-275	1-275	1-275	1-275
엑손 2	4471-4578	4471-4578	4471-4578	건너뜀	4471-4578
엑손 3	5684-5815	5684-5815	5684-5815	5684-5815	5684-5815
엑손 3'	존재하지 않음	존재하지 않음	존재하지 않음	존재하지 않음	6210-6281
엑손 4	7308-7414	7308-7414	7308-7414	7308-7414	7308-7414
엑손 5	8947-9084	8947-9084	8947-9084	8947-9084	8947-9084
엑손 6	건너뜀	12548-12665^	12548-13502*	12548-12665^	12548-12665^
엑손 7	17600-19119	17600-19119	건너뜀	17600-19119	17600-19119

[0134]

[0135] ^ 전사물 A와 비교하여 3' 단부에 추가적인 잔기 12665를 포함함

[0136] * 엑손 6으로부터 인트론 6까지의 초과번역을 포함함; 초과번역 = 위치 12665-13502

[0137] 본원의 다른 곳에서 보다 상세하게 설명되는 바, *HSD17B13* rs72613567 변종은 감소된 알라닌 및 아스파테이트 아미노전이효소 수준 및 비알코올성 및 알코올성 간 지방간 질환, 간경변, 및 간세포 암종을 포함하는 만성 간 질환의 위험의 감소와 관련된다. *HSD17B13* rs72613567 변종은 또한 단순 지방증으로부터 만성 간 질환의 보다 임상적으로 진전된 단계로의 진행의 감소와도 관련된다.

[0138] **A. 핵산**

[0139] 본원에는 *HSD17B13* 변종 및 변종 HSD17B13 전사물과 관련된 분리된 핵산이 제공된다. 또한 엄격한 또는 보통의 조건 하에서 본원에 개시된 핵산들 중 임의의 핵산과 혼성화하는 분리된 핵산이 개시된다. 이러한 핵산은, 예를 들어, *HSD17B13* 변종 단백질을 발현시키기 위해 또는 프라이머, 프로브, 외인성 도너 서열, 가이드 RNA, 안티센스 RNA, shRNA, 및 siRNA로서 유용할 수 있고, 이것들은 각각 본원의 다른 곳에서 보다 상세하게 기술된다.

[0140] 또한 개시된 폴리뉴클레오타이드와 상호작용할 수 있는 기능성 핵산이 개시된다. 기능성 핵산은 특이적 기능, 예컨대 표적 분자에 대한 결합 또는 특이적 반응의 촉매작용을 가지는 핵산 분자들이다. 기능성 핵산의 예로는 안티센스 분자, 앱타머, 리보자임, 트리플렉스 형성 분자, 및 외부 가이드 서열을 들 수 있다. 기능성 핵산 분자는 표적 분자가 갖고 있는 특이적 활성의 이펙터, 억제제, 조절제, 및 자극제로서 작용할 수 있거나, 또는 기능성 핵산 분자는 임의의 다른 분자들과는 무관한 드노보 활성을 가질 수 있다.

[0141] 안티센스 분자는 규범적인 또는 비규범적인 염기쌍 형성 어느 한 쪽을 통해 표적 핵산 분자와 상호작용하도록 설계된다. 안티센스 분자와 표적 분자의 상호작용은, 예를 들어, RNAe-H-매개된 RNA-DNA 하이브리드 분해를 통해 표적 분자의 파괴를 촉진하도록 설계된다. 대안적으로, 안티센스 분자는 표적 분자상에서 정상적으로 일어날 수 있는 가공 기능, 예컨대 전사 또는 복제를 방해하도록 설계된다. 안티센스 분자는 표적 분자의 서열을 기반으로 설계될 수 있다. 표적 분자의 가장 접근이 쉬운 영역들을 발견함으로써 안티센스 효율을 최적화하기 위한 수많은 방법이 존재한다. 예시적인 방법은 DMS 및 DEPC를 사용하는 시험관내 선택 실험 및 DNA 변형 연구일 것이다. 안티센스 분자는 일반적으로 10^{-6}, 10^{-8}, 10^{-10}, 또는 10^{-12} 이하의 해리 상수 (k_d)로 표적 분자에 결합한다.

안티센스 분자의 설계 및 사용을 보조하는 방법 및 기법의 대표적인 샘플은 다음의 비제한적인 미국 특허 목록에서 찾아볼 수 있다: 제 5,135,917호; 5,294,533호; 5,627,158호; 5,641,754호; 5,691,317호; 5,780,607호; 5,786,138호; 5,849,903호; 5,856,103호; 5,919,772호; 5,955,590호; 5,990,088호; 5,994,320호; 5,998,602호; 6,005,095호; 6,007,995호; 6,013,522호; 6,017,898호; 6,018,042호; 6,025,198호; 6,033,910호; 6,040,296호; 6,046,004호; 6,046,319호; 및 6,057,437호 (이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 안티센스 분자의 실례는 안티센스 RNA, 작은 간섭 RNA (siRNA), 및 짧은 헤어핀 RNA (shRNA)를 포함하고, 이것들은 본원의 다른 곳에서 보다 상세하게 기술된다.

[0142] 본원에 개시된 분리된 핵산은 RNA, DNA, 또는 RNA와 DNA 둘 다를 포함할 수 있다. 분리된 핵산은 또한 벡터에서와 같이 이종성 핵산 서열에, 또는 이종성 표지에 결합되거나 융합될 수 있다. 예를 들어, 본원에 개시된 분리된 핵산은 분리된 핵산 및 이종성 핵산 서열을 포함하는 벡터 또는 외인성 도너 서열에 있을 수 있다. 분리된 핵산은 또한 이종성 표지, 예컨대 형광성 표지에 결합되거나 융합될 수 있다. 표지의 다른 실례는 본원의 다른

곳에서 개시된다.

[0143] 개시된 핵산 분자는, 예를 들어, 뉴클레오타이드 또는 비천연 또는 변형된 뉴클레오타이드, 예컨대 뉴클레오타이드 유사체 또는 뉴클레오타이드 치환체로 구성될 수 있다. 이러한 뉴클레오타이드는 변형된 염기, 당 또는 포스페이트 기를 함유하는, 또는 그것의 구조에 비천연 모이어티를 포함한 뉴클레오타이드를 포함한다. 비천연 뉴클레오타이드의 실례로는 다이데옥시뉴클레오타이드, 비오티닐화된, 아민화된, 탈아민화된, 알킬화된, 벤질화된 및 형광단-표지된 뉴클레오타이드를 들 수 있다.

[0144] 본원에 개시된 핵산 분자는 하나 이상의 뉴클레오타이드 유사체 또는 치환체를 포함할 수 있다. 뉴클레오타이드 유사체는 염기, 당, 또는 포스페이트 모이어티 중 어느 하나에 대한 일부 유형의 변형을 함유한 뉴클레오타이드이다. 염기 모이어티에 대한 변형은 A, C, G, 및 T/U뿐만 아니라 상이한 퓨린 또는 피리미딘 염기, 예컨대 슈도우리딘, 우라실-5-일, 하이포크산틴-9-일 (I), 및 2-아미노아데닌-9-일의 천연 및 합성 변형을 포함할 것이다. 변형된 염기로는, 예를 들어, 5-메틸시토신 (5-mg-C), 5-하이드록시메틸 시토신, 크산틴, 하이포크산틴, 2-아미노아데닌, 아데닌 및 구아닌의 6-메틸 및 다른 알킬 유도체, 아데닌 및 구아닌의 2-프로필 및 다른 알킬 유도체, 2-티오우라실, 2-티오티민 및 2-티오시토신, 5-할로우라실 및 시토신, 5-프로피닐 우라실 및 시토신, 6-아조 우라실, 시토신 및 티민, 5-우라실 (슈도우라실), 4-티오우라실, 8-할로, 8-아미노, 8-티올, 8-티오알킬, 8-하이드록실 및 다른 8-치환된 아데닌 및 구아닌, 5-할로, 특히 5-브로모, 5-트라이플루오로메틸 및 다른 5-치환된 우라실 및 시토신, 7-메틸구아닌 및 7-메틸아데닌, 8-아자구아닌 및 8-아자아데닌, 7-데아자구아닌 및 7-데아자아데닌 및 3-데아자구아닌 및 3-데아자아데닌을 들 수 있다. 추가적인 염기 변형은, 예를 들어 미국 특허 제 3,687,808호; Englisch et al. (1991) *Angewandte Chemie, International Edition* 30:613; 및 Sanghvi, Y. S., Chapter 15, Antisense Research 및 Applications, pages 289-302, Crooke, S. T. 및 Lebleu, B. ed., CRC Press, 1993 (이 문헌들은 각 각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨)에서 찾아볼 수 있다. 특정 뉴클레오타이드 유사체, 예컨대 5-치환된 피리미딘, 6-아자피리미딘 및 N-2, N-6 및 O-6 치환된 퓨린, 이를테면 2-아미노프로필아데닌, 5-프로피닐우라실, 5-프로피닐시토신, 및 5-메틸시토신은 듀플렉스 형성의 안정성을 증가시킬 수 있다. 종종 염기 변형은, 예를 들어, 당 변형, 예컨대 2'-O-메톡시에틸과 조합되어 증가된 듀플렉스 안정성과 같은 특유의 특성을 이룰 수 있다. 많은 미국 특허들, 예컨대 제 4,845,205호; 5,130,302호; 5,134,066호; 5,175,273호; 5,367,066호; 5,432,272호; 5,457,187호; 5,459,255호; 5,484,908호; 5,502,177호; 5,525,711호; 5,552,540호; 5,587,469호; 5,594,121호; 5,596,091호; 5,614,617; 및 5,681,941호가 다양한 염기 변형을 상세하게 기술한다. 이 특허들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다.

[0145] 뉴클레오타이드 유사체는 또한 당 모이어티의 변형을 포함할 수 있다. 당 모이어티에 대한 변형은, 예를 들어, 리보오스 및 데옥시 리보오스의 천연 변형뿐만 아니라 합성 변형을 포함할 수 있다. 당 변형으로는, 예를 들어, 2' 위치에서의 다음의 변형을 포함한다: OH; F; O-, S-, 또는 N-알킬; O-, S-, 또는 N-알케닐; O-, S- 또는 N-알키닐; 또는 O-알킬-O-알킬, 여기서 알킬, 알케닐, 및 알키닐은 치환된 또는 비치환된 C1 내지 C10, 알킬 또는 C2 내지 C10 알케닐 및 알키닐일 수 있다. 예시적인 2' 당 변형은 또한, 예를 들어, -O[(CH2)$_n$ O]m CH3, -O(CH2)$_n$OCH3, -O(CH2)$_n$ NH2, -O(CH2)$_n$ CH3, -O(CH2)$_n$ -ONH2, 및 -O(CH2)$_n$ON[(CH2)$_n$CH3)]2를 포함하고, 여기서 n 및 m은 1 내지 약 10이다.

[0146] 2' 위치에서의 다른 변형은, 예를 들어, C_1 내지 C_{10} 저급 알킬, 치환된 저급 알킬, 알카릴, 아르알킬, O-알카릴 또는 O-아르알킬, SH, SCH3, OCN, Cl, Br, CN, CF3, OCF_3, $SOCH_3$, SO_2CH_3, ONO_2, NO_2, N_3, NH_2, 헤테로사이클로알킬, 헤테로사이클로알카릴, 아미노알킬아미노, 폴리알킬아미노, 치환된 실릴, RNA 절단기, 리포터 기, 인터칼레이터(intercalator), 올리고뉴클레오타이드의 약물동역학적 특성을 개선하기 위한 기, 또는 올리고뉴클레오타이드의 약물역학적 특성을 개선하기 위한 기, 및 유사한 특성을 가진 다른 치환체를 포함한다. 유사한 변형은 또한 당의 다른 위치에서, 특히 3' 말단 뉴클레오타이드 상의 또는 2'-5' 연결된 올리고뉴클레오타이드에서 당의 3' 위치에서 및 5' 말단 뉴클레오타이드의 5' 위치에서 이루어질 수 있다. 변형된 당은 또한 가교 고리 산소에서의 변형, 예컨대 CH_2 및 S를 함유한 것들을 포함할 수 있다. 뉴클레오타이드 당 유사체는 또한 펜토푸라노실 당 대신 사이클로부틸 모이어티와 같은 당 모방물을 가질 수 있다. 이러한 변형된 당 구조의 제조를 교시하는 많은 미국 특허가 있는데, 예컨대 제 4,981,957호; 5,118,800호; 5,319,080호; 5,359,044호; 5,393,878호; 5,446,137호; 5,466,786호; 5,514,785호; 5,519,134호; 5,567,811호; 5,576,427호; 5,591,722호; 5,597,909호; 5,610,300호; 5,627,053호; 5,639,873호; 5,646,265호; 5,658,873호; 5,670,633호; 및 5,700,920호가 있고, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다.

[0147] 뉴클레오타이드 유사체는 또한 포스페이트 모이어티에서 변형될 수 있다. 변형된 포스페이트 모이어티로는, 예를 들어, 두 뉴클레오타이드 사이의 결합이 포스포로티오에이트, 키랄 포스포로티오에이트, 포스포로다이티오에이트, 포스포트라이에스테르, 아미노알킬포스포트라이에스테르, 3'-알킬렌 포스포네이트 및 키랄 포스포네이트를 포함하는 메틸 및 다른 알킬 포스포네이트, 포스피네이트, 3'-아미노 포스포라미데이트 및 아미노알킬포스포라미데이트를 포함하는 포스포라미데이트, 티오노포스포라미데이트, 티오노알킬포스포네이트, 티오노알킬포스포트라이에스테르, 및 보라노포스페이트를 함유하도록 변형될 수 있는 것들을 포함한다. 두 뉴클레오타이드 사이의 이런 포스페이트 또는 변형된 포스페이트 결합은 3'-5' 결합 또는 2'-5' 결합을 통해서일 수 있고, 결합은 반전된 극성, 예컨대 3'-5'에서 5'-3'으로 또는 2'-5'에서 5'-2'으로를 함유할 수 있다. 다양한 염, 혼합염 및 유리 산 형태가 또한 포함된다. 많은 미국 특허들이 변형된 포스페이트를 함유하는 뉴클레오티드의 제조 및 사용을 교시하는데, 예를 들어, 미국 특허 제 3,687,808호; 4,469,863호; 4,476,301호; 5,023,243호; 5,177,196호; 5,188,897호; 5,264,423호; 5,276,019호; 5,278,302호; 5,286,717호; 5,321,131호; 5,399,676호; 5,405,939호; 5,453,496호; 5,455,233호; 5,466,677호; 5,476,925호; 5,519,126호; 5,536,821호; 5,541,306호; 5,550,111호; 5,563,253호; 5,571,799호; 5,587,361호; 및 5,625,050호가 있고, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다.

[0148] 뉴클레오타이드 치환체는 뉴클레오타이드에 대해 유사한 기능적 특성을 갖지만, 포스페이트 모이어티를 함유하지 않는 분자, 예컨대 펩타이드 핵산 (PNA)을 포함한다. 뉴클레오타이드 치환체는 왓슨-크릭 또는 후그스틴 방식으로 핵산을 인지하겠지만, 포스페이트 모이어티 이외의 모이어티를 통해 함께 연결되는 분자들을 포함한다. 뉴클레오티드 치환체는 적절한 표적 핵산과 상호작용할 때 이중 나선형 구조를 따를 수 있다.

[0149] 뉴클레오타이드 치환체는 또한 교체된 포스페이트 모이어티 또는 당 모이어티를 가진 뉴클레오타이드 또는 뉴클레오타이드 유사체를 포함한다. 뉴클레오타이드 치환체는 표준 인 원자를 함유하지 않을 수 있다. 포스페이트에 대한 치환체는, 예를 들어, 단쇄 알킬 또는 사이클로알킬 뉴클레오사이드간 결합, 혼합 헤테로원자 및 알킬 또는 사이클로알킬 뉴클레오사이드간 결합, 또는 하나 이상의 단쇄 헤테로원자 또는 헤테로고리형 뉴클레오사이드간 결합일 수 있다. 이것들은 모르폴리노 결합을 가지는 것들 (뉴클레오사이드의 당 부분으로부터 부분적으로 형성됨); 실록산 골격; 황화물, 설폭사이드 및 설폰 골격; 포름아세틸 및 티오포름아세틸 골격; 메틸렌 포름아세틸 및 티오포름아세틸 골격; 알켄 함유 골격; 설파메이트 골격; 메틸렌이미노 및 메틸렌하이드라지노 골격; 설포네이트 및 설폰아미드 골격; 아미드 골격; 및 혼합된 N, O, S 및 CH_2 구성요소 부분을 가지는 기타 골격을 가지는 것들을 포함한다. 많은 미국 특허가 포스페이트 치환의 이런 유형을 제조하고 사용하는 방법을 개시하며, 한정하는 것은 아니지만 다음의 특허들을 포함한다: 미국 특허 제 5,034,506호; 5,166,315호; 5,185,444호; 5,214,134호; 5,216,141호; 5,235,033호; 5,264,562호; 5,264,564호; 5,405,938호; 5,434,257호; 5,466,677호; 5,470,967호; 5,489,677호; 5,541,307호; 5,561,225호; 5,596,086호; 5,602,240호; 5,610,289호; 5,602,240호; 5,608,046호; 5,610,289호; 5,618,704호; 5,623,070호; 5,663,312호; 5,633,360호; 5,677,437호; 및 5,677,439호 (이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0150] 또한 뉴클레오타이드 치환체에서 뉴클레오타이드의 당 및 포스페이트 모이어티 둘 다, 예를 들어, 아미드형 결합 (아미노에틸글리신) (PNA)에 의해 치환될 수 있는 것이 인지된다. 미국 특허 제 5,539,082호; 5,714,331호; 및 5,719,262호는 PNA 분자를 제조하고 사용하는 방법을 교시한다 (이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 또한 Nielsen et al. (1991) *Science* 254:1497-1500 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)을 참조한다.

[0151] 또한 다른 유형의 분자 (콘쥬게이트)를 뉴클레오타이드 또는 뉴클레오타이드 유사체에 연결시켜서, 예를 들어, 세포 흡수를 증강시키는 것도 가능하다. 콘쥬게이트는 뉴클레오타이드 또는 뉴클레오타이드 유사체에 화학적으로 결합될 수 있다. 이러한 콘쥬게이트로는, 예를 들어, 콜레스테롤 모이어티와 같은 지질 모이어티 (Letsinger et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 콜산 (Manoharan et al. (1994) *Bioorg. Med. Chem. Let.* 4:1053-1060, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 헥실-S-트리틸티올과 같은 티오에테르 (Manoharan et al. (1992) *Ann. N.Y. Acad. Sci.* 660:306-309; Manoharan et al. (1993) *Bioorg. Med. Chem. Let.* 3:2765-2770, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 티오콜레스테롤 (Oberhauser et al. (1992) *Nucl. Acids Res.* 20:533-538, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 도데칸다이올 또는 운데실 잔기와 같은 지방족 사슬 (Saison-Behmoaras et al. (1991) *EMBO J.* 10:1111-1118; Kabanov et al. (1990) *FEBS Lett.* 259:327-330; Svinarchuk et al. (1993) *Biochimie* 75:49-54, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 인지질, 예

컨대 다이-헥사데실-rac-글리세롤 또는 트라이에틸암모늄 1,2-다이-O-헥사데실-rac-글리세로-3-H-포스포네이트 (Manoharan et al. (1995) *Tetrahedron Lett.* 36:3651-3654; Shea et al. (1990) *Nucl. Acids Res.* 18:3777-3783, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 폴리아민 또는 폴리에틸렌 글리콜 사슬 (Manoharan et al. (1995) *Nucleosides & Nucleotides* 14:969-973, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 또는 아다만탄 아세트산 (Manoharan et al. (1995) *Tetrahedron Lett.* 36:3651-3654, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 팔미틸 모이어티 (Mishra et al. (1995) *Biochim. Biophys. Acta* 1264:229-237), 또는 옥타데실아민 또는 헥실아미노-카르보닐-옥시콜레스테롤 모이어티 (Crooke et al. (1996) *J. Pharmacol. Exp. Ther.* 277:923-937, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨)를 들 수 있다. 많은 미국 특허가 이러한 콘쥬게이트의 제조를 교시하며, 예를 들어, 다음을 포함한다: 미국 특허 제 4,828,979호; 4,948,882호; 5,218,105호; 5,525,465호; 5,541,313호; 5,545,730호; 5,552,538호; 5,578,717호; 5,580,731호; 5,580,731호; 5,591,584호; 5,109,124호; 5,118,802호; 5,138,045호; 5,414,077호; 5,486,603호; 5,512,439호; 5,578,718호; 5,608,046호; 4,587,044호; 4,605,735호; 4,667,025호; 4,762,779호; 4,789,737호; 4,824,941호; 4,835,263호; 4,876,335호; 4,904,582호; 4,958,013호; 5,082,830호; 5,112,963호; 5,214,136호; 5,082,830호; 5,112,963호; 5,214,136호; 5,245,022호; 5,254,469호; 5,258,506호; 5,262,536호; 5,272,250호; 5,292,873호; 5,317,098호; 5,371,241호; 5,391,723호; 5,416,203호; 5,451,463호; 5,510,475호; 5,512,667호; 5,514,785호; 5,565,552호; 5,567,810호; 5,574,142호; 5,585,481호; 5,587,371호; 5,595,726호; 5,597,696호; 5,599,923호; 5,599,928호 및 5,688,941호 (이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0152] 본원에 개시된 분리된 핵산은 자연적으로 발생하는 *HSD17B13* 유전자 또는 mRNA 전사물의 뉴클레오타이드 서열을 포함하거나, 또는 자연적으로 발생하지 않는 서열을 포함할 수 있다. 한 실례에서, 자연적으로 발생하지 않는 서열은 동의 돌연변이 (동의적 mutation) 또는 암호화된 HSD17B13 단백질에 영향을 미치지 않는 돌연변이로 인해 자연적으로 발생하지 않는 서열과 상이할 수 있다. 예를 들어, 서열은 동의 돌연변이 또는 암호화된 HSD17B13 단백질에 영향을 미치지 않는 돌연변이를 제외하고 동일할 수 있다. 동의 돌연변이 또는 치환은 생성된 아미노산 서열이 변형되지 않도록 하는, 단백질을 암호화하는 유전자의 엑손에서 한 뉴클레오타이드의 또 다른 뉴클레오타이드에 대한 치환이다. 이것은 유전자 암호의 축퇴성 때문에 가능한데, 이때 일부 아미노산이 하나 이상의 3-염기쌍 코돈에 의해 암호화된다. 동의 치환은, 예를 들어, 코돈 최적화 과정에 사용된다.

[0153] 또한 본원에는 본원에 개시된 핵산에 의해 암호화된 단백질 및 본원에 개시된 분리된 핵산 또는 단백질 및 그 분리된 핵산 또는 단백질의 안정성을 증가시키는 (예컨대, 분해 산물이 한계값 아래로 유지되는, 예컨대 출발 핵산 또는 단백질의 0.5 중량% 아래로 유지되어야 하는 주어진 보관 조건 (예컨대, -20℃, 4℃, 또는 주위 온도) 하에서의 기간을 연장시키는; 또는 생체내에서 안정성을 증가시키는) 담체를 포함하는 조성물이 개시된다. 이러한 담체의 비제한적인 실례는 폴리(락트산) (PLA) 마이크로스피어, 폴리(D,L-락틱-코글리콜-산) (PLGA) 마이크로스피어, 리포좀, 미셀, 역미셀, 지질 코클레이트, 및 지질 미세관을 들 수 있다.

[0154] ***(1) HSD17B13 rs72613567 변종의 돌연변이 잔기를 포함하는 핵산***

[0155] 본원에는 *HSD17B13* 유전자의 적어도 15개의 연속 뉴클레오타이드를 포함하고 *HSD17B13* rs72613567 변종과 최적으로 정렬될 때 *HSD17B13* rs72613567 변종 (SEQ ID NO: 2)의 위치 12666에 상응하는 위치에 티민 (또는 위치 12666 및 12667에 상응하는 위치에 티민들)을 가지는 분리된 핵산이 개시된다. 즉, 본원에는 *HSD17B13* 유전자의 적어도 15개의 연속 뉴클레오타이드를 포함하고 야생형 *HSD17B13* 유전자와 최적으로 정렬될 때 야생형 *HSD17B13* 유전자 (SEQ ID NO: 1)의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 가지는 분리된 핵산이 개시된다. 이러한 분리된 핵산은, 예를 들어, *HSD17B13* 변종 전사물 및 단백질을 발현하기 위해 또는 외인성 도너 서열로서 유용할 수 있다. 이러한 분리된 핵산은 또한, 예를 들어, 가이드 RNA, 프라이머, 및 프로브로서도 유용할 수 있다.

[0156] *HSD17B13* 유전자는 임의의 유기체로부터의 *HSD17B13* 유전자일 수 있다. 예를 들어, *HSD17B13* 유전자는 인간 *HSD17B13* 유전자이거나 다른 유기체, 예컨대 비인간 포유류, 설치류, 마우스, 또는 래트로부터의 병렬상동체 (ortholog)일 수 있다.

[0157] 집단 내의 유전자 서열은 단일-뉴클레오타이드 다형성(polymorphisms)과 같은 다형성으로 인해 다를 수 있다. 본원에 제공된 예들은 단지 예시적인 서열들이다. 다른 서열 또한 가능하다. 한 예로서, 적어도 15개의 연속 뉴클레오타이드는 SEQ ID NO: 2와 최저으로 정렬딜 때 SEQ ID NO: 2의 위치 12666 또는 위치 12666 및 12667을

포함하는 *HSD17B13* rs72613567 변종 (SEQ ID NO: 2)의 상응하는 서열에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일할 수 있다. 선택적으로, 분리된 핵산은 SEQ ID NO: 2의 위치 12666 또는 위치 12666 및 12667을 포함하는 SEQ ID NO: 2의 적어도 15개의 연속 뉴클레오타이드를 포함한다. 다른 예로서, 적어도 15개의 연속 뉴클레오타이드는 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666을 포함하는 야생형 *HSD17B13* 유전자 (SEQ ID NO: 1)의 상응하는 서열에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일할 수 있고, 여기서 티민은 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 위치들 사이에 존재한다. 선택적으로, 분리된 핵산은 SEQ ID NO: 1의 위치 12665 및 12666을 포함하는 SEQ ID NO: 1의 적어도 15개의 연속 뉴클레오타이드를 포함하고, 여기서 티민은 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 위치들 사이에 존재한다.

[0158] 분리된 핵산은 *HSD17B13* 유전자의, 예를 들어, 적어도 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개의 연속 뉴클레오타이드를 포함할 수 있다. 대안적으로, 분리된 핵산은 *HSD17B13* 유전자의, 예를 들어, 적어도 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 11000, 12000, 13000, 14000, 15000, 16000, 17000, 18000, 또는 19000개의 연속 뉴클레오타이드를 포함할 수 있다.

[0159] 일부 경우에, 분리된 핵산은 유전자의 하나 이상의 비필수 분절이 상응하는 야생형 *HSD17B13* 유전자와 관련하여 삭제되어 있는 *HSD17B13* 미니유전자를 포함할 수 있다. 한 예로서, 삭제된 분절은 하나 이상의 인트론성 서열을 포함한다. 이러한 *HSD17B13* 미니유전자는, 예를 들어, 전사물 D로부터의 엑손 1 내지 7에 상응하는 엑손 및 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 인트론 6에 상응하는 인트론을 포함할 수 있다. 한 예로서, *HSD17B13* 미니유전자는 SEQ ID NO: 2로부터의 엑손 1 내지 7 및 인트론 6을 포함할 수 있다. 미니유전자는 본원의 다른 곳에서 보다 상세하게 기술된다.

[0160] ***(2) HSD17B13 rs72613567 변종의 돌연변이 잔기에 인접한 또는 돌연변이 잔기를 포함하는 서열에 혼성화하는 핵산***

[0161] 또한 본원에는 *HSD17B13* rs72613567 변종과 최적으로 정렬될 때 *HSD17B13* rs72613567 변종 (SEQ ID NO: 2)의 위치 12666 또는 위치 12666 및 12667에 상응하는 위치의 1000, 500, 400, 300, 200, 100, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5개 뉴클레오타이드 내에 있거나 포함하는 분절에서 *HSD17B13* 유전자 (예컨대, *HSD17B13* 미니유전자)에 혼성화하는 적어도 15개의 연속 뉴클레오타이드를 포함하는 분리된 핵산이 개시된다. 이러한 분리된 핵산은, 예를 들어, 가이드 RNA, 프라이머, 프로브, 또는 외인성 도너 서열로서 유용할 수 있다.

[0162] *HSD17B13* 유전자는 임의의 유기체로부터의 *HSD17B13* 유전자일 수 있다. 예를 들어, *HSD17B13* 유전자는 인간 *HSD17B13* 유전자이거나 다른 유기체, 예컨대 비인간 포유류, 마우스, 또는 래트로부터의 병렬상동체일 수 있다.

[0163] 한 예로서, 적어도 15개의 연속 뉴클레오타이드는 SEQ ID NO: 2와 최적으로 정렬될 때 *HSD17B13* rs72613567 변종 (SEQ ID NO: 2)의 상응하는 서열에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 *HSD17B13* 유전자 또는 *HSD17B13* 미니유전자의 분절에 혼성화할 수 있다. 선택적으로, 분리된 핵산은 SEQ ID NO: 2의 적어도 15개의 연속 뉴클레오타이드에 혼성화할 수 있다. 선택적으로, 분리된 핵산은 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666 또는 위치 12666 및 12667을 포함하는 분절 또는 SEQ ID NO: 2의 위치 12666 또는 위치 12666 및 12667에 상응하는 위치에 혼성화한다.

[0164] 분리된 핵산이 혼성화할 수 있는 분절은 *HSD17B13* 유전자의, 예를 들어, 적어도 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 75, 90, 95, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개의 연속 뉴클레오타이드를 포함할 수 있다. 대안적으로, 분리된 핵산은 *HSD17B13* 유전자의, 예를 들어, 적어도 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 11000, 12000, 13000, 14000, 15000, 16000, 17000, 18000, 또는 19000개의 연속 뉴클레오타이드를 포함할 수 있다. 대안적으로, 분리된 핵산이 혼성화할 수 있는 분절은 *HSD17B13* 유전자의, 예를 들어, 최대 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 75, 90, 95, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개의 연속 뉴클레오타이드일 수 있다. 예를 들어, 분절은 길이가 약 15 내지 100개 뉴클레오타이드, 또는 길이가 약 15 내지 35개 뉴클레오타이드일 수 있다.

[0165] ***(3) HSD17B13 rs72613567 변종에 의해 생성된 cDNA 및 변종 전사물***

[0166] 또한 본원에는 전사물 A 내지 H 중 임의의 하나와 최적으로 정렬될 때, 전사물 A 내지 H (각각 SEQ ID NO: 4 내지 11), 특히 전사물 C 내지 H 중 임의의 하나에 상응하는 mRNA 전사물 또는 cDNA의 전부 또는 일부에 상응하는

핵산이 제공된다. 집단 내의유전자 서열 및 이러한 유전자로부터 전사된 mRNA 서열은 단일-뉴클레오타이드 다형성과 같은 다형성으로 인해 다를 수 있는 것이 인지된다. 각 전사물에 대해 본원에서 제공된 서열은 단지 예시적인 서열이다. 다른 서열 또한 가능하다. 구체적이고, 비제한적인 예들이 하기에 제공된다. 이러한 분리된 핵산은, 예를 들어, *HSD17B13* 변종 전사물 및 단백질을 발현시키기에 유용할 수 있다.

[0167] 분리된 핵산은 임의의 길이를 가질 수 있다. 예를 들어, 분리된 핵산은 HSD17B13 단백질의 전부 또는 일부를 암호화하는 적어도 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 또는 2000개의 연속 뉴클레오타이드를 포함할 수 있다. 일부 경우에, 분리된 핵산은 HSD17B13 단백질의 전부 또는 일부를 암호화하는 연속 뉴클레오타이드를 포함하며, 여기서 연속 뉴클레오타이드는 *HSD17B13* 유전자의 적어도 2개의 상이한 엑손으로부터의 서열 (예컨대, 개재하는 인트론이 없는 *HSD17B13* 유전자의 적어도 하나의 엑손-엑손 경계에 걸쳐 있는)을 포함한다.

[0168] *HSD17B13* 전사물 D (SEQ ID NO: 7), 전사물 G (SEQ ID NO: 10), 및 전사물 H (SEQ ID NO: 11)는 엑손 6의 3' 단부에 구아닌의 삽입을 포함하여, 전사물 A와 비교하여 엑손 7에 프레임시프트 및 엑손 7에 의해 암호화된 HSD17B13 단백질의 영역의 미성숙 절단을 초래한다. 따라서, 본원에는 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 D, G, 및 H (또는 이것의 단편 또는 상동체)에 존재하는 분절 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는 분리된 핵산이 제공된다. 이러한 영역은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 본원에는 HSD17B13 단백질의 전부 또는 일부를 암호화하는 적어도 15개의 연속 뉴클레오타이드 (예컨대, 적어도 20개의 연속 뉴클레오타이드 또는 적어도 30개의 연속 뉴클레오타이드)를 포함하는 분리된 핵산이 제공되며, 여기서 연속 뉴클레오타이드의 분절 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)은 SEQ ID NO: 7, 10, 또는 11의 각각과 최적으로 정렬될 때 SEQ ID NO: 7 (*HSD17B13* 전사물 D), SEQ ID NO: 10 (*HSD17B13* 전사물 G), 또는 SEQ ID NO: 11 (*HSD17B13* 전사물 H)에서 엑손 6-엑손 7 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일하며, 분절은 SEQ ID NO: 7의 엑손 6의 3' 단부에서 잔기 878에 상응하는 잔기에 구아닌을 포함하거나 (즉, 엑손 7이 시작되는 곳의 구아닌에 더불어 전사물 A에 비교하여 엑손 6의 3' 단부에 구아닌이 삽입됨), SEQ ID NO: 10의 엑손 6의 3' 단부에서 잔기 770에 상응하는 잔기에 구아닌을 포함하거나 (즉, 엑손 7이 시작되는 곳의 구아닌에 더불어 전사물 B에 비교하여 엑손 6의 3' 단부에 구아닌이 삽입됨), 또는 SEQ ID NO: 11의 엑손 6의 3' 단부에서 잔기 950에 상응하는 잔기에 구아닌을 포함한다 (즉, 엑손 7이 시작되는 곳의 구아닌에 더불어 전사물 E에 비교하여 엑손 6의 3' 단부에 구아닌이 삽입됨). 이러한 핵산은 삽입된 구아닌을 *HSD17B13* 전사물에서의 다른 특징들 (예컨대, 엑손 7이 시작되는 곳의 구아닌, 전사물 F에서 인트론 6으로의 번역 초과, 또는 전사물 C에서 삭제된 엑손 6)과 구별하기 위하여 엑손 6 및 7의 각각에서 충분한 수의 뉴클레오타이드를 포함할 것으로 인지된다.

[0169] 한 예로서, 분리된 핵산은 엑손 6-엑손 7 경계에 걸쳐 있는 SEQ ID NO: 7의 적어도 15개의 연속 뉴클레오타이드 (예컨대, 적어도 20개의 연속 뉴클레오타이드 또는 적어도 30개의 연속 뉴클레오타이드)를 포함할 수 있고, 선택적으로 SEQ ID NO: 7의 엑손 6 및 7을 포함할 수 있으며, 선택적으로 SEQ ID NO: 7의 전체 서열을 포함할 수 있다.

[0170] 선택적으로, 분리된 핵산은 추가로 전사물 G (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 D (또는 이것의 단편 또는 상동체)에 존재하는 분절을 포함하고, 분리된 핵산은 추가로 전사물 H (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 D (또는 이것의 단편 또는 상동체)에 존재하는 분절을 포함한다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 이러한 분리된 핵산은 전사물 H와 구별하기 위하여 SEQ ID NO: 7과 최적으로 정렬될 때 SEQ ID NO: 7 (*HSD17B13* 전사물 D)의 엑손 3 및 4의 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 연속 뉴클레오타이드 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)의 분절을 포함할 수 있다. 마찬가지로, 이러한 분리된 핵산은 전사물 G와 구별하기 위하여 SEQ ID NO: 7과 최적으로 정렬될 때 SEQ ID NO: 7 (*HSD17B13* 전사물 D)의 엑손 2 내의 영역, 또는 SEQ ID NO: 7의 엑손 1-엑손 2 경계에 걸쳐 있는 영역, 또는 SEQ ID NO: 7의 엑손 2-엑손 3 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 연속 뉴클레오타이드 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)의 분절을 포함할 수 있다. 선택적으로, 분리된 핵산은 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 제시된 서열에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 서열을 포함하고, SEQ ID NO: 15 (HSD17B13 동형 D)에 제시된 서열을 포함하는 HSD17B13 단백질을 암호

화한다. 전사물 D와 같이, 전사물 H (SEQ ID NO: 11)는 전사물 A와 비교하여 엑손 6의 3' 단부에 구아닌의 삽입을 포함한다. 전사물 H는 추가로 전사물 A 및 전사물 D와 비교하여 엑손 3과 4 사이에 추가적인 엑손 (엑손 3')을 포함한다. 따라서, 본원에는 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 D, G, 및 H (또는 이것의 단편 또는 상동체)에 존재하는 분절을 포함하며 추가로 전사물 D (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 H (또는 이것의 단편 또는 상동체)의 분절 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는 상기 기술된 것과 같은 분리된 핵산이 제공된다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 본원에는 전사물 D에 대해 기술된 것과 같은 분리된 핵산이 제공되며, 여기서 연속 뉴클레오타이드 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)의 분절은 SEQ ID NO: 11과 최적으로 정렬될 때 SEQ ID NO: 11 (*HSD17B13* 전사물 H)의 엑손 3' 내의 영역, SEQ ID NO: 11의 엑손 3-엑손 3' 경계에 걸쳐 있는 영역, 또는 SEQ ID NO: 11의 엑손 3'-엑손 4 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일하다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 (예컨대, 엑손 3 및 4의 경계)과 구별하기 위하여 엑손 3 및 3'의 각각 또는 엑손 3' 및 4의 각각에 충분한 수의 뉴클레오타이드를 포함할 것으로 인지된다. 예를 들어, 엑손 3'의 영역은 전체 엑손 3'를 포함할 수 있다. 선택적으로, 분리된 핵산은 SEQ ID NO: 11 (*HSD17B13* 전사물 H)에 제시된 서열에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 서열을 포함하며 SEQ ID NO: 19 (HSD17B13 동형 H)에 제시된 서열을 포함하는 HSD17B13 단백질을 암호화한다.

[0171] 한 예로서, 분리된 핵산은 엑손 3' 내의 영역, 엑손 3-엑손 3' 경계에 걸쳐 있는 영역, 또는 엑손 3'-엑손 4 경계에 걸쳐 있는 영역을 포함하는 SEQ ID NO: 11의 적어도 15개의 연속 뉴클레오타이드 (예컨대, 적어도 20개의 연속 뉴클레오타이드 또는 적어도 30개의 연속 뉴클레오타이드)를 포함할 수 있고, 선택적으로 SEQ ID NO: 11의 전체 엑손 3'를 포함할 수 있으며, 선택적으로 SEQ ID NO: 11의 전체 서열을 포함할 수 있다.

[0172] 전사물 D와 같이, 전사물 G (SEQ ID NO: 10)는 전사물 A와 비교하여 엑손 6의 3'에 구아닌의 삽입을 포함한다. 그러나, 이에 더불어, 전사물 G는 전사물 A 및 전사물 D와 비교하여 엑손 2가 없다 (즉, 전사물 G는 전사물 A 및 전사물 D에 존재하지 않는 엑손 1-엑손 3 경계를 포함한다). 따라서, 본원에는 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 D, G, 및 H (또는 이것의 단편 또는 상동체)에 존재하는 분절을 포함하지만 추가로 전사물 D (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 G (또는 이것의 단편 또는 상동체)로부터의 분절 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는 상기 기술된 바와 같은 분리된 핵산이 제공된다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 본원에는 전사물 D에 대해 기술된 것과 같은 분리된 핵산이 제공되며, 여기서 연속 뉴클레오타이드 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)의 분절은 SEQ ID NO: 10과 최적으로 정렬될 때 SEQ ID NO: 10 (*HSD17B13* 전사물 G)의 엑손 1-엑손 3 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일하다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 ((예컨대, 엑손 1 및 2의 경계 또는 엑손 2 및 3의 경계)과 구별하기 위하여 엑손 1과 3의 각각에 충분한 수의 뉴클레오타이드를 포함할 것으로 인지된다. 예를 들어, 그 영역은 SEQ ID NO: 10의 엑손 1 및 3의 전체를 포함할 수 있다. 선택적으로, 분리된 핵산은 SEQ ID NO: 10 (*HSD17B13* 전사물 G)에 제시된 서열에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 서열을 포함하며 SEQ ID NO: 18 (HSD17B13 동형 G)에 제시된 서열을 포함하는 HSD17B13 단백질을 암호화한다.

[0173] 한 예로서, 분리된 핵산은 엑손 1-엑손 3 경계에 걸쳐 있는 영역을 포함하는 SEQ ID NO: 10의 적어도 15개의 연속 뉴클레오타이드 (예컨대, 적어도 20개의 연속 뉴클레오타이드 또는 적어도 30개의 연속 뉴클레오타이드)를 포함할 수 있고, 선택적으로 SEQ ID NO: 10의 엑손 1 및 3을 포함할 수 있며, 선택적으로 SEQ ID NO: 10의 전체 서열을 포함할 수 있다.

[0174] 또한 본원에는 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 E (또는 이것의 단편 또는 상동체)에 존재하는 분절 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는 분리된 핵산이 제공된다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 전사물 E (SEQ ID NO: 8)는 전사물 A와 비교하여 엑손 3과 4 사이에 추가적인 엑손을 포함한다. 따라서, 본원에는 HSD17B13 단백질의 전부 또는 일부를 암호화하는 적어도 15개의 연속 뉴클레오타이드 (예컨대, 적어도 20개의 연속 뉴클레오타이드 또는 적어도 30개의 연속 뉴클레오타이드)를 포함하는 분리된 핵산이 제공되며, 여기서 연속 뉴클레오타이드 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)의 분

절은 SEQ ID NO: 8과 최적으로 정렬될 때 SEQ ID NO: 8 (*HSD17B13* 전사물 E)의 엑손 3' 내의 영역, SEQ ID NO: 8의 엑손 3-엑손 3' 경계에 걸쳐 있는 영역, 또는 SEQ ID NO: 8의 엑손 3'-엑손 4 경계에 걸쳐 있는 영역에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일하다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 (예컨대, 엑손 3 및 4의 경계)과 구별하기 위하여 엑손 3 및 3'의 각각 또는 엑손 3' 및 4의 각각에 충분한 수의 뉴클레오타이드를 포함할 것으로 인지된다. 예를 들어, 엑손 3'의 영역은 전체 엑손 3'를 포함할 수 있다. 선택적으로, 분리된 핵산은 추가로 전사물 H (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 E (또는 이것의 단편 또는 상동체)로부터의 분절 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함한다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 본원에는 상기 기술된 것과 같은 분리된 핵산이 제공되며, 여기서 연속 뉴클레오타이드 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)의 분절은 SEQ ID NO: 8과 최적으로 정렬될 때, SEQ ID NO: 8 (*HSD17B13* 전사물 E)의 엑손 6-엑손 7 경계에 걸쳐 있는 영역에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일하다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 (특히, 전사물 H에서 엑손 6의 3' 단부에 있는 추가적인 구아닌)과 구별하기 위하여 엑손 6 및 7의 각각에 충분한 수의 뉴클레오타이드를 포함할 것으로 인지된다. 예를 들어, 영역은 SEQ ID NO: 8의 엑손 6 및 7의 전체를 포함할 수 있다. 선택적으로, 분리된 핵산은 SEQ ID NO: 8 (*HSD17B13* 전사물 E)에 제시된 서열에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 서열을 포함하고 SEQ ID NO: 16 (HSD17B13 동형 E)에 제시된 서열을 포함하는 HSD17B13 단백질을 암호화한다.

[0175] 한 예로서, 분리된 핵산은 엑손 3' 내의 영역, 엑손 3-엑손 3' 경계에 걸쳐 있는 영역, 또는 엑손 3'-엑손 4 경계에 걸쳐 있는 영역을 포함하는 SEQ ID NO: 8의 적어도 15개의 연속 뉴클레오타이드 (예컨대, 적어도 20개의 연속 뉴클레오타이드 또는 적어도 30개의 연속 뉴클레오타이드)를 포함할 수 있고, 선택적으로 SEQ ID NO: 8의 전체 엑손 3'를 포함할 수 있으며, 선택적으로 SEQ ID NO: 8의 전체 서열을 포함할 수 있다.

[0176] 또한 본원에는 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 F (또는 이것의 단편 또는 상동체)에 존재하는 분절 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는 분리된 핵산이 제공된다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 전사물 F (SEQ ID NO: 9)는 전사물 A와 비교하여 엑손 6으로부터 인트론 6까지의 초과번역을 포함하며, 초과번역은 *HSD17B13* rs72613567 변종 유전자에 존재하는 삽입된 티민을 포함한다. 따라서, 본원에는 HSD17B13 단백질의 전부 또는 일부를 암호화하는 적어도 15개의 연속 뉴클레오타이드 (예컨대, 적어도 20개의 연속 뉴클레오타이드 또는 적어도 30개의 연속 뉴클레오타이드)를 포함하는 분리된 핵산이 제공되며, 여기서 연속 뉴클레오타이드 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)의 분절은 SEQ ID NO: 9와 최적으로 정렬될 때 SEQ ID NO: 9 (*HSD17B13* 전사물 F)의 인트론 6으로의 초과번역 내의 영역 또는 SEQ ID NO: 9의 인트론 6으로의 초과번역과 엑손 6의 나머지 사이의 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일하다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 (예컨대, 다른 *HSD17B13* 전사물의 엑손 6 및 7의 경계)으로부터의 초과번역과 구별하기 위하여 초과번역에 충분한 수의 뉴클레오타이드일 것으로 인지된다. 선택적으로, 연속 뉴클레오타이드는 전사물 F' (SEQ ID NO: 246)에 존재하지 않는 전사물 F에 존재하는 서열 (즉, 삽입된 티민)을 포함한다. 전사물 F'는 또한 전사물 A와 비교하여 엑손 6으로부터 인트론 6으로의 초과번역을 포함하지만, 초과번역은 *HSD17B13* rs72613567 변종 유전자에 존재하는 삽입된 티민을 포함하지 않는다. 예를 들어, 영역은 SEQ ID NO: 9에서 인트론 6으로의 초과번역 전체일 수 있다. 선택적으로, 분리된 핵산은 SEQ ID NO: 9 (*HSD17B13* 전사물 F)에 제시된 서열에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 서열을 포함하며 SEQ ID NO: 17 (HSD17B13 동형 F)에 제시된 서열을 포함하는 HSD17B13 단백질을 암호화한다.

[0177] 한 예로서, 분리된 핵산은 인트론 6으로의 초과번역 내의 영역 또는 인트론 6으로의 초과번역과 엑손 6의 나머지 사이의 경계에 걸쳐 있는 영역을 포함하는 SEQ ID NO: 9의 적어도 15개의 연속 뉴클레오타이드 (예컨대, 적어도 20개의 연속 뉴클레오타이드 또는 적어도 30개의 연속 뉴클레오타이드)를 포함할 수 있고, 선택적으로 인트론 6으로의 초과번역 전체를 포함할 수 있으며, 선택적으로 SEQ ID NO: 9의 전체 서열을 포함할 수 있다.

[0178] 또한 본원에는 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 F' (또는 이것의 단편 또는 상동체)에 존재하는 분절 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는 분리된 핵산이 제공된다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 전사물 F' (SEQ ID NO: 246)는 전사물 A와 비교하여 엑손 6으로부터 인트론 6으로의 초과번역을 포함하며, 초과번역은 *HSD17B13* rs72613567 변종 유전자에 존재하는 삽입된 티민을 포함하지 않는다. 따라서, 본원에는 HSD17B13 단백질의 전부 또는 일부를 암호화

하는 적어도 15개의 연속 뉴클레오타이드 (예컨대, 적어도 20개의 연속 뉴클레오타이드 또는 적어도 30개의 연속 뉴클레오타이드)를 포함하는 분리된 핵산이 제공되며, 여기서 연속 뉴클레오타이드 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)의 분절은 SEQ ID NO: 246과 최적으로 정렬될 때 SEQ ID NO: 246 (*HSD17B13* 전사물 F')의 인트론 6으로의 초과번역 내의 영역 또는 SEQ ID NO: 246의 인트론 6으로의 초과번역과 엑손 6의 나머지 사이의 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일하다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 (예컨대, 다른 *HSD17B13* 전사물의 엑손 6 및 7의 경계)으로부터의 초과번역과 구별하기 위하여 초과번역에 충분한 수의 뉴클레오타이드일 것으로 인지된다. 선택적으로, 연속 뉴클레오타이드는 전사물 F (SEQ ID NO: 9)에 존재하지 않는 전사물 F'에 존재하는 서열을 포함한다. 전사물 F의 초과번역은 *HSD17B13* rs72613567 변종 유전자에 존재하는 삽입된 티민을 포함하는 반면, 전사물 F'의 초과번역은 그렇지 않다. 예를 들어, 영역은 SEQ ID NO: 246에서 인트론 6으로의 초과번역 전체일 수 있다. 선택적으로, 분리된 핵산은 SEQ ID NO: 246 (*HSD17B13* 전사물 F')에 제시된 서열에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 서열을 포함하며 SEQ ID NO: 247 (HSD17B13 동형 F')에 제시된 서열을 포함하는, 또는 본질적으로 그것으로 구성되는, 또는 그것으로 구성되는 HSD17B13 단백질을 암호화한다.

[0179] 한 예로서, 분리된 핵산은 인트론 6으로의 초과번역 내의 영역 또는 인트론 6으로의 초과번역과 엑손 6의 나머지 사이의 경계에 걸쳐 있는 영역을 포함하는 SEQ ID NO: 246의 적어도 15개의 연속 뉴클레오타이드 (예컨대, 적어도 20개의 연속 뉴클레오타이드 또는 적어도 30개의 연속 뉴클레오타이드)를 포함할 수 있고, 선택적으로 인트론 6으로의 초과번역 전체를 포함할 수 있으며, 선택적으로 SEQ ID NO: 246의 전체 서열을 포함할 수 있다.

[0180] 또한 본원에는 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 C (또는 이것의 단편 또는 상동체)에 존재하는 분절 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는 분리된 핵산이 제공된다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 전사물 C (SEQ ID NO: 6)는 전사물 A와 비교하여 엑손 6이 없다 (즉, 전사물 C는 전사물 A에 존재하지 않는 엑손 5-엑손 7 경계를 포함한다). 따라서, 본원에는 HSD17B13 단백질의 전부 또는 일부를 암호화하는 적어도 15개의 연속 뉴클레오타이드 (예컨대, 적어도 20개의 연속 뉴클레오타이드 또는 적어도 30개의 연속 뉴클레오타이드)를 포함하는 분리된 핵산이 제공되며, 여기서 연속 뉴클레오타이드 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)의 분절은 SEQ ID NO: 6과 최적으로 정렬될 때 SEQ ID NO: 6 (*HSD17B13* 전사물 C)의 엑손 5-엑손 7에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일하다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 (예컨대, 다른 *HSD17B13* 전사물의 엑손 5 및 6 또는 엑손 6 및 7의 경계)과 구별하기 위하여 엑손 5 및 7의 각각에 충분한 수의 뉴클레오타이드를 가질 것으로 인지된다. 예를 들어, 영역은 SEQ ID NO: 6의 엑손 5 및 7의 전체를 포함할 수 있다. 선택적으로, 분리된 핵산은 SEQ ID NO: 6 (*HSD17B13* 전사물 C)에 제시된 서열에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 서열을 포함하며 SEQ ID NO: 14 (HSD17B13 동형 C)에 제시된 서열을 포함하는 HSD17B13 단백질을 암호화한다.

[0181] 한 예로서, 분리된 핵산은 엑손 5-엑손 7 경계에 걸쳐 있는 영역을 포함하는 SEQ ID NO: 6의 적어도 15개의 연속 뉴클레오타이드 (예컨대, 적어도 20개의 연속 뉴클레오타이드 또는 적어도 30개의 연속 뉴클레오타이드)를 포함할 수 있고, 선택적으로 SEQ ID NO: 6의 엑손 5 및 7의 전체를 포함할 수 있으며, 선택적으로 SEQ ID NO: 6의 전체 서열을 포함할 수 있다.

[0182] ***(4) cDNAs 및 변종 HSD17B13 전사물에 혼성화하는 핵산***

[0183] 또한 본원에는 전사물 A 내지 H 중 임의의 하나와 최적으로 정렬될 때 전사물 A 내지 H (각각 SEQ ID NO: 4 내지 11), 특히 전사물 C 내지 H 중 임의의 하나에 상응하는 mRNA 전사물 또는 cDNA의 분절에 혼성화하는 핵산이 제공된다. 구체적이고, 비제한적인 실례들이 하기에 제공된다. 이러한 분리된 핵산은 유용한, 예를 들어, 프라이머, 프로브, 안티센스 RNA, siRNA, 또는 shRNA일 수 있다.

[0184] 분리된 핵산이 혼성화할 수 있는 분절은 HSD17B13 단백질을 암호화하는 핵산의, 예를 들어 적어도 5, 적어도 10, 또는 적어도 15개의 연속 뉴클레오타이드를 포함할 수 있다. 분리된 핵산이 혼성화할 수 있는 분절은 HSD17B13 단백질을 암호화하는 핵산의, 예를 들어, 적어도 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 75, 90, 95, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 또는 2000 연속 뉴클레오타이드를 포함할 수 있다. 대안적으로, 분리된 핵산이 혼성화할 수 있는 분절은 HSD17B13 단백질을 암호화하는 핵산의, 예를 들어, 최대 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 75, 90, 95, 100, 200, 300, 400, 500,

600, 700, 800, 900, 또는 1000개의 연속 뉴클레오타이드일 수 있다. 예를 들어, 분절은 길이가 약 15 내지 100개 뉴클레오타이드, 또는 길이가 약 15 내지 35개 뉴클레오타이드일 수 있다.

[0185] *HSD17B13* 전사물 D (SEQ ID NO: 7), 전사물 G (SEQ ID NO: 10), 및 전사물 H (SEQ ID NO: 11)는 엑손 6의 3' 단부에 구아닌의 삽입을 포함하여, 전사물 A와 비교하여 엑손 7에서의 프레임시프트 및 엑손 7의 미성숙 절단을 초래한다. 따라서, 본원에는 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 D, G, 및 H (또는 이것의 단편 또는 상동체)에 존재하는 분절에 혼성화하는 영역 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는 분리된 핵산이 제공된다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 본원에는 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하는 분리된 핵산이 제공되고, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 7과 최적으로 정렬될 때 SEQ ID NO: 7 (*HSD17B13* 전사물 D)의 엑손 6-엑손 7 경계에 걸쳐 있는 영역에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)을 포함하며, 분절은 SEQ ID NO: 7의 엑손 6의 3' 단부에서 잔기 878에 상응하는 잔기에 구아닌을 포함한다 (즉, 엑손 7이 시작되는 곳의 구아닌에 더불어 전사물 A에 비교하여 엑손 6의 3' 단부에 구아닌이 삽입됨). 대안적으로, 본원에는 HSD17B13 단백질을 암호화하는 핵산의 분절의 적어도 15개의 연속 뉴클레오타이드에 혼성화하는 분리된 핵산이 제공되며, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 10과 최적으로 정렬될 때 SEQ ID NO: 10 (*HSD17B13* 전사물 G)의 엑손 6-엑손 7 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)을 포함하고, 분절은 SEQ ID NO: 10의 엑손 6의 3' 단부에서 잔기 770에 상응하는 잔기에 구아닌을 포함한다 (즉, 엑손 7이 시작되는 곳의 구아닌에 더불어 전사물 B에 비교하여 엑손 6의 3' 단부에 구아닌이 삽입됨). 대안적으로, 본원에는 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하는 분리된 핵산이 제공되며, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 11과 최적으로 정렬될 때 SEQ ID NO: 11 (*HSD17B13* 전사물 H)의 엑손 6-엑손 7 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)을 포함하고, 분절은 SEQ ID NO: 11의 엑손 6의 3' 단부에서 잔기 950에 상응하는 잔기에 구아닌을 포함한다 (즉, 엑손 7이 시작되는 곳의 구아닌에 더불어 전사물 E에 비교하여 엑손 6의 3' 단부에 구아닌이 삽입됨). 이러한 핵산은 삽입된 구아닌을 *HSD17B13* 전사물의 다른 특징 (예컨대, 전사물 F에서 인트론 6으로의 초과번역 또는 전사물 C에서 삭제된 엑손 6)과 구별하기 위하여 엑손 6 및 7의 각각에서 충분한 수의 뉴클레오타이드에 혼성화하도록 설계될 수 있을 것으로 인지된다.

[0186] 한 예로서, 분절은 엑손 6-엑손 7 경계에 걸쳐 있는 (즉, SEQ ID NO: 7의 잔기 878에서 구아닌을 포함하는) SEQ ID NO: 7의 영역을 포함할 수 있다. 다른 예로서, 분절은 엑손 6-엑손 7 경계에 걸쳐 있는 (즉, SEQ ID NO: 10의 잔기 770에서 구아닌을 포함하는) SEQ ID NO: 10의 영역을 포함할 수 있다. 또 다른 예로서, 분절은 엑손 6-엑손 7 경계에 걸쳐 있는 (즉, SEQ ID NO: 11의 잔기 950에서 구아닌을 포함하는) SEQ ID NO: 11의 영역을 포함할 수 있다.

[0187] 선택적으로, 분리된 핵산은 추가로 전사물 G (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 D (또는 이것의 단편 또는 상동체)에 존재하는 분절에 혼성화하는 영역 (예컨대, 15 연속 뉴클레오타이드)을 포함하고, 분리된 핵산은 추가로 전사물 H (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 D (또는 이것의 단편 또는 상동체)에 존재하는 분절에 혼성화하는 영역을 포함한다. 이러한 분절은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 전사물 H (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 D (또는 이것의 단편 또는 상동체)에 존재하는 분절 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)은 전사물 H와 구별하기 위하여 SEQ ID NO: 7과 최적으로 정렬될 때 SEQ ID NO: 7 (*HSD17B13* 전사물 D)의 엑손 3 및 4의 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일할 수 있다. 마찬가지로, 전사물 G (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 D (또는 이것의 단편 또는 상동체)에 존재하는 분절 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)은 전사물 G와 구별하기 위하여 SEQ ID NO: 7과 최적으로 정렬될 때 SEQ ID NO: 7 (*HSD17B13* 전사물 D)의 엑손 2 내의 영역, SEQ ID NO: 7의 엑손 1-엑손 2 경계에 걸쳐 있는 영역, 또는 SEQ ID NO: 7의 엑손 2-엑손 3 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또

는 적어도 99% 동일할 수 있다.

[0188] 전사물 D와 마찬가지로, 전사물 H (SEQ ID NO: 11)는 전사물 A와 비교하여 엑손 6의 3' 단부에 구아닌의 삽입을 포함한다. 전사물 H는 추가로 전사물 A 및 전사물 D와 비교하여 엑손 3과 4 사이에 추가적인 엑손을 포함한다. 따라서, 본원에는 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 D, G, 및 H (또는 이것의 단편들 또는 상동체들)에 존재하는 분절에 혼성화하는 영역을 포함하지만, 추가로 전사물 D (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 H (또는 이것의 단편 또는 상동체)에 존재하는 분절에 혼성화하는 영역 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는, 상기에서 기술된 분리된 핵산이 제공된다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 분절은 SEQ ID NO: 11과 최적으로 정렬될 때 SEQ ID NO: 11 (*HSD17B13* 전사물 H)의 엑손 3' 내의 영역 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드), SEQ ID NO: 11의의 엑손 3-엑손 3' 경계에 걸쳐 있는 영역, 또는 EQ ID NO: 11의의 엑손 3'-엑손 4 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일할 수 있다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 (예컨대, 엑손 3 및 4의 경계)과 구별하기 위하여 엑손 3 및 3'의 각각 또는 엑손 3'과 4의 각각에서 충분한 수의 뉴클레오타이드에 혼성화하도록 설계될 수 있을 것으로 인지된다.

[0189] 한 예로서, 분절은 엑손 3' 내의, 엑손 3-엑손 3' 경계에 걸쳐 있는, 또는 엑손 3'-엑손 4 경계에 걸쳐 있는 SEQ ID NO: 11의 영역을 포함할 수 있다.

[0190] 전사물 D와 마찬가지로, 전사물 G (SEQ ID NO: 10)는 전사물 A와 비교하여 엑손 6의 3' 단부에 구아닌의 삽입을 포함한다. 그러나, 이에 더불어, 전사물 G는 전사물 A 및 전사물 D와 비교하여 엑손 2가 없다 (즉, 전사물 G는 전사물 A 및 전사물 D에 존재하지 않는 엑손 1-엑손 3 경계를 포함한다). 따라서, 본원에는 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 D, G, 및 H (또는 이것의 단편들 또는 상동체들)에 존재하는 분절에 혼성화하는 영역을 포함하지만, 추가로 전사물 D (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 G (또는 이것의 단편 또는 상동체)에 존재하는 분절에 혼성화하는 영역 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는, 상기에서 기술된 분리된 핵산이 제공된다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 분절은 SEQ ID NO: 10과 최적으로 정렬될 때 SEQ ID NO: 10 (*HSD17B13* 전사물 G)의 엑손 1-엑손 3 경계에 걸쳐 있는 영역 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일할 수 있다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 (예컨대, 엑손 1 및 2의 경계 또는 엑손 2 및 3의 경계)과 구별하기 위하여 엑손 1 및 3의 각각에서 충분한 수의 뉴클레오타이드에 혼성화하도록 설계될 수 있을 것으로 인지된다.

[0191] 한 예로서, 분절은 엑손 1-엑손 3 경계에 걸쳐 있는 SEQ ID NO: 10의 영역을 포함할 수 있다.

[0192] 또한 본원에는 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않지만 전사물 E (또는 이것의 단편 또는 상동체)에 존재하는 HSD17B13 단백질을 암호화하는 핵산의 분절에 혼성화하는 영역 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는 분리된 핵산이 제공된다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 전사물 E (SEQ ID NO: 8)는 전사물 A와 비교하여 엑손 3과 4 사이에 추가적인 엑손을 포함한다. 따라서, 본원에는 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하는 분리된 핵산이 제공되며, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 8과 최적으로 정렬될 때 SEQ ID NO: 8 (*HSD17B13* 전사물 E)의 엑손 3' 내의 영역 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드), SEQ ID NO: 8의 엑손 3-엑손 3' 경계에 걸쳐 있는 영역, 또는 SEQ ID NO: 8의 엑손 3'-엑손 4 경계에 걸쳐 있는 영역에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절을 포함한다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 (예컨대, 엑손 3 및 4의 경계)과 구별하기 위하여 엑손 3 및 3'의 각각 또는 엑손 3' 및 4의 각각에서 충분한 수의 뉴클레오타이드와 혼성화하도록 설계될 수 있을 것으로 인지된다.

[0193] 한 예로서, 분절은 엑손 3' 내의, SEQ ID NO: 8의 엑손 3-엑손 3' 경계에 걸쳐 있는, 또는 엑손 3'-엑손 4 경계에 걸쳐 있는 SEQ ID NO: 8의 영역을 포함할 수 있다.

[0194] 선택적으로, 분리된 핵산은 추가로 전사물 H (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 E (또는 이것의 단편 또는 상동체)에 존재하는 분절에 혼성화하는 영역 (예컨대, 15개의 연속 뉴클레오타이드)을 포함한다. 이러한 분절들은 전사물의 서열을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 전사물 H (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 E (또는 이것의 단편 또는 상동체)에 존재하는 분절 (예컨대, 적어

도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)은 전사물 G와 구별하기 위하여 SEQ ID NO: 8과 최적으로 정렬될 때 SEQ ID NO: 8 (*HSD17B13* 전사물 E)의 엑손 6과 7의 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일할 수 있다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 (특히 전사물 H에서 엑손 6의 3' 단부에 있는 추가적인 구아닌)과 구별하기 위하여 엑손 6 및 7의 각각에서 충분한 수의 뉴클레오타이드에 혼성화하도록 설계될 수 있을 것으로 인지된다.

[0195] 또한 본원에는 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않지만 전사물 F (또는 이것의 단편 또는 상동체)에 존재하는 HSD17B13 단백질을 암호화하는 핵산의 분절에 혼성화하는 영역 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는 분리된 핵산이 제공된다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 전사물 F (SEQ ID NO: 9)는 전사물 A와 비교하여 엑손 6으로부터 인트론 6으로의 초과번역을 포함한다. 따라서, 본원에는 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하는 분리된 핵산이 제공되고, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 9와 최적으로 정렬될 때 SEQ ID NO: 9 (*HSD17B13* 전사물 F)의 인트론 6으로의 초과번역 내의 영역 또는 SEQ ID NO: 9의 인트론 6으로의 초과번역과 엑손 6의 나머지 사이의 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)을 포함한다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 (예컨대, 다른 *HSD17B13* 전사물의 엑손 6과 7의 경계)으로부터의 초과번역과 구별하기 위하여 초과번역에서의 충분한 수의 뉴클레오타이드에 혼성화하도록 설계될 수 있을 것으로 인지된다. 선택적으로, 연속 뉴클레오타이드는 전사물 F' (SEQ ID NO: 246)에 존재하지 않는 전사물 F에 존재하는 서열 (즉, 삽입된 티민)을 포함한다. 전사물 F'는 또한 전사물 A와 비교하여 엑손 6으로부터 인트론 6으로의 초과번역을 포함하지만, 초과번역은 *HSD17B13* rs72613567 변종 유전자에 존재하는 삽입된 티민을 포함하지 않는다.

[0196] 한 예로서, 분절은 인트론 6으로의 초과번역 내의 또는 인트론 6으로의 초과번역과 엑손 6의 나머지 사이의 경계에 걸쳐 있는 SEQ ID NO: 9의 영역을 포함할 수 있다.

[0197] 또한 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않지만 전사물 F' (또는 이것의 단편 또는 상동체)에 존재하는 HSD17B13 단백질을 암호화하는 핵산의 분절에 혼성화하는 영역 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는 분리된 핵산이 제공된다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 전사물 F' (SEQ ID NO: 246)는 전사물 A와 비교하여 엑손 6으로부터 인트론 6으로의 초과번역을 포함한다. 따라서, 본원에는 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하는 분리된 핵산이 제공되고, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 246과 최적으로 정렬될 때 SEQ ID NO: 246 (*HSD17B13* 전사물 F')의 인트론 6으로의 초과번역 내의 영역 또는 SEQ ID NO: 246의 인트론 6으로의 초과번역과 엑손 6의 나머지 사이의 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)을 포함한다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 (예컨대, 다른 *HSD17B13* 전사물의 엑손 6과 7의 경계)으로부터의 초과번역과 구별하기 위하여 초과번역에서의 충분한 수의 뉴클레오타이드에 혼성화하도록 설계될 수 있을 것으로 인지된다. 선택적으로, 연속 뉴클레오타이드는 전사물 F (SEQ ID NO: 9)에 존재하지 않는 전사물 F'에 존재하는 서열을 포함한다. 전사물 F의 초과번역은 *SD17B13* rs72613567 변종 유전자에 존재하는 삽입된 티민을 포함하는 반면, 전사물 F'의 초과번역은 그렇지 않다.

[0198] 한 예로서, 분절은 인트론 6으로의 초과번역 내의 또는 인트론 6으로의 초과번역과 엑손 6의 나머지 사이의 경계에 걸쳐 있는 SEQ ID NO: 246의 영역을 포함할 수 있다.

[0199] 또한 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않지만 전사물 C (또는 이것의 단편 또는 상동체)에 존재하는 HSD17B13 단백질을 암호화하는 핵산의 분절에 혼성화하는 영역 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함하는 분리된 핵산이 제공된다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. 전사물 C (SEQ ID NO: 6)는 전사물 A와 비교하여 엑손 6이 없다 (즉, 전사물 C는 전사물 A에 존재하지 않는 엑손 5-엑손 7 경계를 포함한다). 따라서, 본원에는 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하는 분리된 핵산이 제공되고, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 6과 최적으로 정렬될 때 SEQ ID NO: 6 (*HSD17B13* 전사물 C)의 엑손 5-엑손 7 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)을

포함한다. 이러한 핵산은 *HSD17B13* 전사물의 다른 특징 (예컨대, 다른 *HSD17B13* 전사물의 엑손 5와 6의 경계 또는 엑손 6과 7의 경계)과 구별하기 위하여 엑손 5 및 7의 충분한 수의 뉴클레오타이드에 혼성화하도록 설계될 수 있을 것으로 인지된다.

[0200] 한 예로서, 분절은 엑손 5-엑손 7 경계에 걸쳐 있는 SEQ ID NO: 6으로부터의 영역을 포함할 수 있다.

[0201] 또한 본원에는 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하는 분리된 핵산 (예컨대, 안티센스 RNA, siRNA, 또는 shRNA)이 제공되고, 여기서 연속 뉴클레오타이드는 *HSD17B13* 전사물 D (SEQ ID NO: 7)의 영역에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)을 포함한다. 분리된 핵산은 전사물 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 D (또는 이것의 단편 또는 상동체)에 존재하는 분절에 혼성화하는 영역 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함할 수 있다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. *HSD17B13* 전사물 D (SEQ ID NO: 7)는 엑손 6의 3' 단부에 구아닌의 삽입을 포함하여, 전사물 A (SEQ ID NO: 4)와 비교하여 엑손 7에 프레임시프트 및 엑손 7의 미성숙 절단을 초래한다. 예를 들어, 본원에는 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하는 분리된 핵산이 제공되고, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 7과 최적으로 정렬될 때 SEQ ID NO: 7 (*HSD17B13* 전사물 D)의 엑손 6-엑손 7 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)을 포함한다. 분절은 SEQ ID NO: 7의 엑손 6의 3' 단부에서 잔기 878에 상응하는 잔기에서 구아닌을 포함할 수 있다 (즉, 엑손 7이 시작되는 곳의 구아닌에 더불어 전사물 A와 비교하여 엑손 6의 3' 단부에서 구아닌이 삽입됨). 이러한 핵산은 *HSD17B13* 전사물에서 다른 특징 (예컨대, 전사물 F에서 인트론 6으로의 초과번여 또는 전사물 C에서 삭제된 엑손 6)과 삽입된 구아닌을 구별하기 위하여 엑손 6 및 7의 각각에서 충분한 수의 뉴클레오타이드에 혼성화하도록 설계될 수 있을 것으로 인지된다.

[0202] 또한 본원에는 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하는 분리된 핵산 (예컨대, 안티센스 RNA, siRNA, 또는 shRNA)이 제공되며, 여기서 연속 뉴클레오타이드는 *HSD17B13* 전사물 A (SEQ ID NO: 4)의 영역에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)을 포함한다. 분리된 핵산은 전사물 D (또는 이것의 단편 또는 상동체)에 존재하지 않는 전사물 A (또는 이것의 단편 또는 상동체)에 존재하는 분절에 혼성화하는 영역 (예컨대, 적어도 15개의 연속 뉴클레오타이드)을 포함할 수 있다. 이러한 영역들은 전사물의 서열들을 비교함으로써 쉽게 확인될 수 있다. *HSD17B13* 전사물 D (SEQ ID NO: 7)는 엑손 6의 3' 단부에 구아닌의 삽입을 포함하여, 전사물 A (SEQ ID NO: 4)와 비교하여 엑손 7에서 프레임시프트 및 엑손 7의 미성숙 절단을 초래한다. 예를 들어, 본원에는 HSD17B13 단백질을 암호화하는 핵산의 적어도 15개의 연속 뉴클레오타이드에 혼성화하는 분리된 핵산이 제공되며, 여기서 연속 뉴클레오타이드는 SEQ ID NO: 4와 최적으로 정렬될 때 SEQ ID NO: 4 (*HSD17B13* 전사물 A)의 엑손 6-엑손 7 경계에 걸쳐 있는 영역에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 분절 (예컨대, 적어도 5개의 연속 뉴클레오타이드, 적어도 10개의 연속 뉴클레오타이드 또는 적어도 15개의 연속 뉴클레오타이드)을 포함한다.

[0203] ***(5) 벡터***

[0204] 또한 본원에 개시된 핵산들 중 임의의 핵산 및 이종성 핵산을 포함하는 벡터가 제공된다. 벡터는 핵산을 수송할 수 있는 바이러스 또는 비바이러스 벡터일 수 있다. 일부 경우에, 벡터는 플라스미드 (예컨대, 추가적인 DNA 분절이 그 안으로 결찰될 수 있는 원형의 이중 가닥 DNA)일 수 있다. 일부 경우에, 벡터는 바이러스 벡터일 수 있고, 이때 추가적인 DNA 분절은 바이러스 게놈 안으로 결찰될 수 있다. 일부 경우에, 벡터는 그것이 도입되는 숙주 세포에서 자율적으로 복제할 수 있다 (예컨대, 박테리아 복제 기원을 가지는 박테리아 벡터 및 에피솜성 포유류 벡터). 다른 경우에, 벡터 (예컨대, 비에피솜성 포유류 벡터)는 숙주 세포에 도입될 때 숙주 세포의 게놈으로 통합될 수 있고 그로써 숙주 게놈과 함께 복제된다. 더욱이, 특정 벡터는 그것들이 작동 가능하게 연결된 유전자들의 발현을 지시할 수 있다. 이러한 벡터는 "재조합 발현 벡터" 또는 "발현 벡터"로서 언급될 수 있다. 이러한 벡터는 또한 본원의 다른 곳에서 개시되는 바 표적화 벡터 (즉, 외인성 도너 서열)일 수 있다.

[0205] 일부 경우에, 개시된 유전자 변종들에 의해 암호화된 단백질은 유전자들이 전사 및 번역 제어 서열과 같은 필수 발현 제어 서열들에 작동 가능하게 연결되도록 개시된 유전자 변종들을 암호화하는 핵산을 발현 벡터 안으로 삽

입함으로써 발현된다. 발현 벡터로는, 예를 들어, 플라스미드, 레트로바이러스, 아데노바이러스, 아데노-관련 바이러스 (AAV), 콜리플라워 모자이크 바이러스, 담배 모자이크 바이러스와 같은 식물 바이러스, 코스미드, YAC, EBV 유래 에피솜, 등을 들 수 있다. 일부 경우에, 개시된 유전자 변종을 포함하는 핵산은 벡터 내의 전사 및 번역 제어 서열들이 유전자 변종의 전사 및 번역을 조절하는 그것들의 의도된 기능을 실시하도록 벡터 안으로 결찰될 수 있다. 발현 벡터 및 발현 제어 서열들은 사용된 발현 숙주 세포와 부합하도록 선택된다. 개시된 유전자 변종을 포함하는 핵산 서열은 별도의 벡터 안으로 또는 동일한 발현 벡터 안으로 삽입될 수 있다. 개시된 유전자 변종을 포함하는 핵산 서열은 표준 방법 (예컨대, 개시된 유전자 변종 및 벡터를 포함하는 핵산 상의 상보하는 제한 부위들의 결찰, 또는 제한 부위가 없는 경우 블런트 단부 결찰)에 의해 발현 벡터 안으로 삽입될 수 있다.

[0206] 개시된 유전자 변종을 포함하는 핵산 서열에 더불어, 재조합 발현 벡터는 숙주 세포에서 유전자 변종의 발현을 제어하는 조절 서열들을 포함할 수 있다. 조절 서열들의 선택을 포함한, 발현 벡터의 설계는 형질전환될 숙주 세포의 선택, 원하는 단백질의 발현 수준, 등과 같은 인자들에 좌우된다. 포유류 숙주 세포 발현을 위한 바람직한 조절 서열로는, 예를 들어, 포유류 세포에서 고수준의 단백질 발현을 지시하는 바이러스 요소들, 예컨대 레트로바이러스 LTR로부터 유래된 프로모터 및/또는 인핸서, 사이토메갈로바이러스 (CMV)로부터 유래된 프로모터 및/또는 인핸서 (예컨대 CMV 프로모터/인핸서), 유인원 바이러스 40 (SV40)으로부터 유래된 프로모터 및/또는 인핸서 (예컨대 SV40 프로모터/인핸서), 아데노바이러스로부터 유래된 프로모터 및/또는 인핸서 (예컨대, 아데노바이러스 주요 후기 프로모터 (AdMLP)), 폴리오마바이러스 및 천연 면역글로불린 및 액틴 프로모터와 같은 강력한 포유류 프로모터를 들 수 있다. 바이러스 조절 요소들, 이것들의 서열에 대한 추가의 설명은 미국 특허 제 5,168,062호; 4,510,245호; 및 4,968,615호에서 제공되며, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다. 박테리아 세포 또는 진균 세포 (예컨대 효모 세포)에서 폴리펩타이드를 발현시키는 방법 또한 잘 알려져 있다.

[0207] 개시된 유전자 변종 및 조절 서열을 포함하는 핵산 서열에 더불어, 재조합 발현 벡터는 추가적인 서열, 예컨대 숙주 세포에서 벡터의 복제를 조절하는 서열 (예컨대 복제 기원) 및 선택 가능한 마커 유전자를 포함할 수 있다. 선택 가능한 마커 유전자는 벡터가 도입되어 들어가는 숙주 세포의 선택을 용이하게 할 수 있다 (예컨대, 미국 특허 제 4,399,216호; 4,634,665호; 및 5,179,017호 참조, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 예를 들어, 선택 가능한 마커 유전자는 벡터가 도입되는 숙주 세포에 대해 약물, 예컨대 G418, 하이그로마이신, 또는 메토트렉세이트에 대한 내성을 부여할 수 있다. 예시적인 선택 가능한 마커 유전자로는 다이하이드로폴레이트 환원효소 (DHFR) 유전자 (메토트렉세이트 선택/증폭으로 dhfr-숙주 세포에서 사용하기 위함), 네오 유전자 (G418 선택을 위함), 및 글루타메이트 합성효소 (GS) 유전자를 들 수 있다.

[0208] **B. 단백질**

[0209] 본원에는 분리된 *HSD17B13* 단백질 및 이것의 단편, 특히 *HSD17B13* rs72613567 변종에 의해 생성된 *HSD17B13* 단백질 및 이것의 단편이 개시된다.

[0210] 본원에 개시된 분리된 단백질은 자연적으로 발생하는 HSD17B13 단백질의 아미노산 서열을 포함하거나, 또는 자연적으로 발생하지 않는 서열을 포함할 수 있다. 한 예에서, 자연적으로 발생하지 않는 서열은 보존성 아미노산 치환으로 인해 자연적으로 발생하지 않는 서열과 상이할 수 있다. 예를 들어, 서열은 보존성 아미노산 치환을 제외하고 동일할 수 있다.

[0211] 본원에 개시된 분리된 단백질은 이종성 폴리펩타이드 또는 이종성 분자 또는 표지에 연결되거나 융합될 수 있고, 이것의 수많은 예가 본원의 다른 곳에서 개시된다. 예를 들어, 단백질은 증가된 또는 감소된 안정성을 제공하는 이종성 폴리펩타이드에 융합될 수 있다. 융합된 도메인 또는 이종성 폴리펩타이드는 단백질의 N-말단에, C-말단에, 또는 단백질 내부에 위치할 수 있다. 융합 파트너는, 예를 들어, T 헬퍼 에피토프를 제공하는 것을 보조할 수 있거나 (면역학적 융합 파트너), 또는 천연 재조합 단백질보다 높은 수율로 단백질을 발현시키는 것을 보조할 수 있다 (발현 인핸서). 특정 융합 파트너는 면역학적 및 발현 증강 융합 파트너 둘 다이다. 다른 융합 파트너는 폴리펩타이드의 용해도를 증가시키기 위하여 또는 폴리펩타이드가 원하는 세포내 구획으로 표적화되는 것을 가능하게 하기 위하여 선택될 수 있다. 또 다른 융합 파트너로는 친화성 태그를 들 수 있고, 이것은 폴리펩타이드의 정제를 용이하게 한다.

[0212] 융합 단백질은 이종성 분자에 직접 융합되거나 링커, 예컨대 펩타이드 링커를 통해 이종성 분자에 연결될 수 있다. 적합한 펩타이드 링커 서열은, 예를 들어, 다음의 인자들을 기반으로 선택될 수 있다: (1) 가요성 확대된 형태를 채택하는 능력; (2) 제1 및 제2 폴리펩타이드 상의 기능성 에피토프들과 상호작용할 수 있는 이차 구조

를 채택하지 못하는 무능력; 및 (3) 폴리펩타이드 기능성 에피토프와 반응할 소수성 또는 대전된 잔기의 결핍. 예를 들어, 펩타이드 링커 서열은 Gly, Asn 및 Ser 잔기를 함유할 수 있다. 다른 거의 중성 아미노산, 예컨대 Thr 및 Ala가 또한 링커 서열에 사용될 수 있다. 링커로서 유용하게 사용될 수 있는 아미노산 서열로는 Maratea et al. (1985) Gene 40:39-46; Murphy et al. (1986) *Proc. Natl. Acad. Sci. USA* 83:8258-8262; 미국 특허 제 4,935,233호; 및 미국 특허 제 4,751,180호 (이들 문헌은 각각 그 전문이 본원에 참조로 포함됨)에 개시된 것들을 들 수 있다. 링커 서열은 일반적으로 길이가, 예를 들어, 1 내지 약 50개 아미노산일 수 있다. 링커 서열은 일반적으로 제1 및 제2 폴리펩타이드가 기능성 도메인을 분리하고 입체적 간섭을 방지하기 위해 사용될 수 있는 비필수 N-말단 아미노산 영역을 가지고 있을 때 필요하지 않다.

[0213] 단백질은 또한 세포-투과 도메인에 작동 가능하게 연결될 수 있다. 예를 들어, 세포-투과 도메인은 HIV-1 TAT 단백질, 인간 B형 간염 바이러스로부터의 TLM 세포-투과 모티프, MPG, Pep-1, VP22, 단순 포진 바이러스로부터의 세포-투과 펩타이드, 또는 폴리아르기닌 펩타이드 서열로부터 유래될 수 있다. 예컨대, WO 2014/089290 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)을 참조한다. 세포-투과 도메인은 단백질의 N-말단에, C-말단에, 또는 단백질 내의 어느 곳에든지 위치할 수 있다.

[0214] 단백질은 또한 추적 또는 정제를 용이하게 하기 위해 이종성 폴리펩타이드, 예컨대 형광 단백질, 정제 태그, 또는 에피토프 태그에 작동 가능하게 연결될 수 있다. 형광 단백질의 예로는 녹색 형광 단백질 (예컨대, GFP, GFP-2, tagGFP, turboGFP, eGFP, 에메랄드(Emerald), 아자미 그린(Azami Green), 모노머 아자미 그린, CopGFP, AceGFP, ZsGreen1), 황색 형광 단백질 (예컨대, YFP, eYFP, 시트린(Citrine), 비너스(Venus), YPet, PhiYFP, ZsYellow1), 청색 형광 단백질 (예컨대, eBFP, eBFP2, 아즈라이트(Azurite), mKalama1, GFPuv, 사파이어(Sapphire), T-사파이어), 청록색 형광 단백질 (예컨대, eCFP, 세룰리언(Cerulean), CyPet, AmCyan1, Midoriishi-Cyan), 적색 형광 단백질 (예컨대, mKate, mKate2, mPlum, DsRed 모노머, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-모노머, HcRed-탠덤, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred), 오렌지색 형광 단백질 (예컨대, mOrange, mKO, Kusabira-Orange, 모노머 Kusabira-Orange, mTangerine, tdTomato), 및 임의의 기타 적합한 형광 단백질을 들 수 있다. 태그의 예로는 글루타티온-S-트랜스퍼라제 (GST), 키틴 결합 단백질 (CBP), 말토오스 결합 단백질, 티오레독신 (TRX), 폴리(NANP), 탠덤 친화도 정제 (TAP) 태그, myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, 헤마글루티닌 (HA), nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 히스티딘 (His), 비오틴 카르복실 담체 단백질 (BCCP), 및 칼모듈린을 들 수 있다.

[0215] 본원의 분리된 단백질은 또한 비천연 또는 변형된 아미노산 또는 펩타이드 유사체를 포함할 수 있다. 예를 들어, 자연적으로 발생하는 아미노산과 상이한 기능성 치환체를 가진 수많은 D-아미노산 또는 아미노산들이 있다. 자연적으로 발생하는 펩타이드의 반대되는 입체 이성질체뿐만 아니라, 펩타이드 유사체의 입체 이성질체들이 개시된다. 이런 아미노산들은 tRNA 분자를 선택된 아미노산으로 대전시키고, 예를 들어, 앰버 코돈을 활용하여 유전자 구성물을 조작하여 유사 아미노산을 펩타이드 사슬에 부위-특이적 방식으로 삽입시킴으로써 쉽게 폴리펩타이드 사슬에 포함될 수 있다 (Thorson et al. (1991) *Methods Molec. Biol.* 77:43-73; Zoller (1992) *Current Opinion in Biotechnology* 3:348-354; Ibba, (1995) *Biotechnology & Genetic Engineering Reviews* 13:197-216; Cahill et al. (1989) *TIBS* 14(10):400-403; Benner (1993) *TIB Tech* 12:158-163; 및 Ibba 및 Hennecke (1994) *Biotechnology* 12:678-682, 이들 문헌은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0216] 펩타이드를 닮았지만, 천연 펩타이드 결합을 통해 연결되지 않은 분자들이 생성될 수 있다. 예를 들어, 아미노산 또는 아미노산 유사체에 대한 결합으로는 CH2NH--, --CH2S--, --CH2-- --, --CH=CH-- (시스 및 트랜스), --COCH2 --, --CH(OH)CH2--, 및 --CHH2SO--를 들 수 있다 (예컨대, Spatola, A.F. in Chemistry 및 Biochemistry of Amino Acids, Peptides, 및 Proteins, B. Weinstein, eds., Marcel Dekker, New York, p. 267 (1983); Spatola, A. F., Vega Data (March 1983), Vol. 1, Issue 3, Peptide Backbone Modifications (General review); Morley (1994) *Trends Pharm Sci* 15(12):463-468; Hudson et al. (1979) *Int J Pept Prot Res* 14:177-185; Spatola et al. (1986) *Life Sci* 38:1243-1249; Hann (1982) *Chem. Soc Perkin Trans.* I 307-314; Almquist et al. (1980) *J. Med. Chem.* 23:1392-1398; Jennings-White et al. (1982) *Tetrahedron Lett* 23:2533); Szelke et al. European Appln, EP 45665 CA (1982): 97:39405 (1982); Holladay et al. (1983) *Tetrahedron. Lett* 24:4401-4404; 및 Hruby (1982) *Life Sci* 31:189-199 참조; 이들 문헌은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). b-알라닌, 가미노부티르산, 등과 같은 펩타이드 유사체는 결합 원자들 사이에 하나 이상의 원자를 가질 수 있다.

[0217] 아미노산 유사체 및 펩타이드 유사체는 종종 증강된 또는 바람직한 특성, 예컨대, 보다 경제적인 생산, 더 큰 화학적 안정성, 증강된 약물학적 특성 (반감기, 흡수, 효능, 효율, 등등), 변경된 특이성 (예컨대, 광범위한 생물학적 활성), 감소된 항원성, 및 기타 바람직한 특성을 가진다.

[0218] D-아미노산은, D-아미노산이 펩티다제 및 그런 것들에 의해 인지되지 않기 때문에 보다 안정적인 펩타이드를 생성하기 위해 사용될 수 있다. 공통 서열의 하나 이상의 아미노산의 동일 유형의 D-아미노산 (예컨대, L-리신 대신 D-리신)으로의 체계적 치환은 보다 안정적인 펩타이드를 생성하기 위해 사용될 수 있다. 시스테인 잔기는 둘 이상의 펩타이드를 함께 고리화하기 위해 또는 부착시키기 위해 사용될 수 있다. 이것은 펩타이드를 특정 형태로 제한하기 위해 유익할 수 있다 (예컨대, Rizo and Gierasch (1992) *Ann. Rev. Biochem.* 61:387 참조, 모든 목적에 대해 그 전문이 본원에 참조됨).

[0219] 또한 본원에는 본원에 개시된 임의의 단백질을 암호화하는 핵산이 개시된다. 이것은 특이적 폴리펩타이드 서열에 관련된 모든 축퇴성 서열을 포함한다 (즉, 한 특정 폴리펩타이드 서열을 암호화하는 서열을 가지는 모든 핵산뿐만 아니라, 단백질 서열의 개시된 변종 및 유도체를 암호화하는, 축퇴성 핵산을 포함한, 모든 핵산). 그러므로, 각각의 특정 핵산 서열이 본원에서 다 표기될 수 없는 한편, 각각의 및 모든 서열은 개시된 폴리펩타이드 서열을 통해 실제로 본원에 개시되고 기술된다.

[0220] 또한 본원에는 본원에 개시된 분리된 폴리펩타이드 또는 단백질 및 분리된 폴리펩타이드의 안정성을 증가시키는 담체를 포함하는 조성물이 개시된다. 이러한 담체의 비제한적인 예로는 폴리(락트산) (PLA) 마이크로스피어, 폴리(D,L-락틱-코글리콜-산) (PLGA) 마이크로스피어, 리포좀, 미셀, 역미셀, 지질 코클레이트, 및 지질 미세관을 들 수 있다.

[0221] ***(1) HSD17B13 단백질 및 단편***

[0222] 본원에는 *HSD17B13* rs72613567 변종에 의해 생성된 분리된 HSD17B13 단백질 및 이것의 단편, 특히 HSD17B13 동형 C, D, E, F, F', G, 및 H가 개시된다. 이러한 단백질은, 예를 들어, HSD17B13 동형 C, D, E, F, F', G, 또는 H 또는 이것들의 단편의 적어도 5, 6, 8, 10, 12, 14, 15, 16, 18, 20, 22, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 또는 300개의 연속 아미노산을 포함하는 분리된 폴리펩타이드를 포함할 수 있다. 집단 내의 유전자 서열 및 이러한 유전자에 의해 암호화된 단백질은 단일-뉴클레오타이드 다형성과 같은 다형성으로 인해 다를 수 있는 것으로 인지된다. 각 HSD17B13 동형에 대해 본원에서 제공된 서열은 단지 예시적인 서열이다. 다른 서열들 또한 가능하다. 예를 들어, 분리된 폴리펩타이드는 각각 동형 C, D, E, F, F', G, 또는 H와 최적으로 정렬될 때 HSD17B13 동형 C, D, E, F, F', G, 또는 H에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 아미노산 서열 (예컨대, 연속 아미노산들의 서열)을 포함한다. 선택적으로, 분리된 폴리펩타이드는 HSD17B13 동형 C, D, E, F, F', G, 또는 H에 동일한 서열을 포함한다.

[0223] 한 예로서, 분리된 폴리펩타이드는 동형 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 동형 D, G, 및 H (또는 이것의 단편 또는 상동체)에 존재하는 분절 (예컨대, 적어도 8개의 연속 아미노산)을 포함할 수 있다. 이러한 영역들은 동형의 서열들을 비교함으로써 쉽게 확인될 수 있다. 동형 D, G, 및 H에서 엑손 7에 의해 암호화된 영역은 동형 A에서 엑손 7에 의해 암호화된 영역과 비교하여 프레임시프트되고 절단된다. 그러므로, 이러한 분리된 폴리펩타이드는 HSD17B13 단백질의 적어도 5, 6, 8, 10, 12, 14, 15, 16, 18, 20, 22, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 또는 200개의 연속 아미노산 (예컨대, HSD17B13 단백질의 적어도 8개의 연속 아미노산, 적어도 10개의 연속 아미노산, 또는 적어도 15개의 연속 아미노산)을 포함할 수 있고, 여기서 연속 아미노산 (예컨대, 적어도 3개의 연속 아미노산, 적어도 5개의 연속 아미노산, 적어도 8개의 연속 아미노산, 적어도 10개의 연속 아미노산, 또는 적어도 15개의 연속 아미노산)의 분절은 분리된 폴리펩타이드가 각각 SEQ ID NO: 15, 18, 또는 19와 최적으로 정렬될 때 SEQ ID NO: 15 (HSD17B13 동형 D), SEQ ID NO: 18 (HSD17B13 동형 G), 또는 SEQ ID NO: 19 (HSD17B13 동형 H)에서 엑손 7에 의해 암호화된 영역의 적어도 일부를 포함하는 분절에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다.

[0224] 이러한 분리된 폴리펩타이드는 추가로 동형 G (또는 이것의 단편 또는 상동체)에 존재하지 않는 동형 D (또는 이것의 단편 또는 상동체)에 존재하는 분절을 포함할 수 있고 추가로 동형 H (또는 이것의 단편 또는 상동체)에 존재하지 않는 동형 D (또는 이것의 단편 또는 상동체)에 존재하는 분절을 포함할 수 있다. 이러한 영역들은 동형의 서열들을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 이러한 분리된 폴리펩타이드는 동형 H와 구별하기 위하여 SEQ ID NO: 15와 최적으로 정렬될 때 SEQ ID NO: 15 (HSD17B13 동형 D)에서 엑손 3 및 4에 의해 암

호화된 영역들의 경계에 걸쳐 있는 분절에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 연속 아미노산 (예컨대, 적어도 3개의 연속 아미노산, 적어도 5개의 연속 아미노산, 적어도 8개의 연속 아미노산, 적어도 10개의 연속 아미노산, 또는 적어도 15개의 연속 아미노산)의 분절을 포함할 수 있다. 마찬가지로, 이러한 분리된 폴리펩타이드는 동형 G와 구별하기 위하여 SEQ ID NO: 15와 최적으로 정렬될 때 SEQ ID NO: 15 (HSD17B13 동형 D)에서 엑손 2에 의해 암호화된 영역 내의 분절, SEQ ID NO: 15에서 엑손 1 및 2에 의해 암호화된 영역들의 경계에 걸쳐 있는 분절, 또는 SEQ ID NO: 15에서 엑손 2 및 3에 의해 암호화된 영역들의 경계에 걸쳐 있는 분절에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 연속 아미노산 (예컨대, 적어도 3개의 연속 아미노산, 적어도 5개의 연속 아미노산, 적어도 8개의 연속 아미노산, 적어도 10개의 연속 아미노산, 또는 적어도 15개의 연속 아미노산)의 분절을 포함할 수 있다.

[0225] 동형 D와 마찬가지로, 동형 H (SEQ ID NO: 19)에서 엑손 7에 의해 암호화된 영역은 동형 A와 비교하여 프레임시프트되고 절단된다. 그러나, 이에 더불어, 동형 H는 동형 A 및 D와 비교하여 엑손 3과 4 사이에 추가적인 엑손 (엑손 3')에 의해 암호화된 영역을 포함한다. 따라서, 이러한 분리된 폴리펩타이드는 동형 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 동형 D, G, 및 H (또는 이것의 단편 또는 상동체)에 존재하는 분절을 포함하지만 동형 D (또는 이것의 단편 또는 상동체)에 존재하지 않는 동형 H (또는 이것의 단편 또는 상동체)로부터의 분절 (예컨대, 적어도 8개의 연속 아미노산)을 포함하는 상기에서 기술된 것과 같은 것일 수 있다. 이러한 영역들은 동형의 서열들을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 이러한 분리된 폴리펩타이드는 추가로 분리된 폴리펩타이드가 SEQ ID NO: 19와 최적으로 정렬될 때 SEQ ID NO: 19 (HSD17B13 동형 H)에서 엑손 3'에 의해 암호화된 영역의 적어도 일부를 포함하는 분절에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 연속 아미노산 (예컨대, 적어도 3개의 연속 아미노산, 적어도 5개의 연속 아미노산, 적어도 8개의 연속 아미노산, 적어도 10개의 연속 아미노산, 또는 적어도 15개의 연속 아미노산)의 분절을 포함할 수 있다.

[0226] 동형 D와 마찬가지로, 동형 G (SEQ ID NO: 18)에서 엑손 7에 의해 암호화된 영역은 동형 A와 비교하여 프레임시프트되고 절단된다. 그러나, 이에 더불어, 동형 G는 동형 A 및 D와 비교하여 엑손 2에 의해 암호화된 영역이 빠져 있고 따라서 동형 A 및 D에 존재하지 않는 엑손 1-엑손 3 경계를 포함한다. 따라서, 이러한 분리된 폴리펩타이드는 동형 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 동형 D, G, 및 H (또는 이것의 단편 또는 상동체)에 존재하는 분절을 포함하지만 추가로 동형 D (또는 이것의 단편 또는 상동체)에 존재하지 않는 동형 G (또는 이것의 단편 또는 상동체)로부터의 분절 (예컨대, 적어도 8개의 연속 아미노산)을 포함하는 상기에서 기술된 것과 같은 것일 수 있다. 이러한 영역들은 동형의 서열들을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 이러한 분리된 폴리펩타이드는 추가로 분리된 폴리펩타이드가 SEQ ID NO: 18과 최적으로 정렬될 때 SEQ ID NO: 18 (HSD17B13 동형 G)에서 엑손 1 및 3에 의해 암호화된 영역들의 경계에 걸쳐 있는 분절에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 연속 아미노산 (예컨대, 적어도 3개의 연속 아미노산, 적어도 5개의 연속 아미노산, 적어도 8개의 연속 아미노산, 적어도 10개의 연속 아미노산, 또는 적어도 15개의 연속 아미노산)의 분절을 포함할 수 있다.

[0227] 또한 본원에는 동형 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 동형 E (또는 이것의 단편 또는 상동체)에 존재하는 분절 (예컨대, 적어도 8개의 연속 아미노산)을 포함하는 분리된 폴리펩타이드가 제공된다. 동형 E는 동형 A에 존재하지 않는 엑손 3과 4 사이의 추가적인 엑손 9엑손 3')에 의해 암호화된 영역을 포함한다. 이러한 영역들은 동형의 서열들을 비교함으로써 쉽게 확인될 수 있다. 따라서, 분리된 폴리펩타이드는 HSD17B13 단백질의 적어도 5, 6, 8, 10, 12, 14, 15, 16, 18, 20, 22, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 또는 200개의 연속 아미노산 (예컨대, HSD17B13 단백질의 적어도 8개의 연속 아미노산, 적어도 10개의 연속 아미노산, 또는 적어도 15개의 연속 아미노산)을 포함할 수 있고, 여기서 연속 아미노산 (예컨대, 적어도 3개의 연속 아미노산, 적어도 5개의 연속 아미노산, 적어도 8개의 연속 아미노산, 적어도 10개의 연속 아미노산, 또는 적어도 15개의 연속 아미노산)의 분절은 분리된 폴리펩타이드가 각각 SEQ ID NO: 16 또는 19와 최적으로 정렬될 때 SEQ ID NO: 16 (HSD17B13 동형 E) 또는 SEQ ID NO: 19 (HSD17B13 동형 H)에서 엑손 3'에 의해 암호화된 영역의 적어도 일부를 포함하는 분절에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다. 선택적으로, 이러한 분리된 폴리펩타이드는 추가로 동형 H (또는 이것의 단편 또는 상동체)에 존재하지 않는 동형 E (또는 이것의 단편 또는 상동체)로부터의 분절 (예컨대, 적어도 8개의 연속 아미노산)을 포함할 수 있다. 이러한 영역들은 동형의 서열들을 비교함으로써 쉽게 확인될 수 있다. 예를 들어, 이러한 분리된 폴리펩타이드는 추가로 분리된 폴리펩타이드가 SEQ ID NO: 16과 최적으로 정렬될 때 SEQ ID NO: 16 (HSD17B13 동형 E)에서 엑손 6 및 7에 의해 암호화된 영역들의 경계에 걸쳐 있는

분절에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 연속 아미노산 (예컨대, 적어도 3개의 연속 아미노산, 적어도 5개의 연속 아미노산, 적어도 8개의 연속 아미노산, 적어도 10개의 연속 아미노산, 또는 적어도 15개의 연속 아미노산)의 분절을 포함할 수 있다.

[0228] 또한 동형 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 동형 F (또는 이것의 단편 또는 상동체)에 존재하는 분절 (예컨대, 적어도 8개의 연속 아미노산)을 포함하는 분리된 폴리펩타이드가 제공된다. 동형 F는 동형 A에 존재하지 않는 엑손 6으로부터 인트론 6으로의 초과번역에 의해 암호화된 영역을 포함한다. 이러한 영역들은 동형의 서열들을 비교함으로써 쉽게 확인될 수 있다. 따라서, 분리된 폴리펩타이드는 HSD17B13 단백질의 적어도 5, 6, 8, 10, 12, 14, 15, 16, 18, 20, 22, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 또는 200개의 연속 아미노산 (예컨대, HSD17B13 단백질의 적어도 8개의 연속 아미노산, 적어도 10개의 연속 아미노산, 또는 적어도 15개의 연속 아미노산)을 포함할 수 있고, 여기서 연속 아미노산 (예컨대, 적어도 3개의 연속 아미노산, 적어도 5개의 연속 아미노산, 적어도 8개의 연속 아미노산, 적어도 10개의 연속 아미노산, 또는 적어도 15개의 연속 아미노산)의 분절은 분리된 폴리펩타이드가 SEQ ID NO: 17과 최적으로 정렬될 때 SEQ ID NO: 17 (HSD17B13 동형 F)에서 인트론 6으로의 초과번역에 의해 암호화된 영역의 적어도 일부를 포함하는 분절에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다.

[0229] 또한 동형 A (또는 이것의 단편 또는 상동체)에 존재하지 않는 동형 C (또는 이것의 단편 또는 상동체)에 존재하는 분절 (예컨대, 적어도 8개의 연속 아미노산)을 포함하는 분리된 폴리펩타이드가 제공된다. 동형 C는 동형 A와 비교하여 엑손 6에 의해 암호화된 영역이 빠져 있고 동형 A에 존재하지 않는 엑손 5-엑손 7 경계를 포함한다. 이러한 영역들은 동형의 서열들을 비교함으로써 쉽게 확인될 수 있다. 따라서, 분리된 폴리펩타이드는 HSD17B13 단백질의 적어도 5, 6, 8, 10, 12, 14, 15, 16, 18, 20, 22, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 또는 200개의 연속 아미노산 (예컨대, HSD17B13 단백질의 적어도 8개의 연속 아미노산, 적어도 10개의 연속 아미노산, 또는 적어도 15개의 연속 아미노산)을 포함할 수 있고, 여기서 연속 아미노산 (예컨대, 적어도 3개의 연속 아미노산, 적어도 5개의 연속 아미노산, 적어도 8개의 연속 아미노산, 적어도 10개의 연속 아미노산, 또는 적어도 15개의 연속 아미노산)의 분절은 분리된 폴리펩타이드가 SEQ ID NO: 14와 최적으로 정렬될 때 SEQ ID NO: 14 (HSD17B13 동형 C)에서 엑손 5 및 7에 의해 암호화된 영역들의 경계에 걸쳐 있는 분절에 대해 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일하다.

[0230] 본원에 개시된 임의의 분리된 폴리펩타이드는 이종성 분자 또는 이종성 표지에 연결될 수 있다. 이러한 이종성 분자 또는 표지의 예는 본원의 다른 곳에서 개시된다. 예를 들어, 이종성 분자는 면역글로불린 Fc 도메인, 본원의 다른 곳에서 개시되는 것과 같은 펩타이드 태그, 폴리(에틸렌 글리콜), 폴리시알산, 또는 글리콜산일 수 있다.

[0231] ***(2) HSD17B13 단백질 또는 단편의 제조 방법***

[0232] 또한 본원에 개시된 HSD17B13 단백질 또는 이것의 단편들 중 임의의 것의 제조 방법이 개시된다. 이러한 HSD17B13 단백질 또는 이것의 단편은 임의의 적합한 방법에 의해 제조될 수 있다. 예를 들어, HSD17B13 단백질 또는 이것의 단편은 이러한 HSD17B13 단백질 또는 이것의 단편을 암호화하는 핵산 (예컨대, 재조합 발현 벡터)을 포함하는 숙주 세포로부터 생성될 수 있다. 이러한 방법은 이러한 HSD17B13 단백질 또는 이것의 단편을 암호화하는 핵산 (예컨대, 재조합 발현 벡터)을 포함하는 숙주 세포를 배양하는 단계, 그로써 HSD17B13 단백질 또는 이것의 단편을 생성하는 단계를 포함할 수 있다. 핵산은 숙주 세포에서 활성인 프로모터에 작동 가능하게 연결될 수 있고, 배양 단계는 그로 인해 핵산이 발현되는 조건 하에서 이루어질 수 있다. 이러한 방법은 추가로 발현된 HSD17B13 단백질 또는 이것의 단편을 회수하는 단계를 포함할 수 있다. 회수 단계는 추가로 HSD17B13 단백질 또는 이것의 단편을 정제하는 단계를 포함할 수 있다.

[0233] 단백질 발현에 적합한 시스템의 예로는 박테리아 세포 발현 시스템 (예컨대, 대장균, 락토코쿠스 락티스 (*Lactococcus lactis*)), 효모 세포 발현 시스템 (예컨대, 사카로마이세스 세레비시아에(*Saccharomyces cerevisiae*), 피치아 파스토리스(*Pichia pastoris*)), 곤충 세포 발현 시스템 (예컨대, 배큘로바이러스-매개된 단백질 발현), 및 포유류 세포 발현 시스템을 들 수 있다.

[0234] HSD17B13 단백질 또는 이것의 단편을 암호화하는 핵산의 예는 본원의 다른 곳에서 보다 상세하게 개시된다. 선택적으로, 이러한 핵산은 숙주 세포에서의 발현에 대해 최적화된 코돈이다. 선택적으로, 이러한 핵산은 숙주 세포에서 활성인 프로모터에 작동 가능하게 연결된다. 프로모터는 이종성 프로모터 (즉, 자연적으로 발생하는 *HSD17B13* 프로모터가 아닌 프로모터)일 수 있다. 대장균에 적합한 프로모터의 예로는 아라비노오스, *lac*, *tac*, 및 T7 프로모터를 들 수 있다. 락토코쿠스 락티스에 적합한 프로모터의 예로는 P170 및 니신 프로모터를 들 수

있다. 사카로마이세스 세레비시아에에 적합한 프로모터의 예로는 알코올 탈수소효소 (ADHI) 또는 에놀라아제 (ENO) 프로모터와 같은 구성성 프로모터 또는 PHO, CUP1, GAL1, 및 G10과 같은 유도성 프로모터를 들 수 있다. 피치아 파스토리스에 적합한 프로모터의 예로는 알코올 산화효소 I (AOX I) 프로모터, 글리세르알데하이드 3 포스페이트 탈수소효소 (GAP) 프로모터, 및 글루타티온 의존성 포름알데하이드 탈수소효소 (FLDI) 프로모터를 들 수 있다. 배큘로바이러스-매개된 시스템에 적합한 프로모터의 예는 후기 바이러스 강력한 폴리헤드린 프로모터이다.

[0235] 선택적으로, 핵산은 추가로 단백질 정제를 용이하게 하기 위하여 HSD17B13 단백질 또는 이것의 단편과 프레임에 있는 태그를 암호화한다. 태그의 예는 본원의 다른 곳에서 개시된다. 이러한 태그는, 예를 들어, 태그가 달린 단백질이 모든 다른 단백질 (예컨대 숙주 세포 단백질)로부터 분리될 수 있도록 파트너 리간드 (예컨대, 수지에 고정된 리간드)에 결합할 수 있다. 친화도 크로마토그래피, 고성능 액체 크로마토그래피 (HPLC), 및 크기 축출 크로마토그래피 (SEC)가 발현된 단백질의 순도를 개선하기 위해 사용될 수 있는 방법의 예이다.

[0236] 다른 방법들도 또한 HSD17B13 단백질 또는 이것의 단편을 제조하기 위해 사용될 수 있다. 예를 들어, 둘 이상의 펩타이드 또는 폴리펩타이드가 단백질 화학 기법에 의해 함께 연결될 수 있다. 예를 들어, 펩타이드 또는 폴리펩타이드는 Fmoc (9-플루오레닐메틸옥시카르보닐) 또는 Boc (*tert*-부틸옥시카르보닐) 화학 중 어느 하나를 사용하여 화학적으로 합성될 수 있다. 이러한 펩타이드 또는 폴리펩타이드는 표준 화학 반응에 의해 합성될 수 있다. 예를 들어, 펩타이드 또는 폴리펩타이드는 합성될 수 있고 이것의 합성 수지로부터 절단되지 않을 수 있는 반면, 펩타이드 또는 단백질의 다른 단편은 합성되고 계속해서 수지로부터 절단될 수 있어서, 다른 단편에 대해 기능적으로 차단된 말단 기를 노출시킬 수 있다. 펩타이드 축합 반응에 의해, 이런 두 단편은 그것들의 카르복실 및 아미노 말단에서 각각 펩타이드 결합을 통해 공유적으로 연합될 수 있다 (Grant GA (1992) Synthetic Peptides: A User Guide. W.H. Freeman and Co., N.Y. (1992); 및 Bodansky M and Trost B., Ed. (1993) Principles of Peptide Synthesis. Springer-Verlag Inc., NY, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 대안적으로, 펩타이드 또는 폴리펩타이드는 본원에서 기술되는 것과 같이 독립적으로 생체내에서 합성될 수 있다. 일단 분리된 후에, 이런 독립적인 펩타이드 또는 폴리펩타이드는 유사한 펩타이드 축합 반응을 통해 펩타이드 또는 이것의 단편을 형성하기 위해 연결될 수 있다.

[0237] 예를 들어, 클로닝된 또는 합성 펩타이드 분절의 효소적 결찰은 상대적으로 더 짧은 펩타이드 단편이 더 큰 펩타이드 단편, 폴리펩타이드, 또는 전체 단백질 도메인을 생성하기 위해 연합될 수 있도록 허용한다 (Abrahmsen L et al. (1991) *Biochemistry* 30:4151, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 대안적으로, 합성 펩타이드의 천연 화학적 결찰은 더 짧은 펩타이드 단편으로부터 큰 펩타이드 또는 폴리펩타이드를 합성적으로 구성하기 위해 활용될 수 있다. 이 방법은 2-단계 화학 반응으로 구성될 수 있다 (Dawson et al. (1994) *Science* 266:776-779, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 제1 단계는 보호되지 않은 합성 펩타이드-티오에스테르와 아미노-말단 Cys 잔기를 함유한 또 다른 보호되지 않은 펩타이드 분절과의 화학적 선택 반응일 수 있고, 이것으로 초기 공유 생성물로서 티오에스테르-결합 중간체가 생성된다. 반응 조건에 변화 없이, 이 중간체는 자발적인, 빠른 분자내 반응이 진행되어 결찰 부위에서 천연 펩타이드 결합이 형성될 수 있다 (Baggiolini et al. (1992) *FEBS Lett* 307:97-101; Clark-Lewis et al. (1994) *J Biol Chem* 269:16075; Clark-Lewis et al. (1991) *Biochemistry* 30:3128; 및 Rajarathnam et al. (1994) *Biochemistry* 33:6623-6630, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0238] 대안적으로, 보호되지 않은 펩타이드 분절은 화학적으로 연결될 수 있는데, 이때 화학적 결찰의 결과로서 펩타이드 분절 사이에 형성된 결합은 비천연 (비-펩타이드) 결합이다 (Schnolzer et al. (1992) *Science* 256:221, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 이 기법은 단백질 도메인의 유사체뿐만 아니라 전체 생물학적 활성을 가진 대량의 상대적으로 순수한 단백질을 합성하기 위하여 사용되어 왔다 (deLisle Milton RC et al., Techniques in Protein Chemistry IV. Academic Press, New York, pp. 257-267 (1992), 모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0239] **C. 세포**

[0240] 또한 본원에는 본원에 개시된 핵산 및 단백질 중 임의의 것을 포함하는 세포 (예컨대, 재조합 숙주 세포)가 제공된다. 세포는 시험관내, 생체외, 또는 생체내의 것일 수 있다. 핵산이 암호화된 단백질을 생성하기 위해 발현되도록 핵산은 프로모터 및 기타 조절 서열에 연결될 수 있다. 임의의 유형의 세포가 제공된다.

[0241] 세포는, 예를 들어, 전능성 세포(totipotent cell) 또는 만능 세포(pluripotent cell) (예컨대, 배아 줄기 (ES) 세포, 예컨대 설치류 ES 세포, 마우스 ES 세포, 또는 래트 ES 세포)일 수 있다. 전능성 세포로는 임의의

세포 유형을 야기할 수 있는 미분화 세포를 포함하고, 만능 세포는 하나 이상의 분화된 세포 유형으로 발달할 수 있는 능력은 가진 미분화 세포를 포함한다. 이러한 만능 및/또는 전능성 세포로는, 예를 들어, ES 세포 또는 ES-유사 세포, 예컨대 유도된 만능 줄기 (iPS) 세포를 들 수 있다. ES 세포는 배아 안으로 도입될 때 발생하는 배아의 임의의 조직에 기여할 수 있는 배아-유래 전능성 또는 만능 세포를 포함한다. ES 세포는 배반포(blastocyst)의 내세포 집단(inner cell mass)으로부터 유래될 수 있고 척추동물 3배엽 (내배엽, 외배엽 및 중배엽) 중 임의의 것의 세포로 분화할 수 있다.

[0242] 세포는 또한 일차 체세포성 세포, 또는 일차 체세포성 세포가 아닌 세포일 수 있다. 체세포성 세포는 생식세포, 배세포, 생식모세포, 또는 미분화된 줄기 세포가 아닌 임의의 세포를 포함할 수 있다. 세포는 또한 일차 세포일 수 있다. 일차 세포는 유기체, 기관, 또는 조직으로부터 직접 분리된 세포 또는 세포 배양물을 포함한다. 일차 세포는 형질전환되지 않았거나 불멸화되지 않은 세포를 포함한다. 이런 세포들은 이전에 조직 배양에서 계대되지 않았거나 또는 이전에 조직 배양에서 계대되었지만 조직 배양에서 무한하게 계대될 수 없는, 유기체, 기관, 또는 조직으로부터 얻어진 임의의 세포를 포함한다. 이러한 세포들은 종래 기법에 의해 분리될 수 있고, 예를 들어, 체세포성 세포, 조혈 세포, 내피 세포, 상피 세포, 섬유모세포, 중간엽 세포, 케라틴생성세포, 멜라닌 생성세포, 단구세포, 단핵 세포, 지방 세포, 지방전구세포, 뉴런, 신경교 세포, 간세포, 골격 근모세포, 및 평활근 세포를 포함한다. 예를 들어, 일차 세포는 결합 조직, 근육 조직, 신경계 조직, 또는 상피 조직으로부터 유래될 수 있다.

[0243] 이러한 세포는 또한 정상적으로는 무한하게 증식할 수 없지만, 돌연변이 또는 변경으로 인해, 정상적인 세포 서열을 피할 수 있게 되고 대신 계속해서 분할을 진행할 수 있는 것들을 포함한다. 이러한 돌연변이 또는 변경은 자연적으로 발생하거나 또는 의도적으로 유도될 수 있다. 불멸화된 세포의 예로는 차이니즈 햄스터 난소 (CHO) 세포, 인간 배아 신장 세포 (예컨대, HEK 293 세포), 및 마우스 배아 섬유모세포 (예컨대, 3T3 세포)를 들 수 있다. 수많은 유형의 불멸화된 세포가 잘 알려져 있다. 불멸화된 또는 일차 세포는 재조합 유전자 또는 단백질을 배양하기 위하여 또는 발현시키기 위하여 전형적으로 사용되는 세포를 포함한다.

[0244] 세포는 또한 분화된 세포, 예컨대 간세포 (예컨대, 인간 간세포)일 수 있다.

[0245] 세포는 임의의 공급원으로부터 유래될 수 있다. 예를 들어, 세포는 진핵 세포, 동물 세포, 식물 세포, 또는 진균 (예컨대 효모) 세포일 수 있다. 이러한 세포는 어류 세포 또는 조류 세포일 수 있거나, 또는 이러한 세포는 포유류 세포, 예컨대 인간 세포, 비인간 포유류 세포, 설치류 세포, 마우스 세포 또는 래트 세포일 수 있다. 포유류는, 예를 들어, 인간, 비인간 영장류, 원숭이, 유인원, 고양이, 개, 말, 황소, 사슴, 비숑, 양, 설치류 (예컨대 마우스, 래트, 햄스터, 기니아 피그), 가축 (예컨대, 황소, 수송아지, 등과 같은 소 종; 양, 염소, 등과 같은 양 종; 및 돼지 및 야생돼지 같은 돼지 종)을 포함한다. 조류는, 예를 들어, 닭, 칠면조, 타조, 거위, 오리, 등을 포함한다. 가축 및 경작용 동물이 또한 포함된다. 용어 "비인간 동물"은 인간을 배제한다.

[0246] 마우스 세포에 대해, 마우스는, 예를 들어, 129 계통, C57BL/6 계통, BALB/c 계통, 스위스 웹스터 (Swiss Webster) 계통, 129 및 C57BL/6 계통의 혼합, BALB/c 및 C57BL/6 계통의 혼합, 129 및 BALB/c 계통의 혼합, 및 BALB/c, C57BL/6, 및 129 계통의 혼합을 포함하여, 임의의 계통일 수 있다. 예를 들어, 마우스는 적어도 부분적으로 BALB/c 계통으로부터 유래될 수 있다 (예컨대, BALB/c 계통으로부터 적어도 약 25%, 적어도 약 50%, 적어도 약 75% 유래됨, 또는 BALB/c 계통으로부터 약 25%, 약 50%, 약 75%, 또는 약 100% 유래됨). 한 예에서, 마우스는 50% BALB/c, 25% C57BL/6, 및 25% 129를 포함하는 계통이다. 대안적으로, 마우스는 BALB/c를 배제한 계통 또는 계통 조합을 포함한다.

[0247] 129 계통의 예로는 129P1, 129P2, 129P3, 129X1, 129S1 (예컨대, 129S1/SV, 129S1/Svlm), 129S2, 129S4, 129S5, 129S9/SvEvH, 129S6 (129/SvEvTac), 129S7, 129S8, 129T1, 및 129T2를 들 수 있다. 예컨대, Festing et al. (1999) *Mammalian Genome* 10(8):836을 참조한다 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨). C57BL 계통의 예로는 C57BL/A, C57BL/An, C57BL/GrFa, C57BL/Kal_wN, C57BL/6, C57BL/6J, C57BL/6ByJ, C57BL/6NJ, C57BL/10, C57BL/10ScSn, C57BL/10Cr, 및 C57BL/Ola를 들 수 있다. 마우스 세포는 또한 상기 언급된 129 계통과 상기 언급된 C57BL/6 계통의 혼합으로부터 유래될 수 있다 (예컨대, 50% 129 및 50% C57BL/6). 마찬가지로, 마우스 세포는 상기 언급된 129 계통들의 혼합으로부터 또는 상기 언급된 BL/6 계통들의 혼합으로부터 유래될 수 있다 (예컨대, 129S6 (129/SvEvTac) 계통).

[0248] 래트 세포에 대해, 래트는, 예를 들어, ACI 래트 계통, Dark Agouti (DA) 래트 계통, Wistar 래트 계통, LEA 래트 계통, Sprague Dawley (SD) 래트 계통, 또는 Fischer 래트 계통, 예컨대 Fisher F344 또는 Fisher F6을 포함하여, 임의의 래트 계통일 수 있다. 래트는 또한 상기 인용된 계통의 둘 이상의 혼합으로부터 유래된 계통

으로부터 유래될 수 있다. 예를 들어, 래트는 DA 계통 또는 ACI 계통으로부터 유래될 수 있다. ACI 래트 계통은 백색의 배 및 발과 *RT1*^{av1} 하플로타입을 가진, 흑색 아구티를 가지는 것으로 특징지어진다. 이러한 계통은 Harlan 실험실을 포함한 다양한 공급처로부터 입수 가능하다. 다크 아구티 (DA) 래트 계통은 아구티 외피 및 *RT1*^{av1} 하플로타입을 가지는 것으로서 특징지어진다. 이러한 래트는 Charles River 및 Harlan 실험실을 포함한 다양한 공급처로부터 입수 가능하다. 일부 경우에, 래트는 근친교배된 래트 계통으로부터 유래된다. 예컨대, US 2014/0235933 A1 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)을 참조한다.

[0249] ***III. HSD17B13의 발현을 변형 또는 변경시키는 방법***

[0250] 뉴클레아제 제제, 외인성 도너 서열, 전사 활성화제, 전사 억제제, 안티센스 분자, 예컨대 안티센스 RNA, siRNA, 및 shRNA, HSD17B13 단백질 또는 이것의 단편, 및 재조합 *HSD17B13* 유전자 또는 HSD17B13 단백질을 암호화하는 핵산을 발현시키기 위한 발현 벡터의 임의의 조합의 사용을 통해 세포를 변형시키는 다양한 방법이 제공된다. 방법은 시험관내에서, 생체외에서, 또는 생체내에서 일어날 수 있다. 뉴클레아제 제제, 외인성 도너 서열, 전사 활성화제, 전사 억제제, 안티센스 분자, 예컨대 안티센스 RNA, siRNA, 및 shRNA, HSD17B13 단백질 또는 이것의 단편, 및 발현 벡터는 세포에 임의의 형태로 및 본원의 다른 곳에서 기술되는 것과 같은 임의의 수단에 의해 도입될 수 있고, 모든 것 또는 일부는 임의의 조합으로 동시에 또는 순차적으로 도입될 수 있다. 일부 방법은 세포에서 내인성 *HSD17B13* 유전자를 변경시키는 것만을 포함한다. 일부 방법은 전사 활성화제 또는 억제제의 사용을 통해 또는 안티센스 RNA, siRNA, 및 shRNA와 같은 안티센스 분자의 사용을 통해 내인성 *HSD17B13* 유전자를 변경시키는 것만을 포함한다. 일부 방법은 재조합 *HSD17B13* 유전자 또는 HSD17B13 단백질을 암호화하는 핵산 또는 이것의 단편을 세포에 도입시키는 것만을 포함한다. 일부 방법은 HSD17B13 단백질 또는 이것의 단편 (예컨대 본원에 개시된 HSD17B13 단백질 또는 이것의 단편 중 임의의 하나 또는 임의의 조합 또는 본원에 개시된 HSD17B13 동형 A 내지 H 또는 이것들의 단편들 중 임의의 하나 또는 임의의 조합)을 세포에 도입시키는 것만을 포함한다. 예를 들어, 이러한 방법은 HSD17B13 동형 C, D, F, G, 및 H (또는 이것들의 단편) 중 하나 이상을 세포에 도입시키는 단계를 포함할 수 있다. 대안적으로, 이러한 방법은 HSD17B13 동형 A, B, 및 E 또는 동형 A, B, E, 및 F' (또는 이것들의 단편) 중 하나 이상을 세포에 도입시키는 단계를 포함할 수 있다. 다른 방법은 내인성 *HSD17B13* 유전자를 세포에서 변경시키는 단계 및 HSD17B13 단백질 또는 이것의 단편 또는 재조합 *HSD17B13* 유전자 또는 HSD17B13 단백질을 암호화하는 핵산 또는 이것의 단편을 세포에 도입시키는 단계를 둘 다 포함할 수 있다. 또 다른 방법은 세포에서 내인성 *HSD17B13* 유전자의 발현을 변경시키는 단계 및 HSD17B13 단백질 또는 이것의 단편 또는 재조합 *HSD17B13* 유전자 또는 HSD17B13 단백질을 암호화하는 핵산 또는 이것의 단편을 세포에 도입시키는 단계를 둘 다 포함할 수 있다.

[0251] **A. *HSD17B13* 핵산을 변형시키는 방법**

[0252] 세포 내 게놈에서 뉴클레아제 제제 및/또는 외인성 도너 서열의 사용을 통하여 세포 (예컨대, 전능 세포 또는 분화된 세포, 예컨대 간세포) 내의 게놈에서 *HSD17B13* 유전자를 변형시키기 위한 다양한 방법이 제공된다. 방법은 시험관내에서, 생체외에서, 또는 생체내에서 일어날 수 있다. 뉴클레아제 제제는 단독으로 또는 외인성 도너 서열과 함께 사용될 수 있다. 대안적으로, 외인성 도너 서열은 단독으로 또는 뉴클레아제 제제와 함께 사용될 수 있다.

[0253] 이중 가닥 끊김 (DSB)에 대한 반응으로서의 복구(repair)는 원칙적으로 2개의 보존된 복구 경로를 통해 일어난다: 비-상동성 단부 연합 (NHEJ) 및 상동성 재조합 (HR). Kasparek & Humphrey (2011) *Seminars in Cell & Dev. Biol.* 22:886-897 참조 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨). NHEJ는 서로에 대한 또는 상동성 주형에 대한 요구 없이 외인성 서열에 대한 끊김 단부의 직접적인 결찰에 의한 핵산의 이중 가닥 끊김의 복구를 포함한다. NHEJ에 의한 비연속성 서열의 결찰은 종종 이중 가닥 끊김 부위 가까이에서의 삭제, 삽입, 또는 전좌(translocation)를 초래한다.

[0254] 외인성 도너 서열에 의해 매개된 표적 핵산 (예컨대, *HSD17B13* 유전자)의 복구는 2개의 폴리뉴클레오타이드 사이의 유전자 정보 교환의 임의의 과정을 포함할 수 있다. 예를 들어, NHEJ는 또한 외인성 도너 서열의 단부와 끊김 단부의 직접적인 결찰 (즉, NHEJ-기반 포획)을 통한 외인성 도너 서열의 표적화된 통합을 초래할 수 있다. 이러한 NHEJ-매개된 표적화된 통합은 상동성 지시된 복구 (HDR) 경로가 쉽게 사용될 수 없을 때 (예컨대 비분할 세포, 일차 세포, 및 상동성-기반 DNA 복구가 빈약하게 수행되는 세포에서) 외인성 도너 서열의 삽입에 바람직할 수 있다. 이에 더불어, 상동성-의존 복구와는 대조적으로, 절단 부위를 플랭킹하는 서열 동일성 (Cas-매개된 절단에 의해 생성된 오버행을 뛰어넘는)을 가진 큰 영역에 관련된 지식이 필요하지 않으며, 이런 사실은 게놈 서열에 대한 지식이 제한된 게놈을 가진 유기체 안으로 표적화된 삽입을 시도할 때에 유익할 수 있다. 통합은

외인성 도너 서열과 절단된 게놈 서열 사이의 블런트 단부의 결찰을 통해, 또는 절단된 게놈 서열에서 Cas 단백질에 의해 생성된 것들과 부합하는 오버행에 의해 플랭킹된 외인성 도너 서열을 사용하는 스티키 단부 (즉, 5' 또는 3' 오버행을 가진 단부)의 결찰을 통해 진행될 수 있다. 예컨대, US 2011/020722, WO 2014/033644, WO 2014/089290, 및 Maresca *et al*. (2013) *Genome Res*. 23(3):539-546을 참조한다 (이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 만약 블런트 단부들이 결찰된다면, 표적 및/또는 도너 절제는 단편 연합에 필요한 미세상동성의 생성 영역에 대해 필요할 수 있고, 이것은 표적 서열의 원하지 않는 변경을 유발할 수 있다.

[0255] 복구는 또한 상동성 지시된 복구 (HDR) 또는 상동성 재조합 (HR)을 통해 일어날 수 있다. HDR 또는 HR은 뉴클레오타이드 서열 상동성을 필요로 할 수 있고, "표적" 분자 (즉, 이중 가닥 끊김을 겪은 분자)의 복구를 위한 주형으로서 "도너" 분자를 사용하며, 도너로부터 표적으로의 유전자 정보의 전달을 주도하는 핵산 복구의 형태를 포함한다. 임의의 특정 이론에 의해 구속되는 것을 바라지 않으면서, 이러한 전달은 파괴된 표적과 도너 사이에 형성되는 이종듀플렉스 DNA의 미스매치 보정, 및/또는 합성-의존성 가닥 어닐링을 포함할 수 있고, 이때 도너는 표적, 및/또는 관련된 과정의 일부가 될 유전자 정보를 재합성하기 위해 사용된다. 일부 경우에, 도너 폴리뉴클레오타이드, 도너 폴리뉴클레오타이드의 일부, 도너 폴리뉴클레오타이드의 복사물, 또는 도너 폴리뉴클레오타이드의 복사물의 일부는 표적 DNA로 통합된다. Wang *et al*. (2013) *Cell* 153:910-918; Mandalos *et al*. (2012) *PLOS ONE* 7:e45768:1-9; 및 Wang *et al*. (2013) *Nat Biotechnol*. 31:530-532를 참조한다 (이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0256] 게놈에서 *HSD17B13* 유전자에 대한 표적화된 유전자 변형은 세포를, *HSD17B13* 유전자 내의 표적 게놈 유전자좌에서 5' 표적 서열에 혼성화하는 5' 상동성 팔 및 *HSD17B13* 유전자 내의 표적 게놈 유전자좌에서 3' 표적 서열에 혼성화하는 3' 상동성 팔을 포함하는 외인성 도너 서열과 접촉시킴으로써 생성될 수 있다. 외인성 도너 서열은 표적 게놈 유전자좌와 재조합하여 *HSD17B13* 유전자에 대한 표적화된 유전자 변형을 생성할 수 있다. 한 예로서, *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 5' 상동성 팔은 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 표적 서열 5'에 혼성화할 수 있고, 3' 상동성 팔은 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 표적 서열 3'에 혼성화할 수 있다. 이러한 방법은, 예를 들어, 티민이 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입되는 (또는 아데닌이 반대쪽 가닥에서 상응하는 위치에 삽입됨) *HSD17B13* 유전자를 초래할 수 있다. 다른 예로서, 5' 및 3' 상동성 팔은 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 엑손 6에 플랭킹하는 위치들에 상응하는 위치에서, 각각 표적 서열의 5' 및 3'에 혼성화할 수 있다. 이러한 방법은, 예를 들어, *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 엑손 6에 상응하는 위치가 삭제되어 있는 *HSD17B13* 유전자를 초래할 수 있다. 또 다른 예로서, 5' 및 3' 상동성 팔은 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 엑손 2에 플랭킹하는 위치들에 상응하는 위치에서, 각각 표적 서열의 5' 및 3'에 혼성화할 수 있다. 이러한 방법은, 예를 들어, *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 엑손 2에 상응하는 서열이 삭제되어 있는 *HSD17B13* 유전자를 초래할 수 있다. 또 다른 예로서, 5' 및 3' 상동성 팔은 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 엑손 6/인트론 6 경계에 상응하는 위치에서, 각각 표적 서열의 5' 및 3'에 혼성화할 수 있다. 또 다른 예로서, 5' 및 3' 상동성 팔은 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 엑손 6 및 엑손 7에 상응하는 위치에서, 각각 표적 서열의 5' 및 3'에 혼성화할 수 있다. 이러한 방법은, 예를 들어, 티민이 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입되는 (또는 아데닌이 반대쪽 가닥에서 상응하는 위치에 삽입됨) *HSD17B13* 유전자를 초래할 수 있다. 또 다른 예로서, 5' 및 3' 상동성 팔은 SEQ ID NO: 1의 인트론 6의 도너 스플라이스 부위에 상응하는 영역 (즉, SEQ ID NO: 1의 인트론 6의 5' 단부에서의 영역)을 플랭킹하는 또는 그 영역 내에 있는 것들에 상응하는 위치에서, 각각 표적 서열의 5' 및 3'에 혼성화할 수 있다. 이러한 방법은, 예를 들어, 인트론 6의 도너 스플라이스 부위가 파괴된 *HSD17B13* 유전자를 초래할 수 있다. 외인성 도너 서열의 예는 본원의 다른 곳에서 개시된다.

[0257] 게놈에서 *HSD17B13* 유전자에 대한 표적화된 유전자 변형은 또한 *HSD17B13* 유전자 내의 표적 게놈 유전자좌에서 표적 서열에서 하나 이상의 닉(nick) 또는 이중 가닥 끊김을 포함하는 뉴클레아제 제제와 세포를 접촉시킴으로써 생성될 수 있다. 이러한 방법은, 예를 들어, SEQ ID NO: 1의 인트론 6에서 도너 스플라이스 부위에 상응하는 영역 (즉, SEQ ID NO: 1의 인트론 6의 5' 단부에 있는 영역)이 파괴된 *HSD17B13* 유전자를 초래할 수 있다. 방법에 사용될 수 있는 뉴클레아제 제제의 예 및 변이는 본원의 다른 곳에서 기술된다.

[0258] 예를 들어, 게놈에서 *HSD17B13* 유전자에 대한 표적화된 유전자 변형은 세포 또는 세포의 게놈을, Cas 단백질 및

HSD17B13 유전자의 표적 게놈 유전자좌 내에서 하나 이상의 가이드 RNA 인식 서열에 혼성화하는 하나 이상의 가이드 RNA와 접촉시킴으로써 생성될 수 있다. 즉, 게놈에서 *HSD17B13* 유전자에 대한 표적화된 유전자 변형은 세포 또는 세포의 게놈을, Cas 단백질 및 *HSD17B13* 유전자의 표적 게놈 유전자좌 내에서 하나 이상의 가이드 RNA 표적 서열을 표적화하는 하나 이상의 가이드 RNA와 접촉시킴으로써 생성될 수 있다. 예를 들어, 이러한 방법은 세포를 Cas 단백질 및 *HSD17B13* 유전자 내의 가이드 RNA 표적 서열을 표적화하는 가이드 RNA와 접촉시키는 단계를 포함할 수 있다. 한 예로서, 가이드 RNA 표적 서열은 HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 엑손 6 및/또는 인트론 6에 상응하는 영역 내에 있다. 한 예로서, 가이드 RNA 표적 서열은 HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 엑손 6 및/또는 인트론 6 및/또는 엑손 7 (예컨대, 엑손 6 및/또는 인트론 6, 또는 엑손 6 및/또는 엑손 7)에 상응하는 영역 내에 있다. 다른 예로서, 가이드 RNA 표적 서열은 HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치를 포함하거나 그 위치에 근접할 수 있다. 예를 들어, 가이드 RNA 표적 서열은 HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 약 1000, 500, 400, 300, 200, 100, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5개 뉴클레오타이드 내에 있을 수 있다. 또 다른 예로서, 가이드 RNA 표적 서열은 *HSD17B13* 유전자의 개시 코돈 또는 *HSD17B13* 유전자의 종결 코돈을 포함하거나 이것들에 근접할 수 있다. 예를 들어, 가이드 RNA 표적 서열은 개시 코돈 또는 종결 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있을 수 있다. Cas 단백질 및 가이드 RNA는 복합체를 형성할 수 있고, Cas 단백질은 가이드 RNA 표적 서열을 절단한다. Cas 단백질에 의한 절단은 이중 가닥 끊김 또는 단일 가닥 끊김 (예컨대, 만약 Cas 단백질이 닉카아제(nickase)인 경우)를 생성할 수 있다. 이러한 방법은, 예를 들어, SEQ ID NO: 1의 인트론 6에서 도너 스플라이스 부위에 상응하는 영역 (즉, SEQ ID NO: 1에서 인트론 6의 5' 단부에서의 영역)이 파괴되거나, 개시 코돈이 파괴되거나, 종결 코돈이 파괴되거나, 또는 코딩 서열이 삭제된 *HSD17B13* 유전자를 초래할 수 있다. 방법에 사용될 수 있는 Cas (예컨대, Cas9) 단백질 및 가이드 RNA의 예 및 변이는 본원의 다른 곳에서 기술된다.

[0259] 일부 경우에, 둘 이상의 뉴클레아제 제제가 사용될 수 있다. 예를 들어, 2개의 뉴클레아제 제제가 사용될 수 있는데, 각각 HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 엑손 6 및/또는 인트론 6, 또는 엑손 6 및/또는 엑손 7에 상응하는 영역 내의 뉴클레아제 표적 서열을 표적화하거나, 또는 HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치를 포함하거나 그 위치에 근접한다 (예컨대, HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 약 1000, 500, 400, 300, 200, 100, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5개 뉴클레오타이드 내의 위치). 예를 들어, 두 뉴클레아제 제제가 사용될 수 있는데, 각각 HSD17B13 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 엑손 6 및/또는 인트론 6 및/또는 엑손 7에 상응하는 영역 내의 뉴클레아제 표적 서열을 표적화한다. 다른 예로서, 둘 이상의 뉴클레아제 제제가 사용될 수 있고, 각각은 개시 코돈을 포함하거나 그것에 근접한 뉴클레아제 표적 서열을 표적화한다. 또 다른 예로서, 두 뉴클레아제 제제가 사용될 수 있는데, 하나는 개시 코돈을 포함하거나 그것에 근접한 뉴클레아제 표적 서열을 표적화하고, 하나는 종결 코돈을 포함하거나 그것에 근접한 뉴클레아제 표적 서열을 표적화하며, 여기서 뉴클레아제 제제에 의한 절단은 두 뉴클레아제 표적 서열 사이의 코딩 영역의 삭제를 초래할 수 있다. 또 다른 예로서, 3개 이상의 뉴클레아제 제제가 사용될 수 있고, 이때 하나 이상 (예컨대 2개)은 개시 코돈을 포함하거나 그것에 근접한 뉴클레아제 표적 서열을 표적화하고, 하나 이상 (예컨대 2개)은 종결 코돈을 포함하거나 그것에 근접한 뉴클레아제 표적 서열을 표적화하며, 여기서 뉴클레아제 제제에 의한 절단은 개시 코돈을 포함하거나 그것에 근접한 뉴클레아제 표적 서열과 종결 코돈을 포함하거나 그것에 근접한 뉴클레아제 표적 서열 사이의 코딩 영역의 삭제를 초래할 수 있다.

[0260] 선택적으로, 세포는 *HSD17B13* 유전자의 표적 게놈 유전자좌 내에서 추가적인 가이드 RNA 표적 서열을 표적화하는 하나 이상의 추가적인 가이드 RNA와 추가로 접촉될 수 있다. 세포를 하나 이상의 추가적인 가이드 RNA (예컨대, 제2 가이드 RNA 표적 서열을 표적화하는 제2 가이드 RNA)와 접촉시킴으로써, Cas 단백질에 의한 절단은 둘 이상의 이중 가닥 끊김 또는 둘 이상의 이중 가닥 끊김 (예컨대, Cas 단백질이 닉카아제인 경우)를 생성할 수 있다.

[0261] 선택적으로, 세포는 표적화된 유전자 변형을 생성하기 위하여 *HSD17B13* 유전자의 표적 게놈 유전자좌와 재조합하는 하나 이상의 외인성 도너 서열과 추가적으로 접촉될 수 있다. 방법에 사용될 수 있는 외인성 도너 서열의 예 및 변이는 본원의 다른 곳에서 개시된다.

[0262] Cas 단백질, 가이드 RNA(들), 및 외인성 도너 서열(들)은 세포에 임의의 형태로 및 본원의 다른 곳에서 기술된 임의의 수단에 의하여 도입될 수 있고, Cas 단백질, 가이드 RNA(들), 및 외인성 도너 서열(들)의 전부 또는 일

부는 임의의 조합으로 동시에 또는 순차적으로 도입될 수 있다.

[0263] 일부 이러한 방법에서, 외인성 도너 서열에 의한 표적 핵산 (예컨대, *HSD17B13* 유전자)의 복구는 상동성-의존 복구 (HDR)를 통해 일어난다. 상동성-의존 복구는 Cas 단백질이 이중 가닥 끊김을 생성하기 위하여 *HSD17B13* 유전자의 DNA의 양 가닥을 절단할 때, Cas 단백질이 단일 가닥 끊김을 생성하기 위하여 표적 핵산의 DNA의 한 가닥을 절단하는 닉카아제일 때, 또는 두 상쇄 닉에 의해 형성된 이중 가닥 끊김을 생성하기 위해 Cas 닉카아제가 사용될 때 일어날 수 있다. 이러한 방법에서, 외인성 도너 서열은 5' 및 3' 표적 서열에 상응하는 5' 및 3' 상동성 팔을 포함한다. 가이드 RNA 표적 서열(들) 또는 절단 부위(들)은 5' 표적 서열에 인접하거나, 3' 표적 서열에 인접하거나, 5' 표적 서열 및 3' 표적 서열 둘 다에 인접하거나, 또는 5' 표적 서열이나 3' 표적 서열에 인접하지 않을 수 있다. 선택적으로, 외인성 도너 서열은 추가로 5' 및 3' 상동성 팔에 의해 플랭킹된 핵산 삽입물을 포함할 수 있고, 핵산 삽입물은 5' 및 3' 표적 서열 사이에 삽입된다. 만약 핵산 삽입물이 존재하지 않는다면, 외인성 도너 서열은 5' 및 3' 표적 서열 사이의 게놈 서열을 삭제하는 기능을 할 수 있다. 외인성 도너 서열의 예는 본원의 다른 곳에서 개시된다.

[0264] 대안적으로, 외인성 도너 서열에 의해 매개된 *HSD17B13* 유전자의 복구는 비상동성 단부 연합 (NHEJ)-매개된 결찰을 통해 일어날 수 있다. 이러한 방법에서, 외인성 도너 서열의 적어도 한 단부는 *HSD17B13* 유전자에서 Cas-매개된 절단에 의해 생성된 적어도 하나의 오버행에 상보하는 짧은 단일-가닥 영역을 포함한다. 외인성 도너 서열의 상보하는 단부는 핵산 삽입물을 플랭킹할 수 있다. 예를 들어, 외인성 도너 서열의 각 단부는 *HSD17B13* 유전자에서 Cas-매개된 절단에 의해 생성된 하나의 오버행에 상보하는 짧은 단일-가닥 영역을 포함할 수 있고, 외인성 도너 서열의 이런 상보하는 영역들은 핵산 삽입물을 플랭킹할 수 있다.

[0265] 오버행 (즉, 엇갈린 단부)은 Cas-매개된 절단에 의해 생성된 이중 가닥 끊김의 블런트 단부들의 절제에 의해 생성될 수 있다. 이러한 절제는 단편 연합에 필요한 미세상동성 영역을 생성할 수 있지만, 이것은 *HSD17B13* 유전자에서 원하지 않는 또는 제어할 수 없는 변경을 생성할 수 있다. 대안적으로, 이러한 오버행은 쌍을 이룬 Cas 닉카아제를 사용함으로써 생성될 수 있다. 예를 들어, 세포는 DNA의 반대쪽 가닥을 절단하는 제1 및 제2 닉카아제와 접촉될 수 있고, 그로써 게놈은 이중 닉킹을 통해 변형된다. 이것은 세포를 제1 Cas 단백질 닉카아제, *HSD17B13* 유전자의 표적 게놈 유전자좌 내의 제1 가이드 RNA 표적 서열을 표적화하는 제1 가이드 RNA, 제2 Cas 단백질 닉카아제, 및 *HSD17B13* 유전자의 표적 게놈 유전자좌 내의 제2 가이드 RNA 표적 서열을 표적화하는 제2 가이드 RNA와 접촉시킴으로써 이루어질 수 있다. 제1 Cas 단백질 및 제1 가이드 RNA는 제1 복합체를 형성하고, 제2 Cas 단백질 및 제2 가이드 RNA는 제2 복합체를 형성한다. 제1 Cas 단백질 닉카아제는 제1 가이드 RNA 표적 서열 내의 게놈 DNA의 제1 가닥을 절단하고, 제2 Cas 단백질 닉카아제는 제2 가이드 RNA 표적 서열 내의 게놈 DNA의 제2 가닥을 절단하며, 선택적으로 외인성 도너 서열은 *HSD17B13* 유전자의 표적 게놈 유전자좌와 재조합하여 표적화된 유전자 변형을 생성시킨다.

[0266] 제1 닉카아제는 게놈 DNA의 제1 가닥 (즉, 상보하는 가닥)을 절단하고, 제2 닉카아제는 게놈 DNA의 제2 가닥 (즉, 비-상보하는하는 가닥)을 절단한다. 제1 및 제2 닉카아제는, 예를 들어, Cas9의 RuvC 도메인의 촉매적 잔기를 돌연변이시킴으로써 (예컨대, 본원의 다른 곳에서 기술된 D10A 돌연변이) 또는 Cas9의 HNH 도메인에서 촉매적 잔기를 돌연변이시킴으로써 (예컨대, 본원의 다른 곳에 기술되는 H804A 돌연변이) 생성될 수 있다. 이러한 방법에서, 이중 닉킹이 엇가린 단부 (즉, 오버행)를 가지는 이중 가닥 끊김을 생성하기 위해 사용될 수 있다. 제1 및 제2 가이드 RNA 표적 서열은 DNA의 제1 및 제2 가닥 상에서 제1 및 제2 닉카아제에 의해 생성된 닉이 이중 가닥 끊김을 생성하도록 절단 부위를 생성하기 위해 위치할 수 있다. 오버행은 제1 및 제2 CRISPR RNA 표적 서열 내의 닉들이 상쇄적일 때 생성된다. 상쇄 창은, 예를 들어, 적어도 약 5 bp, 10 bp, 20 bp, 30 bp, 40 bp, 50 bp, 60 bp, 70 bp, 80 bp, 90 bp, 100 bp 또는 그 이상일 수 있다. 예컨대, Ran *et al.* (2013) *Cell* 154:1380-1389; Mali *et al.* (2013) *Nat. Biotech.*31:833-838; 및 Shen *et al.* (2014) *Nat. Methods* 11:399-404를 참조한다 (이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0267] ***(1) 표적화된 유전자 변형의 유형***

[0268] 다양한 유형의 표적화된 유전자 변형이 본원에 기술된 방법들을 사용하여 도입될 수 있다. 이러한 표적화된 변형으로는, 예를 들어, 하나 이상의 뉴클레오타이드의 첨가, 하나 이상의 뉴클레오타이드의 삭제, 하나 이상의 뉴클레오타이드의 치환, 점 돌연변이, 또는 이것들의 조합을 들 수 있다. 예를 들어, 표적화된 게놈 변형을 형성하기 위하여 적어도 1, 2, 3, 4, 5, 7, 8, 9, 10개 또는 그 이상의 뉴클레오타이드가 변화 (예컨대, 삭제, 삽입, 또는 치환)될 수 있다. 삭제, 삽입, 또는 치환은, 본원의 다른 곳에서 개시되는 것과 같이, 임의의 크기를 가질 수 있다. 예컨대, Wang et al. (2013) *Cell* 153:910-918; Mandalos et al. (2012) *PLOS ONE* 7:e45768:1-

9; 및 Wang et al. (2013) *Nat Biotechnol*. 31:530-532를 참조한다 (이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0269] 이러한 표적화된 유전자 변형은 표적 게놈 유전자좌의 파괴를 초래할 수 있다. 파괴는 조절 요소 (예컨대, 프로모터 또는 인핸서)의 변경, 미스센스 돌연변이, 넌센스 돌연변이, 프레임-시프트 돌연변이, 절단 돌연변이, 널 (null) 돌연변이, 또는 적은 수의 뉴클레오타이드의 삽입 또는 삭제 (예컨대, 프레임시프트 돌연변이를 유발함)를 포함할 수 있고, 이것들은 비활성화 (즉, 기능의 상실) 또는 대립유전자의 상실을 초래할 수 있다. 예를 들어, 표적화된 변형은 개시 코돈이 더 이상 기능할 수 없도록 하는 *HSD17B13* 유전자의 개시 고돈의 파괴를 포함할 수 있다.

[0270] 특정 실례에서, 표적화된 변형은 제1 및 제2 가이드 RNA 표적 서열 사이 또는 Cas 절단 부위들 사이의 삭제를 포함할 수 있다. 만약 외인성 도너 서열 (예컨대, 복구 주형 또는 표적화 벡터)이 사용된다면, 변형은 제1 및 제2 가이드 RNA 표적 서열 사이 또는 Cas 절단 부위들 사이의 삭제뿐만이 아니라 5' 및 3' 표적 서열 사이에 핵산 삽입물의 삽입을 포함할 수 있다.

[0271] 대안적으로, 만약 외인성 도너 서열이 단독으로 또는 뉴클레아제 제제와 함께 사용된다면, 변형은 5' 및 3' 표적 서열 사이의 결실뿐만 아니라 제1 및 제2 상동성 염색체의 쌍에서 5' 및 3' 표적 서열 사이에 핵산 삽입물의 삽입을 포함할 수 있고, 그로써 동형접합성 변형 게놈을 초래할 수 있다. 대안적으로, 만약 외인성 도너 서열이 핵산 삽입물 없이 5' 및 3' 상동성 팔을 포함한다면, 변형은 5' 및 3' 표적 서열 사이의 삭제를 포함할 수 있다.

[0272] 제1 및 제2 가이드 RNA 표적 서열 사이의 삭제 또는 5' 및 3' 표적 서열 사이의 삭제는 삭제된 핵산이 제1 및 제2 뉴클레아제 절단 부위 사이의 핵산 서열로마 구성되거나 또는 5' 및 3' 표적 서열 사이의 핵산 서열로만 구성되어서 변형된 게놈 표적 유전자좌에 추가적인 삭제 또는 삽입이 없는 경우에 정확한 삭제일 수 있다. 제1 및 제2 가이드 RNA 표적 서열 사이의 삭제는 또한 제1 및 제2 뉴클레아제 절단 부위를 지나 연장된 부정확한 삭제일 수 있고, 이것은 비상동성 단부 연합 (NHEJ)에 의한 부정확한 복구와 일치하며, 변형된 게놈 유전자좌에서 추가적인 삭제 및/또는 삽입을 초래할 수 있다. 예를 들어, 삭제는 제1 및 제2 Cas 단백질 절단 부위를 지나 약 1 bp, 약 2 bp, 약 3 bp, 약 4 bp, 약 5 bp, 약 10 bp, 약 20 bp, 약 30 bp, 약 40 bp, 약 50 bp, 약 100 bp, 약 200 bp, 약 300 bp, 약 400 bp, 약 500 bp, 또는 그 이상 연장될 수 있다. 마찬가지로, 변형된 게놈 유전자좌는 NHEJ에 의한 부정확한 복구와 일치하는 추가적인 삽입, 예컨대 약 1 bp, 약 2 bp, 약 3 bp, 약 4 bp, 약 5 bp, 약 10 bp, 약 20 bp, 약 30 bp, 약 40 bp, 약 50 bp, 약 100 bp, 약 200 bp, 약 300 bp, 약 400 bp, 약 500 bp, 또는 그 이상의 삽입을 포함할 수 있다.

[0273] 표적화된 유전자 변형은, 예를 들어, 이중대립유전자 변형 또는 단일대립유전자 변형일 수 있다. 이중대립유전자 변형은 동일한 변형이 (예컨대 이배체 세포에서) 상응하는 상동성 염색체 상의 동일한 유전자좌에 대해 만들어지거나, 또는 상이한 변형이 상응하는 상동성 염색체 상의 동일한 유전자좌에 대해 만들어지는 사건을 포함한다. 일부 방법에서, 표적화된 유전자 변형은 단일대립유전자 변형이다. 단일대립유전자 변형은 변형이 오직 한 대립유전자에 대해서만 만들어지는 사건 (즉, 두 상동성 염색체 중 하나에서만 *HSD17B13* 유전자에 대한 변형)을 포함한다. 상동성 염색체는 동일한 유전자좌에서 동일한 유전자를 갖지만 가능한 상이한 대립유전자를 갖는 염색체들 (예컨대, 감수분열 중에 쌍을 이루는 염색체들)을 포함한다. 용어 대립유전자는 유전자 서열의 하나 이상의 대체 형태 중 임의의 것을 포함한다. 이배체 세포 또는 유기체에서, 주어진 서열의 두 대립유전자는 전형적으로 한 쌍의 상동성 염색체에서 상응하는 유전자좌를 차지한다.

[0274] 단일대립유전자 돌연변이는 표적화된 *HSD17B13* 변형에 대해 이형접합성인 세포를 초래할 수 있다. 이형접합성은 *HSD17B13* 유전자의 오직 하나의 대립유전자 (즉, 두 상동성 염색체 상의 상응하는 대립유전자)만이 표적화된 변형을 가지는 상황을 포함한다.

[0275] 이중대립유전자 변형은 표적화된 변형에 대해 동형접합성을 초래할 수 있다. 동형접합성은 *HSD17B13* 유전자의 두 대립유전자 (즉, 두 상동성 염색체 상의 상응하는 대립유전자)가 모두 표적화된 변형을 가지는 상황을 포함한다. 대안적으로, 이중대립유전자 변형은 표적화된 변형에 대해 화합물 이형접합성 (예컨대, 반접합성)을 초래할 수 있다. 화합물 이형접합성은 표적 유전자좌의 두 대립유전자 (즉, 두 상동성 염색체 상의 대립유전자)가 모두 변형되었지만, 상이한 방식으로 변형된 (예컨대, 한 대립유전자에서는 표적화된 변형 및 다른 대립유전자의 비활성화 또는 파괴) 상황을 포함한다. 예를 들어, 표적화된 변형이 없는 대립유전자에서, Cas 단백질에 의해 생성된 이중 가닥 끊김은 비상동성 단부 연합 (NHEJ)-매개된 DNA 복구에 의해 복구될 수 있고, 이것은 핵산 서열의 삽입 또는 삭제를 포함하는 돌연변이 대립유전자를 생성하며 그로써 그 게놈 유전자좌의 파괴를 유발한

다. 예를 들어, 이중대립유전자 변형은 만약 세포가 표적화된 변형을 가진 하나의 대립유전자 및 발현될 수 없는 또 다른 대립유전자를 가진다면 화합물 이형접합성을 초래할 수 있다. 화합물 이형접합성은 반접합성을 포함한다. 반접합성은 표적 유전자좌의 오직 하나만의 대립유전자 (즉, 두 상동성 염색체 중 하나의 대립유전자)가 존재하는 상황을 포함한다. 예를 들어, 이중대립유전자 변형은 만약 표적화된 변형이 한 대립유전자에서 일어나고 다른 대립유전자는 상응하여 상실 또는 삭제되는 경우에 표적화된 변형에 대해 반접합성을 초래할 수 있다.

[0276] ***(2) 표적화된 유전자 변형을 가진 세포 확인***

[0277] 본원에 개시된 방법은 추가로 변형된 *HSD17B13* 유전자를 가지는 세포를 확인하는 단계를 포함할 수 있다. 다양한 방법, 예컨대 삭제 또는 삽입이 표적화된 유전자 변형을 가지는 세포를 확인하기 위하여 사용될 수 있다. 이러한 방법들은 *HSD17B13* 유전자에 표적화된 유전자 변형을 가지는 한 세포를 확인하는 단계를 포함할 수 있다. 스크리닝은 변형된 게놈 유전자좌를 가지는 이러한 세포를 확인하기 위하여 사용될 수 있다.

[0278] 스크리닝 단계는 모(parental) 염색체의 대립유전자의 변형을 평가하기 위한 정량적 검정 (예컨대, 대립유전자 상실 (LOA) 및/또는 대립유전자 획득 (GOA) 검정)을 포함할 수 있다. 예를 들어, 정량 검정은 정량적 PCR, 예컨대 실시간 PCR (qPCR)을 통해 수행될 수 있다. 실시간 PCR은 표적 게놈 유전자좌를 인식하는 제1 프라이머 세트 및 비-표적화된 참조 유전자좌를 인식하는 제2 프라이머 세트를 활용할 수 있다. 프라이머 세트는 증폭된 서열을 인식하는 형광 프로브를 포함할 수 있다. 대립유전자 상실 (LOA) 검정은 종래의 스크리닝 로직을 반전시키고 돌연변이가 지시된 천연 유전자좌의 복사물의 수를 정량화한다. 정확하게 표적화된 세포 클론에서, LOA 검정은 2개의 천연 대립유전자 (X 또는 Y 염색체 상에 없는 유전자에 대한) 중 하나를 검출하고, 다른 대립유전자는 표적화된 변형에 의해 파괴된다. 동일한 원리가 삽입된 표적화 벡터의 복사물 수를 정량화하기 위하여 대립유전자 획득 (GOA) 검정으로서 역으로 적용될 수 있다. 예를 들어, GOA 및 LOA 검정의 조합된 사용은 정확하게 표적화된 이형접합성 클론을 천연 표적 유전자의 한 복사물을 잃어버리고 약물 내성 유전자 또는 다른 삽입된 마커의 한 복사물을 얻은 것으로서 드러낼 것이다.

[0279] 실례로서, 정량적 중합효소 연쇄 반응 (qPCR)이 대립유전자 정량화의 방버브로서 사용될 수 있지만, 표적 유전자의 제로, 하나, 및 두 복사물 사이의 차이 또는 핵산 삽입물의 제로, 하나, 및 두 복사물 사이의 차이를 쉽게 구별할 수 있는 임의의 방법이 MOA 검정을 개발하기 위해 사용될 수 있다. 예를 들어, TAQMAN®은 게놈 DNA 샘플 중의 DNA 주형의 복사물의 수를, 특히 참조 유전자에 대한 비교에 의하여 정량화하기 위해 사용될 수 있다 (예컨대, US 6,596,541 참조, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 참조 유전자는 동일한 게놈 DNA에서 표적 유전자(들) 또는 유전자좌(loci)로서 정량화된다. 그러므로, 두 TAQMAN® 증폭 (각각 그것의 각각의 프로브를 이용함)이 수행된다. 한 TAQMAN® 프로브는 참조 유전자의 "Ct" (한계값 사이클)을 측정하는 한편, 다른 프로브는 성공적인 표적화에 의해 치환된 표적화된 유전자(들) 또는 유전자좌(loci)의 영역의 Ct를 측정한다 (즉, LOA 검정). Ct는 TAQMAN® 프로브의 각각에 대한 출발 DNA의 양을 반영하는 양이다, 즉, 덜 풍부한 한계값 사이클에 도달하기 위해 PCR의 더 많은 사이클을 필요로 한다. TAQMAN® 반응에 대한 주형 서열의 복사물의 수를 절반으로 감소시키는 것은 약 1 Ct 단위의 증가를 초래할 것이다. 표적 유전자(들) 또는 유전자좌(loci) 중 한 대립유전자가 상동성 재조합에 의해 치환되어 있는 세포에서 TAQMAN® 반응은 비표적화 세포로부터의 DNA와 비교할 때 참조 유전자에 대한 Ct의 증가 없이 표적 TAQMAN® 반응에 대한 1 Ct의 증가를 초래할 것이다. GOA 검정의 경우, 또 다른 TAQMAN® 프로브가 성공적인 표적화에 의해 표적화된 유전자(들) 또는 유전자좌(loci)를 대체한 핵산 삽입물의 Ct를 측정하기 위해 사용될 수 있다.

[0280] 적합한 정량 검정의 다른 실례는 고정된 프로브(들), INVADER® 프로브, TAQMAN® Molecular Beacon 프로브, 또는 ECLIPSE™ 프로브에 대한 형광-매개된 제자리 혼성화 (FISH), 비교 게놈 혼성화, 항온성 DNA 증폭, 정량적 혼성화 기술을 포함한다 (예컨대, US 2005/0144655 참조, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 표적화된 변형을 위한 스크리닝에 대한 종래 검정, 예컨대 긴-범위 PCR, 서던 블롯팅, 또는 Sanger 서열분석이 또한 사용될 수 있다. 이러한 검정은 전형적으로 삽입된 표적화 벡터 및 표적화된 게놈 유전자좌 사이의 결합에 대한 증거를 얻기 위해 사용된다. 예를 들어, 긴-범위 PCR 검정의 경우, 한 프라이머는 삽입된 DNA 내의 서열을 인식할 수 있는 한편, 다른 프라이머는 표적화 벡터의 상동성 팔의 단부들을 넘어간 표적 게놈 유전자좌 서열을 인식한다.

[0281] 차세대 서열분석 (NGS)은 또한 스크리닝을 위해 사용될 수 있다. 차세대 서열분석은 또한 "NGS" 또는 "대량 병

렬 서열분석" 또는 "고출력량 서열분석"으로서 언급될 수 있다. 본원에 개시된 방법에서, 선택 마커를 사용한 표적화된 세포에 대한 스크리닝은 필요하지 않다. 예를 들어, 본원에 개시된 MOA 및 NGS는 선택 카세트를 사용하지 않는 것에 의존할 수 있다.

[0282] B. *HSD17B13 핵산의 발현을 변경시키는 방법*

[0283] HSD17B13 단백질을 암호화하는 핵산의 발현을 변경시키기 위한 다양한 방법이 제공된다. 일부 방법에서, 발현은 본원의 다른 곳에서 한층 더 상세하게 기술되는 바와 같이, HSD17B13 단백질을 암호화하는 핵산의 파괴를 유발하기 위해 뉴클레아제 제제로의 절단을 통해 변경된다. 일부 방법에서, 발현은 전사 활성화 도메인 또는 전사 억제 도메인에 융합된 또는 연결된 DNA-결합 단백질의 사용을 통해 변경된다. 일부 방법에서, 발현은 안티센스 RNA, shRNA, 또는 siRNA와 같은, RNA 간섭 조성물의 사용을 통해 변경된다.

[0284] 한 예에서, *HSD17B13* 유전자 또는 HSD17B13 단백질을 암호화하는 핵산의 발현은 세포 또는 세포 내의 게놈을 *HSD17B13* 유전자 또는 HSD17B13 단백질을 암호화하는 핵산 내의 표적 게놈 유전자좌에서 표적 서열에서의 하나 이상의 닉 또는 이중 가닥 끊김을 유도하는 뉴클레아제 제제와 접촉시킴으로써 변형될 수 있다. 이러한 절단은 *HSD17B13* 유전자 또는 HSD17B13 단백질을 암호화하는 핵산의 발현의 파괴를 초래할 수 있다. 예를 들어, 뉴클레아제 표적 서열은 *HSD17B13* 유전자의 개시 코돈을 포함하거나 개시 코돈에 근접할 수 있다. 예를 들어, 표적 서열은 개시 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있을 수 있고, 뉴클레아제 제제에 의한 절단은 개시 코돈을 파괴할 수 있다. 다른 예로서, 둘 이상의 뉴클레아제 제제가 사용될 수 있는데, 각각의 뉴클레아제를 표적화하는 표적 서열은 개시 코돈을 포함하거나 개시 코돈에 근접할 수 있다. 또 다른 예로서, 두 개의 뉴클레아제 제제가 사용될 수 있는데, 하나의 뉴클레아제를 표적화하는 표적 서열은 개시 코돈을 포함하거나 개시 코돈에 근접하고, 하나의 뉴클레아제를 표적화하는 표적 서열은 종결 코돈을 포함하거나 개시 코돈에 근접하며, 여기서 뉴클레아제 제제에 의한 절단은 두 뉴클레아제 제제 서열 사이의 코딩 영역의 삭제를 초래할 수 있다. 또 다른 예로서, 3개 이상의 뉴클레아제 제제가 사용될 수 있는데, 하나 이상 (예컨대 2개)의 뉴클레아제를 표적화하는 표적 서열은 개시 코돈을 포함하거나 개시 코돈에 근접하고, 하나 이상 (예컨대 2개)의 뉴클레아제를 표적화하는 표적 서열은 종결 코돈을 포함하거나 개시 코돈에 근접하며, 여기서 뉴클레아제 제제에 의한 절단은 개시 코돈을 포함하거나 개시 코돈에 근접한 뉴클레아제 표적 서열과 종결 코돈을 포함하거나 종결 코돈에 근접한 뉴클레아제 표적 서열 사이의 코딩 영역의 삭제를 초래할 수 있다. *HSD17B13* 유전자 또는 HSD17B13 단백질을 암호화하는 핵산을 변형시키는 다른 예들은 본원의 다른 곳에서 개시된다.

[0285] 다른 예에서, *HSD17B13* 유전자 또는 HSD17B13 단백질을 암호화하는 핵산의 발현은 세포 또는 세포 내의 게놈을 *HSD17B13* 유전자 내의 표적 게놈 유전자좌에 결합하는 DNA-결합 단백질과 접촉시킴으로써 변형될 수 있다. DNA-결합 단백질은, 예를 들어, 전사 활성화제 도메인 또는 전사 억제자 도메인에 융합된 뉴클레아제-비활성 Cas 단백질일 수 있다. DNA-결합 단백질의 다른 예는 전사 활성화제 도메인 또는 전사 억제자 도메인에 융합된 아연 핑거 단백질, 또는 전사 활성화제 도메인 또는 전사 억제자 도메인에 융합된 전사 활성화제-유사 이펙터 (TALE) 단백질을 포함한다. 이러한 단백질의 예들은 본원의 다른 곳에서 개시된다. 예를 들어, 일부 방법에서, 전사 억제자는 야생형 *HSD17B13* 유전자 또는 rs72613567 변종이 아닌 *HSD17B13* 유전자의 발현을 감소시키기 위해 (예컨대, *HSD17B13* 전사물 또는 동형 A의 발현을 감소시키기 위해) 사용될 수 있다. 마찬가지로, 일부 방법에서, 전사 활성화제는 *HSD17B13* 유전자 rs72613567 변종 유전자의 발현을 증가시키기 위해 (예컨대, *HSD17B13* 전사물 또는 동형 D의 발현을 증가시키기 위해) 사용될 수 있다.

[0286] DNA-결합 단백질에 대한 표적 서열 (예컨대, 가이드 RNA 표적 서열)은 발현을 변경시키기에 적합한 *HSD17B13* 유전자 또는 HSD17B13 단백질을 암호화하는 핵산 내의 어느 곳에든 있을 수 있다. 한 예로서, 표적 서열은 조절 요소, 예컨대 인핸서 또는 프로모터 내에 있거나, 또는 조절 요소와 근접한 곳에 있을 수 있다. 예를 들어, 표적 서열은 *HSD17B13* 유전자의 개시 코돈을 포함하거나 개시 코돈에 근접할 수 있다. 예를 들어, 표적 서열은 개시 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있을 수 있다.

[0287] 다른 예에서, 안티센스 분자가 *HSD17B13* 유전자 또는 HSD17B13 단백질을 암호화하는 핵산의 발현을 변경시키기 위하여 사용될 수 있다. 안티센스 분자의 예로는 안티센스 RNA, 작은 간섭 RNA (siRNA), 및 짧은 헤어핀 RNA (shRNA)를 들 수 있다. 이러한 안티센스 RNA, siRNA, 또는 shRNA는 mRNA의 임의의 영역을 표적화하기 위하여 설계될 수 있다. 예를 들어, 안티센스 RNA, siRNA, 또는 shRNA는 본원에 개시된 하나 이상의 HSD17B13 전사물에 특유한 영역, 또는 본원에 개시된 하나 이상의 HSD17B13 전사물에 공통적인 영역을 표적화하도록 설계될 수 있

다. cDNA 및 변종 *HSD17B13* 전사물에 혼성화하는 핵산의 예는 본원의 다른 곳에서 보다 상세하게 개시된다. 예를 들어, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A) 내의 서열에 혼성화할 수 있다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 세포에서 *HSD17B13* 전사물 A의 발현을 감소시킬 수 있다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 존재하지 않는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하는 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)의 엑손 7 내의 서열 또는 엑손 6-엑손 7 경계에 걸쳐 있는 서열에 혼성화한다.

[0288] 또 다른 예로서, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 7 (*HSD17B13* 전사물 D) 내의 서열에 혼성화할 수 있다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하지 않는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 존재하는 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)의 엑손 7 내의 서열 또는 엑손 6-엑손 7 경계에 걸쳐 있는 서열에 혼성화한다.

[0289] **C. 핵산 및 단백질의 세포 안으로의 도입**

[0290] 본원에 개시된 핵산 및 단백질은 임의의 수단에 의하여 세포 안으로 도입될 수 있다. "도입"은 서열이 세포의 내부에 접근하게 되는 방식으로 핵산 또는 단백질을 세포에 제공하는 것을 포함한다. 도입은 임의의 수단에 의해 이루어질 수 있고, 하나 이상의 구성요소 (예컨대, 구성요소들 중 둘, 또는 모든 구성요소)는 임의의 조합으로 세포에 동시에 또는 순차적으로 도입될 수 있다. 예를 들어, 외인성 도너 서열은 뉴클레아제 제제의 도입 전에 도입될 수 있고, 또는 뉴클레아제 제제의 도입 후에 도입될 수 있다 (예컨대, 외인성 도너 서열은 뉴클레아제 제제의 도입 전 또는 후 약 1, 2, 3, 4, 8, 12, 24, 36, 48, 또는 72시간에 투여될 수 있다). 예컨대, US 2015/0240263 및 US 2015/0110762를 참조한다 (이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 세포의 게놈을 뉴클레아제 제제 또는 외인성 도너 서열과 접촉시키는 것은 하나 이상의 뉴클레아제 제제 또는 뉴클레아제 제제를 암호화하는 핵산 (예컨대, 하나 이상의 Cas 단백질 또는 하나 이상의 Cas 단백질을 암호화하는 핵산, 및 하나 이상의 가이드 RNA 또는 하나 이상의 가이드 RNA를 암호화하는 핵산 (즉, 하나 이상의 하나 이상의 CRISPR RNA 및 하나 이상의 tracrRNA)) 및/또는 하나 이상의 외인성 도너 서열을 세포에 도입시키는 것을 포함할 수 있다. 세포의 게놈을 접촉시키는 것 (즉, 세포를 접촉시키는 것)은 상기 구성요소들 중 오직 하나, 하나 이상의 구성요소, 또는 모든 구성요소를 세포에 도입시키는 것을 포함할 수 있다.

[0291] 뉴클레아제 제제는 단백질의 형태로 또는 뉴클레아제 제제를 암호화하는 핵산, 예컨대 RNA (예컨대, 메신저 RNA (mRNA)) 또는 DNA의 형태로 세포 안으로 도입될 수 있다. DNA의 형태로 도입될 때, DNA는 세포에서 활성인 프로모터에 작동 가능하게 연결될 수 있다. 이러한 DNA는 하나 아싱의 발현 구성물에 있을 수 있다.

[0292] 예를 들어, Cas 단백질은 단백질, 예컨대 gRNA와 복합체가 형성된 Cas 단백질의 형태로, 또는 Cas 단백질을 암호화하는 핵산, 예컨대 RNA (예컨대, 메신저 RNA (mRNA)) 또는 DNA의 형태로 세포 안으로 도입될 수 있다. 가이드 RNA는 RNA의 형태로 또는 가이드 RNA를 암호화하는 DNA의 형태로 세포 안으로 도입될 수 있다. DNA의 형태로 도입될 때, Cas 단백질 및/또는 RNA를 암호화하는 DNA는 세포에서 활성인 프로모터에 작동 가능하게 연결될 수 있다. 이러한 DNA는 하나 이상의 발현 구성물에 있을 수 있다. 예를 들어, 이러한 발현 구성물은 단일 핵산 분자의 구성요소일 수 있다. 대안적으로, 이것들은 둘 이상의 핵산 분자 중에서 임의의 조합으로 분리될 수 있다 (즉, 하나 이상의 CRISPR RNA를 암호화하는 DNA, 하나 이상의 tracrRNA를 암호화하는 DNA, 및 Cas 단백질을 암호화하는 DNA는 별도의 핵산 분자의 구성요소일 수 있다).

[0293] 일부 방법에서, 뉴클레아제 제제 (예컨대, Cas 단백질 및 가이드 RNA)를 암호화하는 DNA 및/또는 외인성 도너 서열을 암호화하는 DNA는 DNA 미니서클을 통해 세포 안으로 도입될 수 있다. 예컨대, WO 2014/182700을 참조한다 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨). DNA 미니서클은 복제 기원이나 항생물질 선택 마커를 갖지 않은 비바이러스 유전자 전달에 사용될 수 있는 수퍼코일 DNA 분자이다. 그러므로, DNA 미니서클은 전형적으로 플라스미드 벡터보다 크기가 더 작다. 이런 DNA는 박테리아 DNA가 없고, 따라서 박테리아 DNA에서 발견되는 메틸화되지 않은 CpG 모티프가 결핍되어 있다.

[0294] 본원에 제공된 방법은 핵산 또는 단백질을 세포 안으로 도입하는 특정 방법에 좌우되지 않으며, 오직 핵산 또는 단백질이 최소한 한 세포의 내부에 접근하게 한다. 핵산 및 단백질을 다양한 세포 유형 안으로 도입시키는 방법은 알려져 있고, 예를 들어, 안정적인 트랜스펙션 방법, 일시적 트랜스펙션 방법, 및 바이러스-매개된 방법을 들 수 있다.

[0295] 핵산 또는 단백질을 세포 안으로 도입시키기 위한 프로토콜뿐만 아니라 트랜스펙션 프로토콜은 다를 수 있다. 비제한적인 트랜스펙션 방법으로는 리포좀; 나노입자; 칼슘 포스페이트를 사용하는 화학적-기반 트랜스펙션 방법 (Graham et al. (1973) *Virology* 52 (2): 456-67, Bacchetti et al. (1977) *Proc. Natl. Acad. Sci. USA* 74 (4): 1590-4, 및 Kriegler, M (1991). Transfer and Expression: A Laboratory Manual. New York: W. H. Freeman and Company. pp. 96-97); 덴드라이머; 또는 양이온성 폴리머, 예컨대 DEAE-덱스트란 또는 폴리에틸렌이민을 사용하는 화학적-기반 트랜스펙션 방법을 들 수 있다. 비화학적 방법으로는 전기천공법, 초음파 천공법(Sono-poration), 및 광학적 트랜스펙션을 들 수 있다. 입자-기반 트랜스펙션은 유전자 총의 사용, 또는 자기-보조 트랜스펙션(magnet-assisted transfection)을 포함한다 (Bertram (2006) *Current Pharmaceutical Biotechnology* 7, 277-28). 바이러스 방법 또한 트랜스펙션에 사용될 수 있다.

[0296] 핵산 또는 단백질의 세포 안으로의 도입은 또한 전기천공법에 의해, 세포질내 주입에 의해, 바이러스 감염에 의해, 아데노바이러스에 의해, 아데노-관련 바이러스에 의해, 렌티바이러스에 의해, 레트로바이러스에 의해, 트랜스펙션에 의해, 지질-매개된 트랜스펙션에 의해, 또는 뉴클레오펙션(nucleofection)에 의해 매개될 수 있다. 뉴클레오펙션은 핵산 기질이 세포질로 전달될뿐만 아니라 핵 막을 통해 핵 안으로 전달되는 것을 가능하게 하는 개량된 전기천공 기술이다. 이에 더불어, 본원에 개시된 방법에서 뉴클레오펙션의 사용은 전형적으로 정규 전기천공법보다 훨씬 더 적은 수의 세포 (예컨대, 정규 전기천공법에 의한 7백만개와 비교하여 단지 약 2백만개)를 필요로 한다. 한 예에서, 뉴클레오펙션은 LONZA® NUCLEOFECTOR™ 시스템을 사용하여 수행된다.

[0297] 핵산 또는 단백질의 세포 안으로의 도입은 또한 미세주입에 의해 이루어질 수 있다. mRNA의 미세주입은 바람직하게는 세포질 안으로 주입되는 한편 (예컨대, mRNA를 번역 기계로 직접 전달하기 위하여), 단백질 또는 Cas 단백질을 암호화하는 DNA의 미세주입은 바람직하게는 핵 안으로 주입된다. 대안적으로, 미세주입은 핵 및 세포질 둘 다로의 주입에 의해 수행될 수 있다: 바늘이 먼저 핵 안으로 도입되고 첫 번째 양이 주입될 수 있으며, 한편 바늘이 세포로부터 제거되면서 두 번째 양이 세포질 안으로 주입될 수 있다. 만약 뉴클레아제 제제 단백질이 세포질 안으로 주입된다면, 단백질은 바람직하게 핵/전핵으로의 전달을 확실히 하기 위해 핵 국지화 신호를 포함한다. 미세주입을 수행하기 위한 방법은 잘 알려져 있다. 예컨대, Nagy et al. (Nagy A, Gertsenstein M, Vintersten K, Behringer R., 2003, Manipulating the Mouse Embryo. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press); Meyer et al. (2010) *Proc. Natl. Acad. Sci. USA* 107:15022-15026 및 Meyer et al. (2012) *Proc. Natl. Acad. Sci. USA* 109:9354-9359를 참조한다.

[0298] 핵산 또는 단백질을 세포 안으로 도입하기 위한 방법으로는, 예를 들어, 벡터 전달, 입자-매개 전달, 엑소좀-매개 전달, 지질-나노입자-매개 전달, 세포 투과 펩타이드-매개 전달, 또는 이식가능한 장치-매개 전달을 들 수 있다. 핵산 또는 단백질을 생체내에서 세포를 변형시키기 위해 대상체에 전달하는 방법은 본원의 다른 곳에서 개시된다.

[0299] 핵산 및 단백질의 세포 안으로의 도입은 또한 유체역학적 전달 (HDD)에 의하여 이루어질 수 있다. 유체역학적 전달은 생체내에서 세포내 DNA 전달을 위한 거의 완벽한 방법으로서 출현하였다. 실질 세포로의 유전자 전달을 위해, 필수 DNA 서열만이 선택된 혈관을 통해 주입될 필요가 있고, 이것으로 현재 바이러스 및 합성 벡터와 관련된 안전성 문제가 제거된다. 혈류로 주입될 때, DNA는 혈액에 접근할 수 있는 상이한 조직에서 세포에 도달할 수 있다. 유체역학적 전달은 대용량의 용액을 크고 막-불투과성인 화합물이 실질 세포로 유입되는 것을 막는 내피세포 및 세포막의 물리적 장벽을 극복하기 위해 순환중인 비압축성 혈액으로의 신속한 주입에 의해 생성된 힘을 사용한다. DNA의 전달에 더불어, 이 방법은 RNA, 단백질, 및 기타 작은 화합물의 생체내에서의 효율적인 세포내 전달에 유용하다. 예컨대, Bonamassa et al. (2011) *Pharm. Res.* 28(4):694-701을 참조한다 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0300] 핵산 또는 단백질을 세포 안으로 도입하기 위한 다른 방법으로는, 예를 들어, 벡터 전달, 입자-매개 전달, 엑소좀-매개 전달, 지질 나노입자-매개 전달, 세포 투과 펩타이드-매개 전달, 또는 이식가능한 장치-매개 전달을 들 수 있다. 구체적인 예로서, 핵산 또는 단백질은 폴리(락트산) (PLA) 마이크로스피어, 폴리(D,L-락틱-코글리콜-산) (PLGA) 마이크로스피어, 리포좀, 미셀, 역미셀, 지질 코클레이트, 또는 a 지질 미세관과 같은 담체로 세포 안으로 도입될 수 있다.

[0301] 핵산 또는 단백질의 세포 안으로의 도입은 한번에 또는 일정 시간의 기간 동안 여러 번 수행될 수 있다. 예를 들어, 도입은 일정 시간의 기간 동안 적어도 2회, 일정 시간의 기간 동안 적어도 3회, 일정 시간의 기간 동안 적어도 4회, 일정 시간의 기간 동안 적어도 5회, 일정 시간의 기간 동안 적어도 6회, 일정 시간의 기간 동안 적어도 7회, 일정 시간의 기간 동안 적어도 8회, 일정 시간의 기간 동안 적어도 9회, 일정 시간의 기간 동안 적어

도 10회, 일정 시간의 기간 동안 적어도 11회, 일정 시간의 기간 동안 적어도 12회, 일정 시간의 기간 동안 적어도 13회, 일정 시간의 기간 동안 적어도 14회, 일정 시간의 기간 동안 적어도 15회, 일정 시간의 기간 동안 적어도 16회, 일정 시간의 기간 동안 적어도 17회, 일정 시간의 기간 동안 적어도 18회, 일정 시간의 기간 동안 적어도 19회, 또는 일정 시간의 기간 동안 적어도 20회 수행될 수 있다.

[0302] 일부 경우에, 방법 및 조성물에 사용된 세포는 그것의 게놈 안으로 안정적으로 포함된 DNA를 가진다. 이러한 경우에, 접촉 단계는 그것의 게놈에 이미 안정적으로 포함된 구성물을 가진 세포를 제공하는 단계를 포함할 수 있다. 예를 들어, 본원에 개시된 방법에 사용된 세포는 그것의 게놈에 안정적으로 포함된 기존의 Cas-암호화 유전자를 가질 수 있다 (즉, Cas-준비 세포). "안정적으로 통합된" 또는 "안정적으로 도입된" 또는 "안정적으로 통합된"은 뉴클레오티드 서열이 세포의 게놈 안으로 통합되고 그것의 자손에 의해 유전될 수 있도록 하는 폴리뉴클레오타이드의 세포 안으로의 도입을 포함한다. 임의의 프로토콜은 DNA 구성물 또는 표적화된 게놈 통합 시스템의 다양한 구성요소들의 안정적인 포함에 사용될 수 있다.

[0303] **D. 뉴클레아제 제제 및 DNA-결합 단백질**

[0304] 원하는 표적 서열 또는 원하는 표적 서열에 결합하는 임의의 DNA-결합 단백질 안으로 닉 또는 이중 가닥 끊김을 유도하는 임의의 뉴클레아제 제제가 본원에 개시된 방법 및 조성물에 사용될 수 있다. 자연적으로 발생하는 또는 천연 뉴클레아제 제제는 뉴클레아제 제제가 원하는 표적 서열에서 닉 또는 이중 가닥 끊김을 유도하는 한 사용될 수 있다. 마찬가지로, 자연적으로 발생하는 또는 천연 DNA-결합 단백질은 DNA-결합 단백질이 원하는 표적 서열에 결합하는 한 사용될 수 있다. 대안적으로, 변형된 또는 조작된 뉴클레아제 제제 또는 DNA-결합 단백질이 사용될 수 있다. "조작된 뉴클레아제 제제 또는 DNA-결합 단백질"은 원하는 표적 서열을 특이적으로 인식하기 위해 그것의 천연 형태로부터 조작된 (변형된 또는 유래된) 뉴클레아제 제제 또는 DNA-결합 단백질을 포함한다. 그러므로, 조작된 뉴클레아제 제제 또는 DNA-결합 단백질은 천연, 자연적으로 발생하는 뉴클레아제 제제 또는 DNA-결합 단백질로부터 유래될 수 있고 또는 그것은 인공적으로 생성되거나 합성될 수 있다. 조작된 뉴클레아제 제제 또는 DNA-결합 단백질은 표적 서열을 인식할 수 있고, 예를 들어, 표적 서열은 천연 (조작되지 않은 또는 변형되지 않은) 뉴클레아제 제제 또는 DNA-결합 단백질에 의해 인식될 수 있을 서열이 아니다. 뉴클레아제 제제 또는 DNA-결합 단백질의 변형은 단백질 절단제에서 한 아미노산 또는 핵산 절단제에서 한 뉴클레오타이드와 같이 아주 작은 것일 수 있다. 표적 서열 또는 다른 DNA에서 닉 또는 이중 가닥 끊김을 생성하는 것은 본원에서 표적 서열 또는 다른 DNA를 "컷팅" 또는 "절단"으로서 언급될 수 있다.

[0305] 뉴클레아제 제제 또는 DNA-결합 단백질의 활성 변종 및 단편 (즉, 조작된 뉴클레아제 제제 또는 DNA-결합 단백질)이 또한 제공된다. 이러한 활성 변종들은 천연 뉴클레아제 제제 또는 DNA-결합 단백질에 대해 적어도 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상의 서열 동일성을 포함할 수 있고, 여기서 활성 변종은 원하는 표적 서열을 커팅하는 능력을 보유하며 따라서 닉 또는 이중 가닥-끊김-유도 활성을 보유하거나 또는 원하는 표적 서열에 결합하는 능력을 보유한다. 예를 들어, 본원에 기술된 임의의 뉴클레아제 제제는 천연 엔도뉴클레아제 서열로부터 변형될 수 있고 천연 뉴클레아제 제제에 의해 인식되지 않은 표적 서열을 인식하고 거기에서 닉 또는 이중 가닥 끊김을 유도하도록 설계될 수 있다. 그러므로, 일부 조작된 뉴클레아제는 상응하는 천연 뉴클레아제 제제 표적 서열과 상이한 표적 서열에서 닉 또는 이중 가닥 끊김을 유도하는 특이성을 가진다. 닉 또는 이중 가닥-끊김-유도 활성에 대한 검정은 알려져 있고 일반적으로 표적 서열을 함유한 DNA 기질에 대한 전체적인 활성 및 특이성을 측정한다.

[0306] 용어 "뉴클레아제 제제에 대한 표적 서열"은 닉 또는 이중 가닥 끊김이 뉴클레아제 제제에 의해 유도되는 DNA 서열을 포함한다. 마찬가지로, 용어 "DNA-결합 단백질에 대한 표적 서열"은 DNA-결합 단백질이 결합할 DNA 서열을 포함한다. 표적 서열은 세포에 대해 내인성 (또는 천연)이거나 세포에 대해 외인성일 수 있다. 세포에 대해 외인성인 표적 서열은 세포의 게놈에서 자연적으로 발생하지 않는다. 표적 서열은 또한 당업자가 표적 유전자좌에서 위치할 것으로 기대하는 관심이 있는 폴리뉴클레오타이드에 대해 외인성일 수 있다. 일부 경우에, 표적 서열은 숙주 세포의 게놈에서 단 한번 존재한다.

[0307] 예시된 표적 서열의 활성 변종 및 단편이 또한 제공된다. 이러한 활성 변종은 주어진 표적 서열에 대해 적어도 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상의 서열 동일성을 포함할 수 있고, 여기서 활성 변종은 생물학적 활성을 보유하며, 따라서 서열-특이적 방식으로 뉴클레아제 제제에 의해 인식 및 절단될 수 있다. 뉴클레아제 제제에 의한 표적 서열의 이중 가닥 끊김을 측정하기 위한 검정은 알려져 있다 (예컨대, TAQMAN® qPCR 검정, Frendewey et al. (2010) *Methods in Enzymology* 476:295-307, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0308] 표적 서열의 길이는 다를 수 있고, 예를 들어, 아연 핑거 단백질 또는 아연 핑거 뉴클레아제 (ZFN) 쌍에 대해 약 30 내지 36 bp (즉, 각 ZFN에 대해 약 15 내지 18 bp)인 표적 서열, 전사 활성화제-유사 이펙터 (TALE) 단백질 또는 전사 활성화제-유사 이펙터 뉴클레아제 (TALEN)에 대해 약 36 bP인 표적 서열, 또는 CRISPR/Cas9 가이드 RNA에 대해 약 20 bp인 표적 서열을 들 수 있다.

[0309] DNA-결합 단백질 또는 뉴클레아제 제제의 표적 서열은 표적 게놈 유전자좌에 또는 가까이에 위치할 수 있다. 표적 서열은 유전자 (예컨대, *HSD17B13* 유전자)의 코딩 영역 내에, 또는 유전자의 발현에 영향을 미치는 조절 영역 내에 위치할 수 있다. DNA-결합 단백질 또는 뉴클레아제 제제의 표적 서열은 인트론, 엑손, 프로모터, 인핸서, 조절 영역, 또는 임의의 비-단백질 코딩 영역에 위치할 수 있다.

[0310] 본원에 개시된 다양한 방법 및 조성물에서 사용될 수 있는 DNA-결합 단백질의 한 유형은 전사 활성화제-유사 이펙터 (TALE)이다. TALE은, 예를 들어, 에피유전자 변형 도메인, 전사 활성화 도메인, 또는 전사 억제자 도메인에 융합 또는 연결될 수 있다. 이러한 도메인의 예는 하기에서 Cas 단백질과 관련하여 기술되며, 또한 예를 들어, WO 2011/145121 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)에서 찾아볼 수 있다. 이에 상응하여, 본원에 개시된 다양한 방법 및 조성물에서 사용될 수 있는 뉴클레아제 제제의 한 유형은 전사 활성화제-유사 이펙터 뉴클레아제 (TALEN)이다. TAL 이펙터 뉴클레아제는 원핵 또는 진핵 유기체의 게놈에서의 특이적 표적 서열에서 이중 가닥 끊김을 만들기 위해 사용될 수 있는 서열-특이적 뉴클레아제의 한 부류이다. TAL 이펙터 뉴클레아제는 천연 또는 조작된 전사 활성화제-유사 (TAL) 이펙터, 또는 이것의 기능성 단편을 *FokI*과 같은 엔도뉴클레아제의 촉매성 도메인에 융합시킴으로써 생성된다. 특유한, 모듈식 TAL 이펙터 DNA 결합 도메인은 잠재적으로 임의의 주어진 DNA 인식 특이성을 가진 단백질의 설계를 가능하게 한다. 그러므로, TAL 이펙터 뉴클레아제의 DNA 결합 도메인은 특이적 DNA 표적 부위를 인식하기 위해 조작될 수 있고, 그러므로 원하는 표적 서열에서 이중 가닥 끊김을 만들기 위해 사용될 수 있다. WO 2010/079430; Morbitzer et al. (2010) *Proc. Natl. Acad. Sci. U.S.A.* 107(50:21617-21622; Scholze & Boch (2010) *Virulence* 1:428-432; Christian et al. (2010) *Genetics* 186:757-761; Li et al. (2011) *Nucleic Acid Res.* 39(1):359-372; 및 Miller et al. (2011) *Nature Biotechnology* 29:143-148 (이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨)을 참조한다.

[0311] 적합한 TAL 뉴클레아제, 및 적합한 TAL 뉴클레아제를 제조하는 방법의 예들은, 예컨대, US 2011/0239315 A1, US 2011/0269234 A1, US 2011/0145940 A1, US 2003/0232410 A1, US 2005/0208489 A1, US 2005/0026157 A1, US 2005/0064474 A1, US 2006/0188987 A1, 및 US 2006/0063231 A1에 개시되어 있고, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다. 다양한 구체예에서, 예를 들어, 관심이 있는 게놈 유전자좌의 표적 핵산 서열에서 또는 가까이에서 절단되는 TAL 이펙터 뉴클레아제가 조작되며, 여기서 표적 핵산 서열은 외인성 도너 서열에 의해 변형될 서열에 또는 가까이에 있다. 본원에 제공된 다양한 방법 및 조성물과 사용하기에 적합한 TAL 뉴클레아제로는 본원의 다른 곳에서 기술되는 것과 같이 외인성 도너 서열에 의해 변형된 표적 핵산 서열에서 또는 가까이에서 결합하도록 특이적으로 설계되는 것들을 들 수 있다.

[0312] 일부 TALEN에서, TALEN의 각 모노머는 2개의 초가변성 잔기를 통해 단일 염기쌍을 인식하는 33 내지 35개의 TAL 반복부를 포함한다. 일부 TALEN에서, 뉴클레아제 제제는 FokI 엔도뉴클레아제와 같은 독립적인 뉴클레아제에 작동 가능하게 연결된 TAL-반복부-기반 DNA 결합 도메인을 포함하는 키메릭 단백질이다. 예를 들어, 뉴클레아제 제제는 제1 TAL-반복부-기반 DNA 결합 도메인 및 제2 TAL-반복부-기반 DNA 결합 도메인을 포함할 수 있고, 여기서 제1 및 제2 TAL-반복부-기반 DNA 결합 도메인의 각각은 FokI 뉴클레아제에 작동가능하게 연결되며, 제1 및 제2 TAL-반복부-기반 DNA 결합 도메인은 다양한 길이 (12 내지 20 bp)의 스페이서 서열에 의해 분리된 표적 서열 DNA 서열의 각 가닥에서 2개의 연속적인 표적 DNA 서열을 인식하며, FokI 뉴클레아제 하위유닛은 표적 서열에서 이중 가닥 끊김을 만드는 활성 뉴클레아제를 생성하기 위해 다이머화한다.

[0313] DNA-결합 단백질의 다른 예는 아연 핑거 단백질이다. 이러한 아연 핑거 단백질은, 예를 들어, 에피유전자 변형 도메인, 전사 활성화 도메인, 또는 전사 억제자 도메인에 결합되거나 융합될 수 있다. 이러한 도메인의 예는 하기에서 Cas 단백질과 관련하여 기술되며, 또한 예를 들어, WO 2011/145121 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)에서 찾아볼 수 있다. 이에 상응하여, 본원에 개시된 다양한 방법 및 조성물에서 사용될 수 있는 뉴클레아제 제제의 또 다른 예는 아연-핑거 뉴클레아제 (ZFN)이다. 일부 ZFN에서, ZFN의 각 모노머는 3개 이상의 아연 핑거-기반 DNA 결합 도메인을 포함하여, 여기서 각각의 아연 핑거-기반 DNA 결합 도메인은 3 bp 하위부위에 결합한다. 다른 ZFN에서, ZFN은 FokI 엔도뉴클레아제와 같은 독립적인 뉴클레아제에 작동가능하게 연결된 아연 핑거-기반 DNA 결합 도메인을 포함하는 키메릭 단백질이다. 예를 들어, 뉴클레아제 제제는 제1 ZFN 및 제2 ZFN을 포함할 수 있고, 제1 ZFN 및 제2 ZFN의 각각은 FokI 뉴클레아제 하위유닛에 작동 가능하게 연결되며, 제1 ZFN 및 제2 ZFN은 약 5 내지 7 bp 스페이서에 의해 분리된 표적 DNA 서열의 각 가닥에서 2개의 연속적인 표적

DNA 서열을 인식하고, 여기서 FokI 뉴클레아제 하위유닛은 이중 가닥 끊김을 만드는 활성 뉴클레아제를 생성하기 위해 다이머화한다. 예컨대, US 2006/0246567; US 2008/0182332; US 2002/0081614; US 2003/0021776; WO 2002/057308 A2; US 2013/0123484; US 2010/0291048; WO 2011/017293 A2; 및 Gaj et al. (2013) *Trends in Biotechnology* 31(7):397-405을 참조하며, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다.

[0314] 본원에 기술된 방법 및 조성물에 사용하기 위한 다른 적합한 DNA-결합 단백질 및 뉴클레아제 제제로는, 본원의 다른 곳에서 기술되는 CRISPR-Cas 시스템을 들 수 있다.

[0315] DNA-결합 단백질 또는 뉴클레아제 제제는 임의의 공지된 수단에 의해 세포 안으로 도입될 수 있다. DNA-결합 단백질 또는 뉴클레아제 제제를 암호화하는 폴리펩타이드는 직접 세포 안으로 도입될 수 있다. 대안적으로, DNA-결합 단백질 또는 뉴클레아제 제제를 암호화하는 폴리뉴클레오타이드는 세포 안으로 도입될 수 있다. DNA-결합 단백질 또는 뉴클레아제 제제를 암호화하는 폴리뉴클레오타이드가 세포 안으로 도입될 때, DNA-결합 단백질 또는 뉴클레아제 제제는 세포 내에서 일시적으로, 조건부로, 또는 구성적으로 발현될 수 있다. 예를 들어, DNA-결합 단백질 또는 뉴클레아제 제제를 암호화하는 폴리뉴클레오타이드는 발현 카세트에 함유될 수 있고 조건부 프로모터, 유도성 프로모터, 구성성 프로모터, 또는 조직-특이적 프로모터에 작동 가능하게 연결될 수 있다. 이러한 프로모터는 본원의 다른 곳에서 한층 더 상세하게 논의된다. 대안적으로, DNA-결합 단백질 또는 뉴클레아제 제제는 DNA-결합 단백질 또는 뉴클레아제 제제를 암호화하는 mRNA로서 세포 안으로 도입될 수 있다.

[0316] DNA-결합 단백질 또는 뉴클레아제 제제를 암호화하는 폴리뉴클레오타이드는 세포의 게놈에 안정적으로 통합될 수 있고 세포에서 활성인 프로모터에 작동 가능하게 연결될 수 있다. 대안적으로, DNA-결합 단백질 또는 뉴클레아제 제제를 암호화하는 폴리뉴클레오타이드는 표적화 벡터에 또는 삽입 폴리뉴클레오타이드를 포함하는 표적화 벡터로부터 분리된 벡터 또는 플라스미드에 있을 수 있다.

[0317] DNA-결합 단백질 또는 뉴클레아제 제제가 DNA-결합 단백질 또는 뉴클레아제 제제를 암호화하는 폴리뉴클레오타이드의 도입을 통해 세포에 제공될 때, 이러한 DNA-결합 단백질 또는 뉴클레아제 제제를 암호화하는 폴리뉴클레오타이드는 DNA-결합 단백질 또는 뉴클레아제 제제를 암호화하는 자연적으로 발생하는 폴리뉴클레오타이드에 비교하여, 관심이 있는 세포에서 고빈도로 사용되는 코돈을 치환하기 위하여 변형될 수 있다. 예를 들어, DNA-결합 단백질 또는 뉴클레아제 제제를 암호화하는 폴리뉴클레오타이드는 관심이 있는 주어진 원핵 또는 진핵 세포, 이를테면 박테리아 세포, 효모 세포, 인간 세포, 비인간 세포, 포유류 세포, 설치류 세포, 마우스 세포, 래트 세포 또는 관심이 있는 임의의 다른 숙주 세포에서, 자연적으로 발생하는 폴리뉴클레오타이드 서열에 비교하여, 고빈도로 사용되는 코돈을 치환하기 위해 변형될 수 있다.

[0318] **E. CRISPR-Cas 시스템**

[0319] 본원에 개시된 방법은 세포 내에서 게놈을 변형시키기 위하여 규칙적인 간격을 갖는 짧은 회문구조 반복단위(CRISPR)/CRISPR-관련 (Cas) 시스템 또는 이러한 시스템의 구성요소들을 이용할 수 있다. CRISPR-Cas 시스템은 Cas 유전자의 발현, 또는 활성의 지시에 관련된 전사물 및 다른 요소들을 포함한다. CRISPR-Cas 시스템은 제 I형, 제 II형, 또는 제 III형 시스템일 수 있다. 대안적으로 CRISPR/Cas 시스템은, 예를 들어, 제 V형 시스템(예컨대, 아형 V-A 또는 아형 V-B)일 수 있다. 본원에 개시된 방법 및 조성물은 핵산의 부위-특정 절단을 위해 CRISPR 복합체 (Cas 단백질과 복합체를 형성한 가이드 RNA (gRNA)를 포함함)를 이용함으로써 CRISPR-Cas 시스템을 사용할 수 있다.

[0320] 본원에 개시된 방법에서 사용된 CRISPR-Cas 시스템은 자연적으로 발생하지 않을 수 있다. "자연적으로 발생하지 않는" 시스템은 인간에 의한 발달을 가리키는 무엇, 예컨대, 그것들이 본질상 자연적으로 관련되는 적어도 하나의 다른 구성요소가 적어도 실질적으로 없거나, 또는 그것들이 자연적으로 관련되지 않는 적어도 하나의 다른 구성요소와 관련되는 것인, 그것들의 자연적으로 발생하는 상태로부터 변경되었거나 돌연변이된 시스템의 하나 이상의 구성요소를 포함한다. 예를 들어, 자연적으로 발생하지 않는 CRISPR/Cas 시스템은 자연적으로는 함께 발생하지 않는 gRNA 및 Cas 단백질, 자연적으로 발생하지 않는 Cas 단백질, 또는 자연적으로 발생하지 않는 gRNA를 포함하는 CRISPR 복합체를 사용할 수 있다.

[0321] ***(1) Cas 단백질 및 Cas 단백질을 암호화하는 폴리뉴클레오타이드***

[0322] Cas 단백질은 일반적으로 가이드 RNA (gRNA, 하기에서 보다 상세하게 기술됨)와 상호작용할 수 있는 적어도 하나의 RNA 인식 또는 결합 도메인을 포함한다. Cas 단백질은 또한 뉴클레아제 도메인 (예컨대, DNase 또는 RNase 도메인), DNA 결합 도메인, 헬리카제 도메인, 단백질-단백질 상호작용 도메인, 다이머화 도메인, 및 기타 도메

인을 포함할 수 있다. 뉴클레아제 도메인은 핵산 절단을 위한 촉매 활성을 가지며, 핵산 분자의 공유 결합의 파손을 포함한다. 절단은 블런트 단부 또는 엇갈린 단부를 생성할 수 있고, 단일 가닥 또는 이중 가닥일 수 있다. 예를 들어, 야생형 Cas9 단백질은 전형적으로 블런트 절단 생성물을 생성할 것이다. 대안적으로, 야생형 Cpf1 단백질 (예컨대, FnCpf1)은 5-뉴클레오타이드 5' 오버행을 가진 절단 생성물을 초래할 수 있는데, 이때 절단은 비표적화 가닥 상의 PAM 서열로부터 18번째 염기쌍 뒤에서 및 표적화된 가닥 상에서 23번째 염기 뒤에서 일어난다. Cas 단백질은 *HSD17B13* 유전자에서 이중 가닥 끊김 (예컨대, 블런트 단부를 가진 이중 가닥 끊김)을 생성하기 위해 전체 절단 활성을 가질 수 있거나, 또는 *HSD17B13* 유전자에서 단일 가닥 끊김을 생성하는 닉카아제일 수 있다.

[0323] Cas 단백질의 예로는 Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas5e (CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a1, Cas8a2, Cas8b, Cas8c, Cas9 (Csn1 또는 Csx12), Cas10, Cas10d, CasF, CasG, CasH, Csy1, Csy2, Csy3, Cse1 (CasA), Cse2 (CasB), Cse3 (CasE), Cse4 (CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, 및 Cu1966, 및 이것들의 상동체 또는 변형된 버전을 들 수 있다.

[0324] 예시적인 Cas 단백질은 제 II형 CRISPR/Cas 시스템으로부터의 Cas9 단백질 또는 Cas9 단백질로부터 유래된 단백질이다. Cas9 단백질은 제 II형 CRISPR/Cas 시스템으로부터 유래되고 전형적으로 보존된 구조와 4개의 핵심 모티프를 공유한다. 모티프 1, 2, 및 4는 RuvC-유사 모티프이고, 모티프 3은 HNH 모티프이다. 예시적인 Cas9 단백질은 스트렙토코쿠스 피오게네스(*Streptococcus pyogenes*), 스트렙토코쿠스 써모필루스(*Streptococcus thermophilus*), 스트렙토코쿠스 종(*Streptococcus sp.*), 스타필로코쿠스 아우레우스(*Staphylococcus aureus*), 노카르디옵시스 다쏜빌레이(*Nocardiopsis dassonvillei*), 스트렙토마이세스 프리스티나에스피랄리스(*Streptomyces pristinaespiralis*), 스트렙토마이세스 비리도크로모게네스(*Streptomyces viridochromogenes*), 스트렙토마이세스 비리도크로모게네스(*Streptomyces viridochromogenes*), 스트렙토스포랑기움 로제움(*Streptosporangium roseum*), 스트렙토스포랑기움 로제움(*Streptosporangium roseum*), 알리사이클로바실루스 아시도알다리우스(*Alicyclobacillus acidocaldarius*), 바실루스 슈도미코이데스(*Bacillus pseudomycoides*), 바실루스 셀레니티레두센스(*Bacillus selenitireducens*), 엑시구오박테리움 시비리쿰(*Exiguobacterium sibiricum*), 락토바실루스 델브루엑키이(*Lactobacillus delbrueckii*), 락토바실루스 살리바리우스(*Lactobacillus salivarius*), 마이크로실라 마리나(*Microscilla marina*), 부르콜데리알레스 박테리움(*Burkholderiales bacterium*), 폴라로모나스 나프탈레니보란스(*Polaromonas naphthalenivorans*), 폴라로모나스 종(*Polaromonas sp.*), 크로코스파에라 왓소니이(*Crocosphaera watsonii*), 시아노테스 종(*Cyanothece sp.*), 마이크로시스티스 아에루기노사(*Microcystis aeruginosa*), 시네코코쿠스 종(*Synechococcus sp.*), 아세토할로비움 아라바티쿰(*Acetohalobium arabaticum*), 암모니펙스 데겐시이(*Ammonifex degensii*), 칼디셀룰로시룹터 벡시이(*Caldicelulosiruptor becscii*), 칸디다투스 데술포루디스(*Candidatus Desulforudis*), 클로스트리디움 보툴리눔(*Clostridium botulinum*), 클로스트리디움 디피실레(*Clostridium difficile*), 피네골디아 마그나(*Finegoldia magna*), 나트라나에로비우스 써모필루스(*Natranaerobius thermophilus*), 펠로토마쿨룸 써모프로피오니쿰(*Pelotomaculum thermopropionicum*), 아시디티오바실루스 칼두스(*Acidithiobacillus caldus*), 아시디티오바실루스 페로옥시단스(*Acidithiobacillus ferrooxidans*), 알로크로마티움 비노숨(*Allochromatium vinosum*), 마리노박터 종(*Marinobacter sp.*), 니트로소코쿠스 할로필루스(*Nitrosococcus halophilus*), 니트로소코쿠스 왓소니(*Nitrosococcus watsoni*), 슈도알테로모나스 할로플란크티스(*Pseudoalteromonas haloplanktis*), 크테도노박터 라세미페르(*Ktedonobacter racemifer*), 메타노할로비움 에베스티가툼(*Methanohalobium evestigatum*), 아나바에나 바리아빌리스(*Anabaena variabilis*), 노둘라리아 스퍼미게나(*Nodularia spumigena*), 노스톡 종(*Nostoc sp.*), 아르트로스피라 막시마(*Arthrospira maxima*), 아르트로스피라 플라텐시스(*Arthrospira platensis*), 아르트로스피라 종(*Arthrospira sp.*), 링비아 종(*Lyngbya sp.*), 마이크로콜레우스 크토노플라스테스(*Microcoleus chthonoplastes*), 오실라토리아 종(*Oscillatoria sp.*), 페트로토가 모빌리스(*Petrotoga mobilis*), 써모시포 아프리카누스(*Thermosipho africanus*), 또는 아카리오클로리스 마리나(*Acaryochloris marina*)로부터 유래된다. Cas9 패밀리 구성원의 추가적인 예는 WO 2014/131833 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)에 기술된다. S. 피오게네스로부터의 Cas9 (SpCas9) (SwissProt 등록 번호 Q99ZW2가 배정됨)가 예시적인 Cas9 단백질이다. S. 아우레우스로부터의 Cas9 (SaCas9) (UniProt 등록 번호 J7RUA5로 배정됨)가 또 다른 예시적인 Cas9 단백질이다.

[0325] Cas 단백질의 또 다른 예는 Cpf1 (프레보텔라(*Prevotella*) 및 프란시셀라(*Francisella*) 1로부터의 CRISPR) 단백질이다. Cpf1은 Cas9의 특징적인 아르기닌-풍부 클러스터에 대한 대응부분과 함께 Cas9의 상응하는 도메인에 상동하는 RuvC-유사 뉴클레아제 도메인을 함유하는 큰 단백질 (약 1300개 아미노산)이다. 그러나, Cpf1은, HNH

도메인을 포함하는 긴 삽입물을 함유하는 Cas9와는 대조적으로, Cas9 단백질에 존재하는 HNH 뉴클레아제 도메인이 부족하고, RuvC-유사 도메인은 Cpf1 서열에서 연속적이다. 예컨대, Zetsche et al. (2015) *Cell* 163(3):759-771 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)을 참조한다. 예시적인 Cpf1 단백질은 프란시셀라 툴라렌시스 1(*Francisella tularensis 1*), 프란시셀라 툴라렌시스 아종 노비시다(*Francisella tularensis subsp. novicida*), 프레보텔란 알벤시스(*Prevotella albensis*), 라크노스피라세아에 박테리움(*Lachnospiraceae bacterium*) *MC2017 1*, 부티리비브리오 프로테오클라스티쿠스(*Butyrivibrio proteoclasticus*), 페레그리니박테리아 박테리움(*Peregrinibacteria bacterium*) *GW2011_GWA2_33_10*, 파르쿠박테리아 박테리움(*Parcubacteria bacterium*) *GW2011_GWC2_44_17*, 스미텔라 종(*Smithella sp.*) *SCADC*, 아시다미노코쿠스 종(*Acidaminococcus sp.*) *BV3L6*, 라크노스피라세아에 박테리움(*Lachnospiraceae bacterium*) *MA2020*, 칸디다투스 메타노플라스마 테르미툼(*Candidatus Methanoplasma termitum*), 유바크테리움 엘리겐스(*Eubacterium eligens*), 모락셀라 보보쿨리 237(*Moraxella bovoculi 237*), 렙토스피라 이나다이(*Leptospira inadai*), 라크노스피라세아에 박테리움(*Lachnospiraceae bacterium*) *ND2006*, 포르피로모나스 크레비오리카니스(*Porphyromonas crevioricanis*) *3*, 프레보텔라 디시엔스(*Prevotella disiens*), 및 포르피로모나스 마카카에(*Porphyromonas macacae*)로부터 유래된다. (*Francisella novicida*) U112로부터의 Cpf1 (FnCpf1; UniProt 등록번호 A0Q7Q2로 배정됨)이 예시적인 Cpf1 단백질이다.

[0326] Cas 단백질은 야생형 단백질 (즉, 자연계에서 발생하는 것들), 변형된 Cas 단백질 (즉, Cas 단백질 변종), 또는 야생형 또는 변형된 Cas 단백질의 단편일 수 있다. Cas 단백질은 또한 야생형 또는 변형된 Cas 단백질의 촉매 활성과 관련하여 활성 변종 또는 단편일 수 있다. 촉매 활성과 관련하여 활성 변종 또는 단편은 야생형 또는 변형된 Cas 단백질 또는 이것들의 일부에 대해 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상의 서열 동일성을 포함할 수 있고, 여기서 활성 변종은 원하는 절단 부위에서 커팅되는 능력을 보유하며 따라서 닉-유도 또는 이중 가닥 끊김-유도 활성을 보유한다. 닉-유도 또는 이중 가닥 끊김-유도 활성에 대한 검정은 알려져 있고 일반적으로 절단 부위를 함유한 DNA 기질에 대한 Cas 단백질의 전체적인 활성 및 특이성을 측정한다.

[0327] Cas 단백질은 핵산 결합 친화도, 핵산 결합 특이성, 및 효소 활성 중 하나 이상을 증가 또는 감소시키기 위하여 변형될 수 있다. Cas 단백질은 또한 단백질의 임의의 다른 활성 또는 특성, 예컨대 안정성을 변화시키기 위해 변형될 수 있다. 예를 들어, Cas 단백질의 하나 이상의 뉴클레아제 도메인은 변형되거나, 삭제되거나, 또는 비활성화될 수 있고, 또는 Cas 단백질은 단백질의 기능에 필수적이지 않은 도메인을 제거하기 위하여 또는 Cas 단백질의 활성을 최적화 (예컨대, 증강 또는 감소)시키기 위하여 절단될 수 있다.

[0328] Cas 단백질은 적어도 하나의 뉴클레아제 도메인, 예컨대 DNase 도메인을 포함할 수 있다. 예를 들어, 야생형 Cpf1 단백질은 일반적으로, 아마도 다이머 형태의 표적 DNA의 양 가닥을 절단하는 RuvC-유사 도메인을 포함한다. Cas 단백질은 또한 적어도 2개의 뉴클레아제 도메인, 예컨대 DNase 도메인을 포함할 수 있다. 예를 들어, 야생형 Cas9 단백질은 일반적으로 RuvC-유사 뉴클레아제 도메인 및 HNH-유사 뉴클레아제 도메인을 포함한다. RuvC 및 HNH 도메인은 각각 DNA에 이중 가닥 끊김을 만들기 위하여 이중 가닥 DNA의 상이한 가닥을 자를 수 있다. 예컨대, Jinek et al. (2012) *Science* 337:816-821 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)을 참조한다.

[0329] 하나 이상의 뉴클레아제 도메인이 더 이상 기능성이지 않도록 또는 감소된 뉴클레아제 활성을 갖도록 삭제되거나 돌연변이될 수 있다. 예를 들어, 한 뉴클레아제 도메인이 Cas9 단백질에서 삭제되거나 돌연변이되면, 그 결과로 얻어지는 Cas9 단백질은 닉카아제로서 언급될 수 있고 이중 가닥 끊김이 아닌 이중 가닥 DNA 내의 가이드 RNA 표적 서열에서 단일 가닥 끊김을 생성할 수 있다 (즉, 두 가닥 전부가 아닌, 상보하는 가닥 또는 비-상보하는하는 가닥을 절단할 수 있다). 만약 뉴클레아제 도메인이 둘 다 삭제되거나 돌연변이되며, 그 결과로 얻어지는 Cas 단백질 (예컨대, Cas9)는 이중 가닥 DNA의 양 가닥을 절단하는 능력이 감소될 것이다 (예컨대, 뉴클레아제-널 또는 뉴클레아제-비활성 Cas 단백질, 또는 촉매적으로 활성이 없는 Cas 단백질 (dCas)). Cas9를 닉카아제로 전환시키는 돌연변이의 예는 S. 피오게네스로부터의 Cas9의 RuvC 도메인에서의 D10A (Cas9의 위치 10에서의 아스파테이트의 알라닌으로의) 돌연변이이다. 마찬가지로, S. 피오게네스로부터의 Cas9의 HNH 도메인에서 H939A (아미노산 위치 839에서 히스티딘의 알라닌으로) 또는 H840A (아미노산 위치 840에서 히스티딘의 알라닌으로)는 Cas9를 닉카아제로 전환시킬 수 있다. Cas9를 닉카아제로 전환시키는 돌연변이의 또 다른 예로는 S. 써모필루스로부터의 Cas9에 대한 상응하는 돌연변이를 들 수 있다. 예컨대, Sapranauskas et al. (2011) *Nucleic Acids Research* 39:9275-9282 및 WO 2013/141680을 참조하며, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다. 이러한 돌연변이는 부위-특정 돌연변이생성, PCR-매개된 돌연변이생성, 또는 총 유전자 합성

과 같은 방법을 사용하여 생성될 수 있다. 닉카아제를 생성하는 다른 돌연변이의 예는, 예를 들어, WO 2013/176772 및 WO 2013/142578에서 찾아볼 수 있고, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다.

[0330] Cas 단백질 (예컨대, 뉴클레아제-활성 Cas 단백질 또는 뉴클레아제-비활성 Cas 단백질)은 또한 융합 단백질로서 이종성 폴리펩타이드에 작동 가능하게 연결될 수 있다. 예를 들어, Cas 단백질은 절단 도메인, 에피유전자 변형 도메인, 전사 활성화 도메인, 또는 전사 억제자 도메인에 융합될 수 있다. WO 2014/089290 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)을 참조한다. 전사 활성화 도메인의 예로는 단순 포진 바이러스 VP16 활성화 도메인, VP64 (VP16의 테트라머 유도체임), NFκB p65 활성화 도메인, p53 활성화 도메인 1 및 2, CREB (cAMP 반응 요소 결합 단백질) 활성화 도메인, E2A 활성화 도메인, 및 NFAT (활성화된 T-세포의 핵 인자) 활성화 도메인을 들 수 있다. 다른 예로는 Oct1, Oct-2A, SP1, AP-2, CTF1, P300, CBP, PCAF, SRC1, PvALF, ERF-2, OsGAI, HALF-1, C1, AP1, ARF-5, ARF-6, ARF-7, ARF-8, CPRF1, CPRF4, MYC-RP/GP, TRAB1PC4, 및 HSF1로부터의 활성화 도메인을 들 수 있다. 예컨대, US 2016/0237456, EP3045537, 및 WO 2011/145121을 참조하며, 이것들은 각각 그 전문이 모든 목적에 대해 참조로 포함된다. 일부 경우에, MS2-p65-HSF1과 쌍을 이룬 dCas9-VP64 융합 단백질을 포함한 전사 활성화 시스템이 사용될 수 있다. 이러한 시스템에서 가이드 RNA는 다이머화된 MS2 박테리오파지 코트 단백질에 결합하도록 설계된 sgRNA 테트라루프 및 스템-루프 2에 첨부된 압타머 서열로 설계될 수 있다. 예컨대, Konermann et al. (2015) *Nature* 517(7536):583-588을 참조하며. 이 문헌은 모든 목적에 대해 그 전문이 참조로 포함된다. 전사 억제자 도메인의 예로는 유도성 cAMP 초기 억제자 (ICER) 도메인, 크루펠-관련 박스 A (KRAB-A) 억제자 도메인, YY1 글리신 풍부 억제자 도메인, Sp1-유사 억제자, E(spl) 억제자, IκB 억제자, 및 MeCP2를 들 수 있다. 다른 예로는 A/B, KOX, TGF-베타-유도성 초기 유전자 (TIEG), v-erbA, SID, SID4X, MBD2, MBD3, DNMT1, DNMG3A, DNMT3B, Rb, ROM2로부터의 전사 억제자 도메인을 들 수 있다. 예컨대, EP3045537 및 WO 2011/145121을 참조하며, 이 문헌들은 각각 모든 목적에 대해 그 전문이 참조로 포함된다. Cas 단백질은 또한 증가된 또는 감소된 안정성을 제공하는 이종성 폴리펩타이드에 융합될 수 있다. 융합된 도메인 또는 이종성 폴리펩타이드는 Cas 단백질의 N-말단에, C-말단에, 또는 Cas 단백질 내부에 위치할 수 있다.

[0331] 한 예로서, Cas 단백질은 세포내 국지화를 제공하는 이종성 폴리펩타이드에 융합될 수 있다. 이러한 이종성 폴리펩타이드로는, 예를 들어, 핵으로 표적화하기 위한 SV40 NLS와 같은 하나 이상의 핵 국지화 신호 (NLS), 미토콘드리아로 표적화하기 위한 미토콘드리아 국지화 신호, ER 보유 신호, 등을 들 수 있다. 예컨대, Lange et al. (2007) *J. Biol. Chem.* 282:5101-5105을 참조하며, 이 문헌은 모든 목적에 대해 그 전문이 본원에 참조로 포함된다. 이러한 세포내 국지화 신호는 Cas 단백질의 N-말단에, C-말단에, 또는 Cas 단백질 내 어느 곳에든지 위치할 수 있다. NLS는 기본적인 아미노산 구간을 포함할 수 있고, 모노파타이트 서열(monopartite sequence) 또는 바이파타이트 서열(bipartite sequence)일 수 있다.

[0332] Cas 단백질은 또한 세포-투과 도메인에 작동 가능하게 연결될 수 있다. 예를 들어, 세포-투과 도메인은 HIV-1 TAT 단백질, 인간 B형 간염 바이러스로부터의 TLM 세포-투과 모티프, MPG, Pep-1, VP22, 단순 포진 바이러스로부터의 세포 투과 펩타이드, 또는 폴리아르기닌 펩타이드 서열로부터 유래될 수 있다. 예컨대, WO 2014/089290 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)을 참조한다. 세포-투과 도메인은 Cas 단백질의 N-말단에, C-말단에, 또는 Cas 단백질 내 어느 곳에서든지 위치할 수 있다.

[0333] Cas 단백질은 또한 추적 또는 정제의 용이성을 위한 이종성 폴리펩타이드, 예컨대 형광 단백질, 정제 태그, 또는 에피토프 태그에 작동가능하게 연결될 수 있다. 형광 단백질의 예로는 녹색 형광 단백질 (예컨대, GFP, GFP-2, tagGFP, turboGFP, eGFP, 에메랄드, 아자미 그린, 모노머 아자미 그린, CopGFP, AceGFP, ZsGreen1), 황색 형광 단백질 (예컨대, YFP, eYFP, 시트린, 비너스, YPet, PhiYFP, ZsYellow1), 청색 형광 단백질 (예컨대, eBFP, eBFP2, 아즈라이트, mKalama1, GFPuv, 사파이어, T-사파이어), 청록색 형광 단백질 (예컨대, eCFP, 세룰리언, CyPet, AmCyan1, Midoriishi-Cyan), 적색 형광 단백질 (예컨대, mKate, mKate2, mPlum, DsRed 모노머, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-모노머, HcRed-탠덤, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred), 오렌지색 형광 단백질 (예컨대, mOrange, mKO, Kusabira-Orange, 모노머 Kusabira-Orange, mTangerine, tdTomato), 및 임의의 기타 적합한 형광 단백질을 들 수 있다. 태그의 예로는 글루타티온-S-트랜스퍼라제 (GST), 키틴 결합 단백질 (CBP), 말토오스 결합 단백질, 티오레독신 (TRX), 폴리(NANP), 탠덤 친화도 정제 (TAP) 태그, myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, 헤마글루티닌 (HA), nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 히스티딘 (His), 비오틴 카르복실 담체 단백질 (BCCP), 및 칼모둘린을 들 수 있다.

[0334] Cas 단백질은 또한 외인성 도너 서열 또는 표지된 핵산에 묶일 수 있다. 이러한 묶음 (즉, 물리적 연결)은 공유

상호작용 또는 비공유 상호작용을 통해 이루어질 수 있고, 묶음은 직접적이거나 (예컨대, 단백질 상의 시스테인 또는 리신 잔기의 변형에 의해 또는 인테인 변형에 의해 이루어질 수 있는 직접 융합 또는 화학적 콘쥬게이션을 통해), 또는 하나 이상의 개재하는 링커 또는 스트렙트아비딘 또는 압타머와 같은 어댑터 분자를 통해 이루어질 수 있다. 예컨대, Pierce et al. (2005) *Mini Rev. Med. Chem.* 5(1):41-55; Duckworth et al. (2007) *Angew. Chem. Int. Ed. Engl.* 46(46):8819-8822; Schaeffer and Dixon (2009) *Australian J. Chem.* 62(10):1328-1332; Goodman et al. (2009) *Chembiochem.* 10(9):1551-1557; 및 Khatwani et al. (2012) *Bioorg. Med. Chem.* 20(14):4532-4539을 참조하며, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다. 단백질-핵산 콘쥬게이트를 합성하기 위한 비공유 전략은 비오틴-스트렙트아비딘 및 니켈-히스티딘 방법을 포함한다. 공유 단백질-핵산 콘쥬게이트는 적절하게 기능화된 핵산 및 단백질을 광범위한 화학을 사용하여 연결시킴으로써 합성될 수 있다. 이런 화학의 일부는 올리고뉴클레오타이드의 단백질 표면 상의 아미노산 잔기 (예컨대 리신 아민 또는 시스테인 티올)로의 직접적인 부착을 포함하는 한편, 다른 더 복잡한 계획은 단백질의 번역후 변형 또는 촉매성 또는 반응성 단백질 도메인의 포함을 필요로 한다. 단백질의 핵산으로의 공유 부착을 위한 방법으로는, 예를 들어, 올리고뉴클레오타이드의 단백질 리신 또는 시스테인 잔기로의 화학적 가교 결합, 발현된 단백질 결찰, 화학효소적 방법, 및 광압타머(photoaptamer)의 사용을 들 수 있다. 외인성 도너 서열 또는 표지된 핵산은 Cas 단백질의 N-말단에, C-말단에, 또는 Cas 단백질 내의 내부 영역에 묶일 수 있다. 바람직하게, 외인성 도너 서열 또는 표지된 핵산은 Cas 단백질의 N-말단 또는 C-말단에 묶인다. 마찬가지로, Cas 단백질은 외인성 도너 서열 또는 표지된 핵산의 5' 단부, 3' 단부, 또는 외인성 도너 서열 또는 표지된 핵산 내의 내부 영역에 묶일 수 있다. 즉, 외인성 도너 서열 또는 표지된 핵산은 임의의 방향 및 극성으로 묶일 수 있다. 바람직하게, Cas 단백질은 외인성 도너 서열 또는 표지된 핵산의 5' 단부 또는 3' 단부에 묶인다.

[0335] Cas 단백질은 임의의 형태로 제공될 수 있다. 예를 들어, Cas 단백질은 단백질, 예컨대 gRNA와 복합체를 형성한 Cas 단백질의 형태로 제공될 수 있다. 대안적으로, Cas 단백질은 Cas 단백질을 암호화하는 핵산, 예컨대 RNA (예컨대, 메신저 RNA (mRNA)) 또는 DNA의 형태로 제공될 수 있다. 선택적으로, Cas 단백질을 암호화하는 핵산은 특정 세포 또는 유기체에서 단백질로의 효율적인 번역을 위해 코돈 최적화될 수 있다. 예를 들어, Cas 단백질을 암호화하는 핵산은 박테리아 세포, 효모 세포, 인간 세포, 비인간 세포, 포유류 세포, 설치류 세포, 마우스 세포, 래트 세포 또는 관심이 있는 임의의 다른 숙주 세포에서, 자연적으로 발생하는 폴리뉴클레오타이드 서열에 비교하여, 고빈도로 사용되는 코돈을 치환하기 위해 변형될 수 있다. Cas 단백질을 암호화하는 핵산이 세포 안으로 도입될 때, Cas 단백질은 일시적으로, 조건부로, 또는 구성적으로 세포에서 발현될 수 있다.

[0336] Cas 단백질을 암호화하하는 핵산은 세포의 게놈에 안정적으로 통합될 수 있고 세포에서 활성인 프로모터에 작동가능하게 연결될 수 있다. 대안적으로, Cas 단백질을 암호화하하는 핵산은 발현 구성물에서 프로모터에 작동가능하게 연결된다. 발현 구성물은 유전자 또는 관심이 있는 다른 핵산 서열 (예컨대, Cas 유전자)의 발현을 지시할 수 있고 관심이 있는 이러한 핵산 서열을 표적 세포로 전달할 수 있는 임의의 핵산 구성물을 포함한다. 예를 들어, Cas 단백질을 암호화하는 핵산은 핵산 삽입물을 포함하는 표적화 벡터 및/또는 gRNA를 암호화하는 DNA를 포함하는 벡터에 있을 수 있다. 대안적으로, 그것은 핵산 삽입물을 포함하는 표적화 벡터와 다른 및/또는 gRNA를 암호화하는 DNA를 포함하는 벡터와 다른 벡터 또는 플라스미드에 있을 수 있다. 발현 구성물에 사용될 수 있는 프로모터는, 예를 들어, 진핵 세포, 인간 세포, 비인간 세포, 포유류 세포, 비인간 포유류 세포, 설치류 세포, 마우스 세포, 래트 세포, 햄스터 세포, 토끼 세포, 만능 세포, 배아 줄기 (ES) 세포, 또는 접합체 중 하나 이상에서 활성인 프로모터를 포함한다. 이러한 프로모터로는, 예를 들어, 조건부 프로모터, 유도성 프로모터, 구성성 프로모터, 또는 조직-특이적 프로모터일 수 있다. 선택적으로, 프로모터는 한 방향으로 Cas 단백질의 발현과 다른 방향으로 가이드 RNA의 발현 둘 다를 구동시키는 양방향성 프로모터일 수 있다. 이러한 양방향성 프로모터는 (1) 3개의 외부 제어 요소: 원거리 서열 요소 (DSE), 근거리 서열 요소 (PSE), 및 TATA 박스를 함유한 완전한, 종래의, 일방향성 Pol III 프로모터; 및 (2) PSE 및 역방향으로 DSE의 5' 말단에 융합된 TATA 박스를 포함하는 제2 기본 Pol III 프로모터로 구성될 수 있다. 예를 들어, H1 프로모터에서, DSE는 PSE 및 TATA 박스에 인접하고, 프로모터는 역방향의 전사가 U6 프로모터로부터 유래된 PSE 및 TATA 박스를 첨부시킴으로써 제어되는 하이브리드 프로모터를 생성함으로써 양방향성이 될 수 있다. 예컨대, US 2016/0074535 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)를 참조한다. Cas 단백질 및 가이드 RNA를 암호화하는 유전자들을 발현하기 위한 양방향성 프로모터의 사용은 동시에 전달을 용이하게 하기 위한 압축 발현 카세트의 생성을 허용한다.

[0337] ***(2) 가이드 RNA***

[0338] "가이드 RNA" 또는 "gRNA"는 Cas 단백질 (예컨대, Cas9 단백질)에 결합하고 표적 DNA (예컨대, *HSD17B13* 유전자) 내의 특정 위치에 Cas 단백질을 표적화하는 RNA 분자이다. 특히, 본원에는 *HSD17B13* 유전자좌 또는

HSD17B13 유전자에 결합하거나 그것을 절단하기 위해 Cas 효소를 지시하기에 효과적인 가이드 RNA가 개시된다. 한 예시적인 가이드 RNA는 *HSD17B13* 유전자에 결합하거나 그것을 절단하도록 Cas 효소를 지시하기에 효과적인 가이드 RNA이고, 여기서 가이드 RNA는 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치를 포함하거나 그 위치에 근접한 *HSD17B13* 유전자 내의 가이드 RNA 인식 서열에 혼성화하는 (즉, 가이드 RNA 표적 서열을 표적화하는) DNA-표적화 분절을 포함한다. 가이드 RNA 표적 서열을 표적화한다는 것은 비-상보하는하는 가닥 상의 가이드 RNA 표적 서열의 역 상보물인 상보하는 가닥 서열에 혼성화하는 것을 의미한다. 예를 들어, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 약 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 LDt을 수 있다. 다른 예시적인 가이드 RNA는 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 엑손 6 및/또는 인트론 6에 상응하는 영역 내에 있는 *HSD17B13* 유전자 내의 가이드 RNA 표적 서열을 표적화하는 DNA-표적화 분절을 포함한다. 다른 예시적인 가이드 RNA는 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 엑손 6 및/또는 인트론 6 및/또는 엑손 7에 상응하는 영역 내에 있는 *HSD17B13* 유전자 내의 가이드 RNA 표적 서열을 표적화하는 DNA-표적화 분절을 포함한다. 다른 예시적인 가이드 RNA는 *HSD17B13* 유전자의 개시 코돈을 포함하거나 그것에 근접한 또는 *HSD17B13* 유전자의 종결 코돈을 포함하거나 그것에 근접한 *HSD17B13* 유전자 내의 가이드 RNA 인식 서열에 혼성화하는 (즉, 가이드 RNA 표적 서열을 표적화하는) DNA-표적화 분절을 포함한다. 예를 들어, 가이드 RNA 표적 서열은 개시 코돈의 약 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 또는 종결 코돈의 약 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있을 수 있다. 예를 들어, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 1 또는 2와 최적으로 정렬될 때 SEQ ID NO: 1 또는 2의 엑손 1에 상응하는 영역 내에 있을 수 있다. 마찬가지로, 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 1 또는 2와 최적으로 정렬될 때 SEQ ID NO: 1 또는 2의 엑손 7에 상응하는 영역 내에 있을 수 있다. *HSD17B13* 유전자는 임의의 유기체로부터의 *HSD17B13* 유전자일 수 있다. 예를 들어, *HSD17B13* 유전자는 인간 *HSD17B13* 유전자이거나 다른 유기체, 예컨대 비인간 포유류, 설치류, 마우스, 또는 래트로부터의 병렬상동체일 수 있다.

[0339] 인간 *HSD17B13* 유전자의 5' 단부에서의 가이드 RNA 표적 서열의 예는 SEQ ID NO: 20 내지 81에 제시되고 하기 표에 제시된 서열을 포함하거나, 본질적으로 그것들로 구성되거나, 또는 그것들로 구성된다. SEQ ID NO: 20 내지 81에 상응하는 가이드 RNA DNA-표적화 분절의 예는 하기 표에 제시되고 티민 대신 우라실인 것을 제외하고 SEQ ID NO: 20 내지 81에 동일하다. 가이드 RNA DNA-표적화 분절은 하기 표에 제시된 DNA-표적화 분절 중 임의의 것을 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있다. 인간 *HSD17B13* 유전자의 전사 개시 부위 (TSS)에 인접한 가이드 RNA 표적 서열의 예는 SEQ ID NO: 20 내지 41에 제시되고 하기 표에 제시된 서열을 포함하거나, 본질적으로 그것들로 구성되거나, 또는 그것들로 구성된다. TSS에 인접한 예시적인 가이드 RNA 표적 서열은 SEQ ID NO: 21 내지 23, 33, 및 35를 포함한다. SEQ ID NO: 33 및 35가 TSS에 가장 가깝다. 인간 *HSD17B13* 유전자의 5' 단부에서의 가이드 RNA 표적 서열에 상응하는 예시적인 crRNA 및 sgRNA (스캐폴드 버전 1, 2, 3, 또는 4를 포함함)는 하기 표에 제시된 서열들 중 임의의 것을 포함하거나, 본질적으로 그것으로 구성되거나, 또는 구성된다.

[0340] 인간 *HSD17B13* 유전자의 5' 단부에서의 가이드 RNA 표적 서열

가닥	PAM	가이드 RNA 표적 서열	SEQ ID NO						
			gRNA 표적 서열	DNA-표적화 분절	crRNA	sgRNA			
						v1	v2	v3	v4
-	GGG	TGTCAGGTTAGTTAGATGAA	42	1423	270	500	730	960	1190
-	AGG	GTGTCAGGTTAGTTAGATGA	43	1424	271	501	731	961	1191
+	AGG	CCTGACACATATACAGACTA	44	1425	272	502	732	962	1192
+	GGG	CTGACACATATACAGACTAA	45	1426	273	503	733	963	1193
-	AGG	CCTTAGTCTGTATATGTGTC	46	1427	274	504	734	964	1194
+	AGG	CATATACAGACTAAGGGACC	47	1428	275	505	735	965	1195
+	GGG	ATATACAGACTAAGGGACCA	48	1429	276	506	736	966	1196
-	TGG	TCAAAGTTTGATAAATTCCC	49	1430	277	507	737	967	1197
+	TGG	AAAATACAAAGATAAGTAGA	50	1431	278	508	738	968	1198
+	TGG	ACTCTGTGACTTTAAAAAGT	51	1432	279	509	739	969	1199
-	AGG	GGTTCTGTGGGATATTAATA	52	1433	280	510	740	970	1200
-	GGG	ACAGAGCATATTGGTTCTGT	53	1434	281	511	741	971	1201
-	TGG	GACAGAGCATATTGGTTCTG	54	1435	282	512	742	972	1202
-	TGG	TGCAAAACGACAGAGCATAT	55	1436	283	513	743	973	1203

[0341]

-	AGG	GAGCTGGGCATGGAATAGGC	56	1437	284	514	744	974	1204
-	AGG	ACTGGAGCTGGGCATGGAAT	57	1438	285	515	745	975	1205
-	TGG	CTCATTACTGGAGCTGGGCA	58	1439	286	516	746	976	1206
-	GGG	TTGTTCTCATTACTGGAGCT	59	1440	287	517	747	977	1207
-	TGG	ATTGTTCTCATTACTGGAGC	60	1441	288	518	748	978	1208
-	TGG	GGGGAGATTGTTCTCATTAC	61	1442	289	519	749	979	1209
-	GGG	GAGGAGAAAATCTGTGGCTG	62	1443	290	520	750	980	1210
-	GGG	AGAGGAGAAAATCTGTGGCT	63	1444	291	521	751	981	1211
-	TGG	CAGAGGAGAAAATCTGTGGC	64	1445	292	522	752	982	1212
-	TGG	TCCTCAGAGGAGAAAATCTG	65	1446	293	523	753	983	1213
-	AGG	TGAAGTTTTTCATTCCTCAG	20	1447	294	524	754	984	1214
+	AGG	CTTCACCAACGACTCCAAGT	21	1448	295	525	755	985	1215
-	TGG	CTACTCCTACTTGGAGTCGT	22	1449	296	526	756	986	1216
+	TGG	CTCCAAGTAGGAGTAGATGA	23	1450	297	527	757	987	1217
-	TGG	CACCATCATCTACTCCTACT	24	1451	298	528	758	988	1218
+	AGG	TGATGGTGATCAGAAGCAGA	25	1452	299	529	759	989	1219
+	AGG	TCAGAAGCAGAAGGATTTCT	26	1453	300	530	760	990	1220
+	TGG	GATTTCTAGGATGATGTTCA	27	1454	301	531	761	991	1221
+	TGG	TTGCTCTGTCCTCTTCCTTC	28	1455	302	532	762	992	1222
-	AGG	AGGACTGAACCAGAAGGAAG	29	1456	303	533	763	993	1223
-	AGG	TACACAAGGACTGAACCAGA	30	1457	304	534	764	994	1224
+	AGG	TTCAGTCCTTGTGTAGTCCT	31	1458	305	535	765	995	1225
+	GGG	TCAGTCCTTGTGTAGTCCTA	32	1459	306	536	766	996	1226
+	AGG	GTCCTTGTGTAGTCCTAGGG	33	1460	307	537	767	997	1227
+	AGG	CTTGTGTAGTCCTAGGGAGG	34	1461	308	538	768	998	1228
-	AGG	CTCCTCCCTAGGACTACACA	35	1462	309	539	769	999	1229
-	AGG	GTAGACAGTACCTCCTCCCT	36	1463	310	540	770	1000	1230
+	AGG	TACTGTCTACACAGAGCTCT	37	1464	311	541	771	1001	1231
+	GGG	ACTGTCTACACAGAGCTCTA	38	1465	312	542	772	1002	1232
+	AGG	TCTACACAGAGCTCTAGGGA	39	1466	313	543	773	1003	1233
+	GGG	CTACACAGAGCTCTAGGGAA	40	1467	314	544	774	1004	1234
+	GGG	TACACAGAGCTCTAGGGAAG	41	1468	315	545	775	1005	1235
+	TGG	GGGGTGTGCCCAGTTGTTAA	66	1469	316	546	776	1006	1236
+	GGG	GGGTGTGCCCAGTTGTTAAT	67	1470	317	547	777	1007	1237
-	GGG	TGGTAGTCCCATTAACAACT	68	1471	318	548	778	1008	1238
-	TGG	CTGGTAGTCCCATTAACAAC	69	1472	319	549	779	1009	1239
+	TGG	TTGTTAATGGGACTACCAGA	70	1473	320	550	780	1010	1240
+	TGG	TACCAGATGGAAGCCAGCTT	71	1474	321	551	781	1011	1241
-	TGG	TTCCAAAGCTGGCTTCCATC	72	1475	322	552	782	1012	1242
+	AGG	TGGAAGCCAGCTTTGGAAGC	73	1476	323	553	783	1013	1243
-	TGG	ACAAGGCCTGCTTCCAAAGC	74	1477	324	554	784	1014	1244
+	TGG	GCCTTGTTCACGTGTTCTAA	75	1478	325	555	785	1015	1245
+	GGG	CCTTGTTCACGTGTTCTAAT	76	1479	326	556	786	1016	1246
-	AGG	CCCATTAGAACACGTGAACA	77	1480	327	557	787	1017	1247
-	AGG	TTGGCATCACTTCATATTTG	78	1481	328	558	788	1018	1248
-	TGG	CTTGTGCTCTTGGCATCACT	79	1482	329	559	789	1019	1249
-	TGG	AGCACACTCTCTTGTGCTCT	80	1483	330	560	790	1020	1250

[0342]

[0343]

+	TGG	GCACAAGAGAGTGTGCTCTC	81	1484	331	561	791	1021	1251

[0344] 인간 *HSD17B13* 유전자의 3' 단부에서의 가이드 RNA 표적 서열의 예는 SEQ ID NO: 82 내지 225에 제시되고 하기 표에 제시된 서열들을 포함하거나, 본질적으로 그것들로 구성되거나, 또는 구성된다. SEQ ID NO: 82 내지 225에 상응하는 가이드 RNA DNA-표적화 분절의 예는 각각, 티민 대신 우라실인 것을 제외하고 SEQ ID NO: 82 내지 225와 동일한 SEQ ID NO: 1485 내지 1628에 제시된다. 가이드 RNA DNA-표적화 분절은 SEQ ID NO: 1485 내지 1628에 제시된 임의의 서열을 포함하거나, 본질적으로 그것으로 구성되거나, 또는 구성될 수 있다. 인간 *HSD17B13* 유전자의 3' 단부에서의 가이드 RNA 표적 서열에 상응하는 예시적인 crRNA 및 sgRNA (스캐폴드 버전 1, 2, 3, 또는 4를 포함함)는 하기 표에 제시된 서열들 중 임의의 것을 포함하거나, 본질적으로 그것으로 구성되거나, 또는 구성된다.

[0345] 인간 *HSD17B13* 유전자의 3' 단부에서의 가이드 RNA 표적 서열

가닥	PAM	가이드 RNA 표적 서열	SEQ ID NO					
			gRNA 표적 서열	crRNA	sgRNA			
					v1	v2	v3	v4
+	AGG	GCTTAATCTCACACATAGAA	82	332	562	792	1022	1252
+	GGG	CTTAATCTCACACATAGAAA	83	333	563	793	1023	1253
+	GGG	TTAATCTCACACATAGAAAG	84	334	564	794	1024	1254
-	TGG	AGGAGTGCTGGTTTATCAAC	85	335	565	795	1025	1255
-	TGG	TTCTTTGACAGCAGGAGTGC	86	336	566	796	1026	1256
-	AGG	ACTCTGGTTTCTTTGACAGC	87	337	567	797	1027	1257
+	TGG	ACCAGAGTTGAGAAAACCCC	88	338	568	798	1028	1258
-	TGG	TCCAGGGGTTTTCTCAACTC	89	339	569	799	1029	1259
-	GGG	CAGTTATTAAATGAATCCAG	90	340	570	800	1030	1260
-	GGG	GCAGTTATTAAATGAATCCA	91	341	571	801	1031	1261
-	AGG	GGCAGTTATTAAATGAATCC	92	342	572	802	1032	1262
-	TGG	TGGATGGTAACAGCTACATC	93	343	573	803	1033	1263
+	TGG	GCTGTTACCATCCACATCCT	94	344	574	804	1034	1264
-	TGG	TCAAGAACCAAGGATGTGGA	95	345	575	805	1035	1265
-	TGG	TCCTTCAAGAACCAAGGATG	96	346	576	806	1036	1266
-	AGG	TGAGTGTCCTTCAAGAACCA	97	347	577	807	1037	1267

[0346]

+	AGG	TTTTATTTTATAACTACAAG	98	348	578	808	1038	1268
+	AGG	TTGTTTTTAATAAAAACAAG	99	349	579	809	1039	1269
-	TGG	TATTATAGAATGCTTTTGCA	100	350	580	810	1040	1270
+	TGG	CAAGATTAGTCTTGATGTAG	101	351	581	811	1041	1271
+	GGG	AAGATTAGTCTTGATGTAGT	102	352	582	812	1042	1272
+	CGG	AGTCTTGATGTAGTGGGAGT	103	353	583	813	1043	1273
+	AGG	TTTTTCTATTAAAAAAAAAA	104	354	584	814	1044	1274
+	TGG	TCTATTAAAAAAAAAAAGGC	105	355	585	815	1045	1275
+	GGG	CTATTAAAAAAAAAAAGGCT	106	356	586	816	1046	1276
+	CGG	AAAAAAAAAAAGGCTGGGCA	107	357	587	817	1047	1277
+	TGG	AAAAAAAAGGCTGGGCACGG	108	358	588	818	1048	1278
+	TGG	CACCCGTAATCCCAGCACTT	109	359	589	819	1049	1279
+	GGG	ACCCGTAATCCCAGCACTTT	110	360	590	820	1050	1280
+	AGG	CGTAATCCCAGCACTTTGGG	111	361	591	821	1051	1281
-	GGG	TCCCAAAGTGCTGGGATTAC	112	362	592	822	1052	1282
-	CGG	CTCCCAAAGTGCTGGGATTA	113	363	593	823	1053	1283
+	AGG	CCCAGCACTTTGGGAGGCCG	114	364	594	824	1054	1284
-	GGG	CCTCGGCCTCCCAAAGTGCT	115	365	595	825	1055	1285
+	AGG	GCACTTTGGGAGGCCGAGGC	116	366	596	826	1056	1286
+	TGG	CTTTGGGAGGCCGAGGCAGG	117	367	597	827	1057	1287
+	AGG	GCCGAGGCAGGTGGATCACG	118	368	598	828	1058	1288
-	CGG	ACCTCGTGATCCACCTGCCT	119	369	599	829	1059	1289
+	AGG	GGCAGGTGGATCACGAGGTC	120	370	600	830	1060	1290
+	TGG	TCAGGAGATCGAGACCATCT	121	371	601	831	1061	1291
+	TGG	CGAGACCATCTTGGCTAACA	122	372	602	832	1062	1292
-	TGG	TTTCACCATGTTAGCCAAGA	123	373	603	833	1063	1293
-	GGG	TTGTATTTTTTGTAGAGACG	124	374	604	834	1064	1294
-	GGG	TTTGTATTTTTTGTAGAGAC	125	375	605	835	1065	1295
-	CGG	TTTTGTATTTTTTGTAGAGA	126	376	606	836	1066	1296
+	CGG	AAAAAATACAAAAAATTAGC	127	377	607	837	1067	1297
+	GGG	AAAAATACAAAAAATTAGCC	128	378	608	838	1068	1298
+	TGG	TACAAAAAATTAGCCGGGTG	129	379	609	839	1069	1299
+	TGG	AAAAAATTAGCCGGGTGTGG	130	380	610	840	1070	1300
+	CGG	AAATTAGCCGGGTGTGGTGG	131	381	611	841	1071	1301
+	GGG	AATTAGCCGGGTGTGGTGGC	132	382	612	842	1072	1302
-	CGG	CAGGCGCCCGCCACCACACC	133	383	613	843	1073	1303
+	AGG	GCCTGTAGTCCCAGCTACTC	134	384	614	844	1074	1304
+	AGG	TGTAGTCCCAGCTACTCAGG	135	385	615	845	1075	1305
-	AGG	TCCTGAGTAGCTGGGACTAC	136	386	616	846	1076	1306
+	AGG	CCCAGCTACTCAGGAGGCTG	137	387	617	847	1077	1307
-	GGG	CCTCAGCCTCCTGAGTAGCT	138	388	618	848	1078	1308
-	TGG	GCCTCAGCCTCCTGAGTAGC	139	389	619	849	1079	1309
+	TGG	AGGAGGCTGAGGCAGGAGAA	140	390	620	850	1080	1310
+	CGG	GCAGGAGAATGGCGTGAACC	141	391	621	851	1081	1311
+	GGG	CAGGAGAATGGCGTGAACCC	142	392	622	852	1082	1312
+	AGG	GAGAATGGCGTGAACCCGGG	143	393	623	853	1083	1313
+	TGG	AATGGCGTGAACCCGGGAGG	144	394	624	854	1084	1314

[0347]

-	GGG	CACTGCAAGCTCCACCTCCC	145	395	625	855	1085	1315
-	CGG	TCACTGCAAGCTCCACCTCC	146	396	626	856	1086	1316
+	TGG	CATACCACTGCACTCCAGCC	147	397	627	857	1087	1317
+	GGG	ATACCACTGCACTCCAGCCT	148	398	628	858	1088	1318
-	TGG	TCGCCCAGGCTGGAGTGCAG	149	399	629	859	1089	1319
-	TGG	TCTCACTCTTTCGCCCAGGC	150	400	630	860	1090	1320
-	AGG	GGAGTCTCACTCTTTCGCCC	151	401	631	861	1091	1321
-	TGG	TGTTTTTTGTTTTTTTGAGA	152	402	632	862	1092	1322
-	TGG	AGGAAGAAAGAAAGGTTTTT	153	403	633	863	1093	1323
-	AGG	AGAAGAAAAGGAAGAAAGAA	154	404	634	864	1094	1324
+	TGG	CTTTCTTCCTTTTCTTCTCT	155	405	635	865	1095	1325
+	GGG	TTTCTTCCTTTTCTTCTCTT	156	406	636	866	1096	1326
-	AGG	AATGGACCCAAGAGAAGAAA	157	407	637	867	1097	1327
-	TGG	GGCTATTACATAAGAAACAA	158	408	638	868	1098	1328
-	TGG	CACAGGAAAAGGAACTGTAC	159	409	639	869	1099	1329
-	AGG	ATTAAAGCTAACACAGGAAA	160	410	640	870	1100	1330
-	AGG	TCAAAAATTAAAGCTAACAC	161	411	641	871	1101	1331
+	TGG	TAAAATTGTCTAAACATCTC	162	412	642	872	1102	1332
-	AGG	AGAGATGTTTAGACAATTTT	163	413	643	873	1103	1333
+	AGG	TCTAAACATCTCTGGGACCA	164	414	644	874	1104	1334
-	TGG	TTTATGCTTTCATATATCCT	165	415	645	875	1105	1335
+	AGG	AGCATAAATTACAAAGAAAA	166	416	646	876	1106	1336
+	TGG	TACAAAGAAAAAGGTTATCA	167	417	647	877	1107	1337
+	GGG	ACAAAGAAAAAGGTTATCAT	168	418	648	878	1108	1338
+	GGG	CAAAGAAAAAGGTTATCATG	169	419	649	879	1109	1339
+	CGG	TCTGAGATTTAAAATAGAGT	170	420	650	880	1110	1340
-	AGG	CTTATAAGATACATTATGAA	171	421	651	881	1111	1341
+	AGG	TATCTTATAAGACTATAAAA	172	422	652	882	1112	1342
+	GGG	ATCTTATAAGACTATAAAAA	173	423	653	883	1113	1343
+	AGG	TTATAAGACTATAAAAAGGG	174	424	654	884	1114	1344
+	AGG	TAAAAAGGGAGGAAATATAG	175	425	655	885	1115	1345
+	GGG	AAAAAGGGAGGAAATATAGA	176	426	656	886	1116	1346
+	TGG	AAATATAGAGGGTCCACTTT	177	427	657	887	1117	1347
+	TGG	TATAGAGGGTCCACTTTTGG	178	428	658	888	1118	1348
-	TGG	ACTCTGAAGTCCACCAAAAG	179	429	659	889	1119	1349
+	TGG	AGAATAGAGTTGCACCGTTT	180	430	660	890	1120	1350
-	TGG	AAAACGGTGCAACTCTATTC	181	431	661	891	1121	1351
+	AGG	CCGTTTTGGGCTAATGAAAA	182	432	662	892	1122	1352
-	CGG	CCTTTTTCATTAGCCCAAAA	183	433	663	893	1123	1353
+	AGG	TGGGCTAATGAAAAAGGAAG	184	434	664	894	1124	1354
+	AGG	TAATGAAAAAGGAAGAGGCT	185	435	665	895	1125	1355
+	GGG	AATGAAAAAGGAAGAGGCTA	186	436	666	896	1126	1356
+	AGG	CTGAATCTTAAAATATGTCC	187	437	667	897	1127	1357
-	TGG	CAGGCAGCTTTATCTCAACC	188	438	668	898	1128	1358
-	AGG	CTAAGAGATCAAGTTTCAGC	189	439	669	899	1129	1359
+	TGG	GTGTTCTTGTTGATATTCTG	190	440	670	900	1130	1360
+	TGG	CTTGTTGATATTCTGTGGCA	191	441	671	901	1131	1361

[0348]

+	TGG	TCTGTGGCATGGCTACAGAT	192	442	672	902	1132	1362
-	AGG	AGAACTTATTTACACAGGGA	193	443	673	903	1133	1363
-	GGG	AAAGAGAACTTATTTACACA	194	444	674	904	1134	1364
-	AGG	CAAAGAGAACTTATTTACAC	195	445	675	905	1135	1365
+	AGG	TTCTCTTTGTATTTACTTTT	196	446	676	906	1136	1366
+	GGG	TCTCTTTGTATTTACTTTTA	197	447	677	907	1137	1367
+	AGG	CTTTGTATTTACTTTTAGGG	198	448	678	908	1138	1368
+	TGG	AGCTTTTGTCCACCTTTAAA	199	449	679	909	1139	1369
-	TGG	TTTATTTTTCCATTTAAAGG	200	450	680	910	1140	1370
-	AGG	TATTTTATTTTTCCATTTAA	201	451	681	911	1141	1371
-	AGG	CTTACATAAACATACTTAAA	202	452	682	912	1142	1372
+	AGG	TAAGCACAGAAGTTTTTAAG	203	453	683	913	1143	1373
+	AGG	AAGTTTTTAAGAGGCATGAA	204	454	684	914	1144	1374
-	AGG	ATATTTACGTAGTTTTTCAT	205	455	685	915	1145	1375
+	AGG	CGTAAATATTCTTGAGAAAC	206	456	686	916	1146	1376
+	AGG	TTCTTGAGAAACAGGAAGAC	207	457	687	917	1147	1377
-	TGG	TAATATTAAAAACATTGGTT	208	458	688	918	1148	1378
+	AGG	CCAATGTTTTTAATATTATC	209	459	689	919	1149	1379
-	TGG	CCTGATAATATTAAAAACAT	210	460	690	920	1150	1380
+	TGG	CATTATCATGCATACATCTC	211	461	691	921	1151	1381
+	TGG	ATCATGCATACATCTCTGGC	212	462	692	922	1152	1382
+	TGG	TTCATTTCATTTTGATTTTG	213	463	693	923	1153	1383
-	TGG	ATTCAATTTGAAGCAGTGGT	214	464	694	924	1154	1384
-	TGG	GAATATTCAATTTGAAGCAG	215	465	695	925	1155	1385
+	AGG	CATACGATTTAAAATCGCTG	216	466	696	926	1156	1386
+	AGG	AAAATCGCTGAGGCGCGTTC	217	467	697	927	1157	1387
-	AGG	TTTTTTTTTCTTTTTTGTAC	218	468	698	928	1158	1388
-	TGG	CTGTTGTCAAAGATTTTAAA	219	469	699	929	1159	1389
+	TGG	TGACAACAGAGTTCTGTTTT	220	470	700	930	1160	1390
+	TGG	AGAATACGCTGAGAGTTATC	221	471	701	931	1161	1391
-	AGG	GCAAGAGAAGAAAAGAACGG	222	472	702	932	1162	1392
-	CGG	GTTGCAAGAGAAGAAAAGAA	223	473	703	933	1163	1393
-	TGG	ATGCACACGTAAAAGAGAGG	224	474	704	934	1164	1394
-	AGG	AAGATGCACACGTAAAAGAG	225	475	705	935	1165	1395

[0349]

[0350] SEQ ID NO: 2의 위치 12666에 상응하는 위치에 근접한 가이드 RNA 표적 서열의 예는 SEQ ID NO: 226 내지 239에 제시되고 하기 표에 제시된 서열들을 포함하거나, 본질적으로 그것들로 구성되거나, 또는 구성된다. SEQ ID NO: 226 내지 239에 상응하는 가이드 RNA DNA-표적화 분절의 예는, 티민 대신 우라실인 것을 제외하고 SEQ ID NO: 226 내지 239와 동일한, SEQ ID NO: 1629 내지 1642에서 각각 제시된다. 가이드 RNA DNA-표적화 분절은 SEQ ID NO: 1629 내지 1642에 제시된 임의의 서열을 포함하거나, 본질적으로 그것으로 구성되거나, 또는 구성될 수 있다. SEQ ID NO: 2의 위치 12666에 상응하는 위치에 근접한 예시적인 가이드 RNA 표적 서열은 SEQ ID NO: 230 및 231을 포함한다. SEQ ID NO: 2의 위치 12666에 상응하는 위치에 근접한 가이드 RNA 표적 서열에 상응하는 예시적인 crRNA 및 sgRNA (스캐폴드 버전 1, 2, 3, 또는 4를 포함함)는 하기 표에 제시된 서열들 중 임의의 것을 포함하거나, 본질적으로 그것으로 구성되거나, 또는 구성된다.

[0351] rs72613567 변이 근처의 가이드 RNA 표적 서열

가닥	PAM	가이드 RNA 표적 서열	변이까지의 거리 bp)	SEQ ID NO					
				gRNA 표적 서열	crRNA	sgRNA			
						v1	v2	v3	v4
+	TGG	ATCATGCATACATCTCTGGC	107	226	476	706	936	1166	1396
+	TGG	TTCATTTCATTTTGATTTTG	74	227	477	707	937	1167	1397
-	TGG	ATTCAATTTGAAGCAGTGGT	62	228	478	708	938	1168	1398
-	TGG	GAATATTCAATTTGAAGCAG	58	229	479	709	939	1169	1399
+	AGG	CATACGATTTAAAATCGCTG	22	230	480	710	940	1170	1400
+	AGG	AAAATCGCTGAGGCGCGTTC	12	231	481	711	941	1171	1401
-	AGG	TTTTTTTTTCTTTTTTGTAC	22	232	482	712	942	1172	1402
-	TGG	CTGTTGTCAAAGATTTTAAA	40	233	483	713	943	1173	1403
+	TGG	TGACAACAGAGTTCTGTTTT	65	234	484	714	944	1174	1404
+	TGG	AGAATACGCTGAGAGTTATC	94	235	485	715	945	1175	1405
-	AGG	GCAAGAGAAGAAAAGAACGG	121	236	486	716	946	1176	1406
-	CGG	GTTGCAAGAGAAGAAAAGAA	124	237	487	717	947	1177	1407
-	TGG	ATGCACACGTAAAAGAGAGG	146	238	488	718	948	1178	1408
-	AGG	AAGATGCACACGTAAAAGAG	149	239	489	719	949	1179	1409

[0352]

[0353] 마우스 *Hsd17b13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치에 근접한 마우스 *Hsd17b13* 유전자에서의 가이드 RNA 표적 서열의 예는 실시예 4에서 표 12에 제시된 서열들을 포함하거나, 본질적으로 그것들로 구성되거나, 그것들로 구성된다. 마우스 *Hsd17b13* 유전자의 5' 단부에서의 가이드 RNA 표적 서열의 예는 실시예 4에서 표 12에 제시된 서열들을 포함하거나, 본질적으로 그것들로 구성되거나, 그것들로 구성된다. 이런 가이드 RNA 표적 서열에 상응하는 가이드 RNA DNA-표적화 분절의 예 또한 실시예 4에서 표 12에 제시된다. 가이드 RNA DNA-표적화 분절은 임의의 그런 서열들을 포함하거나, 본질적으로 그것들로 구성되거나, 그것들로 구성될 수 있다. 실시예 4에서 표 12의 가이드 RNA 표적 서열에 상응하는 예시적인 crRNA 및 sgRNA (스캐폴드 버전 1, 2, 3, 또는 4를 포함함)는 실시예 4에서 표 12에 제시된 crRNA 또는 sgRNA 서열들 중 임의의 것을 포함하거나, 본질적으로 그것으로 구성되거나, 그것으로 구성될 수 있다.

[0354] 가이드 RNA는 2개의 분절: "DNA-표적화 분절" 및 "단백질-결합 분절"을 포함한다. "분절"은 분자의 부분 또는 영역, 예컨대 RNA에서 뉴클레오타이드의 연속적인 구간을 포함한다. 일부 gRNA, 예컨대 Cas9에 대한 일부 gRNA는 2개의 별도의 RNA 분자: "활성화제-RNA" (예컨대, tracrRNA) 및 "표적대상-RNA" (예컨대, CRISPR RNA 또는 crRNA)를 포함할 수 있다. 다른 gRNA는 또한 "단일-분자 gRNA", "단일-가이드 RNA", 또는 "sgRNA"로도 언급되는 단일 RNA 분자 (단일 RNA 폴리뉴클레오타이드)이다. 예컨대, WO 2013/176772, WO 2014/065596, WO 2014/089290, WO 2014/093622, WO 2014/099750, WO 2013/142578, 및 WO 2014/131833을 참조하며, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다. Cas9의 경우, 예를 들어, 단일-가이드 RNA는 tracrRNA에 (예컨대, 링커를 통해) 융합된 crRNA를 포함할 수 있다. Cpf1의 경우, 예를 들어, crRNA만이 표적 서열에 대한 결합 및/또는 표적 서열의 절단을 이루기 위해 필요하다. 용어 "가이드 RNA" 및 "gRNA"는 이중-분자 (즉, 모듈식) gRNA 및 단일-분자 gRNA를 둘 다 포함한다.

[0355] 예시적인 2-분자 gRNA는 crRNA-유사 ("CRISPR RNA" 또는 "표적대상-RNA" 또는 "crRNA" 또는 "crRNA 반복단위") 분자 및 상응하는 tracrRNA-유사 ("트란스-작용 CRISPR RNA" 또는 "활성화제-RNA" 또는 "tracrRNA") 분자를 포함한다. crRNA는 gRNA의 DNA-표적화 분절 (단일-가닥) 및 gRNA의 단백질-결합 분절의 dsRNA 듀플렉스의 절반을 형성하는 뉴클레오타이드의 구간 (즉, crRNA 꼬리) 둘 다를 포함한다. DNA-표적화 분절의 하류 (3')에 위치한 crRNA 꼬리의 예는, GUUUUAGAGCUAUGCU (SEQ ID NO: 1421)을 포함하거나, 본질적으로 그것으로 구성되거나, 또는 구성된다. 본원에 개시된 DNA-표적화 분절 중 어느 것이든지 SEQ ID NO: 1421의 5' 단부에 연결되어 crRNA를 형성할 수 있다.

[0356] 상응하는 tracrRNA (활성화제-RNA)는 gRNA의 단백질-결합 분절의 dsRNA 듀플렉스의 절반을 형성하는 뉴클레오타이드의 구간을 포함한다. crRNA의 뉴클레오타이드 구간은 tracrRNA의 뉴클레오타이드 구간에 상보하며 그것과 혼성화하여 gRNA의 단백질-결합 도메인의 dsRNA 듀플렉스를 형성한다. 이와 같이, 각각의 crRNA는 상응하는 tracrRNA를 가진다고 말할 수 있다. tracrRNA 서열의 예는 AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU (SEQ ID NO: 1422)를 포함하거나, 본

질적으로 그것으로 구성되거나, 또는 구성된다.

[0357] crRNA 및 tracrRNA가 둘 다 필요한 시스템에서, crRNA 및 상응하는 tracrRNA는 혼성화하여 gRNA를 형성한다. crRNA만이 필요한 시스템에서, crRNA는 gRNA일 수 있다. crRNA는 반대쪽 가닥 (즉, 상보하는 가닥)에 혼성화함으로써 가이드 RNA 표적 서열을 표적화하는 단일-가닥 DNA-표적화 분절을 추가적으로 제공한다. 만약 세포 내에서 변형을 위해 사용된다면, 주어진 crRNA 또는 tracrRNA 분자의 정확한 서열은 RNA 분자가 사용될 종에 대해 특이적이 되도록 설계될 수 있다. 예컨대, Mali et al. (2013) *Science* 339:823-826; Jinek et al. (2012) *Science* 337:816-821; Hwang et al. (2013) *Nat. Biotechnol.* 31:227-229; Jiang et al. (2013) *Nat. Biotechnol.* 31:233-239; 및 Cong et al. (2013) *Science* 339:819-823을 참조하며, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다.

[0358] 주어진 gRNA의 DNA-표적화 분절 (crRNA)은 표적 DNA에서 서열에 상보하는 뉴클레오타이드 서열 (즉, 가이드 RNA 표적 서열의 반대쪽 가닥에서 가이드 RNA 인식 서열의 상보하는 가닥)을 포함한다. gRNA의 DNA-표적화 분절은 혼성화 (즉, 염기쌍 형성)을 통한 서열-특이적 방식으로 표적 DNA (예컨대, *HSD17B13* 유전자)와 상호작용한다. 이와 같이, DNA-표적화 분절의 뉴클레오타이드 서열은 달라질 수 있고 gRNA 및 표적 DNA와 상호작용하게 될 표적 DNA 내에서의 위치를 결정한다. 문제의 gRNA의 DNA-표적화 분절은 표적 DNA 내에서 임의의 원하는 서열에 혼성화하기 위하여 변형될 수 있다. 자연적으로 발생하는 crRNA는 CRISPR/Cas 시스템 및 유기체에 따라 상이하지만 종종 21 내지 46개 뉴클레오타이드 길이의 2개의 직접 반복단위 (DR)에 의해 플랭킹된, 21 내지 72개 뉴클레오타이드 길이의 표적화 분절을 함유한다 (예컨대, WO 2014/131833 참조, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). S. 피오게네스의 경우에, DR은 36개 뉴클레오타이드 길이이고 표적화 분절은 30개 뉴클레오타이드 길이이다. 3' 위치한 DR은 상응하는 tracrRNA에 대해 상보하며 그것과 혼성화하고, 결국 Cas 단백질에 결합한다.

[0359] DNA-표적화 분절은 적어도 약 12개의 뉴클레오타이드, 적어도 약 15개의 뉴클레오타이드, 적어도 약 17개의 뉴클레오타이드, 적어도 약 18개의 뉴클레오타이드, 적어도 약 19개의 뉴클레오타이드, 적어도 약 20개의 뉴클레오타이드, 적어도 약 25개의 뉴클레오타이드, 적어도 약 30개의 뉴클레오타이드, 적어도 약 35개의 뉴클레오타이드, 또는 적어도 약 40개의 뉴클레오타이드의 길이를 가질 수 있다. 이러한 DNA-표적화 분절은 약 12개의 뉴클레오타이드 내지 약 100개의 뉴클레오타이드, 약 12개의 뉴클레오타이드 내지 약 80개의 뉴클레오타이드, 약 12개의 뉴클레오타이드 내지 약 50개의 뉴클레오타이드, 약 12개의 뉴클레오타이드 내지 약 40개의 뉴클레오타이드, 약 12개의 뉴클레오타이드 내지 약 30개의 뉴클레오타이드, 약 12개의 뉴클레오타이드 내지 약 25개의 뉴클레오타이드, 또는 약 12개의 뉴클레오타이드 내지 약 20개의 뉴클레오타이드의 길이를 가질 수 있다. 예를 들어, DNA 표적화 분절은 약 15개의 뉴클레오타이드 내지 약 25개의 뉴클레오타이드 (예컨대, 약 17개의 뉴클레오타이드 내지 약 20개의 뉴클레오타이드, 또는 약 17개의 뉴클레오타이드, 약 18개의 뉴클레오타이드, 약 19개의 뉴클레오타이드, 또는 약 20개의 뉴클레오타이드)일 수 있다. 예컨대, US 2016/0024523을 참조하며, 이 문헌은 모든 목적에 대해 그 전문이 본원에 참조로 포함된다. S. 피오게네스로부터의 Cas9에 대해, 전형적인 DNA-표적화 분절은 16 내지 20개 뉴클레오타이드 길이 또는 17 내지 20개 뉴클레오타이드 길이이다. Cpf1의 경우, 전형적인 DNA-표적화 분절은 적어도 16개 뉴클레오타이드 길이 또는 적어도 18개 뉴클레오타이드 길이이다.

[0360] TracrRNA는 임의의 형태로 있을 수 있고 (예컨대, 전장 tracrRNA 또는 활성 부분 tracrRNA) 다양한 길이의 것일 수 있다. 이것들은 일차 전사물 또는 프로세싱딘 형태를 포함할 수 있다. 예를 들어, tracrRNA (단일-가이드 RNA의 부분으로서 또는 2-분자 gRNA의 부분으로서 별도의 분자로서)는 야생형 tracrRNA 서열의 전부 또는 일부 (예컨대, 야생형 tracrRNA 서열의 약 20, 26, 32, 45, 48, 54, 63, 67, 85개, 또는 그 이상의 뉴클레오타이드 또는 약 20, 26, 32, 45, 48, 54, 63, 67, 85개 이상 또는 그 이상의 뉴클레오타이드)를 포함하거나 그것으로 구성될 수 있다. S. 피오게네스로부터의 야생형 tracrRNA 서열의 예는 171-뉴클레오타이드, 89-뉴클레오타이드, 75-뉴클레오타이드, 및 65-뉴클레오타이드 버전을 포함한다. 예컨대, Deltcheva et al. (2011) *Nature* 471:602-607; WO 2014/093661 참조하며, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다. 단일-가이드 RNA (sgRNA) 내의 tracrRNA의 예로는 sgRNA의 +48, +54, +67, 및 +85 버전 내에서 발견되는 tracrRNA 분절을 들 수 있고, 여기서 "+n"은 야생형 tracrRNA의 최대 +n개 뉴클레오타이드가 sgRNA에 포함되는 것을 가리킨다. US 8,697,359 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)를 참조한다.

[0361] DNA-표적화 서열과 표적 DNA 내의 가이드 RNA 인식 서열의 상보하는 가닥 사이의 퍼센트 상보성은 적어도 60% (예컨대, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%)일 수 있다. DNA-표적화 서열과 표적 DNA 내의 가이드 RNA 인식 서열의 상보하는 가닥 사이의 퍼센트 상보성은 약 20개의 연속 뉴클레오타이드에 걸쳐 적어도 약 60%일 수 있다. 예

로서, DNA-표적화 서열과 표적 DNA 내의 가이드 RNA 인식 서열의 상보하는 가닥 사이의 퍼센트 상보성은 표적 DNA의 상보하는 가닥 내의 가이드 RNA 인식 서열의 상보하는 가닥의 5' 단부에서 14개의 연속 뉴클레오타이드에 걸쳐 100%이고 나머지에 대해서는 0%로 낮다. 이런 경우에, DNA-표적화 서열은 길이가 14개 뉴클레오타이드인 것으로 여겨질 수 있다. 다른 예로서, DNA-표적화 서열과 표적 DNA 내의 가이드 RNA 인식 서열의 상보하는 가닥 사이의 퍼센트 상보성은 표적 DNA의 상보하는 가닥 내의 가이드 RNA 인식 서열의 상보하는 가닥의 5' 단부에서 7개의 연속 뉴클레오타이드에 걸쳐 100%이고 나머지에 대해서는 0%로 낮다. 이런 경우에, DNA-표적화 서열은 길이가 7개 뉴클레오타이드인 것으로 여겨질 수 있다. 일부 가이드 RNA에서, DNA-표적화 서열 내의 적어도 17개의 뉴클레오타이드가 표적 DNA에 상보한다. 예를 들어, DNA-표적화 서열은 길이가 20개 뉴클레오타이드일 수 있고 가이드 RNA 인식 서열의 상보하는 가닥과 1, 2, 또는 3개의 미스매치를 포함할 수 있다. 바람직하게, 미스매치는 프로토스페이서 인접 모티프 (PAM) 서열에 인접하지 않는다 (예컨대, 미스매치는 DNA-표적화 서열의 5' 단부에 있거나, 미스매치는 PAM 서열로부터 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 또는 19개 염기쌍만큼 떨어져 있다).

[0362] gRNA의 단백질-결합 분절은 서로 상보하는 뉴클레오타이드의 두 구간을 포함할 수 있다. 단백질-결합 분절의 상보하는 뉴클레오타이드는 이중 가닥 RNA 듀플렉스 (dsRNA)를 형성하기 위하여 혼성화한다. 대상체 gRNA의 단백질-결합 분절은 Cas 단백질과 상호작용하고, gRNA는 결합된 Cas 단백질이 DNA-표적화 분절을 통해 표적 DNA 내의 특이적 뉴클레오타이드로 향하게(direct) 한다.

[0363] 단일-가이드 RNA는 DNA-표적화 분절 및 스캐폴드 서열 (즉, 가이드 RNA의 단백질-결합 또는 Cas-결합 서열)을 가진다. 예를 들어, 이러한 가이드 RNA는 5' DNA-표적화 분절 및 3' 스캐폴드 서열을 가진다. 예시적인 스캐폴드 서열은 다음을 포함하거나, 본질적으로 그것으로 구성되거나, 구성된다: GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU (버전 1; SEQ ID NO: 1420); GUUGGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (버전 2; SEQ ID NO: 256); GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (버전 3; SEQ ID NO: 257); 및 GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (버전 4; SEQ ID NO: 258). 본원에 개시된 임의의 가이드 RNA 표적 서열을 표적화하는 가이드 RNA (예컨대, SEQ ID NO: 20 내지 239 및 259 내지 268)는, 예를 들어, 가이드 RNA의 3' 단부 상의 예시적인 가이드 RNA 스캐폴드 서열 중 임의의 것에 융합된 가이드 RNA의 5' 단부 상의 DNA-표적화 분절을 포함할 수 있다. 즉, 본원에 개시된 임의의 DNA-표적화 분절이 SEQ ID NO: 1420, 256, 257, 또는 258 중 임의의 하나의 5' 단부에 연결되어 단일 가이드 (키메릭 가이드 RNA)가 형성될 수 있다. 본원의 다른 곳에서 개시되는 바 가이드 RNA 버전 1, 2, 3, 및 4는 각각 스캐폴드 버전 1, 2, 3, 및 4와 연합된 DNA-표적화 분절을 나타낸다.

[0364] 가이드 RNA는 추가적인 바람직한 특징 (예컨대, 변형된 또는 조절된 안정성; 세포내 표적화; 형광 표지로의 추적; 단백질 또는 단백질 복합체에 대한 결합 부위; 등)을 제공하는 변형 또는 서열을 포함할 수 있다. 이러한 변형의 예로는, 예를 들어, 5' 캡 (예컨대, 7-메틸구아닐레이트 캡 (m7G)); 3' 폴리아데닐화된 꼬리 (즉, 3' 폴리(A) 꼬리); 리보스위치 서열 (예컨대, 단백질 및/또는 단백질 복합체에 의해 조절된 안정성 및/또는 조절된 접근성을 허용하기 위함); 안정성 제어 서열; dsRNA 듀플렉스를 형성하는 서열 (즉, 헤어핀); 세포내 국지화 (예컨대, 핵, 미토콘드리아, 엽록체, 등)를 위하여 RNA를 표적화하는 변형 또는 서열; 추적을 제공하는 변형 또는 서열 (예컨대, 형광 분자에 대한 직접 콘쥬게이션, 형광 검출을 용이하게 하는 모이어티에 대한 콘쥬게이션, 형광 검출을 허용하는 서열, 등); 단백질 (예컨대, 전사 활성화제, 전사 억제자, DNA 메틸전이효소, DNA 탈메틸화효소, 히스톤 아세틸전이효소, 히스톤 탈아세틸화효소, 등을 포함한, DNA에 작용하는 단백질)에 대한 결합 부위를 제공하는 변형 또는 서열; 및 이것들의 조합을 들 수 있다. 변형의 다른 예로는 조작된 줄기 루프 듀플렉스 구조, 조작된 벌지(bulge) 영역, 줄기 루프 듀플렉스 구조의 3'의 조작된 헤어핀, 또는 이것들의 임의의 조합을 들 수 있다. 예컨대, US 2015/0376586 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)을 참조한다. 벌지는 crRNA-유사 영역 및 최소 tracrRNA-유사 영역으로 만들어진 듀플렉스 내의 뉴클레오타이드의 쌍을 이루지 않은 영역일 수 있다. 벌지는, 듀플렉스의 한쪽에서, 쌍을 이루지 않은 5'-XXXY-3', 및 듀플렉스의 다른 쪽의 쌍을 이루지 않은 뉴클레오타이드 영역을 포함할 수 있고, 여기서 X는 임의의 퓨린이며 Y는 반대쪽 가닥의 뉴클레오타이드와 워블 쌍 (wobble pair)을 형성할 수 있는 뉴클레오타이드일 수 있다.

[0365] 일부 경우에, MS2-p65-HSF1과 쌍을 이룬 dCas9-VP64 fusion 단백질을 포함하는 전사 활성화 시스템이 사용될 수 있다. 이러한 시스템에서 가이드 RNA는 다이머화된 MS2 박테리오파지 코트 단백질에 결합하도록 설계된 sgRNA 테트라루프 및 줄기-루프 2에 첨부된 압타머 서열로 설계될 수 있다. 예컨대, Konermann et al. (2015) *Nature* 517(7536):583-588 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)을 참조한다.

[0366] 가이드 RNA는 임의의 형태로 제공될 수 있다. 예를 들어, gRNA는 두 분자로서 (별도의 crRNA 및 tracrRNA) 또는 한 분자로서 (sgRNA), RNA의 형태로, 및 선택적으로 Cas 단백질과의 복합체의 형태로 제공될 수 있다. 예를 들어, gRNA는, 예를 들어, T7 RNA 중합효소를 사용하는 시험관내 전사에 의해 제조될 수 있다 (예컨대, WO 2014/089290 및 WO 2014/065596을 참조하며, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다). 가이드 RNA는 또한 화학적 합성에 의해 제조될 수 있다.

[0367] gRNA는 또한 gRNA를 암호화하는 DNA의 형태로 제공될 수 있다. gRNA를 암호화하는 DNA는 단일 RNA 분자 (sgRNA) 또는 별도의 RNA 분자 (예컨대, 별도의 crRNA 및 tracrRNA)를 암호화할 수 있다. 후자의 경우에, gRNA를 암호화하는 DNA는 한 DNA 분자로서 또는 crRNA 및 tracrRNA를 각각 암호화하는 별도의 DNA 분자로서 제공될 수 있다.

[0368] gRNA가 DNA의 형태로 제공되는 경우에, gRNA는 세포에서 일시저긍로, 조건부로, 또는 구성적으로 발현될 수 있다. gRNA를 암호화하는 DNA는 세포의 게놈 안으로 안정적으로 통합될 수 있고 세포에서 활성인 프로모터에 작동가능하게 연결될 수 있다. 대안적으로, gRNA를 암호화하는 DNA는 발현 구성물에서 프로모터에 작동가능하게 연결될 수 있다. 예를 들어, gRNA를 암호화하는 DNA는 이종성 핵산을 포함하는 벡터에 있을 수 있다. 벡터는 추가로 외인성 도너 서열을 포함할 수 있고 및/또는 벡터는 추가로 Cas 단백질을 암호화하는 핵산을 포함할 수 있다. 대안적으로, gRNA를 암호화하는 DNA는 외인성 도너 서열을 포함하는 벡터 및/또는 Cas 단백질을 암호화하는 핵산을 포함하는 벡터와 다른 벡터 또는 플라스미드에 있을 수 있다. 이러한 발현 구성물에 사용될 수 있는 프로모터로는, 예를 들어, 진핵 세포, 인간 세포, 비인간 세포, 포유류 세포, 비인간 포유류 세포, 설치류 세포, 마우스 세포, 래트 세포, 햄스터 세포, 토끼 세포, 만능 세포, 배아 줄기 (ES) 세포, 성인 줄기 세포, 발달상 제한된 조상 세포, 유도된 만능 줄기 (iPS) 세포, 또는 한-세포 단계 배아 중 하나 이상에서 활성인 프로모터를 포함한다. 이러한 프로모터로는, 예를 들어, 조건부 프로모터, 유도성 프로모터, 구성성 프로모터, 또는 조직-특이적 프로모터일 수 있다. 이러한 프로모터는 또한, 예를 들어, 양방향성 프로모터일 수 있다. 적합한 프로모터의 구체적인 예로는 RNA 중합효소 III 프로모터, 예컨대 인간 U6 프로모터, 래트 U6 중합효소 III 프로모터, 또는 마우스 U6 중합효소 III 프로모터를 들 수 있다.

[0369] 또한 본원에는 본원에 개시된 하나 이상의 가이드 RNA (예컨대, 1, 2, 3, 4개, 또는 그 이상의 가이드 RNA) 및 분리된 핵산 또는 단백질의 안정성을 증가시키는 (예컨대, 분해 산물이 한계값 아래로 유지되는, 예컨대 출발 핵산 또는 단백질의 0.5 중량% 아래로 유지되어야 하는 주어진 보관 조건 (예컨대, -20℃, 4℃, 또는 주위 온도) 하에서의 기간을 연장시키는; 또는 생체내에서 안정성을 증가시키는) 담체를 포함하는 조성물이 개시된다. 이러한 담체의 비제한적인 실례로는 폴리(락트산) (PLA) 마이크로스피어, 폴리(D,L-락틱-코글리콜-산) (PLGA) 마이크로스피어, 리포좀, 미셀, 역미셀, 지질 코클레이트, 및 지질 미세관을 들 수 있다. 이러한 조성물은 추가로 Cas 단백질, 예컨대 Cas9 단백질, 또는 Cas 단백질을 암호화하는 핵산을 포함할 수 있다. 이러한 조성물은 추가로 본원의 다른 곳에서 개시된 것과 같은 하나 이상 (예컨대, 1, 2, 3, 4개 또는 그 이상)의 외인성 도너 서열 및/또는 하나 이상 (예컨대, 1, 2, 3, 4개 또는 그 이상)의 표적화 벡터 및/또는 하나 이상 (예컨대, 1, 2, 3, 4개 또는 그 이상)의 발현 벡터를 포함할 수 있다.

[0370] ***(3) 가이드 RNA 인식 서열 및 가이드 RNA 표적 서열***

[0371] 용어 "가이드 RNA 인식 서열"은 gRNA의 DNA-표적화 분절이 결합하게 될 표적 DNA (예컨대, *HSD17B13* 유전자)에 존재하는 핵산 서열을 포함하며, 단 결합에 대한 충분한 조건이 존재해야 한다. 본원에서 사용되는 바 용어 가이드 RNA 인식 서열은 표적 이중 가닥 DNA의 두 가닥 (즉, 가이드 RNA가 혼성화하는 상보하는 가닥 상의 서열 및 프로토스페이서 인접 모티프 (PAM)에 인접한 비-상보하는 가닥 상의 상응하는 서열)을 다 포함한다. 본원에서 사용되는 바 용어 "가이드 RNA 표적 서열"은 구체적으로 PAM에 인접한 비-상보하는 가닥 상의 서열 (즉, PAM의 상류 또는 5')을 나타낸다. 즉, 가이드 RNA 표적 서열은 가이드 RNA가 상보하는 가닥 상에서 혼성화하는 서열에 상응하는 비-상보하는 가닥 상의 서열을 나타낸다. 가이드 RNA 표적 서열은 가이드 RNA의 DNA-표적화 분절과 동등하지만, 우라실 대신 티민을 포함한다. 한 예로서, Cas9 효소에 대한 가이드 RNA 표적 서열은 5'-NGG-3' PAM에 인접한 비-상보하는 가닥 상의 서열을 나타낼 것이다. 가이드 RNA 인식 서열은 그것에 대해 가이드 RNA가 상보성을 갖도록 설계되는 서열을 포함하며, 여기서 가이드 RNA 인식 서열의 상보하는 가닥과 가이드 RNA의 DNA 표적화 서열 사이의 혼성화는 CRISPR 복합체의 형성을 촉진한다. 완전한 상보성은 반드시 필요하지는 않지만, 혼성화를 유발하고 CRISPR 복합체의 형성을 촉진하기 위해서는 충분한 상보성이 있어야 한다. 가이드 RNA 인식 서열 또는 가이드 RNA 표적 서열은 또한, 하기에서 보다 상세하게 기술되는, Cas 단백질에 대한 절단 부위를 포함한다. 가이드 RNA 인식 서열 또는 가이드 RNA 표적 서열은, 예를 들어, 세포의 핵 또는 세포질에 또는 세포의 소기관 내에, 예컨대 미토콘드리온 또는 엽록체 내에 위치할 수 있는 임의의 폴리뉴클레오타이드를 포함할 수

있다.

[0372] 표적 DNA 내의 가이드 RNA 인식 서열은 Cas 단백질 또는 gRNA에 의해 표적화될 수 있다 (즉, 결합되거나, 혼성화하거나, 또는 상보할 수 있다). 적합한 DNA/RNA 결합 조건으로는 세포에 정상적으로 존재하는 생리적 조건을 포함한다. 다른 적합한 DNA/RNA 결합 조건 (예컨대, 세포-유리 시스템에서의 조건)이 알려져 있다 (예컨대, Molecular Cloning: A Laboratory Manual, 3rd Ed. (Sambrook et al., Harbor Laboratory Press 2001 참조, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). Cas 단백질 또는 gRNA에 상보하고 그것과 혼성화하는 표적 DNA의 가닥은 "상보하는 가닥"으로 불리며, "상보하는 가닥"에 상보하는 (및, 그러므로 Cas 단백질 또는 gRNA에는 상보하지 않는) 표적 DNA의 가닥은 "비-상보하는 가닥" 또는 "주형 가닥"으로 불릴 수 있다.

[0373] Cas 단백질은 gRNA의 DNA-표적화 분절이 결합하게 될 표적 DNA에 존재하는 핵산 서열 내에 또는 외부에 있는 부위에서 핵산을 절단할 수 있다. "절단 부위"는 Cas 단백질이 단일 가닥 끊김 또는 이중 가닥 끊김을 생성하는 핵산의 위치를 포함한다. 예를 들어, CRISPR 복합체 (가이드 RNA 인식 서열의 상보하는 가닥에 혼성화되고 Cas 단백질과 복합체를 형성한 gRNA를 포함함)의 형성은 gRNA의 DNA-표적화 분절이 결합하게 될 표적 DNA에 존재하는 핵산 서열에 또는 가까이에 (예컨대, 핵산 서열로부터 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50개, 또는 그 이상의 염기쌍 내에) 있는 하나 또는 양 가닥의 절단을 초래할 수 있다. 만약 절단 부위가 DNA-표적화 분절이 결합하게 될 핵산 서열의 외부에 있다면, 절단 부위는 여전히 "가이드 RNA 인식 서열" 또는 가이드 RNA 표적 서열 내에 있는 것으로 여겨질 것이다. 절단 부위는 핵산의 오직 한 가닥에 또는 양 가닥에 있을 수 있다. 절단 부위는 핵산의 양 가닥 상의 동일한 위치에 있을 수 있거나 (블런트 단부를 생성함) 또는 각 단부 상의 상이한 부위에 있을 수 있다 (엇갈린 단부 (즉, 오버행)를 생성함). 엇갈린 단부는, 예를 들어, 2개의 Cas 단백질을 사용함으로써 생성될 수 있는데, 이것들은 각각 상이한 가닥 상의 상이한 절단 부위에서 단일 가닥 끊김을 생성함으로써, 이중 가닥 끊김을 생성한다. 예를 들어, 제1 닉카아제는 이중 가닥 DNA (dsDNA)의 제1 가닥에서 단일 가닥 끊김을 생성할 수 있고, 제2 닉카아제는 dsDNA의 제2 가닥에서 단일 가닥 끊김을 생성할 수 있어서 오버행 서열이 생성된다. 일부 경우에, 제1 가닥 상의 닉카아제의 가이드 RNA 인식 서열 또는 가이드 RNA 표적 서열은 제2 가닥 상의 닉카아제의 가이드 RNA 인식 서열 또는 가이드 RNA 표적 서열과 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 75, 100, 250, 500, 또는 1,000개의 염기쌍만큼 떨어져 있다.

[0374] Cas 단백질에 의한 표적 DNA의 부위-특이적 결합 및/또는 절단은 (i) gRNA와 표적 DNA 사이의 염기쌍 형성 상보성 및 (ii) 표적 DNA에서, 프로토스페이서 인접 모티프 (PAM)로 불리는, 짧은 모티프 둘 다에 의해 결정된 위치에서 일어날 수 있다. PAM은 가이드 RNA가 혼성화하는 가닥의 반대쪽에 있는 비-상보하는 가닥 상의 가이드 RNA 표적 서열을 플랭킹한다. 선택적으로, 가이드 RNA 표적 서열은 PAM에 의해 3' 단부에서 플랭킹될 수 있다. 대안적으로, 가이드 RNA 표적 서열은 PAM에 의해 5' 단부에서 플랭킹될 수 있다. 예를 들어, Cas 단백질의 절단 부위는 PAM 서열의 상류 또는 하류로 약 1 내지 약 10개 또는 약 2 내지 약 5개 염기쌍 (예컨대, 3개 염기쌍)일 수 있다. 일부 경우에 (예컨대, S. 피오게네스로부터의 Cas9 또는 밀접하게 관련된 Cas9가 사용될 때), 비-상보하는 가닥의 PAM 서열은 5'-N_1GG-3'일 수 있고, 여기서 N_1은 임의의 DNA 뉴클레오타이드이며 표적 DNA의 비-상보하는 가닥의 가이드 RNA 인식 서열의 바로 3' (즉, 가이드 RNA 표적 서열의 바로 3')에 있다. 이와 같이, 상보하는 가닥의 PAM 서열은 5'-CCN_2-3'일 것이고, 여기서 N_2는 임의의 DNA 뉴클레오타이드이며 표적 DNA의 상보하는 가닥의 가이드 RNA 인식 서열의 바로 5' (즉, 가이드 RNA 표적 서열의 바로 5')에 있다. 일부 이러한 경우에, N_1 및 N_2는 상보할 수 있고 N_1-N_2 염기쌍은 임의의 염기쌍일 수 있다 (예컨대, N_1=C 및 N_2=G; N_1=G 및 N_2=C; N_1=A 및 N_2=T; 또는 N_1=T, 및 N_2=A). S. 아우레우스로부터의 Cas9의 경우에, PAM은 NNGRRT 또는 NNGRR일 수 있고, 여기서 N은 A, G, C, 또는 T일 수 있으며, 및 R은 G 또는 A일 수 있다. 일부 경우에 (예컨대, FnCpf1의 경우), PAM 서열은 5' 단부의 상류에 있을 수 있고 서열 5'-TTN-3'를 가진다.

[0375] PAM 외에 가이드 RNA 표적 서열 또는 가이드 RNA 표적 서열의 예는 하기에 제공된다. 예를 들어, 가이드 RNA 표적 서열은 Cas9 단백질에 의해 인식된 NGG 모티프 바로 앞에 있는 20-뉴클레오타이드 DNA 서열일 수 있다. 이러한 가이드 RNA 표적 서열 플러스 PAM 서열의 예는 GN_{19}NGG (SEQ ID NO: 248) 또는 N_{20}NGG (SEQ ID NO: 249)이다. 예컨대, WO 2014/165825 참조한다 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 5' 단부에서 구아닌은 세포에서 RNA 중합효소에 의한 전사를 용이하게 할 수 있다. RNA 표적 서열 플러스 PAM 서열의 다른 예는 시험관내에서 T7 중합효소에 의한 효율적인 전사를 용이하게 하기 위해 5' 단부에 2개의 구아닌 뉴클레오타이드를 포함할 수 있다 (예컨대, GGN_{20}NGG; SEQ ID NO: 250). 예컨대, WO 2014/065596을 참조한다 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 다른 가이드 RNA 표적 서열 플러스 PAM 서열은 5' G 또는 GG 및 3' GG 또는 NGG를 포함하는, SEQ ID NO: 248 내지 250의 4 내지 22개 뉴클레오타이드 길이를 가질 수 있다. 또 다

른 가이드 RNA 표적 서열은 SEQ ID NO: 248 내지 250의 14 내지 20개 뉴클레오타이드 길이를 가질 수 있다.

[0376] 가이드 RNA 인식 서열 또는 가이드 RNA 표적 서열은 세포에 대해 내인성 또는 외인성인 임의의 핵산 서열일 수 있다. 가이드 RNA 인식 서열 또는 가이드 RNA 표적 서열은 유전자 생성물 (예컨대, 단백질)을 암호화하는 서열이거나 비코딩 서열 (예컨대, 조절 서열)일 수 있거나 또는 둘 다를 포함할 수 있다.

[0377] 한 예로서, 가이드 RNA 인식 서열 또는 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 엑손 6 및/또는 인트론 6, 엑손 6 및/또는 엑손 7, 또는 엑손 6 및/또는 인트론 6 및/또는 엑손 7에 상응하는 영역 내에 있을 수 있다. 다른 예로서, 가이드 RNA 인식 서열 또는 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치를 포함하거나 근접할 수 있다. 예를 들어, 가이드 RNA 인식 서열 또는 가이드 RNA 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 약 1000, 500, 400, 300, 200, 100, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5개 뉴클레오타이드 내에 있을 수 있다. 또 다른 예로서, 가이드 RNA 인식 서열 또는 가이드 RNA 표적 서열은 *HSD17B13* 유전자의 개시 코돈 또는 *HSD17B13* 유전자의 종결 코돈을 포함하거나 그것에 근접할 수 있다. 예를 들어, 가이드 RNA 인식 서열 또는 가이드 RNA 표적 서열은 개시 코돈 또는 종결 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있을 수 있다. 이러한 가이드 RNA 표적 서열 및 이러한 가이드 RNA 표적 서열을 표적화하는 가이드 RNA의 예는 본원의 다른 곳에서 개시된다.

[0378] **F. 외인성 도너 서열 또는 표적화 벡터**

[0379] 본원에 개시된 방법 및 조성물은 *HSD17B13* 유전자의 절단 없이 또는 뉴클레아제 제제로의 *HSD17B13* 유전자의 절단 후에, *HSD17B13* 유전자를 변형시키기 위하여 외인성 도너 서열 (예컨대, 표적화 벡터 또는 복구 주형)을 이용할 수 있다. 외인성 도너 서열은 표적 서열과의 부위-특이적 재조합을 가능하게 하기 위해 필요한 요소들을 포함하는 임의의 핵산 또는 벡터를 나타낸다. 뉴클레아제 제제와의 조합에 외인성 도너 서열을 사용하는 것은 상동성-의존 복구를 촉진함으로써 *HSD17B13* 유전자 내에서 보다 정확한 변형을 초래할 수 있다.

[0380] 이러한 방법에서, 뉴클레아제 제제는 *HSD17B13* 유전자를 절단하여 단일 가닥 끊김 (닉) 또는 이중 가닥 끊김을 생성하고, 외인성 도너 서열은 비-상동성 단부 연합 (NHEJ)-매개된 결찰을 통해 또는 상동성-의존 복구 사건을 통해 *HSD17B13* 유전자를 재조합한다. 선택적으로, 외인성 도너 서열로의 복구는 뉴클레아제 절단 부위를 제거하거나 파괴하여서 표적화된 대립유전자가 뉴클레아제 제제에 의해 재표적화될 수 없다.

[0381] 외인성 도너 서열은 데옥시리보핵산 (DNA) 또는 리보핵산 (RNA)을 포함하며, 이것들은 단일 가닥이거나 이중 가닥일 수 있고, 선형 또는 원형 형태로 있을 수 있다. 예를 들어, 외인성 도너 서열은 단일-가닥 올리고데옥시뉴클레오타이드 (ssODN)일 수 있다. 예컨대, Yoshimi et al. (2016) *Nat. Commun.* 7:10431을 참조함 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 예시적인 외인성 도너 서열은 길이가 약 50개 뉴클레오타이드 내지 약 5 kb이거나, 길이가 약 50개 뉴클레오타이드 내지 약 3 kb이거나, 또는 길이가 약 50개 내지 약 1,000개 뉴클레오타이드이다. 다른 예시적인 외인성 도너 서열은 길이가 약 40개 내지 약 200개 뉴클레오타이드이다. 예를 들어, 외인성 도너 서열은 길이가 약 50 내지 약 60개, 약 60 내지 약 70개, 약 70 내지 약 80개, 약 80 내지 약 90개, 약 90 내지 약 100개, 약 100 내지 약 110개, 약 110 내지 약 120개, 약 120 내지 약 130개, 약 130 내지 약 140개, 약 140 내지 약 150개, 약 150 내지 약 160개, 약 160 내지 약 170개, 약 170 내지 약 180개, 약 180 내지 약 190개, 또는 약 190 내지 약 200개 뉴클레오타이드일 수 있다. 대안적으로, 외인성 도너 서열은 길이가 약 50 내지 약 100개, 약 100 내지 약 200개, 약 200 내지 약 300개, 약 300 내지 약 400개, 약 400 내지 약 500개, 약 500 내지 약 600개, 약 600 내지 약 700개, 약 700 내지 약 800개, 약 800 내지 약 900개, 또는 약 900 내지 약 1,000개 뉴클레오타이드일 수 있다. 대안적으로, 외인성 도너 서열은 길이가 약 1 kb 내지 약 1.5 kb, 약 1.5 kb 내지 약 2 kb, 약 2 kb 내지 약 2.5 kb, 약 2.5 kb 내지 약 3 kb, 약 3 kb 내지 약 3.5 kb, 약 3.5 kb 내지 약 4 kb, 약 4 kb 내지 약 4.5 kb, 또는 약 4.5 kb 내지 약 5 kb일 수 있다. 대안적으로, 외인성 도너 서열은 길이가, 예를 들어, 5 kb 이하, 4.5 kb, 4 kb, 3.5 kb, 3 kb, 2.5 kb, 2 kb, 1.5 kb, 1 kb, 900개 뉴클레오타이드, 800개 뉴클레오타이드, 700개 뉴클레오타이드, 600개 뉴클레오타이드, 500개 뉴클레오타이드, 400개 뉴클레오타이드, 300개 뉴클레오타이드, 200개 뉴클레오타이드, 100개 뉴클레오타이드, 또는 50개 뉴클레오타이드일 수 있다.

[0382] 한 예에서, 외인성 도너 서열은 길이가 약 80개 뉴클레오타이드 내지 약 200개 뉴클레오타이드 (예컨대, 길이가 약 120개 뉴클레오타이드)인 ssODN이다. 다른 예에서, 외인성 도너 서열은 길이가 약 80개 뉴클레오타이드 내지 약 3 kb인 ssODN이다. 이러한 ssODN은, 예를 들어, 각각 길이가 약 40개 뉴클레오타이드 내지 약 60개 뉴클레오

타이드인 상동성 팔을 가질 수 있다. 이러한 ssODN은 또한, 예를 들어 각각 길이가 약 30개 뉴클레오타이드 내지 약 100개 뉴클레오타이드인 상동성 팔을 가질 수 있다. 상동성 팔은 대칭일 수 있거나 (예컨대, 각각 길이가 40개 뉴클레오타이드 또는 각각 길이가 60개 뉴클레오타이드임), 또는 비대칭일 수 있다 (예컨대, 길이가 36개 뉴클레오타이드인 한 상동성 팔, 및 길이가 91개 뉴클레오타이드인 한 상동성 팔).

[0383] 외인성 도너 서열은 추가적인 바람직한 특징 (예컨대, 변형된 또는 조절된 안정성; 형광 표지로의 추적 또는 검출; 단백질 또는 단백질 복합체에 대한 결합 부위; 등)을 제공하는 변형 또는 서열을 포함한다. 외인성 도너 서열은 하나 이상의 형광 표지, 정제 태그, 에피토프 태그, 또는 이것들의 조합을 포함할 수 있다. 예를 들어, 외인성 도너 서열은 하나 이상의 형광 표지 (예컨대, 형광 단백질 또는 다른 형광단 또는 염료), 예컨대 적어도 1, 적어도 2, 적어도 3, 적어도 4, 또는 적어도 5개의 형광 표지를 포함할 수 있다. 예시적인 형광 표지로는 플루오레세인 (예컨대, 6-카르복시플루오레세인 (6-FAM)), 텍사스 레드, HEX, Cy3, Cy5, Cy5.5, 퍼시픽 블루 (Pacific Blue), 5-(및-6)-카르복시테트라메틸로다민 (TAMRA), 및 Cy7을 들 수 있다. 광범위한 형광 염료가 올리고뉴클레오타이드를 표지화하기 위해 (예컨대, Integrated DNA Technologies로부터) 상업적으로 입수 가능하다. 이러한 형광 표지 (예컨대, 내부 형광 표지)는, 예를 들어, 외인성 도너 서열의 단부와 부합하는 돌출된 단부를 가지는 절단된 *HSD17B13* 유전자로 직접 통합되어 있는 외인성 도너 서열을 검출하기 위해 사용될 수 있다. 표지 또는 태그는 외인성 도너 서열의 5' 단부에, 3' 단부에, 또는 외인성 도너 서열 내부에 있을 수 있다. 예를 들어, 외인성 도너 서열은 Integrated DNA Technologies로부터의 IR700 형광단으로 5' 단부에서 콘쥬게이션될 수 있다 (5'IRDYE®700).

[0384] 외인성 도너 서열은 또한 *HSD17B13* 유전자에 통합될 DNA의 분절을 포함하는 핵산 삽입물을 포함할 수 있다. 핵산 삽입물의 *HSD17B13* 유전자에의 통합은 *HSD17B13* 유전자에서 관심 있는 핵산 서열의 첨가, *HSD17B13* 유전자에서 관심 있는 핵산 서열의 삭제, 또는 *HSD17B13* 유전자에서 관심 있는 핵산 서열의 교체 (즉, 삭제 및 삽입)을 초래할 수 있다. 일부 외인성 도너 서열은 *HSD17B13* 유전자에서 임의의 상응하는 삭제 없이 *HSD17B13* 유전자에서 핵산 삽입물의 삽입을 위해 설계된다. 다른 외인성 도너 서열은 핵산 삽입물의 임의의 상응하는 삽입 없이 *HSD17B13* 유전자에서 관심 있는 핵산 서열을 삭제하기 위해 설계된다. 또 다른 외인성 도너 서열은 *HSD17B13* 유전자에서 관심 있는 핵산 서열을 삭제하고 그것을 핵산 삽입물로 교체하기 위해 설계된다.

[0385] 삭제되거나 및/또는 교체되는 *HSD17B13* 유전자에서의 핵산 삽입물 또는 상응하는 핵산은 다양한 길이일 수 있다. 삭제되거나 및/또는 교체되는 *HSD17B13* 유전자에서의 예시적인 핵산 삽입물 또는 상응하는 핵산은 길이가 약 1개 뉴클레오타이드 내지 약 5 kb이거나 또는 길이가 약 1개 뉴클레오타이드 내지 약 1,000개 뉴클레오타이드이다. 예를 들어, 삭제되거나 및/또는 교체되는 *HSD17B13* 유전자에서의 핵산 삽입물 또는 상응하는 핵산은 길이가 약 1 내지 약 10개, 약 10 내지 약 20개, 약 20 내지 약 30개, 약 30 내지 약 40개, 약 40 내지 약 50개, 약 50 내지 약 60개, 약 60 내지 약 70개, 약 70 내지 약 80개, 약 80 내지 약 90개, 약 90 내지 약 100개, 약 100 내지 약 110개, 약 110 내지 약 120개, 약 120 내지 약 130개, 약 130 내지 약 140개, 약 140 내지 약 150개, 약 150 내지 약 160개, 약 160 내지 약 170개, 약 170 내지 약 180개, 약 180 내지 약 190개, 또는 약 190 내지 약 200개 뉴클레오타이드일 수 있다. 마찬가지로, 삭제되거나 및/또는 교체되는 *HSD17B13* 유전자에서의 핵산 삽입물 또는 상응하는 핵산은 길이가 약 1 내지 약 100개, 약 100 내지 약 200개, 약 200 내지 약 300개, 약 300 내지 약 400개, 약 400 내지 약 500개, 약 500 내지 약 600개, 약 600 내지 약 700개, 약 700 내지 약 800개, 약 800 내지 약 900개, 또는 약 900 내지 약 1,000개 뉴클레오타이드일 수 있다. 마찬가지로, 삭제되거나 및/또는 교체되는 *HSD17B13* 유전자에서의 핵산 삽입물 또는 상응하는 핵산은 길이가 약 1 kb 내지 약 1.5 kb, 약 1.5 kb 내지 약 2 kb, 약 2 kb 내지 약 2.5 kb, 약 2.5 kb 내지 약 3 kb, 약 3 kb 내지 약 3.5 kb, 약 3.5 kb 내지 약 4 kb, 약 4 kb 내지 약 4.5 kb, 또는 약 4.5 kb 내지 약 5 kb일 수 있다.

[0386] 핵산 삽입물은 게놈 DNA 또는 임의의 다른 유형의 DNA를 포함할 수 있다. 예를 들어, 핵산 삽입물은 cDNA를 포함할 수 있다.

[0387] 핵산 삽입물은 *HSD17B13* 유전자의 전부 또는 일부 (예컨대, 특정 모티프를 암호화하는 유전자의 부분 또는 HSD17B13 단백질의 영역)에 상동하는 서열을 포함할 수 있다. 예를 들어, 핵산 삽입물은 *HSD17B13* 유전자에서 교체를 위해 표적화된 서열과 비교하여 하나 이상의 점 돌연변이 (예컨대, 1, 2, 3, 4, 5개, 또는 그 이상) 또는 하나 이상의 뉴클레오타이드 삽입 또는 삭제를 포함하는 서열을 포함할 수 있다.

[0388] 삭제되거나 및/또는 교체되는 *HSD17B13* 유전자에서의 핵산 삽입물 또는 상응하는 핵산은 엑손과 같은 코딩 영역; 인트론, 미번역 영역, 또는 조절 영역 (예컨대, 프로모터, 인핸서, 또는 전사 억제자-결합 요소)와 같은 비코딩 영역; 또는 이것들의 임의의 조합일 수 있다.

[0389] 핵산 삽입물은 또한 조건부 대립유전자를 포함할 수 있다. 조건부 대립유전자는 US 2011/0104799 (모든 목적에 대해 그 전문이 본원에 참조로 포함됨)에서 기술되는 것과 같이, 다중기능성 대립유전자일 수 있다. 예를 들어, 조건부 대립유전자는 (a) 표적 유전자의 전사와 관련하여 센스 방향으로의 작동 서열; (b) 센스 또는 안티센스 방향으로의 약물 선택 카세트 (DSC); (c) 안티센스 방향으로의 관심 있는 뉴클레오타이드 서열 (NSI); 및 (d) 역 방향으로의 반전 모듈에 의한 조건부 (COIN, 엑손 스플릿팅 인트론 및 반전 가능한 유전자 트랩-유사 모듈을 이용함)를 포함할 수 있다. 예컨대, US 2011/0104799를 참조한다. 조건부 대립유전자는 (i) 작동 서열 및 DSC가 없고; (ii) 센스 방향으로 NSI 및 안티센스 방향으로 COIN을 함유하는 조건부 대립유전자를 형성하기 위하여 제 1 재조합효소에 대한 노출시 재조합하는 재조합 가능한 단위를 추가로 포함한다. 예컨대, US 2011/0104799를 참조한다.

[0390] 핵산 삽입물은 또한 선택 마커를 암호화하는 폴리뉴클레오타이드를 포함할 수 있다. 대안적으로, 핵산 삽입물은 선택 마커를 암호화하는 폴리뉴클레오타이드가 없을 수 있다. 선택 마커는 선택 카세트에 함유될 수 있다. 선택적으로, 선택 카세트는 자가-삭제 카세트일 수 있다. 예컨대, US 8,697,851 및 US 2013/0312129를 참조하며, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다. 예로서, 자가-삭제 카세트는 마우스 *Prm1* 프로모터에 작동가능하게 연결된 Crei 유전자 (Cre 재조합효소를 암호화하는 2개의 엑손을 포함함, 인트론에 의해 분리됨) 및 인간 유비퀴틴 프로모터에 작동가능하게 연결된 네오마이신 내성 유전자를 포함할 수 있다. 예시적인 선택 마커로는 네오마이신 포스포트랜스퍼라제 (neo^r), 하이그로마이신 B 포스포트랜스퍼라제 (hyg^r), 퓨로마이신-N-아세틸트랜스퍼라제 ($puro^r$), 블라스티시딘 S 탈아민효소 (bsr^r), 크산틴/구아닌 포스포리보실 트랜스퍼라제 (gpt), 또는 단순 포진 바이러스 티미딘 키나아제 (HSV-k), 또는 이것들의 조합을 들 수 있다. 선택 마커를 암호화하는 폴리뉴클레오타이드는 표적화되는 세포에서 활성인 프로모터에 작동가능하게 연결될 수 있다. 프로모터의 예는 본원의 다른 곳에서 기술된다.

[0391] 핵산 삽입물은 또한 리포터 유전자를 포함할 수 있다. 예시적인 리포터 유전자로는 루시퍼라제, β-갈락토시다제, 녹색 형광 단백질 (GFP), 증강된 녹색 형광 단백질 (eGFP), 청록색 형광 단백질 (CFP), 황색 형광 단백질 (YFP), 증강된 황색 형광 단백질 (eYFP), 청색 형광 단백질 (BFP), 증강된 청색 형광 단백질 (eBFP), DsRed, ZsGreen, MmGFP, mPlum, mCherry, tdTomato, mStrawberry, J-Red, mOrange, mKO, mCitrine, 비너스, YPet, 에메랄드, CyPet, 세룰리안, T-사파이어, 및 알칼리성 포스파타제를 암호화하는 것들을 들 수 있다. 이러한 리포터 유전자는 표적화되는 세포에서 활성인 프로모터에 작동가능하게 연결될 수 있다. 프로모터의 예는 본원의 다른 곳에서 기술된다.

[0392] 핵산 삽입물은 또한 하나 이상의 발현 카세트 또는 검출 카세트를 포함할 수 있다. 주어진 카세트는 관심 있는 하나 이상의 뉴클레오타이드 서열, 선택 마커를 암호화하는 폴리뉴클레오타이드, 및 리포터 유전자를, 발현에 영향을 미치는 다양한 조절 구성요소들과 함께 포함할 수 있다. 포함될 수 있는 선택가능한 마커 및 리포터 유전자의 예는 본원의 다른 곳에서 상세하게 논의된다.

[0393] 핵산 삽입물은 부위-특이적 재조합 표적 서열로 플랭킹된 핵산을 포함할 수 있다. 대안적으로, 핵산 삽입물은 하나 이상의 부위-특이적 재조합 표적 서열을 포함할 수 있다. 비록 전체적인 핵산 삽입물이 이러한 부위-특이적 재조합 표적 서열에 의해 플랭킹될 수 있긴 하지만, 핵산 삽입물 내의 관심 있는 임의의 영역 또는 개개의 폴리뉴클레오타이드가 또한 이러한 부위에 의해 플랭킹딜 수 있다. 핵산 삽입물 또는 핵산 삽입물에서 관심 있는 임의의 폴리뉴클레오타이드를 플랭킹할 수 있는 부위-특이적 재조합 표적 서열로는, 예를 들어, loxP, lox511, lox2272, lox66, lox71, loxM2, lox5171, FRT, FRT11, FRT71, attp, att, FRT, rox, 또는 이것들의 조합을 들 수 있다. 한 예에서, 부위-특이적 재조합 부위는 핵산 삽입물 내에 함유된 선택 마커 및/또는 리포터 유전자를 암호화하는 폴리뉴클레오타이드를 플랭킹한다. 핵산 삽입물의 *HSD17B13* 유전자에의 통합 후에, 부위-특이적 재조합 부위들 사이의 서열은 제거될 수 있다. 선택적으로, 2개의 외인성 도너 서열이 사용될 수 있고, 이때 각각은 부위-특이적 재조합 부위를 포함하는 핵산 삽입물을 가진다. 외인성 도너 서열은 관심 있는 핵산을 플랭킹하는 5' 및 3' 영역에 대해 표적화될 수 있다. 2개의 핵산 삽입물의 표적 게놈 유전자좌 안으로의 통합 후에, 2개의 삽입된 부위-특이적 재조합 부위 사이의 관심 있는 핵산이 제거될 수 있다.

[0394] 핵산 삽입물은 또한 제 I형, 제 II형, 제 III형, 및 제 IV형 엔도뉴클레아제를 포함하는, 제한 엔도뉴클레아제 (즉, 제한 효소)에 대한 하나 이상의 제한 부위를 포함할 수 있다. 제 I형 및 제 III형 제한 엔도뉴클레아제는 특이적 인식 서열을 인식하지만, 전형적으로 뉴클레아제 결합 부위로부터 가변적인 위치에서 절단하고, 그 위치는 절단 부위 (인식 서열)로부터 수백개의 염기쌍만큼 떨어여 있을 수 있다. 제 II형 시스템에서 제한 활성은 임의의 메틸라아제 활성과 무관하고, 절단은 전형적으로 결합 부위 내에서 또는 가까이에서 특이적 부위에서 일

어난다. 대부분의 제 II형 효소는 회문 서열을 절단하지만, 제 IIa형 효소는 비회문 인식 서열을 인식하여 인식 서열의 외부를 절단하고, 제 IIb형 효소는 인식 서열의 양쪽 외부로 2회 서열을 절단하며, 제 II형 효소는 비대칭 인식 서열을 인식하고 한쪽을, 인식 서열로부터 약 1 내지 20개 뉴클레오타이드의 한정된 거리에서 절단한다. 제 IV형 제한 효소는 메틸화된 DNA를 표적화한다. 제한 효소는 추가로, 예를 들어 REBASE 데이터베이스에서 기술되고 분류된다 (webpage at rebase.neb.com; Roberts et al., (2003) *Nucleic Acids Res.* 31:418-420; Roberts et al., (2003) *Nucleic Acids Res.* 31:1805-1812; 및 Belfort et al. (2002) in Mobile DNA II, pp. 761-783, Eds. Craigie et al., (ASM Press, Washington, DC)).

[0395] ***(1) 비상동성 단부 연합-매개된 삽입을 위한 도너 서열***

[0396] 일부 외인성 도너 서열은 표적 게놈 유전자좌 (예컨대, *HSD17B13* 유전자의)에서 뉴클레아제-매개된 또는 Cas-단백질-매개된 절단에 의해 생성된 하나 이상의 오버행에 상보하는 5' 단부 및/또는 3' 단부에서 짧은 단일 가닥 영역을 가진다. 이런 오버행들은 또한 5' 및 3' 상동성 팔로서 언급될 수 있다. 예를 들어, 일부 외인성 도너 서열은 표적 게놈 유전자좌에서 5' 및/또는 3' 표적 서열에서의 Cas-단백질-매개된 절단에 의해 생성된 하나 이상의 오버행에 상보하는 5' 단부 및/또는 3' 단부에서 짧은 단일-가닥 영역을 가진다. 일부 이러한 외인성 도너 서열은 5' 단부에서만 또는 3' 단부에서만 상보하는 가닥을 가진다. 예를 들어, 일부 이러한 외인성 도너 서열은 표적 게놈 유전자좌에서 5' 표적 서열에서 생성된 오버행에 상보하는 5' 단부에서만 상보하는 영역을 갖거나 표적 게놈 유전자좌에서 3' 표적 서열에서 생성된 오버행에 상보하는 3' 단부에서만 상보하는 영역을 갖는다. 다른 이러한 외인성 도너 서열은 5' 및 3' 단부에서 모두 상보하는 영역을 가진다. 예를 들어, 다른 이러한 외인성 도너 서열은 5' 및 3' 단부 둘 다에서 상보하는, 예컨대, 각각 표적 게놈 유전자좌에서의 Cas-매개된 절단에 의해 생성된 제1 및 제2 오버행에 상보하는 영역을 가진다. 예를 들어, 만약 외인성 도너 서열이 이중 가닥이면, 단일 가닥 상보 영역은 도너 서열의 상부 가닥의 5' 단부로부터 및 도너 서열의 하부 가닥의 5' 단부로부터 연장되어 각 단부에서 5' 오버행이 생성될 수 있다. 대안적으로, 단일 가닥 상보 영역은 도너 서열의 상부 가닥의 3' 단부로부터 및 주형의 하부 가닥의 3' 단부로부터 연장되어 3' 오버행이 생성될 수 있다.

[0397] 상보하는 영역은 외인성 도너 서열과 *HSD17B13* 유전자 사이의 결찰을 촉진하기에 충분한 임의의 길이의 것일 수 있다. 예시적인 상보하는 영역은 길이가 약 1 내지 약 5개 뉴클레오타이드, 길이가 약 1 내지 약 25개 뉴클레오타이드, 또는 길이가 약 5 내지 약 150개 뉴클레오타이드이다. 예를 들어, 상보하는 영역은 길이가 적어도 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 또는 25개 뉴클레오타이드일 수 있다. 대안적으로, 상보하는 영역은 길이가 약 5 내지 약 10개, 약 10 내지 약 20개, 약 20 내지 약 30개, 약 30 내지 약 40개, 약 40 내지 약 50개, 약 50 내지 약 60개, 약 60 내지 약 70개, 약 70 내지 약 80개, 약 80 내지 약 90개, 약 90 내지 약 100개, 약 100 내지 약 110개, 약 110 내지 약 120개, 약 120 내지 약 130개, 약 130 내지 약 140개, 약 140 내지 약 150개 뉴클레오타이드이거나, 더 길 수 있다.

[0398] 이러한 상보하는 영역은 두 쌍의 닉카아제에 의해 생성된 오버행에 상보할 수 있다. 엇갈린 단부를 가진 2개의 이중 가닥 끊김은 DNA의 반대쪽 가닥을 절단하여 제1 이중 가닥 끊김을 생성하는 제1 및 제2 닉카아제, 및 DNA의 반대쪽 가닥을 절단하여 제2 이중 가닥 끊김을 생성하는 제 3 및 제 4 닉카아제를 사용함으로써 생성될 수 있다. 예를 들어, Cas 단백질은 제1, 제2, 제 3, 및 제 4 가이드 RNA와 상응하는 제1, 제2, 제 3, 및 제 4 가이드 RNA 표적 서열을 자르기 위해 사용될 수 있다. 제1 및 제2 가이드 RNA 표적 서열은 DNA의 제1 및 제2 가닥에서 제1 및 제2 닉카아제에 의해 생성된 닉이 이중 가닥 끊김을 생성하도록 제1 절단 부위를 생성하기 위해 위치될 수 있다 (즉, 제1 절단 부위는 제1 및 제2 가이드 RNA 표적 서열 내의 닉을 포함한다). 마찬가지로, 제 3 및 제 4 가이드 RNA 표적 서열은 DNA의 제1 및 제2 가닥에서 제 3 및 제 4 닉카아제에 의해 생성된 닉이 이중 가닥 끊김을 생성하도록 제2 절단 부위를 생성하기 위해 위치될 수 있다 (즉, 제2 절단 부위는 제 3 및 제 4 가이드 RNA 표적 서열 내의 닉을 포함한다). 바람직하게, 제1 및 제2 가이드 RNA 표적 서열 및/또는 제 3 및 제 4 가이드 RNA 표적 서열 내의 닉은 오버행을 생성하는 오프셋 (off-set) 닉일 수 있다. 오프셋 창은, 예를 들어, 적어도 약 5 bp, 10 bp, 20 bp, 30 bp, 40 bp, 50 bp, 60 bp, 70 bp, 80 bp, 90 bp, 100 bp 또는 그 이상일 수 있다. Ran et al. (2013) *Cell* 154:1380-1389; Mali et al. (2013) *Nat. Biotech.*31:833-838; 및 Shen et al. (2014) *Nat. Methods* 11:399-404를 참조하며, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함된다. 이러한 경우에, 이중 가닥 외인성 도너 서열은 제1 및 제2 가이드 RNA 표적 서열 내의 닉 및 제 3 및 제 4 가이드 RNA 표적 서열 내의 닉에 의해 생성된 오버행에 상보하는 단일-가닥 상보하는 영역으로 설계될 수 있다. 이러한 외인성 도너 서열은 그런 후 비상동성 단부 연합-매개되 결찰에 의해 삽입될 수 있다.

[0399] ***(2) 상동성-의존 복구에 의한 삽입을 위한 도너 서열***

[0400] 일부 외인성 도너 서열 (즉, 표적화 벡터)은 상동성 팔을 포함한다. 만약 외인성 도너 서열이 또한 핵산 삽입물을 포함하면, 상동성 팔은 핵산 삽입물을 플랭킹할 수 있다. 용이한 참조를 위해, 상동성 팔은 본원에서 5' 및 3' (즉, 상류 및 하류) 상동성 팔로서 언급된다. 이 용어는 외인성 도너 서열 내의 핵산 삽입물에 대한 상동성 팔의 상대적인 위치와 관련된다. 5' 및 3' 상동성 팔은, 각각 본원에서 "5' 표적 서열" 및 "3' 표적 서열"로서 언급되는, *HSD17B13* 유전자 내의 영역에 상응한다.

[0401] 상동성 팔 및 표적 서열은 두 영역이 상동성 재조합 반응에 대한 기질로서 작용하기 위하여 서로에 대해 충분한 수준의 서열 동일성을 공유할 때 서로에 대해 "상응한다" 또는 "상응하는" 것이다. 용어 "상동성"은 상응하는 서열에 대해 동일하거나 또는 서열 동일성을 공유하는 DNA 서열을 포함한다. 외인성 도너 서열에서 발견되는 주어진 표적 서열과 상응하는 상동성 팔 사이의 서열 동일성은 상동성 재조합이 발생하는 것을 허용하는 임의의 정도의 서열 동일성일 수 있다. 예를 들어, 외인성 도너 서열 (또는 이것의 단편)과 표적 서열 (또는 이것의 단편)의 상동성 팔에 의해 공유된 서열 동일성의 양은 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성이어서, 서열이 상동성 재조합을 진행할 수 있다. 더욱이, 상동성 팔과 상응하는 표적 서열 사이의 상응하는 상동성 영역은 상동성 재조합을 촉진하기에 충분한 임의의 길이의 것일 수 있다. 예시적인 상동성 팔은 길이가 약 25개 뉴클레오타이드 내지 약 2.5 kb, 길이가 약 25개 뉴클레오타이드 내지 약 1.5 kb, 또는 길이가 약 25개 내지 약 500개 뉴클레오타이드이다. 예를 들어, 주어진 상동성 팔 (또는 각각의 상동성 팔) 및/또는 상응하는 표적 서열은 길이가 약 25 내지 약 30개, 약 30 내지 약 40개, 약 40 내지 약 50개, 약 50 내지 약 60개, 약 60 내지 약 70개, 약 70 내지 약 80개, 약 80 내지 약 90개, 약 90 내지 약 100개, 약 100 내지 약 150개, 약 150 내지 약 200개, 약 200 내지 약 250개, 약 250 내지 약 300개, 약 300 내지 약 350개, 약 350 내지 약 400개, 약 400 내지 약 450개, 또는 약 450 내지 약 500개 뉴클레오타이드인 상응하는 상동성 영역을 포함하여서, 상동성 팔은 *HSD17B13* 유전자 내의 상응하는 표적 서열와 상동성 재조합을 진행하기에 충분한 상동성을 가질 수 있다. 대안적으로, 주어진 상동성 팔 (또는 각각의 상동성 팔) 및/또는 상응하는 표적 서열은 길이가 약 0.5 kb 내지 약 1 kb, 약 1 kb 내지 약 1.5 kb, 약 1.5 kb 내지 약 2 kb, 또는 약 2 kb 내지 약 2.5 kb인 상응하는 상동성 영역을 포함할 수 있다. 예를 들어, 상동성 팔은 각각 길이가 약 750개 뉴클레오타이드일 수 있다. 상동성 팔은 대칭적이거나 (각각 거의 동일한 크기의 길이), 또는 비대칭적일 수 있다 (하나가 다른 하나보다 김).

[0402] 상동성 팔은 세포에 대해 천연인 유전자좌 (예컨대, 표적화된 유전자좌)에 상응할 수 있다. 대안적으로, 예를 들어, 그것들은 예를 들어, 도입유전자, 발현 카세트, 또는 DNA의 이종성 또는 외인성 영역을 포함하는, 세포의 게놈 안으로 통합된 DNA의 이종성 또는 외인성 분절의 영역에 상응할 수 있다. 대안적으로, 표적화 벡터의 상동성 팔은 효모 인공 염색체 (YAC), 박테리아 인공 염색체 (BAC), 인간 인공 염색체의 영역, 또는 적절한 숙주 세포에 함유된 임의의 다른 조작된 영역에 상응할 수 있다. 또한, 표적화 벡터의 상동성 팔은 BAC 라이브러리, 코스미드 라이브러리, 또는 P1 파지 라이브러리의 영역에 상응하거나 그 영역으로부터 유래될 수 있거나, 또는 합성 DNA로부터 유래될 수 있다.

[0403] 뉴클레아제 제제가 외인성 도너 서열과 조합하여 사용될 때, 5' 및 3' 표적 서열은 바람직하게는, 뉴클레아제 절단 부위에서 단일 가닥 끊김 (닉) 또는 이중 가닥 끊김이 일어날 때 표적 서열과 상동성 팔 사이의 상동성 재조합 사건이 발생하는 것을 촉진하기 위하여 뉴클레아제 절단 부위에 충분히 근접하여 위치한다. 용어 "뉴클레아제 절단 부위"는 닉 또는 이중 가닥 끊김이 뉴클레아제 제제 (예컨대, 가이드 RNA와 복합체를 형성한 Cas9 단백질)에 의해 생성되는 DNA 서열을 포함한다. 외인성 도너 서열의 5' 및 3' 상동성 팔에 상응하는 *HSD17B13* 유전자 내의 표적 서열은 거리가 예컨대 뉴클레아제 절단 부위에서 단일 가닥 끊김 또는 이중 가닥 끊김이 있을 때 5' 및 3' 표적 서열과 상동성 팔 사이의 상동성 재조합 사건의 발생을 촉진하기 위한 만큼의 거리인 경우 뉴클레아제 절단 부위에 "충분히 근접하여 위치한" 것이다. 그러므로, 외인성 도너 서열의 5' 및/또는 3' 상동성 팔에 상응하는 표적 서열은, 예를 들어, 주어진 뉴클레아제 절단 부위의 적어도 1개 뉴클레오타이드 내, 또는 주어진 뉴클레아제 절단 부위의 적어도 10개 뉴클레오타이드 내지 약 1,000개 뉴클레오타이드 내에 있을 수 있다. 예로서, 뉴클레아제 절단 부위는 적어도 하나의 또는 둘 다의 표적 서열에 바로 인접할 수 있다.

[0404] 외인성 도너 서열 및 뉴클레아제 절단 부위의 상동성 팔에 상응하는 표적 서열의 공간적 관계는 달라질 수 있다. 예를 들어, 표적 서열은 뉴클레아제 절단 부위에 대해 5'에 위치할 수 있거나, 표적 서열은 뉴클레아제 절단 부위에 대해 3'에 위치할 수 있거나, 또는 표적 서열은 뉴클레아제 절단 부위를 플랭킹할 수 있다.

[0405] ***IV. 치료 및 예방적 적용***

[0406] 또한 내인성 *HSD17B13* 유전자의 발현을 변형시키거나 변경시키기 위해 본원에 개시된 방법을 사용하여 질환에 걸린 또는 질환의 위험이 있는 대상체에서 만성 간 질환의 치료방법 및 치료 또는 예방 방법이 제공된다. 또한 또한 내인성 *HSD17B13* 유전자의 발현을 변형시키거나 변경시키기 위해 본원에 개시된 방법을 사용하여 질환에 걸린 또는 질환의 위험이 있는 대상체에서 알코올성 간 질환 또는 비알코올성 간 질환과 같은 간 질환의 치료 방법 및 치료 또는 예방 방법이 제공된다. 또한 HSD17B13 mRNA 전사물의 발현을 감소시키는 방법을 사용하여 또는 대상체에게 HSD17B13 단백질을 암호화하는 재조합 핵산을 제공하거나, HSD17B13 단백질을 암호화하는 mRNA를 제공하거나, 또는 HSD17B13 단백질을 제공하는 방법을 사용하여 질환에 걸린 또는 질환의 위험이 있는 대상체에서 만성 간 질환의 치료적 방법 및 치료 또는 예방 방법이 제공된다. 또한 HSD17B13 mRNA 전사물의 발현을 감소시키는 방법을 사용하여 또는 대상체에게 HSD17B13 단백질을 암호화하는 재조합 핵산을 제공하거나, HSD17B13 단백질을 암호화하는 mRNA를 제공하거나, 또는 HSD17B13 단백질을 제공하는 방법을 사용하여 질환에 걸린 또는 질환의 위험이 있는 대상체에서 알코올성 간 질환 또는 비알코올성 간 질환과 같은 간 질환의 치료적 방법 및 치료 또는 예방 방법이 제공된다. 방법은 하나 이상의 핵산 또는 단백질을 대상체 안으로, 대상체의 간 안으로, 또는 대상체의 세포 (예컨대, 간세포) 안으로 도입하는 단계 (예컨대, 생체내에서 또는 생체외에서)를 포함할 수 있다.

[0407] 만성 간 질환은 6개월의 기간에 경쳐 지속되는 간의 질환을 포함하며, 예를 들어, 섬유증 및 간경변으로 이어질 수 있는 간 실질의 점진적인 파괴 및 재생을 포함한 간의 질환을 포함할 수 있다. 만성 간 질환은 알코올성 간 질환 또는 비알코올성 간 질환일 수 있다. 만성 간 질환에 의해 포함되는 간 병리는, 예를 들어, 염증 (예컨대 만성 간염), 간의 간경변, 및 간세포 암종을 포함한다. 만성 간 질환의 유형은 본원의 다른 곳에서 개시되며, 예를 들어, 지방간 질환, 비알코올성 지방간 질환, 알코올성 지방간 질환, 간경변, 및 간세포 암종을 들 수 있다. 만성 간 질환의 증상 및 신호는 알려져 있고, 예를 들어, 거대 간, 피로, 상부 우측 복부의 통증, 복부 팽창 (복수), 피표면 바로 아래의 확대된 혈관, 남성에서 확대된 유방, 확대된 비장, 붉은 손바닥, 및 피부 및 눈의 황화 (황달)을 들 수 있다. 만성 간 질환에 대한 실험은 혈액 검사, 간의 영상화, 및 간의 생검을 포함할 수 있다. 만약 대상체가 위험 인자가 없는 개체보다 질환이 발병할 위험이 통계학적으로 유의하게 더 큰 위험 인자를 개체가 가진 것으로 보는 적어도 하나의 공지된 위험 인자 (예컨대, 질환 유발 돌연변이와 같은 유전적 인자)를 가진다면 개체는 만성 간 질환의 위험이 증가한 것이다. 만성 간 질환에 대한 위험 인자는 또한 알려져 있고, 예를 들어, 과도한 알코올 이용, 비만, 고콜레스테롤, 고수준의 혈중 트라이글리세라이드, 다낭성 난소 증후군, 수면 무호흡, 제2형 당뇨병, 불충분한 활동의 갑상선 (갑상선 기능저하증), 불충분한 활동의 뇌하수체 (뇌하수체 기능부전), 및 상승된 혈중 지질을 포함하는 대사 증후군을 들 수 있다.

[0408] 용어 "대상체"는 예방적 또는 치료적 처리를 받는 인간 및 다른 포유류 대상체 (예컨대, 고양이과, 개과, 설치류, 마우스, 또는 래트) 또는 비포유류 대상체 (예컨대, 가금류)를 포함한다. 이러한 대상체는, 예를 들어, *HSD17B13* rs72613567 변종의 보균자가 아닌 (또는 오직 *HSD17B13* rs72613567 변종의 이형접합성 보균자인) 및 만성 간 질환에 걸렸거나 발병하기 쉬운 대상체 (예컨대 인간)일 수 있다. 게놈 DNA를 포함하는 생물학적 샘플에서 *HSD17B13* rs72613567 변종의 존재를 검출하기 위해, mRNA 또는 cDNA를 포함하는 생물학적 샘플에서 *HSD17B13* 전사물 C, D, E, F, F', G, 및 H 중 임의의 하나, 특히 D 또는 이것들의 조합의 존재 또는 수준을 검출하기 위해, 또는 단백질을 포함하는 생물학적 샘플에서 HSD17B13 단백질 동형 C, D, E, F, F', G, 또는 H 중 임의의 하나, 특히 D, 또는 이것들의 조합의 존재 또는 수준을 검출하기 위한 다양한 방법이 가능하다. 게놈 DNA에서 서열의 존재를 검출하기 위한 및 특정 mRNA 전사물 또는 단백질 동형의 존재를 검출하기 위한 방법들은 잘 알려져 있다. 집단 내 유전자 서열 및 이러한 유전자에 의해 암호화된 mRNA 및 단백질은 단일-뉴클레오타이드 다형성과 같은 다형성으로 인해 달라질 수 있는 것이 인지된다. *HSD17B13* 유전자에 대해 및 각각의 *HSD17B13* 전사물 및 HSD17B13 동형에 대해 제공된 서열들은 단지 *HSD17B13* 유전자 및 각각의 *HSD17B13* 전사물 및 HSD17B13 동형에 대한 예시적인 서열들이다. *HSD17B13* 유전자 및 각각의 *HSD17B13* 전사물 및 HSD17B13 동형에 대한 다른 서열들 또한 가능하다.

[0409] 예를 들어, 세포 또는 인간 대상체와 같은 대상체에서 *HSD17B13* rs72613567 변종을 검출하는 방법은, 예를 들어, *HSD17B13* 유전자를 포함하는 대상체로부터 생물학적 샘플을 얻는 단계, 및 *HSD17B13* 유전자 및 SEQ ID NO: 2가 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 *HSD17B13* 유전자의 위치가 티민에 의해 차지된 것을 측정하는 또는 *HSD17B13* 유전자 및 SEQ ID NO: 1이 최적으로 정렬될 때 위치 12665 및 12666에 상으하는 위치들 사이에 삽입되는 것을 측정하는, 생물학적 샘플에 대한 검정을 수행하는 단계를 포함할 수 있다. *HSD17B13* 유전자 및 SEQ ID NO: 2가 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 *HSD17B13* 유전자의 위치가 티민에 의해 차지된 것으로 측정하는 것은 티민이 SEQ ID NO: 1의 위치 12665와 12666에 상응하는

위치 사이에 삽입되는 것으로 측정될 수 있는, SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 위치를 플랭킹하는 위치에서 충분한 수의 뉴클레오타이드의 동일성이 측정되는 것을 의미하는 것으로 인지된다. 이러한 검정은, 예를 들어, *HSD17B13* 유전자 및 SEQ ID NO: 2 (또는 SEQ ID NO: 1)가 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 *HSD17B13* 유전자의 위치 (또는 SEQ ID NO: 1의 위치 12665 및 12666)와 하나 이상의 주변 위치 (예컨대, SEQ ID NO: 2의 위치 12666 또는 SEQ ID NO: 1의의 위치 12665 및 12666의 한쪽 또는 양쪽을 플랭킹하는 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10개의 위치)의 동일성을 측정하는 것을 포함할 수 있다. 마찬가지로, 검정은 *HSD17B13* 유전자 및 SEQ ID NO: 1이 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 위치를 포함하는 *HSD17B13* 유전자의 일부를 서열분석하는 것을 포함할 수 있다. 대안적으로, 이러한 방법에서 검정은 생물학적 샘플을 *HSD17B13* rs72613567 변종에 특이적으로 혼성화하고 상응하는 야생형 *HSD17B13* 서열에는 혼성화하지 않는 (예컨대, 엄격한 조건 하에서) 프라이머 또는 프로브와 접촉시키는 단계, 및 혼성화가 일어났는지의 여부를 측정하는 단계를 포함할 수 있다.

[0410] 이러한 방법은 게놈 편집 또는 유전자 요법을 포함할 수 있다. 예를 들어, *HSD17B13* rs72613567 변종이 아닌 내인성 *HSD17B13* 유전자는 *HSD17B13* rs72613567 변종과 관련된 변이 (즉, *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 티민의 삽입, 또는 반대쪽 가닥에 상응하는 위치에서 아데닌의 삽입)를 포함하도록 변형될 수 있다. 다른 예로서, *HSD17B13* rs72613567 변종이 아닌 내인성 *HSD17B13* 유전자는 녹아웃되거나 비활성화될 수 있다. 마찬가지로, *HSD17B13* rs72613567 변종이 아닌 내인성 *HSD17B13* 유전자는 녹아웃되거나 비활성화될 수 있고, *HSD17B13* rs72613567 변종과 관련된 변형을 포함하는 *HSD17B13* 유전자 (예컨대, 전체 *HSD17B13* rs72613567 변종 또는 변형을 포함하는 미니유전자)는 도입되고 발현될 수 있다. 유사하게, *HSD17B13* rs72613567 변종이 아닌 내인성 *HSD17B13* 유전자는 녹아웃되거나 비활성화될 수 있고, HSD17B13 동형 C, D, F, G, 및 H (또는 이것들의 단편) 중 임의의 하나 또는 임의의 조합을 암호화하는 mRNA가 도입되고 발현될 수 있거나 (예컨대, 세포내 단백질 대체 요법), 또는 HSD17B13 동형 C, D, F, G, 및 H (또는 이것들의 단편) 중 임의의 하나 또는 임의의 조합이 도입될 수 있다 (예컨대, 단백질 대체 요법). 특정 구체예에서, HSD17B13 동형 (또는 이것을 암호화하는 DNA 또는 mRNA)의 조합은 HSD17B13 동형 D를 포함하는 조합이다 (예컨대, D, DC, DF, DG, DH, DCF, DCG, DCH, DFG, DFH, DGH, DCFG, DCFH, DCGH, DFGH, 또는 DCFGH).

[0411] 다른 이러한 방법은 *HSD17B13* rs72613567 변종 (예컨대, 전체 *HSD17B13* rs72613567 변종 또는 변형을 포함하는 미니유전자)과 관련된 변형을 포함하는 재조합 *HSD17B13* 유전자를 도입하고 발현시키는 단계, HSD17B13 동형 C, D, F, G, 및 H 또는 이것들의 단편 중 임의의 하나 또는 임의의 조합을 암호화하는 재조합 핵산 (예컨대, DNA)을 도입하고 발현시키는 단계, HSD17B13 동형 C, D, F, G, 및 H 또는 이것들의 단편 중 임의의 하나 또는 임의의 조합을 암호화하는 하나 이상의 mRNA를 도입하고 발현시키는 단계 (예컨대, 세포내 단백질 대체 요법), 또는 *HSD17B13* rs72613567 변종이 아닌 내인성 *HSD17B13* 유전자를 녹아웃시키거나 비활성화하지 않으면서 HSD17B13 동형 C, D, F, G, 및 H 또는 이것들의 단편 중 임의의 하나 또는 임의의 조합을 도입하는 단계 (예컨대 단백질 대체 요법)를 포함할 수 있다. 특정 구체예에서, HSD17B13 동형 (또는 이것을 암호화하는 DNA 또는 mRNA)의 조합은 HSD17B13 동형 D를 포함하는 조합이다 (예컨대, D, DC, DF, DG, DH, DCF, DCG, DCH, DFG, DFH, DGH, DCFG, DCFH, DCGH, DFGH, 또는 DCFGH). 선택적으로, 이러한 방법은 또한 그것의 발현이 *HSD17B13* rs72613567 변종의 보균자에서 감소하는 *HSD17B13* 전사물 (예컨대, 전사물 A, B, E, 및 F')이 감소된 발현에 대해, 예컨대 안티센스 RNA, siRNA, 또는 shRNA의 사용을 통해 표적화되는 방법과 병용하여 수행될 수 있다. 특정 구체예에서, 감소된 발현에 대해 표적화된 *HSD17B13* 전사물은 전사물 A를 포함하는 조합이다 (예컨대, A, AB, AE, AF', ABE, ABF', AEF', 또는 ABEF').

[0412] *HSD17B13* 유전자 또는 미니유전자 또는 HSD17B13 동형 C, D, F, G, 및 H 또는 이것들의 단편 중 임의의 하나 또는 임의의 조합을 암호화하는 DNA는 게놈을 변형시키지 않는 발현 벡터의 형태로 도입 및 발현될 수 있고, 그것은 *HSD17B13* 유전자좌 안으로 게놈상에 통합되도록 표적화 벡터의 형태로 도입될 수 있거나, 또는 안전한 은닉 유전자좌와 같이, *HSD17B13* 유전자좌 이외의 유전자좌 안으로 게놈상에 통합되도록 도입될 수 있다. 게놈상에 통합된 *HSD17B13* 유전자는 *HSD17B13* 프로모터 또는 다른 프로모터, 예컨대 통합 부위에서 내인성 프로모터에 작동가능하게 연결될 수 있다. 안전한 은닉 유전자좌는 도입유전자가 유전자 구조나 발현에 불리하게 영향을 주지 않으면서 관심 있는 모든 조직에서 안정적으로 및 믿을만하게 발현될 수 있는 염색체 부위이다. 안전한 은닉 유전자좌로는, 예를 들어, 다음 특징들 중 하나 이상 또는 전부를 가질 수 있다: (1) 임의의 유전자의 5' 단부로부터 50 kb 이상의 거리; 임의의 암-관련 유전자로부터 300 kb 이상의 거리; 임의의 microRNA로부터 300 kb 이상의 거리; 유전자 전사 단위 외부; 및 초-보존 영역의 외부. 적합한 안전한 은닉 유전자좌의 예로는 아데노-관련 바이러스 부위 1 (AAVS1), 케모카인 (CC 모티프) 수용체 5 (CCR5) 유전자좌, 및 마우스 ROSA26 유전자좌의

인간 병렬상동체를 들 수 있다.

[0413] 도입되고 발현될 수 있는 HSD17B13 단백질 동형 또는 HSD17B13 단백질 동형을 암호화하는 핵산으로는, 예를 들어, C, D, F, G, H, CD, CF, CG, CH, DF, DG, DH, FG, FH, GH, CDF, CDG, CDH, CFG, CFH, CGH, DFG, DFH, DGH, FGH, CDFG, CDFH, CFGH, DFGH, 및 CDFGH를 들 수 있다. 특정 방법에서, HSD17B13 동형 D 또는 동형 D (단독으로 또는 다른 동형과 함께)를 암호화하는 핵산이 도입되거나 발현된다. 이런 동형 및 전사물의 각각에 대한 예시적인 서열은 본원의 다른 곳에서 제공된다. 그러나, 유전자 서열 및 집단 내에서 이러한 유전자로부터 전사된 mRNA 서열, 및 이러한 mRNA로부터 번역된 단백질은 단일-뉴클레오타이드 다형성과 같은 다형성으로 인해 달라질 수 있다. 각 전사물 및 동형에 대해 본원에서 제공된 서열들은 단지 예시적인 서열들이다. 다른 서열들 또한 가능하다.

[0414] 그것의 발현이 안티센스 RNA, shRNA, 또는 siRNA를 통한 감소에 대해 표적화될 수 있는 HSD17B13 전사물의 조합은, 예를 들어, A, B, E, F', AB, AE, AF', BE, BF', ABE, ABF', AEF', BEF', 및 ABEF'를 포함한다. 특정 방법에서, HSD17B13 전사물 A (단독으로 또는 다른 전사물과 함께)가 표적화된다. 예를 들어, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A) 내의 서열에 혼성화할 수 있다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 존재하지 않는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하는 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)의 엑손 7 내의 서열 또는 엑손 6-엑손 7 경계에 걸쳐 있는 서열에 혼성화한다.

[0415] 예를 들어, 일부 이러한 방법은 *HSD17B13* rs72613567 변종의 보균자가 아닌 (또는 오직 *HSD17B13* rs72613567 변종의 이형접합성 보균자인) 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법을 포함하며, 방법은 대상체 안으로 또는 대상체의 간세포 안으로: (a) *HSD17B13* 유전자 내의 뉴클레아제 표적 서열에 결합하는 뉴클레아제 제제 (또는 그것을 암호화하는 핵산), 여기서 뉴클레아제 표적 서열은 *HSD17B13* 유전자가 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 위치 12666에 상응하는 위치를 포함하거나 근접하며; 및 (b) SEQ ID NO: 2의 위치 12666에 상응하는 위치의 표적 서열 5'에 혼성화하는 5' 상동성 팔, SEQ ID NO: 2의 위치 12666에 상응하는 위치의 표적 서열 3'에 혼성화하는 3' 상동성 팔, 및 5' 상동성 팔 및 3' 상동성 팔에 의해 플랭킹된 티민을 포함하는 핵산 삽입물을 포함하는 외인성 도너 서열을 도입하는 단계를 포함한다. 뉴클레아제 제제는 대상체의 간세포에서 *HSD17B13* 유전자를 절단할 수 있고, 외인성 도너 서열은 간세포에서 *HSD17B13* 유전자와 재조합할 수 있으며, 여기서 외인성 도너 서열과 *HSD17B13* 유전자의 재조합시에, 티민이 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된다. 이러한 방법에 사용될 수 있는 뉴클레아제 제제 (예컨대, Cas9 단백질 및 가이드 RNA)의 예는 본원의 다른 곳에서 개시된다. 적합한 가이드 RNA 및 가이드 RNA 표적 서열의 예는 본원의 다른 곳에서 개시된다. 이러한 방법에 사용될 수 있는 외인성 도너 서열의 예는 본원의 다른 곳에서 개시된다.

[0416] 다른 예로서, 일부 이러한 방법은 *HSD17B13* rs72613567 변종의 보균자가 아닌 (또는 오직 *HSD17B13* rs72613567 변종의 이형접합성 보균자인) 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법을 포함하며, 방법은 대상체 안으로 또는 대상체의 간세포 안으로 SEQ ID NO: 2의 위치 12666에 상응하는 위치의 표적 서열 5'에 혼성화하는 5' 상동성 팔, SEQ ID NO: 2의 위치 12666에 상응하는 위치의 표적 서열 3'에 혼성화하는 3' 상동성 팔, 및 5' 상동성 팔 및 3' 상동성 팔에 의해 플랭킹된 티민을 포함하는 핵산 삽입물을 포함하는 외인성 도너 서열을 도입하는 단계를 포함한다. 외인성 도너 서열은 간세포에서 *HSD17B13* 유전자와 재조합할 수 있고, 여기서 외인성 도너 서열과 *HSD17B13* 유전자의 재조합시에, 티민이 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된다. 이러한 방법에 사용될 수 있는 외인성 도너 서열의 예는 본원의 다른 곳에서 개시된다.

[0417] 일부 이러한 방법은 *HSD17B13* rs72613567 변종의 보균자가 아닌 (또는 오직 *HSD17B13* rs72613567 변종의 이형접합성 보균자인) 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법을 포함하며, 방법은 대상체 안으로 또는 대상체의 간세포 안으로: (a) *HSD17B13* 유전자 내의 뉴클레아제 표적 서열에 결합하는 뉴클레아제 제제 (또는 그것을 암호화하는 핵산)를 도입하는 단계를 포함하며, 여기서 뉴클레아제 표적 서열은 *HSD17B13* 유전자에 대한 개시 코돈을 포함하거나 또는 개시 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있거나 또는 SEQ ID NO: 20 내지 81로부터 선택된다. 뉴클레아제 제제는 대상체의 간세포에서 *HSD17B13* 유전자를 절단하고 유전자의 발현을 파괴할 수 있다. 일부 이러한 방법은 *HSD17B13* rs72613567 변종의 보균자가 아닌 (또는 오직 *HSD17B13* rs72613567 변종의 이형접합성 보균자인) 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체를 치료하는 방법을 포함하며, 방법은 대상체 안으로 또는 대상체의 간세포 안으로; (a) *HSD17B13* 유전자 내의 뉴클레아제 표적 서열에 결합하는 뉴클레아제 제제 (또는 그것을 암호화하는

핵산), 여기서 뉴클레아제 표적 서열은 *HSD17B13* 유전자에 대한 개시 코돈을 포함하거나 또는 개시 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있거나 또는 SEQ ID NO: 20 내지 81로부터 선택되며; 및 (b) 재조합 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는 재조합 *HSD17B13* 유전자를 포함하는 발현 벡터를 도입하는 단계를 포함한다. 발현 벡터는 게놈상으로 통합되지 않는 것일 수 있다. 대안적으로, 재조합 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는 재조합 *HSD17B13* 유전자를 포함하는 표적화 벡터 (즉, 외인성 도너 서열)가 도입될 수 있다. 뉴클레아제 제제는 대상체의 간세포에 *HSD17B13* 유전자를 절단할 수 있고 유전자의 발현을 파괴할 수 있으며, 발현 벡터는 대상체의 간세포에서 재조합 *HSD17B13* 유전자를 발현시킬 수 있다. 대안적으로, 게놈상으로 통합된, 재조합 *HSD17B13* 유전자는 대상체의 간세포에서 발현할 수 있다. 이러한 방법에 사용될 수 있는 뉴클레아제 제제 (예컨대, 뉴클레아제-활성 Cas9 단백질 및 가이드 RNA)의 예는 본원의 다른 곳에서 개시된다. 적합한 가이드 RNA 및 가이드 RNA 표적 서열의 예는 본원의 다른 곳에서 개시된다. 단계 (b)는 대안적으로 HSD17B13 동형 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하는 핵산 (예컨대, DNA)를 포함하거나 및/또는 *HSD17B13* 전사물 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 서열을 포함하는 발현 벡터 또는 표적화 벡터를 도입하는 단계를 포함할 수 있다. 마찬가지로, 단계 (b)는 대안적으로 HSD17B13 동형 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하는 및/또는 *HSD17B13* 전사물 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 상보하는 DNA (또는 이것의 부분)를 가지는 mRNA를 도입하는 단계를 포함할 수 있다. 마찬가지로, 단계 (b)는 대안적으로 HSD17B13 동형 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 서열을 포함하는 단백질을 도입하는 단계를 포함할 수 있다. 특정 방법에서, 전사물은 *HSD17B13* 전사물 D (예컨대, SEQ ID NO: 7)일 수 있거나, 동형은 HSD17B13 동형 D (예컨대, SEQ ID NO: 15)일 수 있다. 다른 특정 방법에서, HSD17B13 동형의 조합, HSD17B13 동형의 조합을 암호화하는 발현 벡터 또는 표적화 벡터, 또는 HSD17B13 동형의 조합을 암호화하하는 mRNA가 도입될 수 있다 (예컨대, D, DC, DF, DG, DH, DCF, DCG, DCH, DFG, DFH, DGH, DCFG, DCFH, DCGH, DFGH, 또는 DCFGH).

[0418] 일부 이러한 방법에서, 제2 뉴클레아제 제제가 또한 대상체 안으로 또는 대상체의 간세포 안으로 도입되는데, 여기서 제2 뉴클레아제 제제는 *HSD17B13* 유전자 내의 제2 뉴클레아제 표적 서열에 결합하며, 제2 뉴클레아제 표적 서열은 *HSD17B13* 유전자에 대한 종결 코돈을 포함하거나 종결 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있거나 또는 SEQ ID NO: 82 내지 225로부터 선택되고, 뉴클레아제 제제는 제1 뉴클레아제 표적 서열 및 제2 뉴클레아제 표적 서열 둘 다의 내부에서 간세포의 *HSD17B13* 유전자를 절단하며, 여기서 간세포는 제1 뉴클레아제 표적 서열 및 제2 뉴클레아제 표적 서열 사이의 삭제를 포함하도록 변형된다. 예를 들어, 제2 뉴클레아제 제제는 Cas9 단백질 및 가이드 RNA일 수 있다. 종결 코돈에 근접한 적합한 가이드 RNA 및 가이드 RNA 표적 서열은 본원의 다른 곳에서 개시된다.

[0419] 이러한 방법은 또한 *HSD17B13* rs72613567 변종의 보균자가 아닌 (또는 오직 *HSD17B13* rs72613567 변종의 이형접합성 보균자인) 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체의 치료 방법을 포함할 수 있고, 방법은 대상체 안으로 또는 대상체의 간세포 안으로: (a) *HSD17B13* 유전자 내의 DNA-결합 단백질 표적 서열에 결합하는 DNA-결합 단백질 (또는 그것을 암호화하는 핵산)을 도입하는 단계를 포함하며, 여기서 DNA-결합 단백질 표적 서열은 *HSD17B13* 유전자에 대한 개시 코돈을 포함하거나 개시 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있거나 또는 SEQ ID NO: 20 내지 81로부터 선택된다. DNA-결합 단백질은 대상체의 간세포에서 *HSD17B13* 유전자의 발현을 변경 (예컨대, 감소)시킬 수 있다. 이러한 방법은 또한 *HSD17B13* rs72613567 변종의 보균자가 아닌 (또는 오직 *HSD17B13* rs72613567 변종의 이형접합성 보균자인) 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체의 치료 방법을 포함할 수 있고, 방법은 대상체 안으로 또는 대상체의 간세포 안으로: (a) *HSD17B13* 유전자 내의 DNA-결합 단백질 표적 서열에 결합하는 DNA-결합 단백질 (또는 그것을 암호화하는 핵산), 여기서 DNA-결합 단백질 표적 서열은 *HSD17B13* 유전자에 대한 개시 코돈을 포함하거나 개시 코돈의 약 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 또는 1,000개 뉴클레오타이드 내에 있거나 또는 SEQ ID NO: 20 내지 81로부터 선택되며; (b) 재조합 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는 재조합 *HSD17B13* 유전자를 포함하는 발현 벡터를 도입하는 단계를 포함한다. 발현 벡터는 게놈상으로 통합되지 않

는 것일 수 있다. 대안적으로, 재조합 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는재조합 *HSD17B13* 유전자 를 포함하는 표적화 벡터 (즉, 외인성 도너 서열)가 도입될 수 있다. DNA-결합 단백질은 대상체의 간세포에서 *HSD17B13* 유전자의 발현을 변경 (예컨대, 감소)시킬 수 있고, 발현 벡터는 대상체의 간세포에서 재조합 *HSD17B13* 유전자를 발현시킬 수 있다. 대안적으로, 게놈상으로 통합된, 재조합 *HSD17B13* 유전자는 대상체의 간세포에서 발현할 수 있다. 이러한 방법에 사용하기에 적합한 DNA-결합 단백질의 예는 본원의 다른 곳에서 개시된다. 이러한 DNA-결합 단백질 (예컨대, Cas9 단백질 및 가이드 RNA)은 전사 억제자 도메인에 작동가능하게 연결될 수 있다. 예를 들어, DNA-결합 단백질은 전사 억제자 도메인에 융합된 촉매적으로 비활성인 Cas9 단백질일 수 있다. 전사 억제자 도메인에 융합된 이러한 DNA-결합 단백질은, 예를 들어, 야생형 *HSD17B13* 유전자 또는 rs72613567 변종이 아닌 *HSD17B13* 유전자의 발현을 감소시키기 위하여 (예컨대, *HSD17B13* 전사물 또는 동형 A의 발현을 감소시키기 위하여) 사용될 수 있다. 적합한 가이드 RNA 및 가이드 RNA 표적 서열의 예는 본원의 다른 곳에서 개시된다. 단계 (b)는 대안적으로 HSD17B13 동형 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하는 핵산 (예컨대, DNA)을 포함하는 및/또는 *HSD17B13* 전사물 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 서열을 포함하는 발현 벡터 또는 표적화 벡터를 도입하는 단계를 포함할 수 있다. 마찬가지로, 단계 (b)는 대안적으로 HSD17B13 동형 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하는 및/또는 *HSD17B13* 전사물 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 상보하는 DNA (또는 이것의 부분)를 가지는 mRNA를 도입하는 단계를 포함할 수 있다. 마찬가지로, 단계 (b)는 대안적으로 HSD17B13 동형 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 서열을 포함하는 단백질을 도입하는 단계를 포함할 수 있다. 특정 방법에서, 전사물은 *HSD17B13* 전사물 D (예컨대, SEQ ID NO: 7)일 수 있거나, 동형은 HSD17B13 동형 D (예컨대, SEQ ID NO: 15)일 수 있다. 다른 특정 방법에서, HSD17B13 동형의 조합, HSD17B13 동형의 조합을 암호화하는 발현 벡터 또는 표적화 벡터, HSD17B13 동형의 조합을 암호화하는 mRNA가 도입될 수 있다 (예컨대, D, DC, DF, DG, DH, DCF, DCG, DCH, DFG, DFH, DGH, DCFG, DCFH, DCGH, DFGH, 또는 DCFGH).

[0420] 이러한 방법은 또한 *HSD17B13* rs72613567 변종의 보균자가 아닌 (또는 오직 *HSD17B13* rs72613567 변종의 이형접합성 보균자인) 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체의 치료 방법을 포함할 수 있고, 방법은 대상체 안으로 또는 대상체의 간세포 안으로: 선택적으로 *HSD17B13* 전사물 C, D, F, G, 및 H 중 하나 이상 (특히 D)에 존재하지 않는 *HSD17B13* 전사물 A, B, E, 및 F' 중 하나 이상 (특히 A)의 영역 내의 서열에 혼성화하는 안티센스 RNA, siRNA, 또는 shRNA를 도입하는 단계를 포함한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A) 내의 서열에 혼성화하고, 안티센스 RNA, siRNA, 또는 shRNA는 세포에서 *HSD17B13* 전사물 A의 발현을 감소시킬 수 있다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 7 (*HSD17B13* 전사물 D)에 존재하지 않는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)에 존재하는 서열에 혼성화한다. 선택적으로, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)의 엑손 7 내의 서열 또는 엑손 6-엑손 7 경계에 걸쳐 있는 서열에 혼성화한다. 예를 들어, 안티센스 RNA, siRNA, 또는 shRNA는 SEQ ID NO: 4 (*HSD17B13* 전사물 A)의 엑손 7의 영역 내의 서열 또는 엑손 6-엑손 7 경계에 걸쳐 있는 영역 내의 서열에 혼성화하고 대상체의 간세포에서 *HSD17B13* 전사물 A의 발현을 감소시킨다. 선택적으로, 이러한 방법은 추가로 재조합 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는 재조합 *HSD17B13* 유전자를 포함하는 발현 벡터를 대상체 안으로 도입하는 단계를 포함할 수 있다. 발현 벡터는 게놈상으로 통합되지 않는 것일 수 있다. 대안적으로, 재조합 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는 재조합 *HSD17B13* 유전자를 포함하는 표적화 벡터 (즉, 외인성 도너 서열)가 도입될 수 있다. 발현 벡터가 사용되는 방법에서, 발현 벡터는 대상체의 간세포에서 재조합 *HSD17B13* 유전자를 발현시킬 수 있다. 대안적으로, 재조합 *HSD17B13* 유전자가 게놈상으로 통합되는 방법에서, 재조합 *HSD17B13* 유전자는 대상체의 간세포에서 발현할 수 있다. 이러한 방법은 대안적으로 HSD17B13 동형 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하는 핵산 (예컨대, DNA)을 포함하는 및/또는 *HSD17B13* 전사물 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 서열을 포함하는 발현 벡터 또는 표적화 벡터를 도입하는 단계를 포함할 수 있다.

마찬가지로, 이러한 방법은 대안적으로 HSD17B13 동형 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하는 및/또는 *HSD17B13* 전사물 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 상보하는 DNA (또는 이것의 일부)를 가지는 mRNA를 도입하는 단계를 포함할 수 있다. 마찬가지로, 이러한 방법은 대안적으로 HSD17B13 동형 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 서열을 포함하는 단백질을 도입하는 단계를 포함할 수 있다. 특정 방법에서, 전사물은 *HSD17B13* 전사물 D (예컨대, SEQ ID NO: 7)일 수 있거나, 동형은 HSD17B13 동형 D (예컨대, SEQ ID NO: 15)일 수 있다. 다른 특정 방법에서, HSD17B13 동형의 조합, HSD17B13 동형의 조합을 암호화하는 발현 벡터 또는 표적화 벡터, 또는 HSD17B13 동형의 조합을 암호화하는 mRNA가 도입될 수 있다 (예컨대, D, DC, DF, DG, DH, DCF, DCG, DCH, DFG, DFH, DGH, DCFG, DCFH, DCGH, DFGH, 또는 DCFGH).

[0421] 다른 이러한 방법은 *HSD17B13* rs72613567 변종의 보균자가 아닌 (또는 오직 *HSD17B13* rs72613567 변종의 이형접합성 보균자인) 및 만성 간 질환을 가졌거나 발병하기 쉬운 대상체의 치료 방법을 포함할 수 있고, 방법은 대상체 안으로 또는 대상체의 간세포 안으로 발현 벡터를 도입하는 단계를 포함하며, 여기서 발현 벡터는 재조합 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는 재조합 *HSD17B13* 유전자를 포함하고, 발현 벡터는 대상체의 간세포에서 재조합 *HSD17B13* 유전자을 발현시킨다. 발현 벡터는 게놈상으로 통합되지 않은 것일 수 있다. 대안적으로, 재조합 *HSD17B13* 유전자가 SEQ ID NO: 1과 최적으로 정렬될 때 SEQ ID NO: 1의 위치 12665 및 12666에 상응하는 뉴클레오타이드 사이에 삽입된 티민을 포함하는 재조합 *HSD17B13* 유전자를 포함하는 표적화 벡터 (즉, 외인성 도너 서열)가 도입될 수 있다. 발현 벡터가 사용되는 방법에서, 발현 벡터는 대상체의 간세포에서 재조합 *HSD17B13* 유전자를 발현시킬 수 있다. 대안적으로, 재조합 *HSD17B13* 유전자가 게놈상으로 통합되는 방법에서, 재조합 *HSD17B13* 유전자는 대상체의 간세포에서 발현할 수 있다. 이러한 방법은 대안적으로 HSD17B13 동형 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하는 핵산 (예컨대, DNA)을 포함하는 및/또는 *HSD17B13* 전사물 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 서열을 포함하는 발현 벡터 또는 표적화 벡터를 도입하는 단계를 포함할 수 있다. 마찬가지로, 이러한 방법은 대안적으로 HSD17B13 동형 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 HSD17B13 단백질을 암호화하는 및/또는 *HSD17B13* 전사물 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 상보하는 DNA (또는 이것의 부분)을 가지는 mRNA을 도입하는 단계를 포함할 수 있다. 마찬가지로, 이러한 방법은 HSD17B13 동형 C, D, F, G, 또는 H 또는 이것들의 단편에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 동일한 서열을 포함하는 단백질을 도입하는 단계를 포함할 수 있다. 특정 방법에서, 전사물은 *HSD17B13* 전사물 D (예컨대, SEQ ID NO: 7)일 수 있거나, 또는 동형은 HSD17B13 동형 D (예컨대, SEQ ID NO: 15)일 수 있다. 다른 특정 방법에서, HSD17B13 동형의 조합, 또는 HSD17B13 동형의 조합을 암호화하는 발현 벡터 또는 표적화 벡터, 또는 HSD17B13 동형의 조합을 암호화하는 mRNA가 도입될 수 있다 (예컨대, D, DC, DF, DG, DH, DCF, DCG, DCH, DFG, DFH, DGH, DCFG, DCFH, DCGH, DFGH, 또는 DCFGH).

[0422] 상기 방법들 중 임의의 방법에 사용하기에 적합한 발현 벡터 및 재조합 *HSD17B13* 유전자는 본원의 다른 곳에서 개시된다. 예를 들어, 재조합 *HSD17B13* 유전자는 전체 rs72613567 변종 유전자일 수 있거나 유전자의 하나 이상의 비필수 분절이 상응하는 야생형 *HSD17B13* 유전자와 관련하여 삭제되어 있는 *HSD17B13* 미니유전자일 수 있다. 예로서, 삭제된 분절은 하나 이상의 인트론성 서열을 포함할 수 있고, 미니유전자는 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2의 인트론 6에 상응하는 인트론을 포함할 수 있다. 전체 rs72613567 변종 유전자의 예는 SEQ ID NO: 2와 최적으로 정렬될 때 SEQ ID NO: 2에 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 것이다.

[0423] 일부 이러한 방법은 만성 간 질환을 가졌거나 발병하기 쉬운 대상체에서 세포 (예컨대, 간세포)를 변형시키는 방법을 포함한다. 이러한 방법에서, 뉴클레아제 제제 및/또는 외인성 도너 서열 및/또는 재조합 발현 벡터는 치료되는 만성 간 질환의 적어도 하나의 신호 또는 증상의 시작을 지연시키거나, 중증도를 감소시키거나, 추가의 악화를 억제하거나, 및/또는 개선시키는 투여량, 투여 경로 및 투여 빈도를 의미하는 효과적인 요법에서 투여를 통해 세포 안으로 도입될 수 있다. 용어 "증상"은 대상체에 의해 인지되는 바 질환의 주관적인 증거를 나타내며, "신호"는 의사에 의해 관찰되는 바 질환의 객관적인 증거를 나타낸다. 만약 대상체가 이미 질환에 걸

렸다면, 요법은 치료적으로 효과적인 요법으로서 언급될 수 있다. 만약 대상체가 일반 집단에 비해 질환에 걸릴 위험이 상승되어 있지만 아직 증상을 나타내지 않는다면, 요법은 예방적으로 효과적인 요법으로서 언급될 수 있다. 일부 경우에, 치료적 또는 예방적 효능은 동일한 대상체에서 과거 대조군 또는 지나간 경험에 비해 개별 환자에서 관찰될 수 있다. 다른 경우에, 미치료 대상체의 대조군 집단에 비해 치료된 대상체 집단에서 전임상 또는 임상 실험으로 증명될 수 있다.

[0424] 전달은 본원의 다른 곳에서 개시되는 것과 같이, 임의의 적합한 방법일 수 있다. 예를 들어, 뉴클레아제 제제 또는 외인성 도너 서열 또는 재조합 발현 벡터는 벡터 전달, 바이러스 전달, 입자-매개된 전달, 나노입자-매개된 전달, 리포좀-매개된 전달, 엑소좀-매개된 전달, 지질-매개된 전달, 지질-나노입자-매개된 전달, 세포 투과-펩타이드-매개된 전달, 또는 임플란트 장치-매개된 전달에 의해 전달될 수 있다. 일부 특정 예로는 유체역학적 전달, 바이러스-매개된 전달, 및 지질-나노입자-매개된 전달을 들 수 있다.

[0425] 투여는, 예를 들어, 비경구, 정맥내, 경구, 피하, 동맥내, 두개내, 척수강내, 복강내, 국소적, 비내, 또는 근육내를 포함한 임의의 적합한 경로에 의할 수 있다. 예를 들어, 단백질 대체 요법에 대해 종종 사용되는 특정 예는 정맥내 주입이다. 투여 빈도 및 투여량 수는 다른 인자들 중에서도 뉴클레아제 제제 또는 외인성 도너 서열 또는 재조합 발현 벡터의 반감기, 대상체의 상태, 및 투여 경로에 따라 좌우될 수 있다. 투여용 제약학적 조성물은 바람직하게는 멸균된 것이고 실질적으로 등장성이며 GMP 조건 하에서 제조된다. 제약학적 조성물은 단위 투여 형태 (즉, 단일 투여용 투여 단위)로 제공될 수 있다. 제약학적 조성물은 하나 이상의 생리적으로 및 제약학적으로 허용되는 담체, 희석제, 부형제 또는 보조제를 사용하여 제제화될 수 있다. 제제는 선택된 투여 경로에 좌우될 수 있다. 용어 "제약학적으로 허용되는"은 담체, 희석제, 부형제 또는 보조제가 제제의 다른 성분들과 부합할 수 있고 실질적으로 그것의 수령체에 유해하지 않은 것을 의미한다.

[0426] 다른 이러한 방법은 만성 간 질환을 가졌거나 발병하기 쉬운 대상체로부터의 세포에서의 생체외 방법을 포함한다. 표적화된 유전자 변형된 세포는 그런 후 대상체로 다시 이식될 수 있다.

[0427] 본원에 개시된 임의의 치료적 또는 예방적 방법은 추가로 만성 간 질환의 보다 임상적으로 진전된 단계들로의 진행 (예컨대, 단순 지방증으로부터 만성 간 질환의 보다 임상적으로 진전된 단계로의 진행, 또는 단순 지방증으로부터 지방간염, 섬유증, 간경변, 및 간세포 암종 중 하나 이상으로의 진행)과 관련된 하나 이상의 증상을 예방 또는 완화시키기 위하여 맞춰진 치료제를 투여하는 단계를 포함할 수 있다. 예를 들어, 이러한 치료는 염증을 예방 또는 감소시키는 것 또는 섬유증을 예방 또는 감소시키는 것에 집중될 수 있다. 개발 중에 있는 이러한 치료제의 예가 하기에 제공된다.

[0428]

약물 (화합물)	단계	유형	유전자 표적	주의사항
OCA - 오베티콜릭산 (Intercept)	제3기	작용물질	NR1H4 (FXR)	개선된 NAS, Iib 단계에서 섬유증을 반전시킴
GS-9674 (Gilead)	제1기			
심투주맙 (Gilead)	제2기	억제제	LOXL2	섬유증을 반전시키는 잠재력 (NASH/PSC)
GS-4997 (Gilead)	제2기	억제제	MAP3K5	산화성 스트레스를 감소시킴
NDI-010976 (Gilead)	제1기	억제제	ACACA ACACB	지방 생성을 예방함
GFT505 / 엘라피브라노 (Genfit)	제3기	작용물질	PPARA PPARD	지방산을 파괴하고, 지방을 차단함 & 글루코오스 생성, dec 염증
아람콜 (Galmed)	제3기	억제제	SCD (ABCA1)	지방산-담즙산 콘쥬게이트; 간 지방 대사를 촉진함
세니크리비록 (Tobira)	제2b기	억제제	CCR2 CCR5	케모카인 수용체가 염증 및 섬유증에 연구됨
GR-MD-02 (Galectin Therapeutics)	제2기	억제제	LGALS3	갈렉틴-3은 섬유증에서 상향조절됨
TD139 (Galecto Biotech)	제1기			
SHP626 (Shire)	제1기	억제제	SLC10A2	담즙산 재순환을 간섭함
PXS4728A - (Boehringer Ingelheim)	제1기	억제제	AOC3	항염증제
RP103 - 중타르타르산 시스테아민(Raptor)	제2기	고갈제	CTNS	시스테인 고갈; 잠재적 항산화제

[0429] 상기 또는 하기에서 인용된 모든 특허 출원, 웹사이트, 기타 공보물, 등록 번호 등은, 각각의 개별적인 항목이 구체적으로 및 개별적으로 참조에 의해 그렇게 포함된 것으로 표시된 것과 동일한 정도로 모든 목적에 대해 그 전문이 참조로 포함된다. 만약 상이한 버전의 서열이 상이한 시점에서의 등록 번호와 관련된다면, 본 출원의 유

효한 출원일의 등록 번호와 관련된 버전이 의미가 있다. 유효한 출원일은 실제 출원일보다 앞선 날이거나 해당되는 경우 등록 번호를 언급한 우선권 출원의 출원일을 의미한다. 마찬가지로, 만약 상이한 버전의 공보물, 웹사이트 등이 상이한 시점에 공개된다면, 다르게 표시되지 않는 한 출원의 유효한 출원일에 가장 가깝게 공개된 버전을 의미한다. 발명의 임의의 특징, 단계, 요소, 구체예, 또는 측면은 구체적으로 다르게 표시되지 않는 한 임의의 다른 것과 조합하여 사용될 수 있다. 비록 본 발명이 명료성 및 이해를 목적으로 어느 정도 상세하게 예시 및 실례에 의해 기술되었지만, 특정 변화 및 변형이 첨부된 청구범위의 범주 내에서 실시될 수 있는 것이 명백할 것이다.

[0430] **서열의 간단한 설명**

[0431] 첨부되는 서열 목록에서 열거된 뉴클레오타이드 및 아미노산 서열은 뉴클레오타이드 염기에 대한 표준 문자 약어, 및 아미노산에 대한 3-문자 암호를 사용하여 제시된다. 뉴클레오타이드 서열은 서열의 5' 단부에서 시작하여 3' 단부를 향해 진행하는 (즉, 각 라인에서 좌측으로부터 우측으로의) 표준 관례를 따른다. 각 뉴클레오타이드 서열의 오직 한 가닥만 제시되지만, 상보하는 가닥은 표시된 가닥의 임의의 언급에 의해 포함되는 것으로 인지된다. 아미노산 서열은 서열의 아미노 말단에서 시작하여 카르복시 말단을 향해 진행하는 (즉, 각 라인에서 좌측으로부터 우측으로의) 표준 관례를 따른다.

[0432]

SEQ ID NO	유형	설명
1	DNA	HSD17B13 야생형 게놈 서열 (인간 게놈 어셈블리 GRCh38) ***야생형 HSD17B13 유전자를 가진 대상체에서 보다 우세한 전사물:*** **전사물 A** - 엑손 1 = 1-275 - 엑손 2 = 4471-4578 - 엑손 3 = 5684-5815 - 엑손 4 = 7308-7414 - 엑손 5 = 8947-9084 - 엑손 6v1 = 12548-12664 - 엑손 7 = 17599-19118 **전사물 B** - 엑손 1 = 1-275 - 엑손 2 = 건너뜀 - 엑손 3 = 5684-5815 - 엑손 4 = 7308-7414 - 엑손 5 = 8947-9084 - 엑손 6v1 = 12548-12664 - 엑손 7 = 17599-19118 **전사물 E** - 엑손 1 = 1-275 - 엑손 2 = 4471-4578 - 엑손 3 = 5684-5815 - 엑손 3' = 6210-6281 - 엑손 4 = 7308-7414 - 엑손 5 = 8947-9084 - 엑손 6v1 = 12548-12664 - 엑손 7 = 17599-19118 **전사물 F'** - 엑손 1 = 1-275 - 엑손 2 = 4471-4578 - 엑손 3 = 5684-5815 - 엑손 4 = 7308-7414 - 엑손 5 = 8947-9084 - 엑손 6v3 = 12548-13501 (엑손 6으로부터 인트론 6 안으로의 초과 번역 = 12665-13501) - 엑손 7 = 건너뜀

2	DNA	HSD17B13 게놈 유전자 변종 (인간 게놈 어셈블리 GRCh38; rs72613567-chr4에서의 T 삽입: 87310241-87310240): 위치 12666에서의 T 삽입 *rs72613567 HSD17B13 유전자 변종을 가진 대상체에서 보다 우세한 전사물:* **전사물 C** - 엑손 1 = 1-275 - 엑손 2 = 4471-4578 - 엑손 3 = 5684-5815 - 엑손 4 = 7308-7414 - 엑손 5 = 8947-9084 - 엑손 6 = 건너뜀 - 엑손 7 = 17600-19119 **전사물 D** - 엑손 1 = 1-275 - 엑손 2 = 4471-4578 - 엑손 3 = 5684-5815 - 엑손 4 = 7308-7414 - 엑손 5 = 8947-9084 - 엑손 6v2 = 12548-12665 (3' 단부에서 추가 잔기 12665를 포함함) - 엑손 7 = 17600-19119 **전사물 F** - 엑손 1 = 1-275 - 엑손 2 = 4471-4578 - 엑손 3 = 5684-5815 - 엑손 4 = 7308-7414 - 엑손 5 = 8947-9084 - 엑손 6v3 = 12548-13502 (엑손 6으로부터 인트론 6 안으로의 초과 번역 = 12665-13502) - 엑손 7 = 건너뜀 **전사물 G** - 엑손 1 = 1-275 - 엑손 2 = 건너뜀 - 엑손 3 = 5684-5815 - 엑손 4 = 7308-7414 - 엑손 5 = 8947-9084 - 엑손 6v2 = 12548-12665 (3' 단부에서 추가 잔기 12665를 포함함) - 엑손 7 = 17600-19119 **전사물 H** - 엑손 1 = 1-275 - 엑손 2 = 4471-4578 - 엑손 3 = 5684-5815 - 엑손 3' = 6210-6281 - 엑손 4 = 7308-7414 - 엑손 5 = 8947-9084 - 엑손 6v2 = 12548-12665 (3' 단부에서 추가 잔기 12665를 포함함) - 엑손 7 = 17600-19119
3	DNA	내인성 HSD17B13 프로모터 (전사 시작 부위 (TSS)에 비해 -499 내지 100)
4	DNA	HSD17B13 전사물 A cDNA
5	DNA	HSD17B13 전사물 B cDNA
6	DNA	HSD17B13 전사물 C cDNA
7	DNA	HSD17B13 전사물 D cDNA
8	DNA	HSD17B13 전사물 E cDNA

[0433]

9	DNA	HSD17B13 전사물 F cDNA
10	DNA	HSD17B13 전사물 G cDNA
11	DNA	HSD17B13 전사물 H cDNA
12	단백질	HSD17B13 단백질 동형 A
13	단백질	HSD17B13 단백질 동형 B
14	단백질	HSD17B13 단백질 동형 C
15	단백질	HSD17B13 단백질 동형 D
16	단백질	HSD17B13 단백질 동형 E
17	단백질	HSD17B13 단백질 동형 F
18	단백질	HSD17B13 단백질 동형 G
19	단백질	HSD17B13 단백질 동형 H
20-41	DNA	인간 *HSD17B13* TSS 가이드 RNA 표적 서열
42-81	DNA	Other 인간 *HSD17B13* 5' 가이드 RNA 표적 서열
82-225	DNA	인간 *HSD17B13* 3' 가이드 RNA 표적 서열
226-239	DNA	rs72613567변이에 가까운 인간 *HSD17B13* 가이드 RNA 표적 서열
240	단백질	인간 HSD17B13 단백질 Q7Z5P4-1
241	단백질	인간 HSD17B13 단백질 Q7Z5P4-2
242	단백질	인간 HSD17B13 단백질 NP_835236.2
243	단백질	인간 HSD17B13 단백질 NP_001129702.1
244	DNA	인간 HSD17B13 cDNA NM_178135.4
245	DNA	인간 HSD17B13 cDNA NM_001136230.2
246	DNA	HSD17B13 전사물 F'
247	단백질	HSD17B13 단백질 동형 F'
248-250	DNA	가이드 RNA 표적 서열 플러스 PAM
251	DNA	PST516 프라이머
252	DNA	PST517 프라이머
253	DNA	DE002 프라이머
254	DNA	HSD17B13 프라이머 1
255	DNA	HSD17B13 프라이머 2
256-258	RNA	가이드 RNA 스캐폴드 v2-v4
259-263	DNA	마우스 5' 가이드 RNA 표적 서열
264-268	DNA	마우스 엑손 6/7 가이드 RNA 표적 서열
269	DNA	마우스 *Hsd17b13* 유전자좌
270-489	RNA	인간 *HSD17B13* crRNA
490-499	RNA	마우스 *Hsd17b13* crRNA
500-719	RNA	인간 *HSD17B13* sgRNA v1
720-729	RNA	마우스 *Hsd17b13* sgRNA v1
730-949	RNA	인간 *HSD17B13* sgRNA v2
950-959	RNA	마우스 *Hsd17b13* sgRNA v2
960-1179	RNA	인간 *HSD17B13* sgRNA v3
1180-1189	RNA	마우스 *Hsd17b13* sgRNA v3
1190-1409	RNA	인간 *HSD17B13* sgRNA v4
1410-1419	RNA	마우스 *Hsd17b13* sgRNA v4
1420	RNA	가이드 RNA 스캐폴드 v1
1421	RNA	crRNA 꼬리
1422	RNA	tracrRNA
1423-1642	RNA	인간 *HSD17B13* 가이드 RNA DNA-표적화 분절

[0434]

[0435]

1643-1652	RNA	마우스 *Hsd17b13* 가이드 RNA DNA-표적화 분절

[0436] **실시예**

[0437] **실시예 1. 변종 17베타-하이드록시스테로이드 탈수소효소 13은 만성 간 질환에 대해 보호한다.**

[0438] 만성 간 질환 및 간경변은 미국에서 이병률 및 사망률의 선두 요인이다 (Kochanek et al. (2016) *Natl Vital Stat Rep* 65:1-122, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 미국에서 간경변의 가장 공통적인 병인은 알코올성 간 질환, 만성 C형 간염, 및 비알코올성 지방간 질환 (NAFLD)이며, 이것들은 함께 간 이식을 기다리고 있는 환자들의 약 80%를 차지한다 (Wong et al. (2015) *Gastroenterology* 148:547-555, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 특히, 미국에서 NAFLD의 추정된 유병률은 19 내지 46 퍼센트이고 (Browning et al. (2004) *Hepatology* 40:1387-1395; Lazo et al. (2013) *Am J Epidemiol* 178:38-45; 및 Williams et al. (2011) *Gastroenterology* 140:124-131, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 비만의 증가 속도와 함께 시간이 지남에 따라 상승하는 가능성이 있다 (Younossi et al. (2011) *Clin Gastroenterol Hepatol* 9:524-530 e1; quiz e60 (2011), 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 현재까지, NAFLD 진행 및 결과에서 개인간 변이에 대해서는 여전히 많은 부분이 불확실하다; 기저

의 유전적 인자들에 대한 지식은 위험 계층화를 개선시키고 신규한 치료 전략에 대한 기초를 제공할 수 있었을 것이다. 본원에서, 본 발명자들은 *HSD17B13*의 스플라이스 변종 (하이드록시스테로이드-17-베타 탈수소효소 13을 암호화함)의 보균자들이 알코올성 및 비알코올성 간 질환의 위험을 감소시켰고, NAFLD 진행의 위험을 감소시켰음을 보여준다. DiscovEHR 연구에서 46,544명의 유럽계 참가자들로부터의 전자 의무 기록과 연계된 전체 엑솜 서열 데이터의 연합 연구는 감소된 알라닌 아미노전이효소 및 아스파테이트 아미노전이효소 수준과 관련된 *HSD17B13*에서의 스플라이스 변종 (rs72613567)의 확인으로 이어졌다; 이런 발견들은 12,528명의 개인을 포함한 3개의 별도 집단에서 반복되었다. 개발 집단에서, 변종 *HSD17B13*은 알코올성 및 비알코올성 간 질환, 간경변, 및 간세포 암종의 감소된 위험과 관련이 있었다. 베리아트릭 수술 집단에서, 변종은 지방증을 가진 개인들에서 조직병리학적 지방간염의 감소된 위험과 관련이 있었다. 베리아트릭 수술 집단으로부터의 인간 간 샘플의 RNA 서열분석은 스플라이스 변종의 동형접합성 보균자들이 절단된 *HSD17B13* 동형을 암호화하는 신규한 전사물을 우세하게 발현하는 것을 나타냈다. 이런 발견은 간 질환 진행을 촉진하는데 있어 *HSD17B13*의 역할, 및 지방간염 및 간경변에 대한 치료 표적으로서의 그것의 잠재력을 새롭게 조명하였다.

[0439] 이전의 전 게놈 관련 분석 (genome wide association studies, GWAS)으로 만성 간 질환과 관련된 제한된 수의 유전자 및 변종이 확인되었다. 오늘날까지 가장 강력하게 입증된 유전자 연관성 연구는 초기에 비알코올성 지방간 질환 (NAFLD)의 증가된 위험과 관련이 있는 것으로 발견되고 (Romeo et al. (2008) *Nat. Genet.* 40:1461-1465 및 Speliotes et al. (2011) *PLoS Genet.* 7:e1001324, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 계속해서 질환의 중증도 (Rotman et al. (2010) *Hepatology* 52:894-903 및 Sookoian et al. (2009) *J. Lipid Res.* 50:2111-2116, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨) 및 진행 (Trepo et al. (2016) *J. Hepatol.* doi:10.1016/j.jhep.2016.03.011, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨)과 관련이 있는 것으로 밝혀진, 3개의 유전자를 함유한 파타틴-유사 포스포리파아제 도메인의 공통 미스센스 변종에 대한 것이다. 경막 6 수퍼패밀리 구성원 2 (*TM6SF2*) 유전자의 변이는 또한 NAFLD에 대한 증가된 위험을 부여하는 것으로 나타났다 (Kozlitina et al. (2014) *Nat. Genet.* 46:352-356; Liu et al. (2014) *Nat. Commun.* 5:4309; 및 Sookoian et al. (2015) *Hepatology* 61:515-525, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 이 두 단백질의 정상 기능은 잘 알려져 있지 않지만, 둘 다 간세포 지질 대사에 관여하는 것으로 제안되어 왔다. *PNPLA3* 및 *TM6SF2*에서의 변종이 어떻게 간 질환의 증가된 위험에 기여하는지는 아직 설명되지 않았다. GWAS는 또한 임상적으로 빈번하게 측정되는 간세포 손상 및 간 지방 축적의 정량적 마커인 혈청 알라닌 아미노전이효소 (ALT) 및 아스파테이트 아미노전이효소 (AST) (Chambers et al. (2011) *Nat. Genet.* 43:131-1138 및 Yuan et al. (2008) *Am. J. Hum. Genet.* 83:520-528, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨)와 관련이 있는 몇몇 유전적 인자를 확인하였다. 오늘날까지, 만성 간 질환에 대한 보호성 유전자 변종은 기술된 것이 없다. 다른 환경에서의 보호성 유전자 변종의 발견, 예컨대 심혈관 질환의 위험을 감소시키는 *PCSK9*에서의 기능-상실 변종의 발견은 새로운 부류의 치료제의 개발에 대한 촉매가 되었다.

[0440] 리제너론 제네틱스 센터 (Regeneron Genetics Center)와 Geisinger Health System (GHS) 사이의 DiscovEHR 공동작업은 유전적 발견 및 정확한 의학을 가능하게 하기 위해 엑솜 서열분석을 누락된 확인된 전자 의무 기록 (EHR) 데이터에 결부시킨다 (Dewey et al. (2016) *Science* 354(6319) doi:10.1126/science.aaf6814, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). DiscovEHR 집단은 간 생검 시편을 제공하는 베리아트릭 수술 환자들을 포함하는, GHS 통합된 건강돌봄 시스템 전체에 걸쳐 일차 및 특수 의료 돌봄 집단으로부터 모집된 환자들로 구성된다 (Gorden et al. (2013) *Hum Hered* 75:34-43, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 이 연구에서, 본 발명자들은 DiscovEHR 집단으로부터의 49,188명의 유럽 혈통의 개인에서 만성 간 질환 및 간경변과 관련된 양적 형질, 질환 진단, 조직병리학적 표현형에 대한 엑솜 서열 변이의 기여를 평가하기 위하여 포괄적인 기능성 유전체학 접근법을 수행하였고, 9,883명의 유럽 혈통의 개인의 전체 엑솜 서열분석을 사용하여 추수 연구를 하였다.

[0441] EHR-유래 표현형에 결부된 전체 엑솜 서열 데이터를 사용하여, 본 발명자들은 먼저 DiscovEHR 집단으로부터 46,544명의 유럽 혈통의 개인 ("GHS 개발 집단")에서 혈청 ALT 및 AST 측정치의 관련 연구를 수행하였다. 이 집단의 임상적 특징들을 **표 1A**에 기술한다. EHR-증명된 아미노전이효소 측정치를 나타낸 개인은 41,908명이었다 (ALT 및 AST 측정치 둘 다 가진 40,561명의 개인을 포함함). 본 발명자들은 $\log_{10}$-변환된 중강 ALT 및 AST 수준

(성별, 연령, 연령^2, 체질량 지수 (BMI) 및 혈통의 제1 4대 주요 구성요소에 대해 조정됨)과 마이너 대립유전자 빈도가 0.1%보다 큰 502,219개의 이중대립유전자 단일 변종 사이의 관련여부를 검출하기 위하여 선형 혼합 모델 (Yang et al. (2011) *Am J Hum Genet* 88:76-82, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨)을 사용하

였다. $P < 1.0x10^{-7}$의 엑솜-전체 유의성 한계값을 사용하여, 본 발명자들은 ALT 및 AST 둘 다와 관련된 7개의 유전자에서 8개의 변종을 포함하여, ALT 또는 AST와 유의하게 관련된 19개의 유전자에서 35개의 변종을 확인하였다 (**도 1** 및 **표 2**).

【표 1A】

개발 및 복제 집단으로부터 서열분석된 유럽계 개인들의 인구통계학적 및 임상적 특징.

특징	개발 집단 (N = 46,544명)	베리아트릭 수술 집단 (N = 2,644명)	델러스 심장 연구 (N = 1,357명)	Penn 약물 바이오뱅크 (N = 8,526명)
연령 (나이) - 중앙값 (IQR)	62.9 (49.6 - 73.8)	52.9 (44.1 - 61.2)	46.0 (38.0 - 54.0)	68.0 (60.0 - 76.0)
여성 - 수 (%)	26,875 (57.7)	2,119 (80.1)	724 (53.4)	3,242 (38.0)
체질량 지수 - 중앙값 (IQR)	29.9 (35.4 - 44.8)	47.4 (42.0 - 53.7)	28 (25-32)	30 (25-32)
아미노전이효소 수준 (U/L) - 중앙값 (IQR)				
알라닌 아미노전이효소 (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 아미노전이효소 (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 (ICD-9 코드에 의한) - 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)	-	-
간 질환 없음	30,628 (65.8)	1 (0.04)	-	-

[0442]

【표 1B】

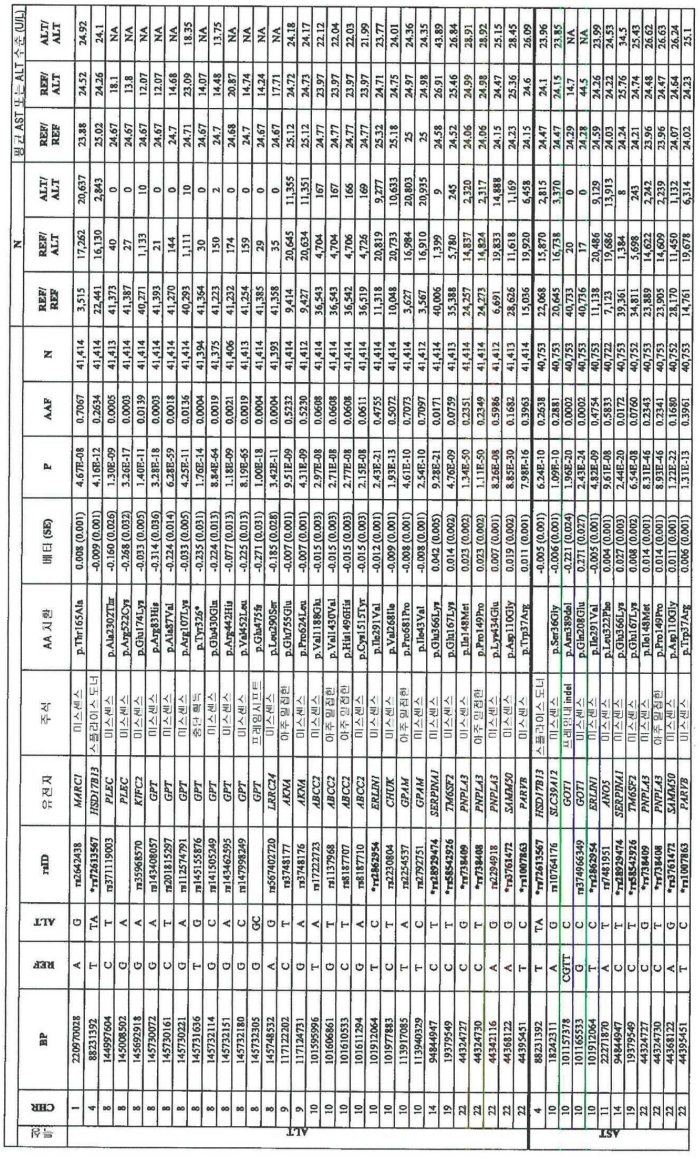
델러스 간 및 소아 간 연구로부터의 유전자형분류된 다민족 케이스 및 대조군의 인구통계학적 및 임상적 특징.

특징	델러스 간 연구 케이스 (N = 517명)	델러스 간 연구 대조군 (N = 4,279명)	델러스 소아 간 연구 케이스 (N = 203명)	델러스 소아 간 연구 대조군 (N = 244명)
연령 (나이) - 중앙값 (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 - 중앙값 (IQR)	30 (27 - 35)	30 (26 - 35)	30 (27 - 34)	31 (28 - 35)
자가-기록 민족				
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 (ICD-9 암호에 의한) - N (%)				
알코올성 간 질환	223 (43)	-	-	-
알코올성 간경변	215 (42)	-	-	-
비알코올성, 비바이러스성 간 질환	212 (20)	-	-	-
비알코올성 간경변	100 (19)	-	-	-
간세포 암종	44 (9)	-	-	-
간 질환 없음	-	4,279 (100)	-	-244 (100)

[0443]

표 2

개발 집단에서 P < 1.0×10^{-7}에서의 혈청 아미노전이효소 수준과 관련된 단일 뉴클레오타이드 변종.



[0444]

[0445] *는 전 엑솜에 걸쳐 ALT 및 AST 둘 다와 유의하게 관련된 변종을 나타낸다.

[0446] 약어: AAF, 대체 대립유전자 빈도; Alt, 대체 대입유전자; ALT, 알라닌 아미노전이효소; AST, 아스파테이트 아미노전이효소; Ref, 참조 대립유전자; SE, 표준 오차.

[0447] 이런 관련성을 복제하기 위하여, 본 발명자들은 3개의 별도의 유럽계 집단: DiscovEHR로부터의 2,644명의 베리아트릭 수술 환자 ("GHS 베리아트릭 수술 집단"), 댈러스 심장 연구(Dallas Heart Study)로부터의 1,357명의 개인, 및 Penn 약물 바이오뱅크(Medicine Biobank)로부터의 8,526명의 개인에서 전체 엑솜 서열분석을 통해 확인된 35개의 AST- 또는 ALT-관련 변종을 분석하였다 (**표 1A**). 복제 집단의 메타-분석에서, 9개의 유전자 중 13개의 변종이 ALT 또는 AST와 유의하게 관련이 있었다 ($P<1.43 \times 10^{-3}$의 본페로니 유의성 한계값) (**표 3**). 이것들은 이전에 보고된 간 질환-관련 유전자 및 변종, 예컨대 *PNPLA3* p.Ile148Met (Romeo et al. (2008) *Nat Genet* 40:1461-1465, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), *TM6SF2* p.Glu167Lys (Kozlitina et al.

(2014) *Nat Genet* 46:352-356, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 및 *SERPINA1* p.Glu366Lys (알파-1-항-트립신 결핍과 관련된 Z 대립유전자) (Brantly et al. (1988) *Am J Med* 84:13-31, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), *SAMM50*, 및 *ERLIN1*을 포함하였다. *SERPINA1*은 알파-1-항트립신을 암호화하며, 이것의 기능성 결핍은 유전적인 간 질환을 유발하는 것으로 알려져 있고; *SAMM50*과의 회합은 *PNPLA3*에서의 변이와의 결합 불균형을 통해 매개될 수 있고, *ERLIN1*은 간 지방 침적에 연루될 수 있다. 각각 ALT 및 AST를 암호화하는 유전자인 *GPT* 및 *GOT1*에서의 여러 변종은 ALT 또는 AST와 유의하게 관련이 있었지만 이전에 간 질환과 관련이 있는 것으로 보고된 적은 없었다. *SLC39A12*는 이전에 아미노전이효소 또는 간 질환과 결부된 적이 없었다. 메타-분석은 또한 본 발명자들의 개발 집단에서 감소된 수준의 ALT (베타 (SE) -0.009 (0.001); $P=4.16x10^{-12}$) 및 AST (베타 (SE) -0.005 (0.001); $P=6.24x10^{-10}$)와 17-베타 하이드록시스테로이드 탈수소효소 패밀리의 특성화되지 않은 구성원인 하이드록시스테로이드 17-베타 탈수소효소 13을 암호화하는 유전자인 *HSD17B13*에서의 스플라이스 변종 사이의 신규한 관련성을 복제하였다. 이 변종, rs72613567은 도너 스플라이스 부위 (TA 대립유전자)에 인접한 A 뉴클레오타이드의 삽입에 상응한다. 이런 관련성에 대한 반복 메타-분석 P-값은 $3.85x10^{-5}$ 및 $9.38x10^{-5}$였고, 연합 메타-분석 P값은 ALT 및 AST 각각에 대해 $1.17x10^{-15}$ 및 $6.82x10^{-13}$이었다 (**표 3**). 선행 GWAS는 4q22에서의 근접한 유전자좌 (rs6834314)를 ALT 수준과 관련이 있는 것으로서 확인하였다 (Chambers et al. (2011) *Nat Genet* 43:1131-1138, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨); 본 발명자들이 알기로, 이전에 rs72613567과의 임의의 관련성을 기술한 연구는 없었다.

표 3

3개의 별도의 유럽계 집단에서 개발 집단으로부터 35개의 엑솜-전체 유의한 단일 뉴클레오타이드 변종의 복제 및 연합 메타-분석.

[0448]

특징	Chr	BP	Ref	Alt	RSID	유전자	Anno	AA 치환	GHS 개별 검단		GHS 베리아트릭 수술 검단		복제 검단						** 복제 메타-분석 (N=3명)		*** 연합 메타-분석 (N = 4명)	
													델러스 성장 연구			U. Penn						
									Beta (SE)	P	Beta (SE)	P	Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	Beta (SE)	P
ALT	1	220970028	A	G	rs2642438	*MARC1*	mis	p.Thr165Ala	0.008 (0.001)	4.67 E-08	0.005 (0.005)	3.10 E-01	0.011 (0.008)	1.76 E-01	1357	0.007 (0.004)	1.02 E-01	6158	0.007 (0.003)	2.31E-02	0.008 (0.001)	3.38E-09
	4	88231392	T	TA	rs72613567	*HSD17B13*	spl		-0.009 (0.001)	4.16 E-12	-0.010 (0.005)	5.57 E-02	-0.016 (0.008)	6.60 E-02	1357	-0.013 (0.004)	1.33 E-03	6158	-0.013 (0.003)	*3.85E-05	-0.010 (0.001)	1.17E-15
	8	144997604	C	T	rs3711119003	*PLEC*	mis	p.Ala2302Thr	-0.160 (0.026)	1.30 E-09	-0.492 (0.165)	2.84 E-03	NA (NA)	NA	NA	-0.051 (0.072)	4.79 E-01	6158	-0.121 (0.066)	6.56E-02	-0.155 (0.025)	2.68E-10
	8	145008502	G	A		*PLEC*	mis	p.Arg522Cys	-0.268 (0.032)	3.26 E-17	-0.161 (0.165)	3.29 E-01	NA (NA)	NA	NA	-0.247 (0.143)	8.48 E-02	6158	-0.210 (0.108)	5.23E-02	-0.264 (0.031)	5.54E-18
	8	145692918	G	A	rs35968570	*KIFC2*	mis	p.Glu174Lys	-0.033 (0.005)	1.40 E-11	-0.009 (0.020)	6.48 E-01	0.032 (0.036)	3.76 E-01	1356	-0.053 (0.018)	3.72 E-03	6158	-0.025 (0.013)	4.69E-02	-0.032 (0.005)	2.25E-12
	8	145730072	G	A	rs143408057	*GPT*	mis	p.Arg83His	-0.314 (0.036)	3.28 E-18	-0.189 (0.165)	2.50 E-01	NA (NA)	NA	NA	-0.298(0.101)	3.26 E-03	6158	-0.268 (0.086)	1.88E-03	-0.308 (0.033)	2.79E-20
	8	145730161	C	T	rs201815297	*GPT*	mis	p.Ala87Val	-0.224 (0.014)	6.28 E-59	-0.341 (0.074)	3.64 E-06	NA (NA)	NA	NA	-0.143 (0.054)	8.50 E-03	6158	-0.213 (0.044)	*1.14E-06	-0.223 (0.013)	4.49E-64
	8	145730221	G	A	rs112574791	*GPT*	mis	p.Arg107Lys	-0.033 (0.005)	4.25 E-11	-0.009 (0.020)	6.45 E-01	0.028 (0.036)	4.37 E-01	1357	-0.060 (0.018)	5.60 E-04	6158	-0.031 (0.013)	1.36E-02	-0.033 (0.005)	1.92E-12
	8	145731636	T	G	rs145155876	*GPT*	stop	p.Tyr326*	-0.235 (0.031)	1.76 E-14	-0.314 (0.165)	5.71 E-02	-0.317 (0.140)	2.35 E-02	1356	-0.148 (0.145)	3.04 E-01	6157	-0.256 (0.086)	2.79E-03	-0.237 (0.029)	1.94E-16
	8	145732114	G	C	rs141505249	*GPT*	mis	p.Glu430Gln	-0.224 (0.013)	8.84 E-64	-0.273 (0.048)	9.83 E-09	-0.240 (0.075)	1.36 E-03	1357	-0.197 (0.041)	1.31 E-06	6157	-0.231 (0.029)	*7.24E-16	-0.225 (0.012)	6.06E-78
	8	145732151	G	A	rs143462595	*GPT*	mis	p.Arg442His	-0.077 (0.013)	1.18 E-09	-0.115 (0.058)	4.82 E-02	-0.106 (0.099)	2.86 E-01	1356	-0.049 (0.041)	2.27 E-01	6157	-0.074 (0.032)	1.88E-02	-0.076 (0.012)	7.03E-11

[0449]

염기	Chr	BP	Ref	Alt	RSID	유전자	Anno	AA 치환	GHS 개발 집단			GHS 베리아트릭 수술집단			복제 집단						** 복제 메타-분석 (N = 3명)		*** 연합 메타-분석 (N = 4명)	
															멜라스 심장 연구			U. Penn					
									Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	Beta (SE)	P
	8	145732180	G	C	rs147998249	*GPT*	mis	p.Val452Leu	-0.225 (0.013)	8.19 E-65	41,413	-0.273 (0.050)	4.26 E-08	2475	-0.191 (0.070)	6.58 E-03	1357	-0.197 (0.041)	1.31 E-06	6158	-0.221 (0.029)	**1.41E-14**	-0.224 (0.012)	1.04E-77
	8	145732305	G	GC		*GPT*	fs	p.Glu475fs	-0.271 (0.031)	1.00 E-18	41,414	-0.161 (0.165)	3.29 E-01	2475	NA (NA)	NA	NA	-0.509 (0.203)	1.21 E-02	6158	-0.299 (0.128)	1.93E-02	-0.273 (0.030)	6.44E-20
	8	145748532	A	G	rs567402720	*LRRC24*	mis	p.Leu290Ser	-0.185 (0.028)	3.42 E-11	41,393	-0.161 (0.165)	3.29 E-01	2475	NA (NA)	NA	NA	-0.307 (0.143)	3.21 E-02	6158	-0.244 (0.108)	2.40E-02	-0.189 (0.027)	2.93E-12
	9	117122202	C	T	rs37481177	*AKNA*	syn	p.Glu755Glu	-0.007 (0.001)	9.51 E-09	41,414	-0.004 (0.005)	4.09 E-01	2475	0.004 (0.008)	6.18 E-01	1357	-0.007 (0.004)	5.29 E-02	6158	-0.005 (0.003)	8.42E-02	-0.007 (0.001)	3.08E-09
	9	117124731	G	A	rs37481176	*AKNA*	mis	p.Pro624Leu	-0.007 (0.001)	4.31 E-09	41,412	-0.004 (0.005)	3.90 E-01	2475	0.003 (0.008)	7.33 E-01	1356	-0.007 (0.004)	4.24 E-02	6158	-0.005 (0.003)	6.15E-02	-0.007 (0.001)	1.00E-09
	10	101595996	T	A	rs17222723	*ABCC2*	mis	p.Val1188Glu	-0.015 (0.003)	2.97 E-08	41,414	-0.002 (0.010)	8.01 E-01	2475	-0.007 (0.017)	6.88 E-01	1357	-0.017 (0.007)	1.55 E-02	6158	-0.012 (0.005)	3.43E-02	-0.014 (0.002)	3.44E-09
	10	101606861	G	T	rs1137968	*ABCC2*	syn	p.Val1430Val	-0.015 (0.003)	2.71 E-08	41,414	-0.003 (0.010)	7.74 E-01	2475	-0.008 (0.017)	6.28 E-01	1357	-0.017 (0.007)	1.70 E-02	6158	-0.012 (0.005)	3.25E-02	-0.014 (0.002)	2.99E-09
	10	101610533	C	T	rs8187707	*ABCC2*	syn	p.His1496His	-0.015 (0.003)	2.77 E-08	41,414	-0.003 (0.010)	7.93 E-01	2475	-0.008 (0.017)	6.28 E-01	1357	-0.017 (0.007)	1.76 E-02	6158	-0.012 (0.005)	3.43E-02	-0.014 (0.002)	3.23E-09
	10	101611294	G	A	rs8187710	*ABCC2*	mis	p.Cys1515Tyr	-0.015 (0.003)	2.15 E-08	41,414	-0.001 (0.010)	9.11 E-01	2475	-0.010 (0.017)	5.40 E-01	1357	-0.016 (0.007)	2.77 E-02	6158	-0.011 (0.005)	5.21E-02	-0.014 (0.002)	4.09E-09
	10	101912064	T	C	rs2862954	*ERLIN1*	mis	p.Ile291Val	-0.012 (0.001)	2.43 E-21	40,834	-0.010 (0.005)	2.91 E-02	2475	-0.006 (0.007)	4.02 E-01	1356	-0.009 (0.004)	2.06 E-02	6158	-0.009 (0.003)	**1.14E-03**	-0.011 (0.001)	1.76E-23
	10	101977883	C	T	rs2230804	*CHUK*	mis	p.Val268Ile	-0.009 (0.001)	1.93 E-13	41,414	-0.006 (0.005)	2.05 E-01	2475	0.0001 (0.008)	9.94 E-01	1357	-0.011 (0.004)	3.91 E-03	6158	-0.008 (0.003)	4.33E-03	-0.009 (0.001)	3.59E-15
	10	113917085	T	A	rs2254537	*GPAM*	syn	p.Pro681Pro	-0.008 (0.001)	4.61 E-10	41,414	-0.003 (0.005)	5.80 E-01	2475	-0.013 (0.008)	1.15 E-01	1357	-0.008 (0.004)	5.12 E-02	6158	-0.007 (0.003)	2.07E-02	-0.008 (0.001)	3.28E-11
	10	113940329	T	C	rs2792751	*GPAM*	mis	p.Ile643Val	-0.008 (0.001)	2.54 E-10	41,412	-0.003 (0.005)	5.61 E-01	2475	-0.013 (0.008)	1.33 E-01	1357	-0.008 (0.004)	4.77 E-02	6158	-0.007 (0.003)	2.00E-02	-0.008 (0.001)	1.77E-11

[0450]

							GHS 개발 집단			복제 집단								** 복제 메타-분석 (N = 3집)		*** 연합 메타-분석 (N = 4집)	
										GHS 메타인아트릭 수술집단			멜리스 성장 연구			U. Penn				
염기	Chr	BP	Ref	Alt	RSID	유전자	Ann	AA 치환	Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	Beta (SE)	P
	14	94844947	C	T	rs28929474	*SERPINA1*	miss	p.Glu366Lys	0.042 (0.005)	9.28 E-21	41,414	0.035 (0.020)	7.97 E-02	2475	0.034 (0.032)	2.92 E-01	1357	0.054 (0.013)	1.63 E-05	6158	0.047 (0.010)	*2.82E-06	0.043 (0.004)	1.59E-25
	19	19379549	C	T	rs58542926	*TM6SF2*	miss	p.Glu167Lys	0.014 (0.002)	4.76 E-09	41,413	0.040 (0.010)	2.40 E-05	2475	0.024 (0.014)	9.50 E-02	1357	0.013 (0.008)	7.51 E-02	6158	0.024 (0.006)	*1.37E-05	0.016 (0.002)	1.15E-12
	22	44324727	C	G	rs738409	*PNPLA3*	miss	p.Ile148Met	0.023 (0.002)	1.34 E-50	41,414	0.019 (0.006)	5.54 E-04	2475	0.006 (0.009)	5.43 E-01	1357	0.016 (0.004)	2.05 E-04	6158	0.016 (0.003)	*7.45E-07	0.021 (0.001)	3.55E-55
	22	44324730	C	T	rs738408	*PNPLA3*	syn	p.Pro149Pro	0.023 (0.002)	1.11 E-50	41,414	0.019 (0.006)	5.51 E-04	2475	0.006 (0.009)	5.43 E-01	1357	0.016 (0.004)	2.14 E-04	6158	0.016 (0.003)	*7.73E-07	0.021 (0.001)	3.10E-55
	22	44342116	A	G	rs2294918	*PNPLA3*	miss	p.Lys434Glu	0.007 (0.001)	8.26 E-08	41,412	0.001 (0.005)	7.77 E-01	2475	0.005 (0.008)	5.18 E-01	1357	0.005 (0.004)	2.16 E-01	6158	0.004 (0.003)	1.91E-01	0.006 (0.001)	6.24E-08
	22	44368122	A	G	rs3761472	*SAMM50*	miss	p.Arg110Gly	0.019 (0.002)	8.85 E-30	41,413	0.009 (0.006)	1.66 E-01	2475	-0.001 (0.01)	9.37 E-01	1357	0.018 (0.005)	4.02 E-04	6158	0.012 (0.004)	*7.69E-04	0.018 (0.002)	1.08E-31
	22	44395451	T	C	rs1007863	*PARVB*	miss	p.Trp37Arg	0.011 (0.001)	7.98 E-16	41,414	0.003 (0.005)	5.22 E-01	2475	0.008 (0.008)	3.13 E-01	1357	0.009 (0.004)	2.50 E-02	6158	0.007 (0.003)	1.78E-02	0.010 (0.001)	1.16E-16
AST	4	88231392	T	TA	rs72613567	*HSD17B13*	spl		-0.005 (0.001)	6.24 E-10	40,753	-0.010 (0.003)	3.12 E-03	2469	-0.012 (0.006)	5.32 E-02	1357	-0.007 (0.004)	5.56 E-02	6166	-0.009 (0.002)	*8.38E-05	-0.006 (0.001)	6.82E-13
	10	18242311	A	G	rs10764176	*SLC39A12*	miss	p.Ser366Gly	-0.006 (0.001)	1.09 E-10	40,753	-0.010 (0.003)	2.91 E-03	2469	-0.003 (0.006)	5.80 E-01	1357	-0.009 (0.004)	1.03 E-02	6166	-0.009 (0.002)	*1.16E-04	-0.006 (0.001)	1.10E-13
	10	101157378	CGTT	C		*GOT1*	inf	p.Asn389del	-0.221 (0.024)	1.96 E-20	40,753	-0.205 (0.062)	8.57 E-04	2469	NA (NA)	NA	NA	-0.243 (0.088)	5.97 E-03	6165	-0.218 (0.051)	*1.66E-05	-0.220 (0.022)	1.68E-24
	10	101165533	G	C	rs374966349	*GOT1*	miss	p.Gln208Glu	0.271 (0.027)	2.43 E-24	40,753	NA (NA)	NA	NA	NA (NA)	NA	NA	0.339 (0.079)	1.85 E-05	6166	0.339 (0.079)	*1.85E-05	0.278 (0.025)	3.25E-28
	10	101912064	T	C	rs2862954	*ERLIN1*	miss	p.Ile291Val	-0.005 (0.001)	4.82 E-09	40,753	-0.004 (0.003)	1.54 E-01	2469	-0.007 (0.006)	2.21 E-01	1357	-0.004 (0.003)	1.94 E-01	6166	-0.005 (0.002)	2.51E-02	-0.005 (0.001)	3.68E-10
	11	22271870	A	T	rs74811951	*ANO5*	miss	p.Leu322Phe	0.004 (0.001)	9.61 E-08	40,722	-0.001 (0.003)	7.85 E-01	2466	0.006 (0.006)	2.85 E-01	1357	-0.002 (0.003)	5.46 E-01	6165	0.000 (0.002)	8.43E-01	0.004 (0.001)	1.13E-06

[0451]

등위	Chr	BP	Ref	Alt	RSID	유전자	Ann	AA 치환	GHS 개발 집단			GHS 베리아트릭 수술 집단			델러스 심장 연구			U. Penn			** 복제 메타-분석 (N = 3명)		** 연합 메타-분석 (N = 4명)	
									Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	Beta (SE)	P
	14	94844947	C	T	rs28929474	*SERPINA1*	mis	p.Glu366Lys	0.027 (0.003)	2.44 E-20	40,753	0.023 (0.013)	7.79 E-02	2469	0.044 (0.024)	6.98 E-02	1357	0.055 (0.011)	4.01 E-07	6166	0.042 (0.008)	*9.54E-08	0.029 (0.003)	6.71E-26
	19	19379549	C	T	rs58542926	*TM6SF2*	mis	p.Glu167Lys	0.008 (0.002)	6.54 E-08	40,192	0.023 (0.006)	1.99 E-04	2469	0.010 (0.011)	3.42 E-01	1356	0.004 (0.007)	5.94 E-01	6166	0.014 (0.004)	*1.20E-03	0.009 (0.002)	5.92E-10
	22	44324727	C	G	rs738409	*PNPLA3*	mis	p.Ile148Met	0.014 (0.001)	8.31 E-46	40,753	0.014 (0.004)	1.27 E-04	2469	0.004 (0.007)	5.44 E-01	1357	0.015 (0.004)	4.87 E-05	6166	0.013 (0.002)	*5.51E-08	0.014 (0.001)	3.14E-52
	22	44324730	C	T	rs738408	*PNPLA3*	syn	p.Pro149Pro	0.014 (0.001)	8.93 E-46	40,753	0.014 (0.004)	1.32 E-04	2469	0.004 (0.007)	5.44 E-01	1357	0.015 (0.004)	4.96 E-05	6166	0.013 (0.002)	*5.81E-08	0.014 (0.001)	3.55E-52
	22	44368122	A	G	rs3761472	*SAMM50*	mis	p.Arg110Gly	0.011 (0.001)	1.22 E-22	40,752	0.008 (0.004)	6.03 E-02	2469	-0.001 (0.008)	9.45 E-01	1357	0.016 (0.004)	2.64 E-04	6166	0.010 (0.003)	*3.40E-04	0.011 (0.001)	1.91E-25
	22	44395451	T	C	rs1007863	*PARVB*	mis	p.Trp37Arg	0.006 (0.001)	1.31 E-13	40,753	0.003 (0.003)	4.12 E-01	2469	0.006 (0.006)	2.95 E-01	1357	0.009 (0.003)	6.17 E-03	6166	0.006 (0.002)	7.34E-03	0.006 (0.001)	3.62E-15

[0452]

[0453] *는 $P < 1.43 \times 10^{-3}$의 본페로니 유의성 한계값을 충족하는 P-값을 나타낸다.

[0454] ** 반복 메타-분석은 3개의 복제 집단: GHS 베리아트릭 수술 집단, 델러스 심장 연구, 및 Penn 약물 바이오뱅크를 포함한다.

[0455] *** 연합 메타-분석은 개발 집단 및 3개의 복제 집단: GHS 베리아트릭 수술 집단, 델러스 심장 연구, 및 Penn 약물 바이오뱅크를 포함한다.

[0456] 약어: AAF, 대체 대립유전자 빈도; Alt, 대체 대립유전자; ALT, 알라닌 아미노전이효소; AST, 아스파테이트 아미노전이효소; Ref, 참조 대립유전자; SE, 표준 오차; ann, 주석; mis, 미스센스; syn, 동의적; spl, 스플라이스 도너; stop, 중단 획득; fs, 프레임시프트; inf, 프레임내 indel.

[0457] *HSD17B13*은 *HSD17B11*의 30 kb 상류에 있는, 동일한 유전자 패밀리의 구성원이며, 두 유전자는 유럽인에서 단일한 큰 하플로타입 블록 내에 있다. 본 발명자들은 *HSD17B11*에서 코딩 또는 스플라이스 변종과 개발 집단에서 아미노전이효소 수준 사이에 (**도 5A** 및 **5B**; 가장 유의한 개발 P-값, ALT의 경우 1.36×10^{-1} 및 AST의 경우 4.32×10^{-2}) 또는 개발 집단과 3개의 복제 집단의 연합 메타-분석에서의 아미노전이효소 수준과의 사이에서 (가장 유의한

P-값, ALT 및 AST에 대해 각각 $6.25x10^{-3}$ 및 $1.17x10^{-5}$) 어떠한 연관성도 관찰하지 못하였다. 나아가, *HSD17B11*에서의 변종과 rs72613567의 결합 불균형은, 주로 본 발명자들의 개발 그룹을 포함하는 유럽계 미국인에서의 결과, 또한 델러스 심장 연구에서 나타난 히스패닉계 및 아프리카계 미국인에서의 결과를 포함하여, 모든 혈통 그룹을 가로질러 대단하지 않았다 (모든 혈통 그룹에서 *HSD17B11*에서 확인된 모든 변종으로 $r^2 < 0.4$; 데이터 미제시). 종합적으로, 이들 발견은 *HSD17B13*을 아미노전이효소 수준과 기능적으로 관련이 있을 것으로 가장 가능성이 높은 게놈 영역에서의 유전자로서 시사한다.

[0458] 다음으로, 본 발명자들은 ALT 또는 AST 수준과 관련된 변종들이 또한 만성 간 질환과도 관련이 있는지의 여부를 확립하고자 하였다. 개발 집단에서, 본 발명자들은 알코올성 및 비알코올성 (비바이러스성) 간 질환의 케이스, 뿐만 아니라 다음의 질환: 알코올성 간경변, 비알코올성 간경변, 및 간세포 암종 (HCC)의 후유증의 케이스를 광범위하게 규정하기 위하여 EHR 진단 암호를 사용하였다. 공통 대조근 그룹 ("간 질환 없음")을 임의의 유형의 간 질환에 대한 진단 암호가 없는 개인으로서 규정하였다 (**표 1**). 본 발명자들은 시험한 13개의 변종 및 2개의 광범위한 만성 간 질환 카테고리 (알코올성 및 비알코올성)를 설명하기 위하여 $P<0.05/24$ ($P<2.08x10^{-3}$)의 본페로니 유의성 한계값을 사용하여, 만성 간 질환과의 연관성에 대해 개발 및 복제 집단으로부터의 12개의 아미노전이효소-관련 변종을 시험하였다 (**표 4**). 전체적으로, 본 발명자들은 5개의 유전자 (*HSD17B13*, *SERPINA1*, *TM6SF2*, *PNPLA3*, 및 *SAMM50*)에서의 6개의 변종과 만성 간 질환 표현형 사이에 유의한 관련성을 발견하였다. *SERPINA1*, *TM6SF2*, *PNPLA3*, 및 *SAMM50* 관련성은 이전에 보고된 관련성을 확인시켜주는 것이다. *GPT*, *GOT1*, *ERLIN1*, 및 *SLC39A12*에서의 변종은 임의의 간 질환 표현형과 유의하게 관련이 있지 않았다. 여기서 보고된 간 질환과의 *HSD17B13* 관련성은 신규한 것이고 기술된 첫 번째의 잠재적 보호성 유전자 변종이다.

표 4

개발 집단에서 12개의 엑솜-전체의 중요한 및 복제하는 단일 뉴클레오타이드

[0459] 변종과 간 질환 표현형과의 연관성.

CHR:BP:Ref:Alt	유전자	rsID	알코올성 간 질환		알코올성 간경변		비알코올성 간 질환		비알코올성 간경변		간세포 암종	
			OR (95% CI)	P-값	OR (95% CI)	P-값	OR (95% CI)	P-값	OR (95% CI)	P-값	OR (95% CI)	P-값
4:88231392:T:TA	*HSD17B13*	rs72613567	0.62 (0.48-0.81)	*1.82E-04	0.56 (0.41-0.78)	*3.35E-04	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	3.83 (1.05-13.94)	8.88E-02	6.33 (1.71-23.43)	2.88E-02	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	0.77 (0.06-10.73)	8.43E-01	1.13 (0.08-15.39)	9.30E-01	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	0.73 (0.05-11.76)	8.17E-01	1.07 (0.07-17.16)	9.60E-01	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
10:182423111:A:G	*SLC39A12*	rs10764176	0.85 (0.68-1.07)	1.64E-01	0.92 (0.70-1.22)	5.80E-01	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*		4.60 (0.25-86.41)	3.93E-01	7.11 (0.38-133.19)	3.00E-01	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	2.20 (0.13-37.68)	6.24E-01	3.47 (0.20 - 59.04)	4.70E-01	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
14:94844947:C:T	*SERPINA1*	rs28929474	2.49 (1.49-4.17)	2.30E-03	3.35 (1.93-5.83)	*3.01E-04	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	1.47 (1.06-2.04)	2.76E-02	1.35 (0.89-2.04)	1.80E-01	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	1.76 (1.43-2.18)	*4.98E-07	2.07 (1.60-2.67)	*1.08E-07	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	1.77 (1.43-2.18)	*4.70E-07	2.07 (1.61-2.67)	*1.03E-07	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	1.90 (1.52-2.38)	*1.36E-07	2.28 (1.75-2.98)	*1.83E-08	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

[0460]

[0461] *는 $P<2.08\times10^{-3}$의 본페로니 유의성 한계값을 만족시키는 P-값을 나타낸다.

[0462] *HSD17B13* rs72613567의 대체물 (TA) 대립유전자를 평가된 임의의 만성 간 질환 표현형을 가진 참가자들과 비교하여 대조군에서 더 높은 빈도로 관찰하였다 (**도 2A** 및 **표 5**). 연령, 연령^2, 성별, BMI, 및 혈통에 대해 조정한 후, 본 발명자들은 TA 대립유전자당 알코올성 간 질환의 38% 더 낮은 가능성 (오즈비 [OR] 0.62; 95% 신뢰도 간격 [CI] 0.48 내지 0.81, $P=1.8\times10^{-4}$) 및 16% 더 낮은 가능성의 비알코올성 (비바이러스성) 간 질환 (OR 0.84, 95% CI 0.78 내지 0.91, $P=1.3\times10^{-5}$)을 관찰하였다. 간경변을 포함한 케이스들로 한정할 때, TA 대립유전자는 44% 더 낮은 가능성의 알코올성 (OR 0.56, 95% CI 0.41 내지 0.78, $P=3.4\times10^{-4}$) 및 26% 더 낮은 가능성의 비알코올성 (OR 0.74, 95% CI 0.62 내지 0.88, $P=4.5\times10^{-4}$) 간경변과 관련이 있었다. TA 대립유전자는 명목상으로 대립유전자당 33% 더 낮은 가능성의 HCC와 관련이 있었다 (OR 0.67, 95% CI 0.45 내지 1.00, $P=4.7\times10^{-2}$). 조정되지 않은 유전자형 OR은 공동-우성 효과를 시사하였다; 예를 들어, 알코올성 간경변의 경우, OR은 이형접합성

T/TA 보균자의 경우 0.59 (95% CI 0.40 내지 0.86) 및 동형접합성 TA/TA 보균자의 경우 0.26 (95% CI 0.08 내지 0.82)이었고, 비알코올성 간경변의 경우, OR은 이형접합성의 경우 0.75 (95% CI 0.61 내지 0.93) 및 동형접합성 보균자의 경우 0.55 (95% CI 0.34 내지 0.91)였다.

[0463] 그러므로, 개발 집단에서, *HSD17B13* rs72613567의 대체물 (TA) 대립유전자는, 일관된 대립유전자 용량-의존 방식으로, 평가된 모든 EHR-유래 만성 간 질환 표현형의 더 낮은 가능성과 관련이 있었다 (**도 2A**): 모든 범주의 알코올성 간 질환, 이형접합성 오즈비 (OR_{het}) [95% 신뢰도 간격] 0.58 [0.42 내지 0.79], 동형접합성 OR (OR_{hom}) 0.46 [0.23 내지 0.94], 대립유전자 OR ($OR_{allelic}$) 0.62 [0.48 내지 0.81], P = $1.82x10^{-4}$; 모든 범주의 비알코올성 간 질환, OR_{het} 0.84 [0.76 내지 0.92], OR_{hom} 0.73 [0.59 내지 0.89], $OR_{allelic}$ 0.84 [0.78 내지 0.91], P = $1.31x10^{-5}$. TA 대립유전자는 또한 이들 만성 간 질환의 가장 진전된 형태 (EHR-유래 진단 암호에 의해 규정되는 바), 즉 알코올성 및 비알코올성 간경변 및 HCC의 더 낮은 가능성과 관련이 있었다. TA 대립유전자는 이형접합체 및 동형접합체의 경우, 각각 42% 및 73% 더 낮은 가능성과 관련이 있었고 (OR_{het} 0.59 [0.40 내지 0.86], OR_{hom} 0.26 [0.08 내지 0.82], $OR_{allelic}$ 0.56 [0.41 내지 0.78], P = $3.35x10^{-4}$), 이형접합체 및 동형접합체의 경우, 각각 26% 및 49% 더 낮은 가능성의 비알코올성 간경변과 관련이 있었다 OR_{het} 0.75 [0.61 내지 0.93], OR_{hom} 0.55 [0.34 내지 0.91], $OR_{allelic}$ 0.74 [0.62 내지 0.88], P = $4.48x10^{-4}$). TA 대립유전자는 또한 명목상으로는 더 낮은 가능성의 HCC와 관련이 있었다.

[0464] 다음으로, 본 발명자들은 아프리카계 미국인, 유럽계 미국인, 및 히스패닉계 미국인 성인 및 아동을 포함한, 다중-민족 댈러스 간 연구 (DLS) 및 댈러스 소아 간 연구 (DPLS)에서 이런 발견을 확인하고 확대시키고자 하였다 (**표 1B**). DLS에서, TA 대립유전자는 대립유전자-투여량 의존 방식으로 더 낮은 가능성의 임의의 간 질환과 관련이 있었다 (OR_{het} 0.74 [0.57 내지 0.97], OR_{hom} 0.41 [0.21 내지 0.83], $OR_{allelic}$ 0.70 [0.5 내지 0.88], P = $1.77x10^{-3}$, **도 8**). 유사한 대립유전자 용량-의존성 효과를 진전된, 간경변 형태의 알코올성 ($OR_{allelic}$ 0.72 [0.53 내지 0.99], P = 4.37x10 내지 ^{2}) 및 비알코올성 ($OR_{allelic}$ 0.65 [0.40 내지 1.07], P = $8.96x10^{-2}$) 간 질환을 포함하여, EHR-유래 간 질환 아형 전체에서 관찰하였다. 자체-보고된 민족성에 의해 분류된 개인들의 하위세트 분석에서, 특히, 히스패닉계 미국인 집단에서의 고비율의 간 질환으로 인해, 이 집단에서 간 질환과의 연관성이 유지되었다 (n = 326 케이스 및 722건의 대조군, $OR_{allelic}$ 0.51 [0.35 내지 0.74], P = $3.98x10^{-4}$); 통계적으로 유의한 수준을 이루지는 않았지만, 유사한 수치 동향이 또한 DLS의 아프리카계 미국인 하위세트에서 (n = 33 케이스 및 2,291건의 대조군, $OR_{allelic}$ 0.74 [0.25 내지 2.47], P = 0.67) 및 유럽계 미국인 하위세트에서 (n = 158 케이스 및 1,266건의 대조군, $OR_{allelic}$ 0.87 [0.65 내지 1.15], P = 0.32) 주목을 끌었다. 히스패닉계 미국인 소아 간 질환 환자 및 비만 대조군의 별도의 연구인 DPLS에서 (**표 1B**), TA 대립유전자는 또한 더 낮은 가능성의 간 질환과 관련이 있었다 ($OR_{allelic}$ 0.59 [0.36 내지 0.97], P = $3.6x10^{-2}$). 그러므로, *HSD17B13* rs72613567:TA 대립유전자는 3개의 독립 집단에서 성인 및 아동에게서, 간경변을 포함하여, 만성 간 질환의 다중 형태의 감소된 가능성과 관련이 있었다.

[0465] NAFLD는 유의한 염증의 증거가 없는 지방 간 (조직병리학적 검사 시 "단순 지방증"으로서 표시됨)으로부터 임상적으로 보다 충격적인 징후 (소엽 염증, 간세포 팽창, 및/또는 섬유증의 조직병리학적 증거를 보이는, "비알코올성 지방간염" (NASH)으로 표시됨)에 이르기까지 다양한 질환의 범위를 기술한다. *HSD17B13* TA 대립유전자와 조직학적으로 규정된 NAFLD 및 NASH 사이의 관계를 이해하기 위하여, 본 발명자들은 GHS 베리아트릭 수술 집단으로부터 간 생검 샘플을 제공한 2,391명의 전체 엑솜 서열분석된 개인에게서 rs72613567의 연관성 검사를 수행하였다. 이들 개인 중에서, 555명 (23%)은 지방증, 지방간염, 또는 섬유증의 증거가 없었고 ("정상"), 830명 (35%)은 단순 지방증이 있었으며, 1006명 (42%)은 NASH (즉, 소엽 염증, 간세포 팽창, 또는 섬유증의 증거)가 있었다. *HSD17B13* TA 대립유전자는 정상 간과 비교하여 단순 지방증 (OR 1.11, 95% CI 0.94 내지 1.32, P=0.21) 또는 NASH (OR 0.86, 95% CI 0.72 내지 1.02, P=0.09)와 유의하게 관련이 있지 않았다 (**도 2B** 및 **표 5**). 정상 간, 단순 지방증, 및 NASH의 유병률을 유전자형에 의해 비교할 때, 정상 간의 유병률은 유전자형에 의해 상이한 것으로 나타나지 않았지만 (T/T, T/TA, 및 TA/TA 보균자에 대해 각각 23%, 24%, 및 23%, 비율로의 경향에 대한 카이 제곱 테스트에 의해 P = 0.5임), TA 대립유전자로 NASH의 유병률은 감소하였고 (T/T, T/TA, 및

TA/TA 보균자에 대해 각각 45%, 40%, 및 31%, P = 1.6x10^{-4}) 단순 지방증의 유병률을 증가하였다 (T/T, T/TA, 및 TA/TA 보균자에 대해 각각 33%, 35%, 및 47%, P = 1.1x10^{-3}) (**도 9**). 지방증을 가진 개인들 중에서, TA 대립유전자는 대립유전자 용량-의존 방식으로, 단순 지방증과 비교하여, 통계적으로 유의하게 더 낮은 가능성의 NASH와 관련이 있었다. 단순 지방증을 배경으로, TA 대립유전자는 23% 더 낮은 가능성의 NASH와 관련이 있었고 (OR 0.77, 95% CI 0.66 내지 0.90, P=6.5x10^{-4}), 이것은 NAFLD의 NASH 및 섬유증의 보다 진전된 단계로의 진행을 매개하는 데 있어 *HSD17B13*의 역할을 시사한다. 유전자형 연관성 결과는 공동 우세 효과와 일치하였다; 단순 지방증 대비 NASH 비교에서, OR은 이형접합성 T/TA 보균자의 경우 0.84 (95% CI 0.69 내지 1.02)였고, 동형접합성 TA/TA 보균자의 경우 0.48 (95% CI 0.34 내지 0.68)이었다.

표 5

HSD17B13 rs72613567은 개발 집단에서 알코올성 및 비알코올성 간 질환 표현형의 감소된 위험, 및 베리아트릭 수술 집단에서 비알코올성 지방간 질환으로부터 비알코올성 지방간염 및 섬유증으로의 진행의 감소된 위험과 관련이 있다.

[0466]

		사례				대조군								
집단	정의	N	REF/REF	REF/ALT	ALT/ALT	정의	N	REF/REF	REF/ALT	ALT/ALT	AAF	Het OR (95% CI)	Hom OR (95% CI)	대립유전자당 OR (95% CI)	P-값
개발 집단	알코올성 간 질환	197	133	56	8	간 질환 없음	30,522	16413	11969	2140	0.266	0.58 (0.42-0.79)	0.46 (0.23-0.94)	0.62 (0.48-0.81)	1.82E-04
	알코올성 간경변	130	89	38	3						0.266	0.59 (0.40-0.86)	0.26 (0.08-0.82)	0.56 (0.41-0.78)	3.35E-04
	비알코올성 간 질환	1930	1131	692	107						0.264	0.84 (0.76-0.92)	0.73 (0.59-0.89)	0.84 (0.78-0.91)	1.31E-05
	비알코올성 간경변	381	235	129	17						0.266	0.75 (0.61-0.93)	0.55 (0.34-0.91)	0.74 (0.62-0.88)	4.48E-04
	간세포 암종	76	49	24	3						0.266	0.67 (0.41-1.10)	0.47 (0.15-1.51)	0.67 (0.45-1.00)	4.66E-02
베리아트릭 수술 집단	단순 지방증	830	421	321	88	정상	555	288	224	43	0.291	0.98 (0.78-1.23)	1.39 (0.94-2.08)	1.11 (0.94-1.32)	2.11E-01
	NASH	1006	578	370	58						0.255	0.82 (0.66-1.02)	0.67 (0.44-1.02)	0.86 (0.72-1.02)	8.53E-02
	NASH	1006	578	370	58	단순 지방증	830	421	321	88	0.268	0.84 (0.69-1.02)	0.48 (0.34-0.68)	0.77 (0.66-0.90)	6.47E-04

[0467] 본 발명자들은 다음으로, *HSD17B13* TA 대립유전자가 유전자의 공지된 및 신규한 전사물의 발현에 어떻게 영향을 미치는지를 이해하고자 하였다. 본 발명자들은 *HSD17B13* rs72613567 스플라이스 변종의 22명의 동형접합성 기준 (T/T), 30명의 이형접합성 (T/TA), 및 17명의 동형접합성 대체물 (TA/TA) 보균자로부터의 조직학적으로 정상인 간 샘플에서 *HSD17B13* mRNA 발현을 평가하기 위해 RNA 서열분석을 사용하였다 (**도 3**). 2개의 공지된 *HSD17B13* 전사물, A 및 B에 더불어, 2개의 신규한 전사물은 동일하였다: 전사물 C는 엑손 6이 결핍되었고, 전사물 D는 엑손 6의 3' 단부에서 G 뉴클레오타이드가 삽입된 것을 특징으로 하여 미성숙 단백질 절단을 초래하였다. 신규한 전사물은 RT-PCR에 의해 확인하였고, D 전사물을 추가로 긴 판독 cDNA 서열분석에 의해 확인하였다. 이들 전사물의 발현 수준은 *HSD17B13* rs72613567 유전자형에 따라 달라졌다; 전사물 A 및 B의 수준은 감소한 한편, 전사물 C 및 D의 수준은 대립유전자-용량-의존 방식으로 T/TA 이형접합체 및 TA/TA 동형접합체에서 증가하였다 (**도 3**). 300개의 아미노산 단백질을 암호화하는 전사물 A는 T/T 동형접합체에서 T/T에서 우세한 전사물인 한편 (**도 3A**), 미성숙하게 절단된 단백질을 암호화하는 전사물 D는 TA/TA 동형접합체에서 우세한 전사물이었다 (**도 3D**). 이들 발현 패턴은 HSD17B13 동형 발현을 결정하는 데 있어 *HSD17B13* rs72613567에 대한 기능성 역할을 시사한다. 매우 낮은 수준의 발현을 나타낸 4개의 추가 전사물 (E 내지 H)을 또한 확인하였다 (**도 6A 내지 6D**). 모든 확인된 *HSD17B13* 동형의 단백질 서열 정렬을 **도 7A** 및 **7B**에 나타낸다.

[0468] HSD17B13은 이전에 인간 간세포에서의 지질-방울 관련 단백질로서 기술되었다 (Su et al. (2014) *Proc Natl Acad Sci USA* 111:11437-11442, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 본 발명자들은 HSD17B13의 공지된 및 신규한 동형 A 내지 D를 발현하는 렌티바이러스로 안정적으로 형질도입된 영구적인 인간 간 세포주 (HepG2 간암종 세포)에서의 단백질 동형 발현 및 국지화를 평가하였다. HSD17B13 동형 A는 미처리 및 올레산-처리 세포에서 지질 방울로 국지화되었다. 동형 A는 주로 막 주위의 BODIPY-표지된 지질 방울에서 검출되었고 지질 방울 코트 단백질 페릴리핀 (PLIN)과 공동-국지화되었다. 유사한 세포내 국지화를 지질 방울 표면에서

HSD17B13 동형 D에 대해 관찰하였다; 그러나, 지질 방울은 올레산 처리 후에 더 큰 것으로 나타났다. 대조적으로, 동형 B와 C는 소포체 마커 칼넥신과 공동 국지화되었다.

[0469] 요약하면, DiscovEHR 연구 집단으로부터의 49,188명의 개인으로부터의 EHR 및 간 생검 데이터와 결부된 엑솜 서열 데이터를 사용하여, 그리고 ALT 및 AST 측정이 포함된 9,883명의 추가된 개인으로부터의 엑솜 서열 데이터의 추수 연구를 사용하여, 본 발명자들은 *HSD17B13*의 스플라이스 변종, 아미노전이효소 수준, 및 만성 간 질환 표현형 사이의 신규한 연관성을 발견하였다. 이것은, 본 발명자들의 지식으로는, 만성 간 질환 표현형과의 보호성 연관성을 가진 엑손 변종의 최초 보고이다. *HSD17B13* TA 대립유전자는 단순 지방증과 관련이 있진 않지만, 지방증을 가진 개인에게서 조직병리학적 지방간염의 위험을 감소시켰고, 이것은 만성 긴 질환의 임상적으로 보다 진전된 단계로의 진행에서 *HSD17B13*에 대한 역할을 시사한다. 간 질환의 조직병리학적 정의뿐만 아니라 EHR 진단 암호를 사용하여 특성화된, 여러 상이한 간 질환 범주를 가로지르는, 4개의 독립적인 집단 (GHS 개발, GHS 베리아트릭, DLS, 및 DPLS)에서 보호성 연관성의 일관성은, 연관성의 현저한 대립유전자 용량-의존성과 함께, 보고된 *HSD17B13* 변종이 만성 간 질환의 임상적으로 보다 진전된 단계로의 진행으로부터 보호한다는 개념을 지지한다. 관찰된 대립유전자 용량-의존성은 또한 HSD17B13 기능의 보다 뚜렷한 조절이 질환 위험 및 진행에 대해 보다 뚜렷한 효과를 초래할 수 있음을 논증한다.

[0470] 다른 17베타-하이드록시스테로이드 탈수소효소 패밀리 구성원들은 성 스테로이드 및 지방산 대사에 관여하는 것으로 알려져 있지만 (Moeller and Adamski (2009) *Mol Cell Endocrinol* 301:7-19, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), HSD17B13의 기능에 대해서는 거의 알려진 것이 없다. *HSD17B13*은 주로 간에서 발현되며 (Liu et al. (2007) *Acta Biochim Pol* 54:213-218, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 간에서 지질 방울에 국지화되는데 (Su et al. (2014) *Proc Natl Acad Sci USA* 111:11437-11442, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 지방간 질환의 발병에서 *HSD17B13*에 대한 역할과 일치하는 것이다. 본 발명자들의 데이터는 *HSD17B13* 과다발현이 마우스 간에서 지질생성을 증가시켰고, 배양된 간세포에서 지질 방울의 수 및 크기를 증가시켰다는 최근의 발견과 일치한다 (Su et al. (2014) *Proc Natl Acad Sci USA* 111:11437-11442, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 2가지 이전 연구들은 또한 HSD17B13 단백질의 간에서의 발현이 지방간 환자에서 증가한 것을 나타냈다 (Su et al. (2014) *Proc Natl Acad Sci USA* 111:11437-11442 및 Kampf et al. (2014) *FASEB J* 28:2901-2914, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 간 질환의 증가된 위험과 관련된 것으로 보고된 변종을 가진 2개 유전자 - *PNPLA3* 및 *TM6SF2* - 또한 간세포 지질 대사에서 생리적 역할을 하였다. 본 발명자들이 본원에서 기술하는 *HSD17B13*의 변종은 간 질환에 대한 제1 보호성 변종이고, 다른 도메인의 새로운 치료제로의 길을 안내한 유전자 변종과 유사한, 만성 간 질환을 표적화하는 새로운 치료 전략에 대한 방안을 제공할 수 있다.

[0471] 전체적인, 본 발명자들의 데이터는 *HSD17B13*을 인간에서 만성 간 질환의 위험을 감소시키기 위한 신규한 치료 표적으로서 지지한다. 중요하게, 본 발명자들의 데이터는 *HSD17B13*의 표적화가 NAFLD로부터, 유의한 유병률 및 사망률과 관련되고, 현재로서는 그것에 대한 효과적인 치료가 없는, NASH, 섬유증, 및 간경변의 후기 단계로의 진행을 감소시킬 수 있을 것임을 나타낸다.

[0472] ***방법***

[0473] **연구 참가자.** 인간 유전학 연구를 리제너론 제네틱스 센터 및 Geisinger Health System (GHS)의 DiscovEHR 공동작업의 일부로서 수행하였다. 연구는 GHS 임상시험 심사위원회 (Institutional Review Board)에 의해 승인받았다. 2개의 DiscovEHR 연구 집단 (개발 집단 및 베리아트릭 수술 집단)은 GHS의 MYCODE® 지역사회 건강 가입자 (Community Health Initiative)로부터 18세 이상의 처음 50,726명의 서면동의를 받은 참가자로부터 비롯되었다 (Dewey et al. (2016) *Science* 354(6319) doi:10.1126/science.aaf6814, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). GHS 개발 집단은, 베리아트릭 수술 집단에 모집된 개인을 전부 배제하고, 2007년과 2016년 사이에 외래환자 일차 진료 및 특수 클리닉으로부터 모집한 46,544명의 유럽 개인으로 구성되었다. GHS 베리아트릭 수술 집단은 베리아트릭 수술에 대해 이관되었던 2,644명의 유럽 개인으로 구성되었다.

[0474] 복제 연구는 댈러스 심장 연구로부터 1,357명의 유럽 개인 및 Penn 약물 바이오뱅크로부터 8,527명의 유럽 개인을 포함하였다. 댈러스 심장 연구는 30세 내지 65세의 댈러스 카운티 거주민들의 확률-기반 인구 집단 연구이다 (Victor et al. (2004) *Am J Cardiol* 93:1473-1480, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). Penn 약물 바이오뱅크는 펜실베니아 대학 건강 시스템으로부터 모집되고 생체시료 저장, EHR 데이터에의 접근, 및 재접촉에 대한 허용에 대해 서면동의를 받은 참가자들을 포함한다.

[0475] 만성 간 질환과의 연관성의 복제 연구는 댈러스 간 연구 (DLS)로부터의 517명의 개인 및 댈러스 소아 간 연구

(DPLS)로부터의 447명의 개인을 포함하였다. DLS는 비바이러스성 병인의 간 질환을 가진 환자들의 바이오뱅크이다. 모집은 2015년 1월에 시작되었고 진행 중이다. 참가자는 댈러스의 UT 남서부 및 파크랜드 건강 및 병원 시스템의 간 클리닉으로부터 모집하였다. 바이오뱅크는 UT 남서부 임상시험 심사위원회에 의해 승인받았다. 참가자들에게는 서면 동의서를 제시하였다. 참가자들은 민족/인종 배경, 병력, 생활양식 요인, 및 간 질환 및 기타 질환의 가족력에 관한 설문지를 작성하였다. 추가적인 임상 정보는 숙련된 전문가에 의해 의료 기록으로부터 얻었다. 본 발명자들은 본 연구 시점에 이용할 수 있는 DNA를 가진 모든 아프리카계 미국인, 유럽계 미국인, 및 히스패닉계 미국인 환자들을 포함시켰다 (n=517명). DPLS는 댈러스의 UT 남서부 및 파크랜드 건강 및 병원 시스템의 소아 간 클리닉으로부터, 및 댈러스의 아동 의료 센터의 비만 클리닉으로부터 모집한 아동의 바이오뱅크이다. 바이오뱅크는 UT 남서부 임상시험 심사위원회에 의해 승인받았다. 참가자들의 법정 후견인들에게 서면 동의서를 제시하였다. 임상 정보는 숙련된 전문가에 의해 의료 기록으로부터 얻었다. 환자의 95% 이상이 히스패닉계 미국인이었기 때문에, 본 발명자들은 본 연구에 오직 히스패닉계 미국인 환자와 대조군을 포함시켰다 (n=203명의 환자 및 244명의 대조군).

[0476] **샘플 제조 및 서열분석.** 샘플 제조 및 전체 엑솜 서열분석을 앞서 기술한 것과 같이 리제너론 제네틱스 센터에서 수행하였다 (Dewey et al. (2016) *Science* 354(6319) doi:10.1126/science.aaf6814, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 간단하게 설명하면, 엑솜 포획을 NimbleGen 프로브를 사용하여 제조사의 권장 프로토콜 (Roche NimbleGen)을 따라 수행하였다. 포획한 DNA를 PCR 증폭시키고 qRT-PCR에 의해 정량화하였다 (Kapa Bio시스템). 다중화된 샘플을 Illumina v4 HiSeq 2500 상에서 75 bp 쌍-단부 서열분석을 사용하여, 96%의 샘플에서 표적화된 염기의 85% 이상의 20배보다 큰 하플로이드 판독 깊이를 제공하기에 충분한 커버리지 깊이로 서열분석을 하였다 (표적화된 염기의 대략 80배의 평균 하플로이드 판독 깊이). 각 Illumina Hiseq 2500 작동으로부터의 미가공 서열 데이터를 서열 판독 정렬 및 변종 확인을 위해 DNAnexus 플랫폼에 업로드하였다 (Reid et al. (2014) *BMC Bioinformatics* 15, 30 doi:10.1186/1471-2105-15-30). 간단하게 설명하면, 미가공 서열 데이터를 BCL 파일로부터 샘플-특이적 FASTQ-파일로 변환시켰고, 그것을 BWA-mem과 인간 기준 빌드 GRCh37.p13에 정렬시켰다 (Li and Durbin (2009) *Bioinformatics* 25:1754-1760, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 단일 뉴클레오타이드 변종 (SNV) 및 삽입/결실 (indel) 서열 변종을 게놈 분석 툴키트 (McKenna et al. (2010) *Genome Res* 20:1297-1303, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨)를 사용하여 확인하였다.

[0477] **댈러스 간 및 소아 간 연구에서 rs72613567의 표적화된 유전자형 분석.** *HSD17B13* rs72613567을 댈러스 간 연구 및 댈러스 소아 간 연구에서 TAQMAN® 검정에 의해, 및 델러스 심장 연구에서 엑솜 서열분석에 의해 유전자형을 분석하였다. TAQMAN® 콜(call)을 각 유전자형을 가진 5명의 개인의 Sanger 서열분석에 의해 확인하였다.

[0478] **개발 집단에서 임상 측정 및 만성 간 질환 정의.** ALT 및 AST에 대한 임상 실험실 측정값을 GHS 개발 집단 및 베리아트릭 수술 집단으로부터의 참가자들의 EHR로부터 얻었다. ALT 및 AST의 중앙값을 모든 참가자에 대해 2회 이상의 측정값에 대해 계산하였고, 연관성 분석 전에 분포를 표준화하기 위하여 $\log_{10}$ 변형하였다.

[0479] *국제 질환 분류, 제9차 개정* (ICD-9) 질환 코드를 EHR로부터 추출하고 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 케이스 정의에 대한 임상 질환 카테고리로 바꾸었다. 단일 진단 코드를 기반으로 한 추가적인 케이스 정의는 알코올성 간경변 (ICD-9 571.2), 비알코올성 간경변 (ICD-9 571.5), 및 HCC (ICD-9 155.0)를 포함하였다. 이들 케이스 정의에 대해, 간 질환이 없는 공통 대조군 그룹을 케이스 참조가 없거나 임의의 유형의 간 질환을 가리키는 단일-충돌 또는 문제-목록 진단 코드가 없는 참가자로서 정의하였다.

[0480] **베리아트릭 수술 집단에서 간 조직병리적 표현형 정의.** GHS 베리아트릭 수술 집단은 유럽 혈통의 2,644명의 개인으로 구성되었고, 이들 개인 중 2,391명으로부터 수술 중에 간 생검 시편을 입수할 수 있었다. 간 생검 시편을 포르말린에 고정하고, 기본적인 조직학을 위해 헤마톡실린 및 에오신으로, 및 앞서 기술된 것과 같이, 섬유증의 평가를 위해 Masson 삼색 염색으로 염색하였다 (Gerhard et al. (2011) *Patient Saf Surg* 5, 1, doi:10.1186/1754-9493-5-1, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 조직학적 진단을 이전에 확립된 기준을 사용하여 간병리학자에 의해 측정하였다 (Brunt et al. (1999) *Am J Gastroenterol* 94:2467-2474, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 조직학적 진단을 사용하여 다음의 표현형을 정의하였다: 1) 정상: 지방증, NASH, 또는 섬유증의 증거가 없음; 2) 단순 지방증: NASH 또는 섬유증의 증거가 없는 지방증 (등급 무시함); 3) NASH/섬유증: 소엽의 염증 또는 간세포 팽창의 임의의 존재 (등급 무시함), 또는 섬유증의

임의의 존재 (등급 무시함); 4) 섬유증: 섬유증의 임의의 존재 (등급 무시함).

[0481] **간 효소의 엑솜-전체 연관성 분석.** GHS 개발 집단에서, 본 발명자들은 502,219개의 이중대립유전자 변종을 테스트하였고, 아미노전이효소 수준과의 연관성에 대해 1% 미만의 데이터 전송 속도, $1.0x10^{-6}$보다 큰 Hardy-Weinberg 평형 p-값, 및 0.1%보다 큰 마이너 대립유전자 빈도를 버렸다. Log_{10}-변형된 중앙값 ALT 및 AST를 연령, 연령^2, 성별, BMI, 및 혈통의 제1 4대 주요 구성요소에 대해 조정하였다. 연구 참가자 중 관련성을 설명하기 위하여, 본 발명자들은 또한 유전적 관련성 기질을 랜덤 효과 공변량으로서 맞추었다. 주요 구성요소 및 유전적 관련성 기질 둘 다를 대략적인 결합 평형으로 및 0.1%보다 큰 대립유전자 빈도로 39,858명의 비-MHC 마커로부터 구성하였다. 본 발명자들은 특성 잔차(trait residual)와 단일 뉴클레오타이드 변종 사이의 연관성을 테스트하기 위하여 GSTA 패키지로 실행된 선형 혼합 모델을 사용하였다 (Yang et al. (2011) *Am J Hum Genet* 88:76-82, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 테스트는 엑솜-전체 분위수 대 분위수 도표 및 게놈 대조군 람다 값에 의해 나타난 것과 같이, 잘 보정되었다 (**도 1**).

[0482] **간 효소 연관성의 반복 메타-분석.** 본 발명자들은 3개의 별도의 유럽-혈통 집단: GHS 베리아트릭 수술 집단, 댈러스 심장 연구, 및 Penn 약물 바이오뱅크 (상기 기술됨)의 GHS 개발 집단에서 연관성을 복제하려는 시도를 하였다. GHS 베리아트릭 수술 집단의 및 Penn 약물 바이오뱅크로부터의 ALT 및 AST 측정치를 log_{10}-변형시키고 연령, 연령^2, 성별, BMI, 및 혈통의 제1 4대 주요 구성요소에 대해 조정하였다. Penn 약물 바이오뱅크 샘플로부터의 ALT 및 AST 측정치를 log_{10}-변형시키고 연령, 연령^2, 성별, BMI, 및 혈통의 제1 4대 주요 구성요소에 대해 조정하였다. 유전적 관련성 기질을 랜덤 효과 공변량으로서 포함시켰고, 분석을 GCTA에서의 선형 혼합 모델을 사용하여 수행하였다. 댈러스 심장 연구에서, log_{10}-변형된 ALT 및 AST 측정치를 연령, 연령^2, 성별, 및 혈통의 제1 10대 주요 구성요소에 대해 조정하고, 분석을 PLINK에서의 선형 혼합 모델을 사용하여 수행하였다. 3개의 복제 집단에 대한 요약 통계를 METAL (반복 메타-분석)을 사용하여 메타-분석하였다 (Willer et al. (2010) *Bioinformatics* 26:2190-2191, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 개발 집단 및 3개의 복제 집단에 대한 요약 통계를 유사하게 메타-분석하였다 (연합 메타-분석).

[0483] **만성 간 질환 표현형과의 연관성 분석.** 본 발명자들은 상기 기술된 것과 같이, GHS 개발 집단으로부터 정의된 2진(binary) 간 질환 표현형과의 연관성에 대해 간 효소 ExWAS로부터의 9개의 중요하고 복제된 단일 뉴클레오타이드 변종을 분석하였다. 본 발명자들은 테스트한 13개의 변종 및 2개의 광범위한 만성 간 질환 카테고리 (알코올성 및 비알코올성)를 설명하기 위하여 $P<0.05/26$ ($P<1.92x10^{-3}$)의 본페로니 유의성 한계값을 사용하였다. 변종 *HSD17B13*을 추가로, 상기 기술된 것과 같이, GHS 베리아트릭 수술 집단으로부터의 조직병리학적으로 정의된 간 표현형과의 연관성에 대해 테스트하였다. 연령, 연령^2, 성별, BMI, 및 혈통의 제1 4대 주요 구성요소에 대해 조정한 후에 로지스틱 회귀의 Firth의 벌점 가능성 방법을 사용하여 오즈비를 추정하였다. 조정하지 않은 유전자형 오즈비를 또한 *HSD17B13* rs72613567에 대해 추정하였다.

[0484] DLS에서 간 질환에 대한 오즈비를 로지스틱 회귀에 의해 추정하고, 연령, 연령^2, 성별, BMI, 및 자가-보고 민족에 대해 조정하였다. 이용가능한 rs72613567 유전자형을 가진 댈러스 심장 연구로부터의 참가자를 정상 대조군으로서 사용하였다 (n=4,279명). DPLS에서의 오즈비를 로지스틱 회귀에 의해 추정하였다.

[0485] **소프트웨어.** 유전적 연관성 분석을 GCTA 소프트웨어, 버전 1.25.0 (Yang et al. (2011) *Am J Hum Genet* 88:76-82, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 및 PLINK, 버전 1.9.0을 사용하여 수행하였다. 분위수-분위수 및 맨하탄 도표를 R 소프트웨어, 버전 3.2.1을 사용하여 생성하였다 (통계적 컴퓨터계산을 위한 R 프로젝트). 위치적 연관성 도표를 LocusZoom을 사용하여 생성하였다 (Pruim et al. (2010) *Bioinformatics* 26:2336-2337, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0486] **RNA 서열분석 연구.** RNA 품질 및 농도를 총 RNA를 Agilent RNA Nano Bioanalyzer 칩 상에서 이동시킴으로써 평가하였다; 모든 샘플은 8보다 큰 RNA 무결성 번호 (RIN)를 가졌다. 폴리아데닐화된 RNA 전사물을 올리고(dT)25 비즈 (Thermo Fisher Scientific)로의 2회기의 풍부화를 사용하여 분리하였다. 샘플을 RNAclean XP 비즈 (Beckman Coulter)로 정제 및 농축하고 대략 140 염기쌍으로 가열-단편화하였다. 제1 가닥 합성을 SuperScript III 역전사효소 (Thermo Fisher Scientific)로 무작위 헥사머를 사용하여 완료하였다; dTTP는 제2 가닥 합성 중에 dUTP로 교체되었다. 샘플을, 가닥-특이적 서열분석 라이브러리를 생성하기 위하여 우라실 DNA-글리코실라

아제 단계를 첨가하여, 엑솜에 대해 상기 참조한 본 발명자들의 표준 DNA 라이브러리 제조 방법에 따라 처리하였다. 샘플을 모으고 75 bp 쌍-단부 서열분석을 사용하여 Illumina v4 HiSeq 2500에서 서열분석을 하였다.

[0487] ***신규한 HSD17B13 전사물의 확인.*** 판독을 두 개의 미스매치를 허용하는 ARRAYSTUDIO® 소프트웨어 (OMICSOFT®, Cary, NC)를 사용하여 인간.B38에 대해 지도화하였다. 2가지 접근법을 사용하여 신규한 *HSD17B13* 전사물을 확인하였다. 신규한 엑손 접합부를 Gencode v24를 기반으로 발견하였다. 데노보 전사물 어셈블리를 디폴트 세팅에서 트리니티(Trinity) (v2.2.0)를 사용하여 이동시켰다. 주문 유전자 모델을 *HSD17B13*의 신규한 전사물을 포함시키기 위하여 구성하였고, 전사물 정량화를 주문 유전자 모델에 대한 판독 정렬에 의해 추정하였다. 모든 확인된 HSD17B13 동형의 단백질 서열 정렬을 **도 7A** 및 **7B**에 나타낸다.

[0488] **신규한 전사물의 RT-PCR 입증.** 인간 간 샘플로부터의 총 RNA에 대한 RT-PCR을 PlatinumTM Taq DNA 중합효소 (Thermofisher)가 포함된 SUPERSCRIPT™ 원스텝 RT-PCR 시스템을 사용하여 수행하였다. 각각 50 uL의 RT-PCR 반응은 1X 반응 믹스, 500 nM의 각각의 전방향 및 역방향 프라이머 (PST516: ATGAACATCATCCTAGAAATCCTTC (SEQ ID NO: 251) 및 PST517: ATCATGCATACATCTCTGGCTGGAG (SEQ ID NO: 252)), 1 μL의 RT/Platinum Taq, 및 75 ng의 RNA를 함유하였다. 사이클 조건은 다음과 같았다; 45℃에서 30분 동안 1 사이클; 94℃에서 2분 동안 1 사이클; 94℃에서 20초 동안, 53℃에서 30초 동안, 및 72℃에서 90초 동안, 40 사이클; 72℃에서 5분 동안 1 사이클; 그런 후 10℃에서 유지함. 생성물을 QIAquick PCR 정제 키트 (Qiagen)를 사용하여 정제하고 프라이머 DE002 (ATCAGAACTTCAGGCCTTGG (SEQ ID NO: 253))를 사용하여 직접 Sanger 서열분석을 위해 제출하였다. B 및 C 전사물을 확인하기 위하여, RT-PCR 생성물을 2% 아가로오스 겔 상에서 이동하도록 하고 SYBR GOLDSYBR® 골드 핵산 겔 염색 (Thermofisher)으로 염색한 후, 예상했던 분자량의 밴드를 절제하여 QIAquick 겔 추출 키트 (Qiagen)를 사용하여 정제한 후, TOPO® TA 클로닝 키트 (Thermofisher)로 클로닝을 수행하였다. TOPO 클론의 서열분석을 M13F 및 M13R 서열분석 프라이머를 사용하여 수행하였다. 서열 분석을 Sequencher DNA 분석 소프트웨어 (Gene Codes Corporation)를 사용하여 수행하였다.

[0489] **신규한 전사물의 PacBio 입증.** 전장 *HSD17B13* 전사물을 50 ng의 총 RNA로부터 백금 Taq 고신뢰도를 가진 SuperScript III 원스텝 RT-PCR 시스템 (Thermo Fisher Scientific)으로 제1 엑손 (GCAAAGCCATGAACATCATCC (SEQ ID NO: 254) 및 마지막 엑손 (TCTTGATGTAGTGGGAGTCGGATT (SEQ ID NO: 255))에서의 유전자-특이적 프라이머를 사용하여 직접 증폭시켜서 약 2.2 kb (최대 예상 크기 전사물)의 암플리콘을 생성하였다. 암플리콘을 Agilent Bioanalyzer 상에서 입증하였다. PacBio-부합하는 바코딩된 어댑터를 암플리콘에 결찰시키고 PacBio PB 비즈 (Pacific Biosciences)로 세정하였다. 라이브러리를 동등한 양으로 모으고 한 SMRT 세포에 대해 180분 동안 PacBio RSII 플랫폼에서 서열분석을 하였다. 데이터를 PacBio 소프트웨어 스마트분석 v2.3 도구 labelzmw를 사용하여 탈다중화한 후 콘센서스 도구 암플리콘 분석으로 분석하였다. 그 결과의 암플리콘을 *HSD17B13* RefSeq 유전자와 비교하여 동형 및 유전자형 상태를 측정하였다.

[0490] **HSD17B13 동형의 세포내 국지화.** HepG2 세포를 10% 소태아 혈청이 첨가된 이글스 최소 필수 배지에서 배양하였다. HSD17B13 전사물 A, B, C, 및 D를 Myc-DDK 골격 렌티바이러스 구성물 안으로 하위클로닝하고, 렌티바이러스를 생성시켰다. HepG2 세포를 다양한 HSD17B13 전사물을 운반하는 렌티바이러스로 감염시켰다. 각각의 HSD17B13 전사물을 발현하는 안정적인 세포주를 완전 배양 배지에서 2주 동안 1 내지 3 mg/ml의 제네티신 G-418 설페이트로 선택하였다. 선택한 HepG2 세포를 200 mM의 올레산으로 밤새 처리하거나 처리하지 않은 후 고정시켰다. HSD17B13 동형을 마우스 항-Myc 항체로 표지하였다. 지질 방울을 BODIPY FL 염료 (Sigma)로 표지하였다. 지질 코트 단백질 및 소포체를 토끼 항-PLIN 항체 (Sigma) 및 토끼 항-칼넥신 항체 (Cell Signaling Technology)로 각각 표지하였다. 면역형광에 대한 이차 항체는 Alexa Fluor 488 당나귀 항-토끼 IgG 및 Alexa Fluor 594 당나귀 항-마우스 IgG (Jackson ImmunoResearch)였다.

[0491] **실시예 2. HSD17B13 mRNA 및 HSD17B13 단백질 발현에 미치는 rs72613567:TA의 영향**

[0492] 유전자의 공지된 및 신규한 전사물의 발현에 미치는 HSD17B13 rs72613567:TA 대립유전자의 영향을 조사하였다. RNA 서열을 사용하여 HSD17B13 rs72613567 스플라이스 변종의 22 T/T 동형접합성, 30 T/TA 이형접합성, 및 17 TA/TA 동형접합성 보균자로부터의 조직학적으로 정상인 간 샘플에서의 HSD17B13 mRNA 발현을 평가하였다. 2개의 공지된 HSD17B13 전사물, A 및 B 외에, 2개의 신규한 전사물: C와 D를 확인하였는데, 이것들은 엑손 6이 결핍되었고, 전사물 D는 엑손 6의 3' 단부에서 구아닌 뉴클레오타이드가 삽입되어 단백질의 미성숙 단백질 절단을 초래하는 것으로 예측할 수 있었다. 전사물을 RT-PCR 및 Sanger 서열분석에 의해 입증하였다 (데이터 미제시). D 전사물을 또한 긴 판독 cDNA 서열분석을 사용하여 입증하였다. 이들 전사물의 발현 수준은 HSD17B13 rs72613567

유전자형에 따라 달라졌다; 전사물 A의 수준은 감소한 한편, 전사물 D의 수준은 각 TA 대립유전자로 대립유전자 용량-의존 방식으로 감소하였다 (**도 3A**, **3D**, 및 **10B** 참조). 전장 300 아미노산 단백질을 암호화하는 전사물 A는 T/T 동형접합체에서 우세한 전사물인 한편, 미성숙하게 절단된 단백질을 암호화하는 전사물 D는 TA/TA 동형접합체에서 우세한 전사물이었다. 인간 간 생검 조직에서, 절단된 동형 D 단백질은 이형접바체 및 TA/TA 동형접합체에서 최소한으로 존재하였고, 동형 A 단백질 풍부성은 대립유전자 용량-의존 방식으로 감소하였다 (**도 10B** 및 **10C**) 참조. 이들 데이터는 실질적으로 인간 간에서의 발현을 감소시키는 단백질의 절단된 형태의 합성을 초래하는, mRNA 스플라이싱을 변경시키는 HSD17B13 rs72613567과 일치한다.

[0493] **도 10A 내지 10E**를 참조하면, 신규한 HSD17B13 전사물의 발현, 세포내 국지화, 및 효소적 활성이 도시된다. HSD17B13 rs72613567 스플라이스 변종의 동형접합성 기준 (T/T), 이형접합성 (T/TA), 및 동형접합성 대체물 (TA/TA) 보균자에서 HSD17B13 전사물 A 및 D의 발현은 **도 3A** 및 **3D**에 도시된다. 유전자 모델의 코딩 영역을 줄무늬 박스에 나타내고 미번역 영역을 검은색 박스에 나타낸다. 전사물 D에서 별표는 엑손 6의 3' 단부에서 rs72613567로부터의 G의 삽입을 나타내고, 이것은 단백질의 미성숙 절단으로 이어진다. mRNA 발현은 FPKM 단위 (백만개의 지도화된 판독당 전사물의 킬로염기당 단편)로 표시된다. HSD17B13 전사물 A 및 D를 과다발현하는 HepG2 세포로부터의 웨스턴 블롯은 HSD17B13 전사물 D가 HSD17B13 전사물 A에 비해 더 낮은 분자량을 가지는 절단된 단백질로 번역되었음을 보여준다 (**도 10A** 참조). 유사한 결과를 신선한 냉동 인간 간 및 HEK293 세포 샘플로부터의 HSD17B13 웨스턴 블롯으로 관찰하였다 (**도 10B** 참조). 인간 간 샘플은 HSD17B13 rs72613567 스플라이스 변종의 동형접합성 기준 (T/T), 이형접합성 (T/TA), 및 동형접합성 대체물 (TA/TA) 보균자로부터 유래하였다. 세포 샘플은 미태그된 HSD17B13 전사물 A 및 D를 과다발현하는 HEK293 세포로부터 유래하였다. HSD17B13 전사물 D를 HSD17B13 IsoA보다 저분자량을 가지는 절단된 단백질 IsoD로 번역하였다. HSD17B13 IsoD 단백질 수준은 인간 간 (좌측) 및 세포 (우측) 샘플 둘 다로부터의 IsoA 단백질 수준보다 낮았다 (**도 10C** 참조). 액틴으로 표준화된 단백질 수준은 **도 10C**에서 막대 칼럼으로 표시한다; ** P<0.001, *P<0.05. HSD17B13 전사물 A 또는 D를 안정적으로 과다발현하는 HepG2에서 지질 방울 막상에 국지화된 HSD17B13 동형 A 및 D를 둘 다 지질 방울 및 항-Myc가 나타나도록 BODIPY로 표지하여 HSD17B13 국지화를 나타냈다 (데이터 미제시). HSD17B13 동형 A 및 D의 17-베타 에스트라다이올 (에스트라다이올), 류코트라이엔 B4 (LTB4), 및 13-하이드록시옥타데카다이엔산 (13(S)-HODE)에 대한 효소적 활성을 또한 평가하였다 (**도 10D** 참조). HSD17B13 동형 D는 동형 A에 대한 상응하는 값의 효소적 활성의 10% 미만을 나타냈다. HSD17B13 동형 D는 HEK293 세포에서 과다발현되었을 때 배양 배지에서 측정될 때 에스트라다이올 (기질)로부터의 에스트론 (생성물)로의 많은 전환을 나타내지 않은 한편, 과다발현된 HSD17B13 동형 A는 활발한 전환을 나타냈다 (**도 10E** 참조).

[0494] HSD17B13은 주로 간에서 발현되며 (Liu et al., *Acta Biochim. Pol.*, 2007, 54, 213-8, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 지방간 질환의 발병에서의 역할과 일치하게, 그곳에서 지질 방울로 국지화된다 (Su et al., *Proc. Natl. Acad. Sci. USA*, 2014, 111, 11437-42, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). HSD171B3의 발현 및 그것의 국지화를 HSD17B13 전사물 A 및 D를 발현하는 렌티바이러스로 안정적으로 형질도입된 불멸화된 인간 간 세포주에서 평가하였다. HSD17B13 동형 A를 주로 막 주위의 BODIPY-표지된 지질 방울에서 검출하였다 (데이터 미제시). 유사한 세포내 국지화를 지질 방울 표면에서 HSD17B13 동형 D에 대해 관찰하였다 (**도 10D** 참조).

[0495] rs72613567:TA로 인한 HSD17B13 단백질의 미성숙 절단의 기능적 결과를 이해하기 위하여, 동형 A 및 D의 효소적 활성을 시험관내에서 재조합 단백질을 사용하여 평가하였다. 300개 이상의 잠정적인 기질을 조사하였고, 그 중 에스트라다이올, 류코트라이엔 B4, 및 13-하이드록시옥타데카다이엔산이 HSD17B13에 의해 효소적으로 전환되어, 하이드록실의 케톤기로의 산화를 초래하였다. HSD17B13 동형 D는 3가지 기질에 대한 크게 감소된 활성을 보였다 (**도 10D** 참조).

[0496] GFP 대조군과 비교하여, HSD17B13-전사물-A-과다발현 세포는 세포 배양 배지에 더 낮은 농도의 에스트라다이올뿐만 아니라 더 높은 농도의 에스트론을 가졌고, 이것은 에스트라다이올에 대한 효소 활성을 시사한다 (**도 10E** 참조). HSD17B13-전사물-D-과다발현 세포는 GFP 대조군과 비교하여 유사한 에스트론/에스트라다이올 비율을 가졌고, 이것은 HSD17B13 전사물 D가 상당한 기능의 손실을 가졌음을 시사한다. 질량 분광 분석은 에스트론의 하이드록시에스트론으로의 신속한 전환 및 소비된 에스트라다이올과 비교하여 에스트론의 저축적을 설명해주는 기타 생성물을 나타냈다.

[0497] 대규모 엑솜 서열분석을 통해, HSD17B13의 스플라이스 변종과 감소된 혈청 아미노전이효소 수준, 뿐만 아니라 간 질환의 진전된 간경변성 형태 및 HCC를 포함하여 간 질환의 비알코올성 및 알코올성 형태의 감소된 위험 사이에서 신규한 연관성을 확인하였다. 본 발명자들의 견해에는, 이것은 간 질환과 보호성 연관성을 가진 단백질-

변경 변종의 최초 보고이다. HSD17B13 rs72613567:TA 대립유전자는 단순 지방증과 관련되지 않았지만, NASH로의 진행 위험을 감소시켰다. 여러 상이한 간 질환 카테고리 및 인종을 가로질러 4개의 독립적인 집단 (DiscovEHR, DiscovEHR에서 독립적인 베리아트릭 수술 집단, DLS, 및 DPLS)에서의 투여량-의존성 보호성 연관성의 일관성은 보고된 HSD17B13 변종이 만성 간 질환의 임상적으로 보다 진전된 단계로의 진행으로부터 보호한다는 개념을 지지한다. 관찰된 대립유전자 용량-의존성은 또한 HSD17B13 기능의 보다 현저한 조절이 질환 위험 및 진행에 대해 보다 현저한 영향을 초래할 수 있음을 논증한다.

[0498] 본원에서 기술된 연관성 발견은 주로 BMI가 상승되어 있는 유럽계 및 히스패닉계 미국인에서의 관찰을 기반으로 하였다. HSD17B13은 HSD17B13에 매우 유사한 서열을 갖지만 더 넓은 조직 분포를 가지는 동일한 유전자 패밀리의 구성원인 HSD17B11과 매우 가깝다. 전체적으로, 본원에서 제시된 데이터는 HSD17B13이 인간의 지방간 질환의 예방 및 치료를 위한 잠재적인 치료 표적이라는 지위를 지지한다. 본원에서 제시된 데이터는 HSD17B13의 표적화가 지방증으로부터, 상당한 유병률 및 사망률과 관련되고, 현재 그것들에 대해 효과적인 치료제가 없는, NASH, 섬유증, 및 간경변의 후기 단계로의 간 질환의 진행을 감소시킬 수 있을 것임을 나타낸다.

[0499] **실시예 3. 변종 17베타-하이드록시스테로이드 탈수소효소 13은 만성 간 질환에 대해 보호한다.**

[0500] 만성 간 질환에 기여하는 유전 인자를 확인하기 위하여, 본 발명자들은 DiscovEHR 인간 유전학 연구에서의 46,544명의 참가자로부터의 엑솜 서열 데이터 및 전자 건강 기록을 활용하였다. 본 발명자들은 만성 간 질환과 관련되었을 후보들을 지명하기 위하여 간 손상의 확립된 생체마커 (혈청 알라닌 아미노전이효소 (ALT) 및 아스파테이트 아미노전이효소 (AST))와 관련된 유전자 변종을 확인하였다. 3개의 추가 집단 (12,527명의 개인)에서 복제하는 후보 변종을 그 후에 DiscovEHR 및 2개의 독립 집단 (총 37,892명의 개인)에서 만성 간 질환의 임상적 진단과의 연관성에 대해 평가하였다. 본 발명자들은 또한 독립적인 베리아트릭 수술 집단에서 간 질환의 조직병리학적 중증도와의 연관성을 조사하였다 (n=2,391개의 인간 간 샘플).

[0501] 간의 지질 방울 단백질 17-베타 하이드록시스테로이드 탈수소효소 13을 암호화하는, HSD17B13에서의 스플라이스 변종 (rs72613567:TA)은 감소된 ALT ($P=4.2\times10^{-12}$) 및 AST ($P=6.2\times10^{-10}$) 수준과 재생적으로 관련되었다. DiscovEHR에서, 이 변종은 대립유전자 용량-의존 방식으로 알코올성 및 비알코올성 간 질환의 감소된 위험 (각 rs72613567:TA 대립유전자에 대해 각각 38%, 95% 신뢰도 간격 (CI) 19% 내지 52%; 및 16%, 95% CI 9% 내지 22%) 및 알코올성 및 비알코올성 간경변의 감소된 위험 (각 rs72613567:TA 대립유전자에 대해, 알코올성 및 비알코올성 간경변에 대해 각각 44%, 95% CI 22 내지 59%; 및 26%, 95% CI 12% 내지 38%)과 관련되었고; 연관성을 2개의 독립 집단에서 확인하였다. rs72613567:TA는 비알코올성 지방간염 (NASH)의 조직학적 특징의 감소된 중증도와 관련이 있었다 (23% 감소, 지방간 질환에 걸린 개인들 중 각 rs72613567:TA 대립유전자에 대해 95% CI 10% 내지 34%). rs72613567:TA는 스테로이드 기질에 대해 감소된 효소적 활성을 가지는, 안정적이지 못하고 절단된 단백질을 초래한다.

[0502] HSD17B13에서 기능 상실 변종은 알코올성 및 비알코올성 간 질환의 감소된 위험, 및 지방증으로부터 NASH로의 진행과 관련이 있었다.

[0503] ***연구 설계 및 참가자***

[0504] 인간 유전자학 연구를 리제너론 제네틱스 센터 및 Geisinger Health System (GHS)의 DiscovEHR 공동작업의 일부로서 수행하였다. 2개의 DiscovEHR 연구 집단 (개발 집단 및 베리아트릭 수술 집단)은 GHS의 MyCode® 지역사회 건강 가입자로부터 18세 이상의 처음 50,726명의 서면동의를 받은 참가자로부터 비롯되었다. GHS 개발 집단은 2007년과 2016년 사이에 외래환자 일차 진료 및 특수 클리닉으로부터 모집한 46,544명의 유럽 개인으로 구성되었고, 베리아트릭 수술 집단에 모집된 개인은 전부 배제하였다. GHS 베리아트릭 수술 집단은 베리아트릭 수술에 대해 이관되었던 2,644명의 유럽 개인으로 구성되었다.

[0505] 간 아미노전이효소와의 연관성의 복제 연구는 댈러스 심장 연구로부터 1,357명의 유럽 개인 및 Penn 약물 바이오뱅크로부터 8,527명의 유럽 개인을 포함하였다. 댈러스 심장 연구는 30세 내지 65세의 댈러스 카운티 거주민들의 확률-기반 인구 집단 연구이다 (Victor et al., *Am. J. Cardiol.*, 2004; 93, 1473-80 , 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). Penn 약물 바이오뱅크는 펜실베니아 대학 건강 시스템으로부터 모집되고 생체시료 저장, EHR 데이터에의 접근, 및 재접촉에 대한 허용에 대해 서면동의를 받은 참가자들을 포함한다.

[0506] 만성 간 질환과의 연관성의 복제 연구는 댈러스 간 연구 (DLS)로부터의 517명의 개인 및 댈러스 소아 간 연구 (DPLS)로부터의 447명의 개인을 포함하였다. DLS는 비바이러스성 병인의 간 질환을 가진 환자들의 바이오뱅크이다. 모집은 2015년 1월에 시작되었고 진행 중이다. 참가자는 댈러스의 UT 남서부 및 파크랜드 건강 및 병원 시

스템의 간 클리닉으로부터 모집하였다. 참가자들은 민족/인종 배경, 병력, 생활양식 요인, 및 간 질환 및 기타 질환의 가족력에 관한 설문지를 작성하였다. 추가적인 임상 정보는 숙련된 전문가에 의해 의료 기록으로부터 얻었다. 본 발명자들은 본 연구 시점에 이용할 수 있는 DNA를 가진 모든 아프리카계 미국인, 유럽계 미국인, 및 히스패닉계 미국인 환자들을 포함시켰고 (n=517명) 댈러스 심장 연구 (Dallas Heart Study)로부터 대조군을 얻었다. DPLS는 댈러스의 UT 남서부 및 파크랜드 건강 및 병원 시스템의 소아 간 클리닉으로부터, 및 댈러스의 아동 의료 센터의 비만 클리닉으로부터 모집한 아동의 바이오뱅크이다. 임상 정보는 숙련된 전문가에 의해 의료 기록으로부터 얻었다. 환자의 95% 이상이 히스패닉계 미국인이었기 때문에, 본 발명자들은 본 연구에 오직 히스패닉계 미국인 환자와 대조군을 포함시켰다 (n=205명의 환자 및 234명의 대조군).

[0507] ***개발 집단에서 임상 측정 및 만성 간 질환 정의***

[0508] ALT 및 AST에 대한 임상 실험 측정값을 GHS 개발 집단 및 베리아트릭 수술 집단으로부터의 참가자의 EHR로부터 추출하였다. 중앙 ALT 및 AST 값을 2회 이상 측정한 모든 참가자에 대해 계산하였고, $\log_{10}$-변형시켜서 연관성 분석 전에 분포를 표준화하였다.

[0509] 질환의 국제 분류, 제9차 개정 (ICD-9) 질환 진단 코드를 EHR로부터 추출하고 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 케이스 정의에 대한 임상 질환 카테고리로 바꾸었다. 단일 진단 코드를 기반으로 한 추가적인 케이스 정의는 알코올성 간경변 (ICD-9 571.2), 비알코올성 간경변 (ICD-9 571.5), 및 HCC (ICD-9 155.0)를 포함하였다. 이들 케이스 정의에 대해, 간 질환이 없는 공통 대조군 그룹을 케이스 참조가 없거나 임의의 유형의 간 질환을 가리키는 단일-충돌 또는 문제-목록 진단 코드가 없는 참가자로서 정의하였다.

[0510] ***베리아트릭 수술 집단에서 간 조직병리학적 표현형 정의***

[0511] GHS 베리아트릭 수술 집단은 유럽 혈통의 2,644명의 개인으로 구성되었다. 간의 웨지 생검 시편을 이들 개인 중 2,391명으로부터 수술 중에 얻었다. 생검 시편을 임의의 간 철회 또는 위에 대한 수술 전에 겸상 인대의 좌측으로 10 cm에서 일관되게 얻었다. 생검 시편을 섹션들로 나누고, 일차 섹션을 간 조직학을 위해 임상 병리사에게 전달하였고 (10% 중성 완충된 포르말린에 고정되고 기본적인 기본적인 조직학을 위해 헤마톡실린 및 에오신으로 염색되고 섬유증의 평가를 위해 Masson 삼색 염색으로 염색됨), 나머지 섹션은 연구 바이오뱅크 내에 보관하였다 (RNAlater 및/또는 액체 질소에 냉동됨). 간 조직학을 숙련된 병리학자에 의해 수행하였고 그런 후에 숙련된 두 번째 병리학자에 의해 NASH 임상 연구망 채점 시스템 (NASH Clinical Research Network scoring system)((Kleiner et al., *Hepatology*, 2005, 41, 1313-21, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨))을 사용하여 다음과 같이 재검토하였다: 지방증 등급 0 (<5% 실질 포함), 1 (5 내지 <33 %), 2 (34 내지 <66 %), 및 3 (>67 %); 서엽 염증 등급 0 (병소 없음), 등급 1 (경미함, 200X 배율당 <2개 병소), 등급 2 (중간, 200X 배율당 2 내지 4개 병소), 등급 3 (중증, 200X 배율당 >4개 병소); 섬유증 단계 0 (없음), 단계 1 (동양혈관주위(perisinusoidal) 또는 문맥주위(periportal) 섬유증), 단계 2 (동양혈관주위 또는 문맥주위 섬유증), 단계 3 (가교성 섬유증), 및 단계 4 (간경변). 이들 조직학적 진단을 다음의 표현형을 정의하는 데 사용하였다: 1) 정상: 지방증, NASH, 또는 섬유증의 증거가 없음; 2) 단순 지방증: NASH 또는 섬유증의 증거가 없는 지방증 (등급 무시함); 3) NASH: 소엽의 염증 또는 간세포 팽창의 임의의 존재 (등급 무시함), 또는 섬유증의 임의의 존재 (등급 무시함); 4) 섬유증: 섬유증의 임의의 존재 (등급 무시함).

[0512] ***샘플 제조, 서열분석, 및 유전자형 분류***

[0513] DiscovEHR 연구, 댈러스 심장 연구, 및 Penn 약물 바이오뱅크의 참가자에 대한 DNA 샘플 제조 및 전 엑솜 서열분석을 리제너론 제네틱스사에서 수행하였다 (Dewey et al., *Science* In Press, 2016, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). HSD17B13 rs72613567을 댈러스 간 연구 및 댈러스 소아 간 연구에서 Taqman 검정에 의해 유전자형을 분류하였다 (그리고 각 유전자형의 5명의 개인에서 Sanger 서열분석에 의해 확인하였다).

[0514] 특히, 엑솜 포획을 NimbleGen 프로브를 사용하여 제조사의 권장 프로토콜 (Roche NimbleGen)을 따라 수행하였다. 포획한 DNA를 PCR 증폭시키고 qRT-PCR에 의해 정량화하였다 (Kapa BioSystem). 다중화된 샘플을 Illumina v4 HiSeq 2500 상에서 75 bp 쌍-단부 서열분석을 사용하여, 96%의 샘플에서 표적화된 염기의 85% 이상의 20배보다 큰 하플로이드 판독 깊이를 제공하기에 충분한 커버리지 깊이로 서열분석을 하였다 (표적화된 염기의 대략 80배의 평균 하플로이드 판독 깊이). 각 Illumina Hiseq 2500 작동으로부터의 미가공 서열 데이터를 서열 판독 정렬 및 변종 확인을 위해 DNAnexus 플랫폼에 업로드하였다 (Reid et al., *BMC Bioinformatics*, 2014, 15, 30, 본원에 모든 목적에 대해 그 전문이 참조로 포함됨). 간단하게 설명하면, 미가공 서열 데이터를

BCL 파일로부터 샘플-특이적 FASTQ-파일로 변환시켰고, 그것을 BWA-mem과 인간 기준 빌드 GRCh37.p13에 정렬시켰다 (Li et al., *Bioinformatics*, 2009, 25:1754-1760, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 단일 뉴클레오타이드 변종 (SNV) 및 삽입/결실 (indel) 서열 변종을 게놈 분석 툴키트 (McKenna et al., *Genome Res.*, 2010, 20, 1297-303, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨)를 사용하여 확인하였다.

[0515] ***간 효소 및 만성 간 질환 표현형에서 엑솜-전체 연관성 분석***

[0516] 본 발명자들은 502,219개의 이중대립유전자 변종을 테스트하기 위하여, 아미노전이효소 수준과의 연관성에 대해 1% 미만의 데이터 전송 속도, $1.0x10^{-6}$보다 큰 Hardy-Weinberg 평형 p-값, 및 0.1%보다 큰 마이너 대립유전자 빈도를 버리는 선형 혼합 모델을 사용하였다. GHS 개발 집단에서 아미노전이효소 ($p<1x10^{-7}$)와의 전체 엑솜 유의성 연관성을 가진 변종에 대하여, 본 발명자들은 상기 기술된 유럽-혈통 복제 연구에서 연관성 분석 및 메타-분석을 수행하였다. 본 발명자들은 복제된 연관성을 정의하기 위하여 테스트한 많은 변종에 의해 측정된 본페로니 유의성 한계값을 사용하였다. 개발 및 복제 연구의 메타-분석을 또한 수행하였다. 모든 P-값을 대립유전자 모델에 대응한 텍스트에 기록하였다.

[0517] 본 발명자들은 그 후에 만성 간 질환 표현형과의 연관성에 대해 아미노전이효소-관련 단일 뉴클레오타이드 변종을 테스트하였다. 본 발명자들은 연관성의 유의성을 측정하기 위하여 테스트된 많은 변종 및 광범위한 만성 간 질환 카테고리에 의해 측정된 본페로니 유의성 한계값을 사용하였다. 본 발명자들은 GHS 베리아트릭 수술 집단으로부터의 조직병리학적으로 정의된 간 표현형과의 연관성에 대해 복제된 신규한 변종을 추가로 테스트하였다. 본 발명자들은 또한 405개의 정량적 임상 측정값 및 3,168개의 임상 진단으로 복제된 신규한 변종의 연관성의 표현형-전체 연구를 수행하였다.

[0518] 특히, 본 발명자들은 502,219개의 이중대립유전자 변종을 테스트하였고, 아미노전이효소 수준과의 연관성에 대해 1% 미만의 데이터 전송 속도, $1.0x10^{-6}$보다 큰 Hardy-Weinberg 평형 p-값, 및 0.1%보다 큰 마이너 대립유전자 빈도를 버렸다. Log_{10}-변형된 중앙 ALT 및 AST를 연령, 연령^2, 성별, BMI, 및 혈통의 제1 4대 주요 구성요소에 대해 조정하였다. 연구 참가자 사이의 관련성을 설명하기 위하여, 본 발명자들은 또한 유전자 관련성 기질을 랜덤 효과 공변량으로서 맞추었다. 주요 구성요소 및 유전자 관련성 기질을 둘 다 대략적인 결합 평형으로 및 빈도가 0.1%보다 큰 마이너 대립유전자로 39,858개의 비-MHC 마커로부터 구성하였다. 본 발명자들은 특성 잔차와 단일 뉴클레오타이드 변종 사이의 연관성에 대해 테스트하기 위하여 GCTA 패키지에서 실행된 것과 같은 선형 혼합 모델을 사용하였다 (Yang et al., *Am. J. Hum. Genet.*, 2011, 88, 76-82, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 모든 P-값을 대립유전자 모델에 대응한 텍스트에 기록하였다.

[0519] 본 발명자들은 3개의 별도의 유럽-혈통 집단: GHS 베리아트릭 수술 집단, 댈러스 심장 연구, 및 Penn 약물 바이오뱅크 (상기 기술됨)의 GHF 개발 집장세어 연관성을 복제하고자 시도하였다. GHS 베리아트릭 수술 집단 및 Penn 약물 바이오뱅크로부터의 ALT 및 AST 측정치를 log_{10}-변형시키고 연령, 연령^2, 성별, BMI, 및 혈통의 제1 4대 주요 구성요소에 대해 조정하였다. 유전자 관련성 기질을 랜덤 효과 공변량으로서 포함시켰고, 분석을 GCTA에서의 선형 혼합 모델을 사용하여 수행하였다. 댈러스 심장 연구에서, log_{10}-변형된 ALT 및 AST 측정치를 연령, 연령^2, 성별, BMI, 및 혈통의 제1 10대 주요 구성요소에 대해 조정하였고, 분석을 PLINK에서 실행된 선형 회귀를 사용하여 수행하였다. 3개의 복제 집단에 대한 요약 통계를 METAL (Willer et al., *Bioinformatics*, 2010, 26, 2190-1, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨)을 사용하여 메타-분석하였다 (반복 메타-분석). 개발 집단 및 3개의 복제 집단에 대한 요약 통계를 유사하게 메타-분석하였다 (연합 메타-분석).

[0520] ***만성 간 질환 표현형과의 연관성 분석***

[0521] 본 발명자들은 상기 기술한 것과 같이, GHS 개발 집단으로부터 정의된 만성 간 질환 표현형과의 연관성에 대해 간 효소 ExWAS로부터의 13개의 중요한 및 복제된 단일 뉴클레오타이드 변종을 분석하였다. 본 발명자들은 테스트한 13개의 변종 및 2개의 광범위한 만성 간 질환 카테고리 (알코올성 및 비알코올성)을 설명하기 위하여 $P<0.05/26$ ($P<1.92x10^{-3}$)의 본페로니 유의성 한계값을 사용하였다. HSD17B13 rs72613567 변종을 상기 기술한 것과 같이, GHS 베리아트릭 수술 집단으로부터의 조직병리학적으로 정의된 간 표현형과의 연관성에 대해 추가로 테스트하였다. 오즈비를 연령, 연령^2, 성별, BMI, 및 혈통의 제1 4대 주요 구성요소에 대한 조정 후에 로지스틱 회귀의 Firth의 벌칙 가능성 방법을 사용하여 추정하였다. 유전자형 오즈비를 동일한 공변량을 사용하여

HSD17B13 rs72613567에 대해 추정하였다.

[0522] DLS에서 간 질환에 대한 오즈비를 로지스틱 회귀에 의해 추정하고, 연령, 연령^2, 성별, 체질량 지수, 및 자가-보고된 민족에 대해 조정하였다. 이용 가능한 rs72613567 유전자형을 가진 댈러스 심장 연구로부터의 참가자를 정상 대조군으로서 사용하였다 (n=4,279명). DPLS에서의 오즈비를 로지스틱 회귀에 의해 추정하였다.

[0523] ***HSD17B13 rs72613567의 표현형-전체 연관성 연구***

[0524] 본 발명자들은 405개의 정량적 EHR-유래 인체측정, 바이탈 TTLS호 실험실, 심전도, 심장초음파검사, 및 골밀도 측정값과, 및 또한 3,168개의 EHR-유래 임상 진단과 HSD17B13 rs72613567의 연관성의 표현형-전체 연구를 수행하였다. 연속적인 외래환자 측정치를 가진 개인에 대한 중앙 실험실 값을 별개의 중앙값 내부로부터 3 표준 편차보다 큰 가능한 거짓 값의 제거 후에 계산하였다; 최대 및 최소 값을 또한 계산하였다. 본 발명자들은 그런 다음 연령, 연령^2, 성별, 및 혈통의 제1 10대 주요 구성요소에 대해 조정 후 모든 실험실 특성에 대해 특성 잔차를 계산하였고, 적절한 형질전환을 적용한 후에 연관성 분석을 하였다. ICD-9 기반 진단 코드를 Denny 등에 의해 제안된 그룹화의 변형 버전을 사용하여 계층에 따른 임상 질환 그룹에 대해 바꾸었다 (Denny et al., *Nature Biotechnology*, 2013, 31, 1102-10 및 Denny et al., *Bioinformatics*, 2010, 26, 1205-10, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). ICD-9 기반 진단은 다음 중 하나 이상을 필요로 하였다: 진단 코드의 문제 목록 유입 또는 별도의 역일(calendar days)에 2개의 별도의 임상 충돌에 대해 유입된 충돌 진단 코드.

[0525] 변형된 정량적 임상 측정값 잔차와의 연관성 분석을 선형 회귀를 사용하여 수행하였고, 임상 진단과의 연관성 분석을 연령, 연령^2, 성별, 및 제1 4대 주요 구성요소에 대해 조정된 로지스틱 회귀를 사용하여 수행하였다. 대립유전자를 부가 (참조 대립유전자 동형접합체의 경우 0, 이형접합체의 경우 1, 및 대체 대립유전자 동형접합체의 경우 2) 및 열성 (참조 대립유전자 동형접합체 및 이형접합체의 경우 0, 대체 대립유전자 동형접합체의 경우 1) 모델 둘 다를 사용하여 코딩하였다.

[0526] ***소프트웨어***

[0527] 유전자 연관성 분석을 GCTA 소프트웨어, 버전 1.25.07 및 PLINK, 버전 1.9.0을 사용하여 수행하였다. 사분위-사분위 및 맨하탄 도표를 R 소프트웨어, 버전 3.2.1 (통계학적 컴퓨터 계산을 위한 R 프로젝트)을 사용하여 생성하였다. 지역적 연관성 도표를 LocusZoom (Pruim et al., *Bioinformatics*, 2010, 26, 2336-7, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨)을 사용하여 생성하였다.

[0528] ***RNA 서열분석 연구***

[0529] RNA 품질 및 농도를 총 RNA를 Agilent RNA Nano Bioanalyzer 칩 상에서 이동시킴으로서 평가하였다; 모든 샘플은 8보다 큰 RNA 무결성 번호 (RIN)를 가졌다. 폴리아데닐화된 RNA 전사물을 올리고(dT)25 비즈 (Thermo Fisher Scientific)로의 2회기의 풍부화를 사용하여 분리하였다. 샘플을 RNAclean XP 비즈 (Beckman Coulter)로 정제 및 농축하고 대략 140 염기쌍으로 가열-단편화하였다. 제1 가닥 합성을 SuperScript III 역전사효소 (Thermo Fisher Scientific)로 무작위 헥사머를 사용하여 완료하였다; dTTP는 제2 가닥 합성 중에 dUTP로 교체되었다. 샘플을, 가닥-특이적 서열분석 라이브러리를 생성하기 위하여 우라실 DNA-글리코실라아제 단계를 첨가하여, 엑솜에 대해 상기 참조한 본 발명자들의 표준 DNA 라이브러리 제조 방법에 따라 처리하였다.

[0530] ***신규한 HSD17B13 전사물의 확인 및 입증***

[0531] 판독을 2개의 미스매치를 허용하는 ArrayStudio® 소프트웨어 (OmicSoft®, Cary, NC)를 사용하여 인간.B38에 대해 지도화하였다. 신규한 HSD17B13 전사물을 확인하기 위하여 2가지 접근법을 사용하엿다. 신규한 엑손 접합부를 ArrayStudio를 사용하여 Gencode v24를 기반으로 발견하였다. 데노보 전사물 어셈블리를 디폴트 세팅에서 트리니티 (v2.2.0)를 사용하여 수행하였다. 주문 유전자 모델을 HSD17B13의 신규한 전사물을 포함시키기 위하여 구성하였고, 전사물 정량화를 주문 유전자 모델에 대한 판독 정렬에 의해 추정하였다. 모든 확인된 HSD17B13 동형의 단백질 서열 정렬을 **도 7A** 및 **7B**에 나타낸다. 인간 간 샘플로부터의 총 RNA에 대한 RT-PCR을 PlatinumTM Taq DNA 중합효소 (Thermofisher)가 포함된 SuperScriptTM 원스텝 RT-PCR 시스템을 사용하여 수행하였다. 각각 50 μL의 RT-PCR 반응은 1X 반응 믹스, 500 nM의 각각의 전방향 및 역방향 프라이머 (PST516: ATGAACATCATCCTAGAAATCCTTC (SEQ ID NO: 251) 및 PST517: ATCATGCATACATCTCTGGCTGGAG (SEQ ID NO: 252)), 1 μL의 RT/Platinum Taq, 및 75 ng의 RNA를 함유하였다. 사이클 조건은 다음과 같았다; 45℃에서 30분 동안 1 사이클; 94℃에서 2분 동안 1 사이클; 94℃에서 20초 동안, 53℃에서 30초 동안, 및 72℃에서 90초 동안, 40 사

이클; 72℃에서 5분 동안 1 사이클; 그런 후 10℃에서 유지함. 생성물을 QIAquick PCR 정제 키트 (Qiagen)를 사용하여 정제하고 프라이머 DE002 (ATCAGAACTTCAGGCCTTGG (SEQ ID NO: 253))를 사용하여 직접 Sanger 서열분석을 위해 제출하였다. B 및 C 전사물을 확인하기 위하여, RT-PCR 생성물을 2% 아가로오스 겔 상에서 이동하도록 하고 SYBR GOLDSYBR® 골드 핵산 겔 염색 (Thermofisher)으로 염색한 후, 예상했던 분자량의 밴드를 절제하여 QIAquick 겔 추출 키트 (Qiagen)를 사용하여 정제한 후, TOPO® TA 클로닝 키트 (Thermofisher)로 클로닝을 수행하였다. TOPO 클론의 서열분석을 M13F 및 M13R 서열분석 프라이머를 사용하여 수행하였다. 서열 분석을 Sequencher DNA 분석 소프트웨어 (Gene Codes Corporation)를 사용하여 수행하였다. 전장 *HSD17B13* 전사물을 50 ng의 총 RNA로부터 백금 Taq 고신뢰도를 가진 SuperScript III 원스텝 RT-PCR 시스템 (ThermoFisher Scientific)으로 제1 엑손 (GCAAAGCCATGAACATCATCC (SEQ ID NO: 254) 및 마지막 엑손 (TCTTGATGTAGTGGGAGTCGGATT (SEQ ID NO: 255))에서의 유전자-특이적 프라이머를 사용하여 직접 증폭시켜서 약 2.2 kb (최대 예상 크기 전사물)의 암플리콘을 생성하였다. 암플리콘을 Agilent Bioanalyzer 상에서 입증하였다. PacBio-부합하는 바코딩된 어댑터를 암플리콘에 결찰시키고 PacBio PB 비즈 (Pacific Biosciences)로 세정하였다. 라이브러리를 동등한 양으로 모으고 한 SMRT 세포에 대해 180분 동안 PacBio RSII 플랫폼에서 서열분석을 하였다. 데이터를 PacBio 소프트웨어 스마트분석 v2.3 도구 labelzmw를 사용하여 탈다중화한 후 콘센서스 도구 암플리콘 분석으로 분석하였다. 그 결과의 암플리콘을 *HSD17B13* RefSeq 유전자와 비교하여 동형 및 유전자형 상태를 측정하였다.

[0532] ***HSD17B13 동형의 세포내 국지화***

[0533] HepG2 세포를 10% 소태아 혈청이 첨가된 이글스 최소 필수 배지에서 배양하였다. HSD17B13 전사물 A 및 D를 Myc-DDK 골격 렌티바이러스 구성물 안으로 하위클로닝하고, 렌티바이러스를 생성시켰다. HepG2 세포를 HSD17B13 전사물을 운반하는 렌티바이러스로 감염시켰다. 각각의 HSD17B13 전사물을 발현하는 안정적인 세포주를 완전 배양 배지에서 2주 동안 1 내지 3 mg/ml의 제네티신 G-418 설페이트로 선택하였다. 고정 후에, HSD17B13 동형을 마우스 항-Myc 항체로 검출하였다. 지질 방울을 BODIPY FL 염료 (Sigma)로 표지하였다. 면역형광에 대한 이차 항체는 Alexa Fluor 488 당나귀 항-토끼 IgG 및 Alexa Fluor 594 당나귀 항-마우스 IgG (Jackson ImmunoResearch)였다.

[0534] ***인간 간 생검 조직 및 안정적인 세포주에서 HSD171B3 단백질 발현의 정량화***

[0535] 인간 간 및 세포 펠릿 샘플을 극저온 1x RIPA 억제제 완충액 (EMD Millipore)에서 프로테아제 및 포스파타아제 억제제 혼합물 (ThermoFisher)의 존재 하에 균질화하였다. 상층액을 수집하고 BCA 단백질 검정 (ThermoFisher)을 사용하는 단백질 농도에 대해 사용하였다. 인간 조직 및 세포 용해물을 SDS/PAGE 겔 (Bio-Rad) 상에 로딩하고 분리시킨 후 PVDF 막 (Bio-Rad)에 옮겼다. 막을 0.1% Tween20 (Bio-Rad)이 첨가된 1 x TBS에서 1시간 동안 5% (중량/부피) 밀크로 차단하였다. 막을 항체와 함께 4℃에서 밤새 HSD17B13에 대해 (1:200, Thermo-Fisher) 및 B-액틴에 대해 (1:500, Cell Signaling Technology) 인큐베이션하였다. 결합한 항체를 HRP-콘쥬게이트된 항-토끼 항체 (1:10,000, Jackson ImmunoResearch)를 사용하여 검출하고 화학 발광 시약 (ThermoFisher)을 사용하여 증강시켰다. 밴드 세기를 Image J 소프트웨어를 사용하여 정량화하였다.

[0536] ***실시간 반정량적 PCR***

[0537] RNA를 세포로부터 TRIzol® (Invitrogen, Carlsbad, CA)을 사용하여 추출하였다. 제1 가닥 cDNA를 Superscript III RT (Invitrogen)을 사용하여 합성하고 인트론-회전 프라이머를 기반으로 반정량적 PCR에 이용하였다. QuantStudio 6 Flex 실시간 PCR 시스템을 사용하여 전사물의 발현 수준을 측정하였다. HSD17B13 및 TBP의 프라이머는 IDT (Integrated DNA Technologies)로부터 주문받았다. 상대적인 유전자 발현을 △△Ct 방법으로 분석하였고, 하우스-키핑 유전자 TBP (△Ct)에 대래 표준화된 발현의 배수-변화를 제공해주었다.

[0538] ***지질 방울 분리 및 웨스턴 블롯팅에 의한 특성화***

[0539] 지질 방울을 앞서 보고된 것과 같이 HSD17B13 전사물 A (IsoA) 또는 전사물 D (IsoD)를 안정적으로 발현하는 HepG2 세포로부터 제조하였다 (Brasaemle DL, Wolins NE. Isolation of lipid droplets from cells by density gradient centrifugation, Current protocols in cell biology 2006; Chapter 3:Unit 3 15 and Ding et al., *Nature Protocols*, 2013, 8, 43-51, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 간단히 설명하면, HSD17B13 IsoA, IsoD를 안정적으로 발현하는 HepG2 세포, 또는 모 라인(parental line)을 1 mM 올레산과 함께 밤새 인큐베이션하였다. 지질 로딩 후에, 세포를 스크랩하고 1X HaltTM 프로테아제/포스파타제 억제제 (Thermo)가 첨가된 저장성 용해 완충액 (20 mM Tris, pH 7.5, 1 mM EDTA)에 재현탁하고,

50 bar에서 8분 동안 공동현상에 의해 용해시켰다. 용해물을 1000 g/4℃에서 원심분리하고, 후핵 상층액 (post-nuclear supernatant, PNS)을 원심분리 튜브에서 2 mL의 최종 부피 및 20%의 농도로 수크로오스와 혼합하였다. 그런 다음 1.5 mL의 5% 수크로오스 및 또 다른 1.5 mL의 저장성 용해 완충액을 용해물 상부에 레이어드하였다. 튜브를 182,000 g/4℃에서 40분 동안 원심분리하고, 지질 방울 (LD) 층을 새로운 튜브로 옮겼다. 튜브 안의 나머지 부피를 흡인하고, 펠릿화된 것 (총 막, TM)을 0.5 mL의 저장성 용해 완충액에 재현탁하였다. PNS, LD, 및 TM 분획을 1x 방사성면역침정 (RIPA) 완충액 (EMD) + NuPAGETM LDS 샘플 완충액 (Thermo) 및 β-메르캅토에탄올과 혼합하고 3시간 동안 37℃에서 음파처리하였다. TM 용해물을 2.5배로 희석하여 PNS를 표준화하였다. 용해물을 4 내지 20%의 SDS-PAGE 겔 (Biorad) 상에서 이동시키고, 트랜스-블롯 (Trans-Blot (Biorad))을 사용하여 저형광 PVDF 막 위로 옮긴 후, 1시간 동안 Odyssey TBS 차단 완충액에서 차단시켰다. 막을 밤새 다음 항체와 함께 인큐베이션하였다: α-HSD17B13 (Abgent, cat # AP5729a 1:500); LD 마커: α-ADRP (Protintech, 152-94-1-AP, 1:2500); LD 마커: α-TIP47 (Proteintech, 10694 1:2000); 리소좀 마커: α-LAMP1 (Novus, NBP2-25183, 1:1000); 세포질 마커: α-GAPDH (Proteintech, 60004-1-Ig, 1:2000); 소포체 마커: α-칼레티쿨린 (Abcam, ab92516, 1:1000); 미토콘드리아 마커: α-COX IV (Abcam, ab33985, 1:500); 세포골격 마커: α-액틴 (Sigma, A5441, 1:4000). 다음날 막을 Tris-완충된 식염수 + 0.1% Tween으로 4회 세척한 후, 1시간 동안 실온에서 각각 1:5,000 및 1:10,000 희석률의 IRDye® α-토끼 (800CW) 및 α-마우스 (680RD) 이차 항체 (Li-Cor)를 함유하는 차단 완충액과 함께 인큐베이션하였다. 겔을 다시 TBST로 세척하고 Odyssey를 사용하여 영상화하였다.

[0540] ***세포내 트라이글리세라이드 함량의 정량화***

[0541] 안정적인 세포로부터의 트라이글리세라이드 (TG) 함량을 TG 정량화 키트 (Abcam)를 사용하여 측정하였다. 검정에서, TG를 유리 지방산 및 글리세롤로 전환시켰다. 그런 다음 글리세롤을 산화시켜서 정량화되는 생성물을 생성하였다 (λ=570 nm에서 분광측정).

[0542] ***정제된 재조합 HSD17B13에 대한 스테로이드 및 생체활성 지질 라이브러리의 기질 스크리닝***

[0543] 500 μM NAD^+, 5 μM 생체활성 지질 또는 50 μM 스테로이드 (모두 최종 농도 5% DMSO 중에 있음), 및 100 ng 재조합 인간 HSD17B13을 함유한 최종 40 μL의 검정 완충액 (0.2 M Tris-HCl, pH 7.5)에서 반응을 수행하였다. 반응을 3시간 동안 23℃에서 인큐베이션한 후, 동일한 부피의 NADH-Glo 검출 시약 (Promega)을 첨가하였다. 1시간 동안 23℃에서 인큐베이션한 후에, 상대적 광 단위 (RLU)를 Envision 플레이트 판독기 (Perkin Elmer) 상에서 측정하였다. 미가공 RLU 값을 다음 식을 사요아여 음성 대조군 (5% DMSO)을 뺀 후에 대조군의 % (50 μM 에스트라다이올)로서 표준화하였다: 대조군의 % (POC) = 100 x (샘플 (RLU) - 음성 CTRL평균) / (양성 CTRL평균 - 음성 CTRL평균).

[0544] ***HSD17B13 효소적 활성의 시험관내 및 세포적 특힝화***

[0545] 재조합 인간 HSD17B13 단백질을 HSD17B13 전사물 A 또는 전사물 D를 은닉한 플라스미드 DNA로 형질전환된 대장균 (Genscript)으로부터 정제하였다. HSD17B13 변종은 C 말단에 10xHis 태그를 함유하였고, 그것을 $Ni2^+$ 친화도 정제를 사용하여 가용성 분획으로부터 정제하였다. 효소적 활성을 NAD(P)H-Glo 검출 시스템 (Promega)을 사용하여 NADH 생성의 측정을 통해 측정하였다. 반응을 3시간 동안 25℃에서 최종 부피 100 μL 중의 0.2 M Tris-HCl, pH 7.5, 0.5mM NAD^+, 75 μM의 기질 (Sigma) 및 500 ng의 정제된 효소에서 수행하였다. 인큐베이션 후에, 20 μL의 반응을 20 μL의 루시페라제 시약 (Promega)과 조합하고, 실온에서 1시간 동안 인큐베이션하고 Envision 플레이트 판독기 (Perkin Elmer) 상에서 판독하였다.

[0546] HSD17B13 전사물 A, 전사물 D 또는 녹색 형광 단백질 (GFP, 대조군)을 과다발현하는 HEK293 세포를 사용하여 세포-기반 검정으로 에스트라다이올에 대한 HSD17B13의 활성을 조사하였다. 에스트라다이올 (1 μM)을 각 세포 유형에 공급하였다. 48시간 후, 배지를 수집하고 에스트라다이올 및 그것의 전환된 생성물인 에스트론의 농도를 확인하고 LC-MS에 의해 정량화하였다.

[0547] ***아스파테이트 및 알라닌 아미노전이효소와 엑손성 변종의 연관성***

[0548] 본 발명자들은 DiscovEHR 연구로부터의 유럽 혈통의 46,544명의 개인에서 혈청 ALT 또는 AST 수준과의 연관성에 대해 502,219개의 이중대립유전자 단일 유전자 변종을 테스트하였다 ("GHS 개발 집단"; **표 6**에서 기본적인 인구통계학). 19개의 유전자에서 총 35개의 변종이 $P<1.0x10^{-7}$에서 ALT 또는 AST와 관련된 것으로 나타났다 (**도 1A** 및 **1B**, 및 **표 7**). 본 발명자들은 유럽 혈통의 세 집단: 1) DiscovEHR로부터의 베리아트릭 수술 환자 (n=2,644명) ("GHS 베리아트릭 수술 집단"); 2) 댈러스 심장 연구로부터의 1,357명의 개인; 및 3) Penn 약물

바이오뱅크로부터의 8,526명의 개인에서 복제 연구를 수행하였다. 복제 집단의 메타-분석에서, 9개의 유전자 중 13개의 변종이 ALT 또는 AST의 혈청 수준과 유의하게 관련이 있었다 (테스트한 35개의 변종에 대해 $P<1.43x10^{-3}$의 본페로니 유의성 한계값, **표 8**). 이것들은 상승된 아미노전이효소 수준, 예컨대 PNPLA37, TM6SF211, SERPINA122, SAMM5023, 및 ERLIN124와 관련된 것으로 이전에 보고된 바 있는 변종을 포함하였다. SERPINA1은 그것의 기능성 결핍이 간 질환을 유발하는 알파-1-항트립신을 암호화한다; SAMM50과의 연관성은 PNPLA3에서의 변이와의 결합 불균형을 통해 매개되며, ERLIN1은 간 지방 침적에 연루되었다. 본 발명자들은 또한 이전에 간 질환과 관련된 것으로 보고된 적이 없는 변종들을 확인하였다. 이것들은 각각 ALT 및 AST를 암호화하는 유전자인 GPT 및 GOT1에서의 몇몇 변종, 및 용질 보균자 패밀리 39 구성원 12를 암호화하는 SLC39A12를 포함하였다.

[0549] 본 발명자들은 또한 HSD17B13에서의 변종, 하이드록시스테로이드 17-베타 탈수소효소 13을 암호화하는 유전자, 17-베타 하이드록시스테로이드 탈수소효소 패밀리의 미확인 구성원, 및 감소된 수준의 ALT (개발 $P=4.2x10^{-12}$, 복제 $P=1.7x10^{-4}$) 및 AST (개발 $P=6.2x10^{-10}$, 복제 $P=1.7x10^{-4}$, **표 8**) 사이의 재생가능한 연관성을 확인하였다. 관련된 변종, rs72613567은 엑손 6 (TA 대립유전자)의 도너 스플라이스 부위에 인접하여 아데닌이 삽입된 것이고, GHS 개발 집단에서 26.0%의 대립유전자 빈도를 가졌다. 앞서, Chambers 등은 ALT 수준과 관련된 4q22 (rs6834314)에서 가가운 유전자좌를 확인하였고 (Chambers et al., *Nat. Genet.*, 2011, 43, 1131-1138, doi:10.1038/ng.970, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨); rs72613567은 지금까지 아미노전이효소 수준과 관련된 것으로 보고되지 않은 것이었다. HSD17B13은 동일한 유전자 패밀리의 또 다른 구성원인 HSD17B11의 30 kb 상류에 있다. 본 발명자들은 HSD17B11의 코딩 또는 스플라이스 변종과 개발 집단에서의 (도 **5A** 및 **5B**) 또는 개발 집단 및 3개의 복제 집단의 연합 메타-분석에서의 아미노전이효소 수준 사이의 엑솜-전체 유의한 연관성을 관찰하지 못하였다. 나아가, HSD17B11에서의 변종과 rs72613567의 결합 불균형은 모든 혈통 그룹을 가로질러 가장 적당하였다 (모든 혈통 그룹에서 HSD17B11에서의 모든 확실한 변종과 $r^2<0.4$). 종합적으로, 이들 발견은 HSD17B13이 아미노전이효소 수준과 기능적으로 관련되어 있을 가능성이 가장 높은 게놈 영역에서의 유전자인 것을 시사한다.

표 6

[0550] **개발 및 복제 집단으로부터의 서열분석된 유럽-혈통 개인의 인구통계학적 및 임상적 특징.**

특성	개발 집단 (N = 46,544명)	베리아트릭 수술 집단 (N = 2,644명)	밸러스 심장 연구 (N = 1,357명)	Penn 약물 바이오뱅크 (N = 8,526명)
연령 (나이) - 중앙 (IQR)	62.9 (49.6 - 73.8)	52.9 (44.1 - 61.2)	46.0 (38.0 - 54.0)	68.0 (60.0 - 76.0)
여성 성별 - 수 (%)	26,875 (57.7)	2,119 (80.1)	724 (53.4)	3,242 (38.0)
체질량 지수 - 중앙 (IQR)	29.9 (35.4 - 44.8)	47.4 (42.0 - 53.7)	28 (25-32)	30 (25-32)
아미노전이효소 수준 (U/L) - 중앙 (IQR)				
알라닌 아미노전이효소 (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 아미노전이효소 (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 (ICD-9 코드에 의함) - 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)	-	-
간 질환 없음	30,628 (65.8)	1 (0.04)	-	-

표 7

[0551] **개발 집단에서 $P < 1.0x10^{-7}$에서의 혈청 아미노전이효소 수준과 관련된 단일 뉴클레오타이드 변종.**

특성	CHR	BP	REF	ALT	rsID	유전자	주석	AA 치환	베타 (SE)
ALT	1	220970028	A	G	rs2642438	*MARC1*	미스센스	p.Thr165Ala	0.008 (0.001)
	4	88231392	T	TA	***rs72613567**	*HSD17B13*	스플라이스 도너		-0.009 (0.001)
	8	144997604	C	T	rs371119003	*PLEC*	미스센스	p.Ala2302Thr	-0.160 (0.026)
	8	145008502	G	A		*PLEC*	미스센스	p.Arg522Cys	-0.268 (0.032)
	8	145692918	G	A	rs35968570	*KIFC2*	미스센스	p.Glu174Lys	-0.033 (0.005)
	8	145730072	G	A	rs143408057	*GPT*	미스센스	p.Arg83His	-0.314 (0.036)
	8	145730161	C	T	rs201815297	*GPT*	미스센스	p.Ala87Val	-0.224 (0.014)
	8	145730221	G	A	rs112574791	*GPT*	미스센스	p.Arg107Lys	-0.033 (0.005)
	8	145731636	T	G	rs145155876	*GPT*	중단 획득	p.Tyr326*	-0.235 (0.031)
	8	145732114	G	C	rs141505249	*GPT*	미스센스	p.Glu430Gln	-0.224 (0.013)
	8	145732151	G	A	rs143462595	*GPT*	미스센스	p.Arg442His	-0.077 (0.013)
	8	145732180	G	C	rs147998249	*GPT*	미스센스	p.Val452Leu	-0.225 (0.013)
	8	145732305	G	GC		*GPT*	프레임시프트	p.Glu475fs	-0.271 (0.031)
	8	145748532	A	G	rs567402720	*LRRC24*	미스센스	p.Leu290Ser	-0.185 (0.028)
	9	117122202	C	T	rs3748177	*AKNA*	동의적	p.Glu755Glu	-0.007 (0.001)
	9	117124731	G	A	rs3748176	*AKNA*	미스센스	p.Pro624Leu	-0.007 (0.001)
	10	101595996	T	A	rs17222723	*ABCC2*	미스센스	p.Val1188Glu	-0.015 (0.003)
	10	101606861	G	T	rs1137968	*ABCC2*	동의적	p.Val1430Val	-0.015 (0.003)
	10	101610533	C	T	rs8187707	*ABCC2*	동의적	p.His1496His	-0.015 (0.003)
	10	101611294	G	A	rs8187710	*ABCC2*	미스센스	p.Cys1515Tyr	-0.015 (0.003)
	10	101912064	T	C	***rs2862954**	*ERLIN1*	미스센스	p.Ile291Val	-0.012 (0.001)
	10	101977883	C	T	rs2230804	*CHUK*	미스센스	p.Val268Ile	-0.009 (0.001)
	10	113917085	T	A	rs2254537	*GPAM*	동의적	p.Pro681Pro	-0.008 (0.001)
	10	113940329	T	C	rs2792751	*GPAM*	미스센스	p.Ile43Val	-0.008 (0.001)
	14	94844947	C	T	***rs28929474**	*SERPINA1*	미스센스	p.Glu366Lys	0.042 (0.005)
	19	19379549	C	T	***rs58542926**	*TM6SF2*	미스센스	p.Glu167Lys	0.014 (0.002)
	22	44324727	C	G	***rs738409**	*PNPLA3*	미스센스	p.Ile148Met	0.023 (0.002)
	22	44324730	C	T	***rs738408**	*PNPLA3*	동의적	p.Pro149Pro	0.023 (0.002)
	22	44342116	A	G	rs2294918	*PNPLA3*	미스센스	p.Lys434Glu	0.007 (0.001)
	22	44368122	A	G	***rs3761472**	*SAMM50*	미스센스	p.Asp110Gly	0.019 (0.002)
	22	44395451	T	C	***rs1007863**	*PARVB*	미스센스	p.Trp37Arg	0.011 (0.001)

AST	4	88231392	T	TA	*rs72613567	*HSD17B13*	스플라이스 도너		-0.005 (0.001)
	10	18242311	A	G	rs10764176	*SLC39A12*	미스센스	p.Ser36Gly	-0.006 (0.001)
	10	101157378	CGT	C		*GOT1*	프레임내 indel	p.Asn389del	-0.221 (0.024)
	10	101165533	G	C	rs374966349	*GOT1*	미스센스	p.Gln208Glu	0.271 (0.027)
	10	101912064	T	C	*rs2862954	*ERLIN1*	미스센스	p.Ile291Val	-0.005 (0.001)
	11	22271870	A	T	rs7481951	*ANO5*	미스센스	p.Leu322Phe	0.004 (0.001)
	14	94844947	C	T	*rs28929474	*SERPINA1*	미스센스	p.Glu366Lys	0.027 (0.003)
	19	19379549	C	T	*rs58542926	*TM6SF2*	미스센스	p.Glu167Lys	0.008 (0.002)
	22	44324727	C	G	*rs738409	*PNPLA3*	미스센스	p.Ile148Met	0.014 (0.001)
	22	44324730	C	T	*rs738408	*PNPLA3*	동의적	p.Pro149Pro	0.014 (0.001)
	22	44368122	A	G	*rs3761472	*SAMM50*	미스센스	p.Asp110Gly	0.011 (0.001)
	22	44395451	T	C	*rs1007863	*PARVB*	미스센스	p.Trp37Arg	0.006 (0.001)

【표 7-계속】

								N			평균 AST 또는 ALT 수준 (U/L)		
특성	CHR	BP	REF	ALT	P	AAF	N	REF/REF	REF/ALT	ALT/ALT	REF/REF	REF/ALT	ALT/ALT
	1	220970028	A	G	4.67E-08	0.7067	41,414	3,515	17,262	20,637	23.88	24.52	24.92
	4	88231392	T	TA	4.16E-12	0.2634	41,414	22,441	16,130	2,843	25.02	24.26	24.1
	8	144997604	C	T	1.30E-09	0.0005	41,413	41,373	40	0	24.67	18.1	NA
	8	145008502	G	A	3.26E-17	0.0003	41,414	41,387	27	0	24.67	13.8	NA
	8	145692918	G	A	1.40E-11	0.0139	41,414	40,271	1,133	10	24.67	12.07	NA

[0552]

ALT	8	145730072	G	A	3.28E-18	0.0003	41,414	41,393	21	0	24.67	12.07	NA
	8	145730161	C	T	6.28E-59	0.0018	41,414	41,270	144	0	24.7	14.68	NA
	8	145730221	G	A	4.25E-11	0.0136	41,414	40,293	1,111	10	24.71	23.09	18.35
	8	145731636	T	G	1.76E-14	0.0004	41,394	41,364	30	0	24.67	14.07	NA
	8	145732114	G	C	8.84E-64	0.0019	41,375	41,223	150	2	24.7	14.48	13.75
	8	145732151	G	A	1.18E-09	0.0021	41,406	41,232	174	0	24.68	20.87	NA
	8	145732180	G	C	8.19E-65	0.0019	41,413	41,254	159	0	24.7	14.74	NA
	8	145732305	G	GC	1.00E-18	0.0004	41,414	41,385	29	0	24.67	14.24	NA
	8	145748532	A	G	3.42E-11	0.0004	41,393	41,358	35	0	24.67	17.71	NA
	9	117122202	C	T	9.51E-09	0.5232	41,414	9,414	20,645	11,355	25.12	24.72	24.18
	9	117124731	G	A	4.31E-09	0.5230	41,412	9,427	20,634	11,351	25.12	24.73	24.17
	10	101595996	T	A	2.97E-08	0.0608	41,414	36,543	4,704	167	24.77	23.97	22.12
	10	101606861	G	T	2.71E-08	0.0608	41,414	36,543	4,704	167	24.77	23.97	22.04
	10	101610533	C	T	2.77E-08	0.0608	41,414	36,542	4,706	166	24.77	23.97	22.03
	10	101611294	G	A	2.15E-08	0.0611	41,414	36,519	4,726	169	24.77	23.97	21.99
	10	101912064	T	C	2.43E-21	0.4755	41,414	11,318	20,819	9,277	25.32	24.71	23.77
	10	101977883	C	T	1.93E-13	0.5072	41,414	10,048	20,733	10,633	25.18	24.75	24.01
	10	113917085	T	A	4.61E-10	0.7073	41,414	3,627	16,984	20,803	25	24.97	24.36
	10	113940329	T	C	2.54E-10	0.7097	41,412	3,567	16,910	20,935	25	24.98	24.35
	14	94844947	C	T	9.28E-21	0.0171	41,414	40,006	1,399	9	24.58	26.91	43.89
	19	19379549	C	T	4.76E-09	0.0759	41,413	35,388	5,780	245	24.52	25.46	26.84
	22	44324727	C	G	1.34E-50	0.2351	41,414	24,257	14,837	2,320	24.06	24.99	28.91
	22	44324730	C	T	1.11E-50	0.2349	41,414	24,273	14,824	2,317	24.06	24.98	28.92
	22	44342116	A	G	8.26E-08	0.5986	41,412	6,691	19,833	14,888	24.15	24.47	25.15
	22	44368122	A	G	8.85E-30	0.1682	41,413	28,626	11,618	1,169	24.23	25.36	28.45
	22	44395451	T	C	7.98E-16	0.3963	41,414	15,036	19,920	6,458	24.15	24.6	26.09
[0553]	4	88231392	T	TA	6.24E-10	0.2638	40,753	22,068	15,870	2,815	24.47	24.1	23.96

AST	10	18242311	A	G	1.09E-10	0.2881	40,753	20,645	16,738	3,370	24.47	24.15	23.85
	10	101157378	CGTT	C	1.96E-20	0.0002	40,753	40,733	20	0	24.29	14.7	NA
	10	101165533	G	C	2.43E-24	0.0002	40,753	40,736	17	0	24.28	44.5	NA
	10	101912064	T	C	4.82E-09	0.4754	40,753	11,138	20,486	9,129	24.59	24.26	23.99
	11	22271870	A	T	9.61E-08	0.5833	40,722	7,123	19,686	13,913	24.03	24.22	24.53
	14	94844947	C	T	2.44E-20	0.0172	40,753	39,361	1,384	8	24.24	25.76	34.5
	19	19379549	C	T	6.54E-08	0.0760	40,752	34,811	5,698	243	24.21	24.74	25.43
	22	44324727	C	G	8.31E-46	0.2343	40,753	23,889	14,622	2,242	23.96	24.48	26.62
	22	44324730	C	T	8.93E-46	0.2341	40,753	23,905	14,609	2,239	23.96	24.47	26.63
	22	44368122	A	G	1.22E-22	0.1680	40,752	28,170	11,450	1,132	24.07	24.64	26.24
	22	44395451	T	C	1.31E-13	0.3961	40,753	14,761	19,678	6,314	24.02	24.23	25.1

[0554]

[0555] * ALT 및 AST 둘 다와 엑솜-전체 유의한 연관성을 가진 변종을 나타낸다.

[0556] 약어: AAF, 대체 대립유전자 빈도; Alt, 대체 대립유전자; ALT, 알라닌 아미노전이효소; AST, 아스파테이트 아미노전이효소; Ref, 참조 대립유전자; SE, 표준 오차.

표 8

[0557] **3개의 별도의 유럽-혈통 집단에서 개발 집단으로부터의 35개의 엑솜-전체**

유의한 단일 뉴클레오타이드 변종의 복제 및 연합 메타-분석.

										GHS 개발 집단		
특성	Chr	BP	Ref	Alt	RSID	유전자	Ann	AA 치환	Beta (SE)	P	N
ALT	1	220970028	A	G	rs2642438	*MARC1*	mis	p.Thr165Ala	0.008 (0.001)	4.67E-08	41,414
	4	88231392	T	TA	rs72613567	*HSD17B13*	spl		-0.009 (0.001)	4.16E-12	41,414
	8	144997604	C	T	rs371119003	*PLEC*	mis	p.Ala2302Thr	-0.160 (0.026)	1.30E-09	41,413
	8	145008502	G	A		*PLEC*	mis	p.Arg522Cys	-0.268 (0.032)	3.26E-17	41,414
	8	145692918	G	A	rs35968570	*KIFC2*	mis	p.Glu174Lys	-0.033 (0.005)	1.40E-11	41,414
	8	145730072	G	A	rs143408057	*GPT*	mis	p.Arg83His	-0.314 (0.036)	3.28E-18	41,414
	8	145730161	C	T	rs201815297	*GPT*	mis	p.Ala87Val	-0.224 (0.014)	6.28E-59	41,414
	8	145730221	G	A	rs112574791	*GPT*	mis	p.Arg107Lys	-0.033 (0.005)	4.25E-11	41,414
	8	145731636	T	G	rs145155876	*GPT*	stop	p.Tyr326*	-0.235 (0.031)	1.76E-14	41,394
	8	145732114	G	C	rs141505249	*GPT*	mis	p.Glu430Gln	-0.224 (0.013)	8.84E-64	41,375
	8	145732151	G	A	rs143462595	*GPT*	mis	p.Arg442His	-0.077 (0.013)	1.18E-09	41,406
	8	145732180	G	C	rs147998249	*GPT*	mis	p.Val452Leu	-0.225 (0.013)	8.19E-65	41,413
	8	145732305	G	GC		*GPT*	fs	p.Glu475fs	-0.271 (0.031)	1.00E-18	41,414
	8	145748532	A	G	rs567402720	*LRRC24*	mis	p.Leu290Ser	-0.185 (0.028)	3.42E-11	41,393
	9	117122202	C	T	rs3748177	*AKNA*	syn	p.Glu755Glu	-0.007 (0.001)	9.51E-09	41,414
	9	117124731	G	A	rs3748176	*AKNA*	mis	p.Pro624Leu	-0.007 (0.001)	4.31E-09	41,412
	10	101595996	T	A	rs17222723	*ABCC2*	mis	p.Val1188Glu	-0.015 (0.003)	2.97E-08	41,414
	10	101606861	G	T	rs1137968	*ABCC2*	syn	p.Val1430Val	-0.015 (0.003)	2.71E-08	41,414
	10	101610533	C	T	rs8187707	*ABCC2*	syn	p.His1496His	-0.015 (0.003)	2.77E-08	41,414
	10	101611294	G	A	rs8187710	*ABCC2*	mis	p.Cys1515Tyr	-0.015 (0.003)	2.15E-08	41,414
	10	101912064	T	C	rs2862954	*ERLIN1*	mis	p.Ile291Val	-0.012 (0.001)	2.43E-21	40,834
	10	101977883	C	T	rs2230804	*CHUK*	mis	p.Val268Ile	-0.009 (0.001)	1.93E-13	41,414
	10	113917085	T	A	rs2254537	*GPAM*	syn	p.Pro681Pro	-0.008 (0.001)	4.61E-10	41,414
	10	113940329	T	C	rs2792751	*GPAM*	mis	p.Ile43Val	-0.008 (0.001)	2.54E-10	41,412

[0558]

									GHS 개발 집단		
특성	Chr	BP	Ref	Alt	RSID	유전자	Ann	AA 치환	Beta (SE)	P	N
	14	94844947	C	T	rs28929474	*SERPINA1*	mis	p.Glu366Lys	0.042 (0.005)	9.28E-21	41,414
	19	19379549	C	T	rs58542926	*TM6SF2*	mis	p.Glu167Lys	0.014 (0.002)	4.76E-09	41,413
	22	44324727	C	G	rs738409	*PNPLA3*	mis	p.Ile148Met	0.023 (0.002)	1.34E-50	41,414
	22	44324730	C	T	rs738408	*PNPLA3*	syn	p.Pro149Pro	0.023 (0.002)	1.11E-50	41,414
	22	44342116	A	G	rs2294918	*PNPLA3*	mis	p.Lys434Glu	0.007 (0.001)	8.26E-08	41,412
	22	44368122	A	G	rs3761472	*SAMM50*	mis	p.Asp110Gly	0.019 (0.002)	8.85E-30	41,413
	22	44395451	T	C	rs1007863	*PARVB*	mis	p.Trp37Arg	0.011 (0.001)	7.98E-16	41,414
AST	4	88231392	T	TA	rs72613567	*HSD17B13*	spl		-0.005 (0.001)	6.24E-10	40,753
	10	18242311	A	G	rs10764176	*SLC39A12*	mis	p.Ser36Gly	-0.006 (0.001)	1.09E-10	40,753
	10	101157378	CGTT	C		*GOT1*	inf	p.Asn389del	-0.221 (0.024)	1.96E-20	40,753
	10	101165533	G	C	rs374966349	*GOT1*	mis	p.Gln208Glu	0.271 (0.027)	2.43E-24	40,753
	10	101912064	T	C	rs2862954	*ERLIN1*	mis	p.Ile291Val	-0.005 (0.001)	4.82E-09	40,753
	11	22271870	A	T	rs7481951	*ANO5*	mis	p.Leu322Phe	0.004 (0.001)	9.61E-08	40,722
	14	94844947	C	T	rs28929474	*SERPINA1*	mis	p.Glu366Lys	0.027 (0.003)	2.44E-20	40,753
	19	19379549	C	T	rs58542926	*TM6SF2*	mis	p.Glu167Lys	0.008 (0.002)	6.54E-08	40,192
	22	44324727	C	G	rs738409	*PNPLA3*	mis	p.Ile148Met	0.014 (0.001)	8.31E-46	40,753
	22	44324730	C	T	rs738408	*PNPLA3*	syn	p.Pro149Pro	0.014 (0.001)	8.93E-46	40,753
	22	44368122	A	G	rs3761472	*SAMM50*	mis	p.Asp110Gly	0.011 (0.001)	1.22E-22	40,752
	22	44395451	T	C	rs1007863	*PARVB*	Mis	p.Trp37Arg	0.006 (0.001)	1.31E-13	40,753

[0559]

【표 8-계속】

			GHS베리아트릭 수술 집단			댈러스 심장 연구			U. Penn		
특성	Chr	BP	Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	N
ALT	1	220970028	0.005 (0.005)	3.10E-01	2475	0.011 (0.008)	1.76E-01	1357	0.007 (0.004)	1.02E-01	6158
	4	88231392	-0.010 (0.005)	5.57E-02	2475	-0.016 (0.008)	6.60E-02	1357	-0.013 (0.004)	1.33E-03	6158
	8	144997604	-0.492 (0.165)	2.84E-03	2475	NA (NA)	NA	NA	-0.051 (0.072)	4.79E-01	6158
	8	145008502	-0.161 (0.165)	3.29E-01	2475	NA (NA)	NA	NA	-0.247 (0.143)	8.48E-02	6158
	8	145692918	-0.009 (0.020)	6.48E-01	2475	0.032 (0.036)	3.76E-01	1356	-0.053 (0.018)	3.72E-03	6158
	8	145730072	-0.189 (0.165)	2.50E-01	2475	NA (NA)	NA	NA	-0.298(0.101)	3.26E-03	6158
	8	145730161	-0.341 (0.074)	3.64E-06	2475	NA (NA)	NA	NA	-0.143 (0.054)	8.50E-03	6158

[0560]

[0561]

			GHS 베리아트릭 수술 집단			델러스 심장 연구			U. Penn		
특성	Chr	BP	Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	N
	8	145730221	-0.009 (0.020)	6.45E-01	2475	0.028 (0.036)	4.37E-01	1357	-0.060 (0.018)	5.60E-04	6158
	8	145731636	-0.314 (0.165)	5.71E-02	2475	-0.317 (0.140)	2.35E-02	1356	-0.148 (0.143)	3.04E-01	6157
	8	145732114	-0.273 (0.048)	9.83E-09	2474	-0.240 (0.075)	1.36E-03	1357	-0.197 (0.041)	1.31E-06	6157
	8	145732151	-0.115 (0.058)	4.82E-02	2475	-0.106 (0.099)	2.86E-01	1356	-0.049 (0.041)	2.27E-01	6157
	8	145732180	-0.273 (0.050)	4.26E-08	2475	-0.191 (0.070)	6.58E-03	1357	-0.197 (0.041)	1.31E-06	6158
	8	145732305	-0.161 (0.165)	3.29E-01	2475	NA (NA)	NA	NA	-0.509 (0.203)	1.21E-02	6158
	8	145748532	-0.161 (0.165)	3.29E-01	2475	NA (NA)	NA	NA	-0.307 (0.143)	3.21E-02	6158
	9	117122202	-0.004 (0.005)	4.09E-01	2475	0.004 (0.008)	6.18E-01	1357	-0.007 (0.004)	5.29E-02	6158
	9	117124731	-0.004 (0.005)	3.90E-01	2475	0.003 (0.008)	7.33E-01	1356	-0.007 (0.004)	4.24E-02	6158
	10	101595996	-0.002 (0.010)	8.01E-01	2475	-0.007 (0.017)	6.88E-01	1357	-0.017 (0.007)	1.55E-02	6158
	10	101606861	-0.003 (0.010)	7.74E-01	2475	-0.008 (0.017)	6.28E-01	1357	-0.017 (0.007)	1.70E-02	6158
	10	101610533	-0.003 (0.010)	7.93E-01	2475	-0.008 (0.017)	6.28E-01	1357	-0.017 (0.007)	1.76E-02	6158
	10	101611294	-0.001 (0.010)	9.11E-01	2475	-0.010 (0.017)	5.40E-01	1357	-0.016 (0.007)	2.77E-02	6158
	10	101912064	-0.010 (0.005)	2.91E-02	2475	-0.006 (0.007)	4.02E-01	1356	-0.009 (0.004)	2.06E-02	6158
	10	101977883	-0.006 (0.005)	2.05E-01	2475	0.0001 (0.008)	9.94E-01	1357	-0.011 (0.004)	3.91E-03	6158
	10	113917085	-0.003 (0.005)	5.80E-01	2475	-0.013 (0.008)	1.15E-01	1357	-0.008 (0.004)	5.12E-02	6158
	10	113940329	-0.003 (0.005)	5.61E-01	2475	-0.013 (0.008)	1.33E-01	1357	-0.008 (0.004)	4.77E-02	6158
	14	94844947	0.035 (0.020)	7.97E-02	2475	0.034 (0.032)	2.92E-01	1357	0.054 (0.013)	1.63E-05	6158
	19	19379549	0.040 (0.010)	2.40E-05	2475	0.024 (0.014)	9.50E-02	1357	0.013 (0.008)	7.51E-02	6158
	22	44324727	0.019 (0.006)	5.54E-04	2475	0.006 (0.009)	5.43E-01	1357	0.016 (0.004)	2.05E-04	6158
	22	44324730	0.019 (0.006)	5.51E-04	2475	0.006 (0.009)	5.43E-01	1357	0.016 (0.004)	2.14E-04	6158
	22	44342116	0.001 (0.005)	7.77E-01	2475	0.005 (0.008)	5.18E-01	1357	0.005 (0.004)	2.16E-01	6158
	22	44368122	0.009 (0.006)	1.66E-01	2475	-0.001 (0.01)	9.37E-01	1357	0.018 (0.005)	4.02E-04	6158
	22	44395451	0.003 (0.005)	5.22E-01	2475	0.008 (0.008)	3.13E-01	1357	0.009 (0.004)	2.50E-02	6158
AST	4	88231392	-0.010 (0.003)	3.12E-03	2469	-0.012 (0.006)	5.32E-02	1357	-0.007 (0.004)	5.56E-02	6166
	10	18242311	-0.010 (0.003)	2.91E-03	2469	-0.003 (0.006)	5.80E-01	1357	-0.009 (0.004)	1.03E-02	6166
	10	101157378	-0.205 (0.062)	8.57E-04	2469	NA (NA)	NA	NA	-0.243 (0.088)	5.97E-03	6165
	10	101165533	NA (NA)	NA	NA	NA (NA)	NA	NA	0.339 (0.079)	1.85E-05	6166
	10	101912064	-0.004 (0.003)	1.54E-01	2469	-0.007 (0.006)	2.21E-01	1357	-0.004 (0.003)	1.94E-01	6166
	11	22271870	-0.001 (0.003)	7.85E-01	2466	0.006 (0.006)	2.85E-01	1357	-0.002 (0.003)	5.46E-01	6165
	14	94844947	0.023 (0.013)	7.79E-02	2469	0.044 (0.024)	6.98E-02	1357	0.055 (0.011)	4.01E-07	6166
	19	19379549	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	44324727	0.014 (0.004)	1.27E-04	2469	0.004 (0.007)	5.44E-01	1357	0.015 (0.004)	4.87E-05	6166
	22	44324730	0.014 (0.004)	1.32E-04	2469	0.004 (0.007)	5.44E-01	1357	0.015 (0.004)	4.96E-05	6166
	22	44368122	0.008 (0.004)	6.03E-02	2469	-0.001 (0.008)	9.45E-01	1357	0.016 (0.004)	2.64E-04	6166
	22	44395451	0.003 (0.003)	4.12E-01	2469	0.006 (0.006)	2.95E-01	1357	0.009 (0.003)	6.17E-03	6166

[0562]

			** 복제 메타-분석 (N=3)		*** 연합 메타-분석 (N=4)	
특성	Chr	BP	Beta (SE)	P	Beta (SE)	P
ALT	1	220970028	0.007 (0.003)	2.31E-02	0.008 (0.001)	3.38E-09
	4	88231392	-0.013 (0.003)	***3.85E-05**	-0.010 (0.001)	1.17E-15
	8	144997604	-0.121 (0.066)	6.56E-02	-0.155 (0.025)	2.68E-10
	8	145008502	-0.210 (0.108)	5.23E-02	-0.264 (0.031)	5.54E-18
	8	145692918	-0.025 (0.013)	4.69E-02	-0.032 (0.005)	2.25E-12

			** 복제 메타-분석 (N=3)		*** 연합 메타-분석 (N=4)	
특성	Chr	BP	Beta (SE)	P	Beta (SE)	P
	8	145730072	-0.268 (0.086)	1.88E-03	-0.308 (0.033)	2.79E-20
	8	145730161	-0.213 (0.044)	***1.14E-06**	-0.223 (0.013)	4.49E-64
	8	145730221	-0.031 (0.013)	1.36E-02	-0.033 (0.005)	1.92E-12
	8	145731636	-0.256 (0.086)	2.79E-03	-0.237 (0.029)	1.94E-16
	8	145732114	-0.231 (0.029)	***7.24E-16**	-0.225 (0.012)	6.06E-78
	8	145732151	-0.074 (0.032)	1.88E-02	-0.076 (0.012)	7.03E-11
	8	145732180	-0.221 (0.029)	***1.41E-14**	-0.224 (0.012)	1.04E-77
	8	145732305	-0.299 (0.128)	1.93E-02	-0.273 (0.030)	6.44E-20
	8	145748532	-0.244 (0.108)	2.40E-02	-0.189 (0.027)	2.93E-12
	9	117122202	-0.005 (0.003)	8.42E-02	-0.007 (0.001)	3.08E-09
	9	117124731	-0.005 (0.003)	6.15E-02	-0.007 (0.001)	1.00E-09
	10	101595996	-0.012 (0.005)	3.43E-02	-0.014 (0.002)	3.44E-09
	10	101606861	-0.012 (0.005)	3.25E-02	-0.014 (0.002)	2.99E-09
	10	101610533	-0.012 (0.005)	3.43E-02	-0.014 (0.002)	3.23E-09
	10	101611294	-0.011 (0.005)	5.21E-02	-0.014 (0.002)	4.09E-09
	10	101912064	-0.009 (0.003)	***1.14E-03**	-0.011 (0.001)	1.76E-23
	10	101977883	-0.008 (0.003)	4.33E-03	-0.009 (0.001)	3.59E-15
	10	113917085	-0.007 (0.003)	2.07E-02	-0.008 (0.001)	3.28E-11
	10	113940329	-0.007 (0.003)	2.00E-02	-0.008 (0.001)	1.77E-11
	14	94844947	0.047 (0.010)	***2.82E-06**	0.043 (0.004)	1.59E-25
	19	19379549	0.024 (0.006)	***1.37E-05**	0.016 (0.002)	1.15E-12
	22	44324727	0.016 (0.003)	***7.45E-07**	0.021 (0.001)	3.55E-55
	22	44324730	0.016 (0.003)	***7.73E-07**	0.021 (0.001)	3.10E-55
	22	44342116	0.004 (0.003)	1.91E-01	0.006 (0.001)	6.24E-08

[0563]

			** 복제 메타-분석 (N=3)		*** 연합 메타-분석 (N=4)	
특성	Chr	BP	Beta (SE)	P	Beta (SE)	P
	22	44368122	0.012 (0.004)	*7.69E-04	0.018 (0.002)	1.08E-31
	22	44395451	0.007 (0.003)	1.78E-02	0.010 (0.001)	1.16E-16
AST	4	88231392	-0.009 (0.002)	*8.38E-05	-0.006 (0.001)	6.82E-13
	10	18242311	-0.009 (0.002)	*1.16E-04	-0.006 (0.001)	1.10E-13
	10	101157378	-0.218 (0.051)	*1.66E-05	-0.220 (0.022)	1.68E-24
	10	101165533	0.339 (0.079)	*1.85E-05	0.278 (0.025)	3.25E-28
	10	101912064	-0.005 (0.002)	2.51E-02	-0.005 (0.001)	3.68E-10
	11	22271870	0.000 (0.002)	8.43E-01	0.004 (0.001)	1.13E-06
	14	94844947	0.042 (0.008)	*9.54E-08	0.029 (0.003)	6.71E-26
	19	19379549	0.014 (0.004)	*1.20E-03	0.009 (0.002)	5.92E-10
	22	44324727	0.013 (0.002)	*5.51E-08	0.014 (0.001)	3.14E-52
	22	44324730	0.013 (0.002)	*5.81E-08	0.014 (0.001)	3.55E-52
	22	44368122	0.010 (0.003)	*3.40E-04	0.011 (0.001)	1.91E-25
	22	44395451	0.006 (0.002)	7.34E-03	0.006 (0.001)	3.62E-15

[0564]

[0565] * P-값이 $P < 1.43 \times 10^{-3}$의 본페로니 유의성 한계값을 충족하는 것을 나타낸다.

[0566] ** 반복 메타-분석은 3개의 복제 집단: GHS 베리아트릭 수술 집단, 댈러스 심장 연구, 및 Penn 약물 바이오뱅크를 포함한다.

[0567] *** 연합 메타-분석은 개발 집단과 3개의 복제 집단: GHS 개발 집단, GHS 베리아트릭 수술 집단, 댈러스 심장 연구, 및 Penn 약물 바이오뱅크를 포함한다.

[0568] 약어: AAF, 대체 대립유전자 빈도; Alt, 대체 대립유전자; ALT, 알라닌 아미노전이효소; AST, 아스파테이트 아미노전이효소; Ref, 참조 대립유전자; SE, 표준 오차; ann, 주석; mis, 미스센스; syn, 동의적; spl, 스플라이스 도너; stop, 중단 획득; fs, 프레임시프트; inf, 프레임내 indel.

[0569] ***엑손 변종의 만성 간 질환의 임상 진단과의 연관성***

[0570] 다음으로, 본 발명자들은 개발 및 복제 집단에서 발견된 9개의 유전자에서의 13개의 아미노전이효소-관련 변종과 만성 간 질환, 이를테면 알코올성 및 비알코올성 (비바이러스성) 간 질환, 뿐만 아니라 가장 진전된 형태의 만성 간 질환: 알코올성 간경변, 비알코올성 간경변, 및 간세포 암종 (HCC) 사이의 관계를 분석하였다. 테스트한 13개의 변종에 대한 $P<1.92\times10^{-3}$의 본페로니 유의성 한계값을 사용하여, 본 발명자들은 5개의 유전자 (HSD17B13, SERPINA1, TM6SF2, PNPLA3, 및 SAMM50)에서의 6개의 변종과 만성 간 질환 표현형 사이의 유의한 연관성을 발견하였다 (**표 9**). SERPINA1, TM6SF2, PNPLA3, 및 SAMM50 연관성은 이전에 보고된 연관성을 확인시켜 준다. 개발 집단에서, HSD17B13 rs72613567:TA는 대립유전자 용량-의존 방식으로 알코올성 및 비알코올성 간 질환 둘 다의 모든 EHR-유래 카테고리의 보다 낮은 가능성과 관련되었다 (**도 2A**): 알코올성 간 질환의 모든 카테고리, 이형접합성 오즈비 (OR_{het}) (95% 신뢰도 간격) 0.58 (0.42 내지 0.80), 동형접합성 OR (OR_{hom}) 0.47 (0.23 내지 0.97), 대립유전자 OR ($OR_{allelic}$) 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), $OR_{allelic}$ 0.84 (0.78 내지 0.91), $P=1.3\times10^{-5}$. HSD17B13 rs72613567:TA는 또한 알코올성 및 비알코올성 간경변의 더 낮은 가능성과도 관련이 있었는데, 이형접합체 및 동형접합체 각각에 대해 42% 및 73% 더 낮은 알코올성 간경변 가능성을 보였고 (OR_{het} 0.58 (0.39 내지 0.86), OR_{hom} 0.27 (0.09 내지 0.85), $OR_{allelic}$ 0.56 (0.41 내지 0.78), $P=3.4\times10^{-4}$) 이형접합체 및

동형접합체 각각에 대해 26% 및 49% 더 낮은 비알코올성 간경변 가능성을 보였다 (OR_{het} 0.74 (0.60 내지 0.93), OR_{hom} 0.51 (0.31 내지 0.85), $OR_{allelic}$ 0.74 (0.62 내지 0.88), $P=4.5x10^{-4}$). HSD17B13 rs72613567:TA는 또한 더 낮은 가능성의 HCC와 명목상으로 관련이 있었다.

[0571] 본 발명자들은 다중-민족 델러스 간 연구 (DLS) 및 델러스 소아 간 연구 (DLPS, **표 10**)에서 이들 연구를 확인하고 확장하고자 하였다. DLS에서, TA 대립유전자는 대립유전자 용량-의존 방식으로 임의의 간 질환의 더 낮은 가능성과 관련되었다 (OR_{het} 0.74 (0.57 내지 0.97), OR_{hom} 0.41 (0.21 내지 0.83), $OR_{allelic}$ 0.70 (0.5 내지 0.88), $P=1.8x10^{-3}$, **도 8**). 유사한 효과를 진전된, 간경변 형태이 알코올성 ($OR_{allelic}$ 0.72 (0.53 내지 0.99), $P=4.4x10^{-2}$) 및 비알코올성 ($OR_{allelic}$ 0.65 (0.40 내지 1.07), $P=9.0x10^{-2}$) 간 질환을 포함하여, EHR-유래 간 질환 아형 전체에서 관찰하였다. 자가-보고 인종에 의해 그룹이 나뉘어진 개인들의 하위세트 분석에서, 간 질환과의 연관성은 히스패닉계 미국인에서 유의하였다 (n=326건의 케이스 및 722명의 대조군, $OR_{allelic}$ 0.51 (0.35 내지 0.74), $P=4.0x10^{-4}$); 통계학적으로 유의하진 않았지만, 유사한 수치 동향이 또한 DLS의 아프리카계 미국인 (n=33건의 케이스 및 2,291명의 대조군, $OR_{allelic}$ 0.74 (0.25 내지 2.47), P=0.67) 및 유럽계 미국인 (n=158건의 케이스 및 1,266명의 대조군$OR_{allelic}$ 0.87 (0.65 내지 1.15), P=0.32) 하위세트에서 주지되었다. DPLS에서, 히스패닉계 미국인 소아 간 질환 환자들 및 비만 대조군의 별도의 연구에서, TA 대립유전자는 또한 간 질환의 더 낮은 가능성과 관련되었다 ($OR_{allelic}$ 0.61 (0.37 내지 0.99), $P=4.6x10^{-2}$). 그러므로, HSD17B13 rs72613567:TA는 3개의 독립적인 집단에서 성인 및 아동에서 간경변을 포함하여 만성 간 질환의 다중 형태의 감소된 가능성과 관련되었다.

표 9

[0572] **개발 집단에서 간 질환 표현형과 12개의 엑솜-전체 중요한 및 복제하는 단일 뉴클레오타이드 변종의 연관성.**

CHR:BP:Ref:Alt	유전자	rsID	알코올성 간 질환		알코올성 간경변	
			OR (95% CI)	P-값	OR (95% CI)	P-값
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
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**

[0573] * P-값이 P < $2.08x10^{-3}$의 본페로니 유의성 한계값을 충족하는 것을 나타낸다.

【표 9-계속】

CHR:BP:Ref:Alt	유전자	rsID	비알코올성 간 질환		비알코올성 간경변		간세포 암종	
			OR (95% CI)	P-값	OR (95% CI)	P-값	OR (95% CI)	P-값
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
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**

[0574]

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

[0575]

표 10

[0576] **델러스 간 및 소아 간 연구로부터의 유전자형분류된 다중-민족 케이스 및 대조군의 인구통계학적 및 임상적 특징.**

특성	델러스 간 연구 케이스 (N = 517명)	델러스 간 연구 대조군 (N = 4,279명)	델러스 소아 간 연구 케이스 (N = 203명)	델러스 소아 간 연구 대조군 (N = 244명)
연령 (나이) - 중앙 (IQR)	55 (48 - 60)	44 (36 - 53)	12 (10 - 15)	12 (11 - 14)
여성 성별 -수 (%)	277 (54)	2,494 (58)	65 (32)	126 (52)
체질량 지수 - 중앙 (IQR)	30 (27 - 35)	30 (26 - 35)	30 (27 - 34)	31 (28 - 35)
자가-기록 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 (ICD-9 코드에 의함) - N (%)				
알코올성 간 질환	223 (43)	-	-	-
알코올성 간경변	215 (42)	-	-	-
비알코올성, 비바이러스성 간 질환	212 (20)	-	-	-
비알코올성 간경변	100 (19)	-	-	-
간세포 암종	44 (9)	-	-	-
간 질환 없음	-	4,279 (100)	-	-244 (100)

[0577] ***HSD17B13 rs72613567:TA의 간 병리와의 연관성***

[0578] NAFLD는 유의한 염증의 증거가 없는 간 지방 축적 (단순 지방증)으로부터, 보다 임상적으로 유해한 NASH에 이르기까지 다양한 질환 스펙트럼을 기술한다. HSD17B13 rs72613567:TA와 EHR-유래 간 질환 진단 코드 사이의 연관성을 확인하기 위하여, 그리고 추가로 그것과 지방증의 NASH로의 조직병리학적 진행과의 연관성을 이해하기 위하여, 본 발명자들은 GHS 베리아트릭 수술 집단에서 연관성 테스트를 수행하였다. 베리아트릭 수술시에 간 생검에 의해 평가된 전체 엑솜 서열분석된 개인 2,391명의 이 집단에서, 총 55명 (23%)의 개인이 지방증, 지방간염, 또는 섬유증의 증거가 없었고 ("정상"), 830명 (35%)이 단순 지방증을 나타냈으며, 1006명 (42%)이 NASH를 나타냈다. 정상 간, 단순 지방증, 및 NASH의 유병률을 유전자형에 의해 비교했을 때, 각 TA 대립유전자로, 정상 간의 유병률은 유전자형에 의해 상이한 것으로 나타나지 않았지만 (T/T, T/TA, 및 TA/TA 보균자에 대해 각각 23%, 24%, 및 23%, 집단에서 동향에 대한 카이-제곱 시험에 의해 P = 0.5), NASH의 유병률은 감소하였고 (T/T, T/TA, 및 TA/TA 보균자에 대해 각각 45%, 40%, 및 31%, P = $1.6x10^{-4}$) 단순 지방증의 유병률은 증가한 것 (T/T, T/TA, 및 TA/TA 보균자에 대해 각각 33%, 35%, 및 47%, P = $1.1x10^{-3}$)을 관찰하였다 (**도 9**).지방증을 가진 개인 중에서, TA 대립유전자는, 대립유전자 용량-의존 방식으로, 단순 지방증과 비교하여 NASH 및 섬유증 둘 다의 통계적으로 유의하게 더 낮은 가능성과 관련되었다 (NASH에 대해 $OR_{allelic}$ 0.77 (0.66-0.90), P=$6.5x10^{-4}$; 섬유증에 대해 $OR_{allelic}$ 0.74 (0.62-0.88), P=$4.15x10^{-4}$; **도 2B**). 이들 데이터는 모두 함께, 단순 지방증으로부터 NASH 및 섬유증의 보다 진전된 단계로의 NAFLD 진행을 매개하는 데 있어 HSD17B13의 역할을 시사한다.

[0579] ***HSD17B13 rs72613567:TA의 임상 정량적 특성 및 진단과의 연관성***

[0580] HSD17B13 스플라이스 변종의 임상 결과를 보다 포괄적으로 조사하기 위하여, 본 발명자들은 HSD17B13 rs72613567:TA와 405개의 정량적 EHR-유래 인체측정, 바이탈 TTLS호 실험실, 심전도, 심장초음파검사, 및 골밀도 측정값, 및 또한 3,168개의 EHR-유래 임상 진단과의 연관성의 표현형-전체 연구를 수행하였다. 정량적 임상 측정값 및 임상 진단과의 연관성에 대해 각각 $1.23x10^{-4}$ 및 $1.58x10^{-5}$의 본페로니 유의성 한계값을 사용하여, 본 발명자들은 간의 아미노전이효소와의 연관성에 더불어, HSD17B13 rs72613567:TA 대립유전자와 더 높은 혈소판 수와의 통계적으로 유의한 연관성을 확인하였다 (**표 11**). 만성 간 질환 이외의 임상 진단과의 통계적으로 유의한 연관성은 없었다 (OR (95% CI) =0.88 (0.84-0.93); P = $9.14x10^{-6}$; AAF = 0.263; N 케이스 총 = 4031, T/T = 2331, T/TA = 1449, TA/TA = 251; N 대조군 총 = 35701, T/T = 19238, T/TA = 13984, TA/TA = 2479).

표 11

[0581] **정량적 임상 측정값과 HSD17B13 rs72613567:TA의 연관성의 표현형-전체 연구.**

표현형	효과	SE	P	AAF	N			
					총	T/T	T/TA	TA/TA
알라닌 아미노전이효소 중앙값: 조정값 (중앙값: 조정값 (잔차 Log)	-0.009	0.001	*1.74E-12*	0.264	44038	23868	17115	3055
아스파테이트 아미노전이효소 중앙값:조정값(잔차 Log)	-0.006	0.001	*2.75E-11*	0.264	43370	23493	16851	3026
알라닌 아미노전이효소 최대:조정값 (잔차 Log)	-0.013	0.002	*1.39E-09*	0.264	43905	23797	17065	3043
아스파테이트 아미노전이효소 최대:조정값 (잔차 Log)	-0.010	0.002	*8.73E-09*	0.264	42733	23145	16609	2979
혈소판 중앙값:조정값(잔차 Log)	0.004	0.001	*1.44E-08*	0.264	46182	25020	17944	3218
알라닌 아미노전이효소 최소:조정값 (잔차 Log)	-0.008	0.002	*2.47E-07*	0.264	44029	23864	17111	3054
혈소판 최소:조정값 (잔차)	1.919	0.443	*1.47E-05*	0.264	46181	25020	17943	3218
혈소판 최대:조정값 (잔차 Log)	0.004	0.001	*3.03E-05*	0.264	46165	25014	17936	3215
아스파테이트 아미노전이효소 최소:조정값 (잔차 Log)	-0.004	0.001	*5.00E-05*	0.264	43327	23471	16831	3025
진한 이탤릭체는 P-값이 $P < 1.23x10^{-4}$의 본페로니 유의성 한계값을 충족했음을 나타낸다.								
약어: AAF, 대체 대립유전자 빈도; SE, 표준 오차.								

[0582] ***HSD17B13 mRNA 및 HSD17B13 단백질 발현에 미치는 HSD17B13 rs72613567:TA의 영향***

[0583] 본 발명자들은 다음에 유전자의 공지된 및 신규한 전사물의 발현에 미치는 HSD17B13 rs72613567:TA 대립유전자의 영향을 조사하였다. 본 발명자들은 HSD17B13 rs72613567 스플라이스 변종의 22 T/T 동형접합성, 30 T/TA 이형접합성, 및 17 TA/TA 동형접합성 보균자로부터의 조직학적으로 정상인 간 샘플에서 HSD17B13 mRNA 발현을 평가하기 위하여 RNA 서열분석을 사용하였다. 2개의 공지된 HSD17B13 전사물, A 및 B 외에, 2개의 신규한 전사물을 확인하였다. 엑손 6이 결핍된 전사물 C, 및 엑손 6의 3' 단부에서 구아닌 뉴클레오타이드의 삽입을 함유하여, 단백질의 미성숙 절단을 초래할 것으로 예상할 수 있는 전사물 D. 4개의 추가 전사물 (E 내지 H)이 매우 낮은 수준으로 발현되었다 (**도 3A 내지 3D** 및 **6A 내지 6D**). 전사물을 RT-PCR 및 Sanger 서열분석에 의해 입증하였다. D 전사물을 또한 긴 판독 cDNA 서열분석을 사용하여 입증하였다. 모든 확인된 HSD17B13 동형 (A 내지 H)의 단백질 서열 정렬을 **도 7A** 및 **7B**에 나타낸다. 이들 전사물의 발현 수준은 HSD17B13 rs72613567 유전자형에 따라 달랐다; 전사물 A 및 B의 수준은 감소한 반면, 전사물 C 및 D의 수준은 각 TA 대립유전자로 대립유전자 용량-의존 방식으로 증가하였다 (**도 3A 내지 3D**). 전장 300개의 아미노산 단백질을 암호화하는 전사물 A는 T/T 동형접합체에서는 우사한 전사물이었던 한편, 미성숙하게 절단된 단백질을 암호화하는 전사물 D는 TA/TA 동형접합체에서 우세한 전사물이었다. 인간 간 생검 조직에서, 절단된 동형 D 단백질은 이형접합체 및 TA/TA 동형접합체에서 최소한으로 존재하였고, 동형 A 단백질 풍부성은 대립유전자 용량-의존 방식으로 감소하였다 (**도 10B** 및 **10C**). HEK293 세포에서 동형 A 및 D의 이종성 발현은 mRNA 발현에 비해 동형 D의 감소된 풍부성을 나타내며, 동형 A와 비교했을 때 동형 D의 불안정성을 시사한다 (**도 11A 내지 11C**). 이들 데이터는 HSD17B13 rs72613567이 mRNA 스플라이싱을 변경시켜서 단백질의 절단 형태의 합성을 초래하여 실질적으로 인간 간에서 발현이 감소되는 것과 일치한다.

[0584] ***인간 간세포에서 HSD17B13의 발현***

[0585] HSD17B13은 주로 간에서 발현되며 (Liu et al., *Acta Biochim. Pol*. 2007, 54, 213-218, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 그곳에서 지질 분말로 국지화되는데 (Su et al., *Proc. Natl. Acad. Sci. USA*, 2014, 111, 11437-11442, doi:10.1073/pnas.1410741111, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), 이것은 지방간 질환의 발병에서의 역할과 일치한다. 본 발명자들은 HSD17B13 전사물 A 또는 D를 발현하는 렌티바이러스로 안정적으로 형질전환된 불멸화된 인간 간 세포주에서 HSD17B13의 발현 및 그것의 국지화를 평가하였다. HSD17B13 동형 A는 주로 막 주위의 BODIPY-표지된 지질 방울에서 검출되었다 (데이터 미제시). 유사한 세포내 국지화를 지질 방울 표면에서 HSD17B13 동형 D에 대해 관찰하였다 (데이터 미제시 및 **도 12**). GFP 대조군 또

는 HSD17B13 동형 A 또는 D를 과다발현하는 세포주의 올레산 처리로 세포내 트라이글리세라이드 함량의 차이는 관찰되지 않았다 (**도 13A 내지 13D**).

[0586] ***시험관내에서 및 세포 모델에서 HSD17B13 활성에 미치는 rs72613567:TA의 영향***

[0587] rs72613567:TA로 인한 HSD17B13 단백질의 미성숙 절단의 기능적 결과를 이해하기 위하여, 본 발명자들은 재조합 단백질 및 니코틴아미드 아데노신 다이뉴클레오타이드를 보조인자로서 사용하여 시험관내에서 동형 A 및 D의 효소적 활성을 평가하였다. 본 발명자들은 265개의 특유한 잠정적 기질을 테스트하였고, 스테로이드 기질 및 생체활성 지질 (예컨대 류코트라이엔 B4)을 HS17B13의 효소적 기질로서 확인하였다. 본 발명자들은, 하이드록실의 케톤 기로의 산화를 초래하는, 에스트라다이올의 효소적 전환에 미치는 HSD17B13 효소적 활성의 후속 특성화에 집중하였다 (**도 14**에서 V_{max} 및 K_m 값). HSD17B13 동형 D는 HSD17B13 동형 A와 비교하여 시험관내에서 (**도 10D**) 및 세포-기반 효소적 전환 검정에서 (**도 10E**) 에스트라다이올로의 크게 감소된 활성을 나타냈다.

[0588] EHR-유래 임상 표현형에 대규모 엑솜 서열분석을 결부시킴으로써, 본 발명자들은 HSD17B13의 스플라이스 변종과 감소된 혈청 아미노전이효소 수준, 뿐만 아니라 간 질환의 비알코올성 및 알코올성 형태의 감소된 위험 사이의 새로운 연관성을 확인하였다. 이들 연관성은 4개의 상이한 집단에서, 및 간 질환 및 HCC의 진전된 간경변성 형태를 포함하여 여러 상이한 간 질환 카테고리를 가로질러 일관되게 관찰되었다. HSD17B13 rs72613567:TA 대립유전자는 단순 지방증과는 관련되지 않았지만, NASH 및 섬유증의 감소된 위험과 관련되었으며, 이것은 이 변종 대립유전자가 만성 간 질환의 보다 임상적으로 진전된 단계로의 진행으로부터 보호하는 것을 시사한다. 표현형-전체 연관성 연구에서, HSD17B13 rs72613567:TA는 만성 간 질환 및 관련된 임상 측정값 (간의 아미노전이효소 및 혈소판 수) 이외의 임상 진단 또는 측정값과 유의하게 관련되지 않았으며, 이것은 변종 대립유전자의 임상 효과가 만성 간 질환에 특이적일 수 있음을 시사한다.

[0589] 다른 하이드록시스테로이드 17-베타 탈수소효소 패밀리 구성원들은 성 스테로이드 및 지방산 대사에 관여하지만 (Moeller, *Mol. Cell. Endocrinol.*, 2009, 301, 7-19, doi:10.1016/j.mce.2008.10.040, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨), HSD17B13의 기능에 대해서 알려진 것은 거의 없다. HSD17B13 과다발현은 이전에 마우스 간에서 지질 생성을 증가시키고, 배양된 간세포에서 지질 방울의 수 및 크기를 증가시키는 것으로 나타났다 (Su et al., *Proc. Natl. Acad. Sci. USA*, 2014, 111, 11437-11442, doi:10.1073/pnas.1410741111, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 2개의 이전 연구는 또한 HSD17B13 단백질의 간에서의 발현이 지방간 환자에서는 증가하는 것을 나타냈다 (Su et al., *Proc. Natl. Acad. Sci. USA*, 2014, 111, 11437-11442, doi:10.1073/pnas.1410741111 및Kampf et al., *FASEB J.*, 2014, 28, 2901-2914, doi:10.1096/fj.14-250555, 이 문헌들은 각각 모든 목적에 대해 그 전문이 본원에 참조로 포함됨). 본 발명자들의 데이터는 두 HSD17B13 동형이 지질 방울 막 상에서 발현되지만, 세포내 중성 지방 함량을 조절하는 것으로 나타나지는 않으며, 이것은 HSD17B13 rs72613567:TA와 인간의 단순 지방증 사이의 연관성의 부족을 비추는 발견인 것을 시사한다. 비록, HSD17B13의 생리적 기질이 알려져 있진 않지만, 효소적 연구는 HSD17B13 rs72613567:TA 대립유전자에 의해 암호화된 HSD17B13 동형이 에스트라다이올에 대해서는 촉매적으로 결핍된 것을 증명한다. 현재 테스트한 임의의 기질이 간 질환에 대해 중요한 것인지는 분명하지 않은 한편으로, HSD17B13이 지질-매개된 염증에 이미 연루되어 있는 여러 생체활성 지질 종 (예컨대 류코트라이엔 B4)에 대해 효소적 활성을 가지는 것은 아주 흥미로운 것이다 (Li et al., *Nature Medicine*, 2015, 21, 239-247, doi:10.1038/nm.3800, 모든 목적에 대해 그 전문이 본원에 참조로 포함됨).

[0590] 이런 HSD17B13 변종은 다른 도메인에서 새로운 치료제로의 길을 안내한 유전자 변종과 유사하게, 만성 간 질환을 표적화하는 새로운 치료 전략으로의 방안을 제공할 수 있다. 본 발명자들의 데이터는 HSD17B13이 지방증으로부터, 상당한 유병률 및 사망률과 관련되고, 현재 그것에 대한 효과적인 치료제가 없는 NASH, 섬유증, 및 간경변의 후기 단계로의 간 질환의 진행을 조절하는 것을 나타낸다.

[0591] **실시예 4. CRISPR/Cas9를 사용한 생체외 및 생체내에서 마우스 *Hsd17b13* 유전자좌의 변형.**

[0592] CRISPR/Cas9 시스템을 사용하는 *Hsd17b13*의 표적화에 대한 개념의 증명으로서, 마우스 *Hsd17b13* 유전자좌의 엑손 1 영역 또는 엑손 6/7 영역 중 어느 하나를 표적화하는 마우스 *Hsd17b13* 가이드 RNA를 테스트하였다. 가이드 RNA 표적 서열을 **표 12**에 제공한다. SEQ ID NO: 259 내지 268에 상응하는 가이드 RNA DNA-표적화 분절을 각각 SEQ ID NO: 1643 내지 1652에 제시하고, 티민 대신 우라실이 있는 것을 제외하고 SEQ ID NO: 259 내지 268과 동일하다. 마우스 *Hsd17b13* (하이드록시스테로이드 (17-베타) 탈수소효소 13에 대한 NCBI 유전자 ID는 243168 (SEQ ID NO: 269)이다. 마우스 게놈 유전자좌는 염색체 5, NC_000071.6 (103955442..103977388, 보체)에 있다.

표 12

[0593] **마우스에 대한 가이드 RNA 표적 서열**

*Hsd17b13*의 영역	#	가이드 RNA 표적 서열	SEQ ID NO					
			gRNA 표적 서열	crRNA	sgRNA v1	sgRNA v2	sgRNA v3	sgRNA v4
엑손 1	1	GGCAGACCGTTCTCATCACG	259	490	720	950	1180	1410
	2	CTTTACCAGTGACTCCAGGT	260	491	721	951	1181	1411
	3	GTCACAGATTTCCTTCTCCG	261	492	722	952	1182	1412
	4	AGATGATGACGCCCACCAGA	262	493	723	953	1183	1413
	5	GGAGAAGGAAATCTGTGACC	263	494	724	954	1184	1414
엑손 6/7	1	TGCGAGGAACTTACTTTTCC	264	495	725	955	1185	1415
	2	AGAGAAATATTGATATAGGA	265	496	726	956	1186	1416
	3	TATCAATATTTCTCTGATCC	266	497	727	957	1187	1417
	4	ATCGCTTTTAAGGCACGCTC	267	498	728	958	1188	1418
	5	TATACGACTGATCGCTTTTA	268	499	729	959	1189	1419

[0594] 가이드 RNA를 먼저 하이브리드 야생형 마우스 (75% C57BL/6NTac 25% 129S6/SvEvTac)로부터 분리한 일차 마우스 간세포에서 생체외에서 테스트하였다. 마우스로부터의 간을 1X PenStrep을 함유한 50 mL의 간 관류 배지로 관류한 후, 50 mL의 간 소화 배지 (HBSS, 100 mM $CaCl_2$, 500 mM HEPES, 콜라게나아제)로 관류하였다. 일단 간이 소화된 것으로 나타나면, 그것을 1X PenStrep 및 L-글루타민을 함유한 세척 배지에 넣었다. 간을 찢어서 간으로부터 부드럽게 흔들어줌으로써 간세포들을 방출시켰다. 일단 세포들이 방출된 후에, 세포를 70 μm의 메시 필터를 통해 놓고 50 g에서 4분 동안 4℃에서 회전시켰다. 펠릿을 세척 완충액으로 2회 세척하였다. 그런 후 펠릿을 20 mL의 38 내지 40% 퍼콜에 재현탁하여 200 g에서 10분 동안 4℃에서 회전시켰다. 펠릿을 2회 세척하고 플레이팅 배지 (Williams E 배지, 1X Penstrep, 1X L-글루타민, 5% FBS)에 재현탁하였다. 세포를 24-웰 콜라겐-코팅 조직 배양 플레이트에 웰당 300,000개 세포로 플레이팅하였다. 세포를 6 내지 18시간 동안 부착되도록 놓아둔 후, 플레이팅 배지를 FBS가 없는 배지로 교체하였다. 사용한 시약을 **표 13**에 나타낸다.

표 13

[0595] **일차 간세포의 분리를 위한 시약.**

물질	카탈로그 번호
간 관류 배지	Gibco [17701-038]
HBSS (1x)	Gibco [14175-079]
간세포 세척 배지	Gibco [17704-024]
Williams E 배지	Gibco [A12176-01]
Penstrep (100x)	Gibco [15140163]
L-글루타민 (200mM)	Gibco [25030081]
FBS 보충물	Gibco [A13450]
HEPES	Gibco [15630080]
콜라겐	Gibco [A1048301]
아세트산	Sigma [A6283]
리버라아제 TM	Roche [TM05401119001]
일차 간세포 해동 및 플레이팅 보충물	Gibco [CM3000]
일차 간세포 유지 보충물	Gibco [CM4000]
퍼콜(Percoll)	GE [17-0891-01]

[0596] Cas9 및 마우스 *Hsd17b13* gRNA를 함유한 리보핵단백질 복합체 (RNP)를 새롭게 분리한 일차 마우스 간세포에 첨가하였다. 일차 마우스 간세포에서의 생체외 실험을 위해, 별도의 crRNA 및 tracrRNA를 가지는 모듈식 가이드 RNA를 사용하였다. crRNA SEQ ID NO를 표 12에 제시하며, tracrRNA 서열을 SEQ ID NO: 1422에 재시한다. 각각의 Cas9/gRNA RNP 복합체를 CRISPRMAX™을 사용하여 2 nM의 최종 농도에서 트랜스펙션하였다. 48시간 후에, DNA 용해물을 세포로부터 제조하고, 차세대 서열분석을 테스트한 각 가이드 RNA에 대해 수행하여 예상된 절단 부위

에 걸친 삽입/삭제 (indel) 빈도를 측정하였다.**도 15**는 일차 마우스 간세포에서, 엑손 1 영역을 표적화하는 5개의 가이드 RNA의 각각 및 엑손 6/7 영역을 표적화하는 5개의 가이드 RNA의 각각을 포함하여, 각각의 가이드 RNA로 마우스 *Hsd17b13* 유전자에서의 편집 수준 (indel의 % 판독)을 나타낸다. 편집 효율은 차세대 서열분석에 의해 측정되는 바 용해된 세포의 풀로부터의 PCR 반응에서 서열 판독의 총 수에 대해 관찰된 삽입 또는 삭제의 총 수를 나타낸다. 거의 모든 가이드 RNA가 적어도 20%의 편집 효율을 나타냈다.

[0597] 다음으로, 5개의 마우스 *Hsd17b13* 가이드 RNA를 게놈상으로 통합된 Cas9 유전자를 가진 마우스 (Cas9-준비 마우스)에서 생체내에서 테스트하였다. 마우스에서의 생체내 실험에 대해, 키메릭 단일 가이드 RNA를 사용하였다. 각 가이드 RNA에 대한 DNA-표적화 서열은 표 12에 제시한 가이드 RNA 표적 서열과 동등하며, 단 우라실이 티민을 대체한다. 각각의 단일 가이드 RNA는 SEQ ID NO: 1420에 제시한 gRNA 스캐폴드의 상류 (5') DNA-표적 서열을 포함하였다. sgRNA SEQ ID NO를 표 12에 제시한다 (sgRNA v1에 대한 칼럼). 상이한 가이드 RNA 스캐폴드를 사용하는 다른 sgRNA 변이를 표 12에 나타내지만 테스트는 하지 않았다. 각 가이드 RNA에 대하여, 3마리의 Cas9-준비 수컷 마우스를 그룹당 배정하였다. 가이드 RNA를 sgRNA 발현 카세트를 운반하는 아데노-관련 바이러스 (AAV8)를 통해 꼬리 정맥 주사에 의해 (마우스당 100 μL PBS 중의 1E11) 유입시켰다. 임의의 Cas9를 발현하지 않는 야생형 마우스를 음성 대조군으로서 모든 5개의 가이드 RNA를 투약하였다. 주사 후 3주 후에, 동물들을 안락사시키고, 간 및 기타 조직과 함께 혈청을 수득하였다. 조직을 DNA 용해물로 처리하고 나중에 NGS 서열분석에 의해 분석하였다.

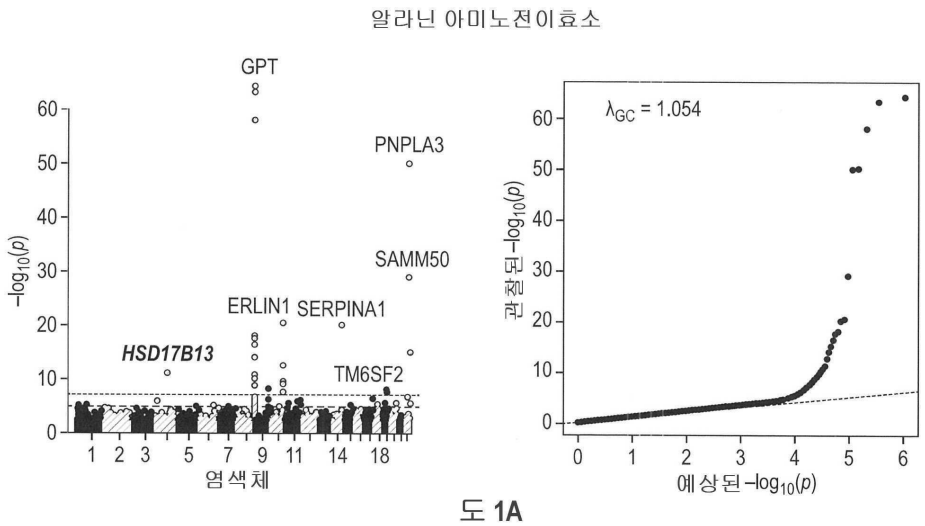
[0598] **도 16**에서 나타낸 것과 같이, NGS 서열분석은 모든 5개의 가이드 RNA에 대해 간에서 유의한 편집을 나타냈다 (각각에 대해 적어도 20%의 편집 퍼센트). 편집 효율은 용해된 세포의 풀로부터의 PCR 반응에서 서열 판독의 총 수에 대해 관찰된 삽입 또는 삭제의 총 수를 나타낸다. 다른 조직에서는 통계적으로 유의한 수준의 유전자 편집이 최소한으로 관찰되었거나 전혀 관찰되지 않았다 (데이터 미제시).

[0599] 간 효소, ALT, AST, 트라이글리세라이드, 총 콜레스테롤, HDL, LDL, 비-에스테르화 지방산 (NEFA), 및 알부민에 대한 혈청 화학 분석은 다양한 처리 그룹 간에 거의 차이를 나타내지 않았다 (데이터 미제시).

[0600] *Hsd17b13* 발현을 RT-qPCR에 의해 간으로부터의 RNA의 동등한 질량을 평가함으로써 평가하였다. 게놈 DNA는 분해되어서 qPCR 반응에 대해 계수할 수 없을 것이다. RNA를 역전사시킨 후 Cas9에 대해 특이적인 검정을 사용하여 Cas9 전사물을 검출하였다. 각각의 개별적인 *Hsd17b13* 가이드 RNA는 *Hsd17b13* mRNA 발현의 적어도 50%의 감소를 나타냈다. **도 17A**를 참조한다. 대조적으로, 비표적 HSD 패밀리 구성원의 발현에서는 유의한 감소를 관찰하지 못하였다. **도 17B**를 참조한다.

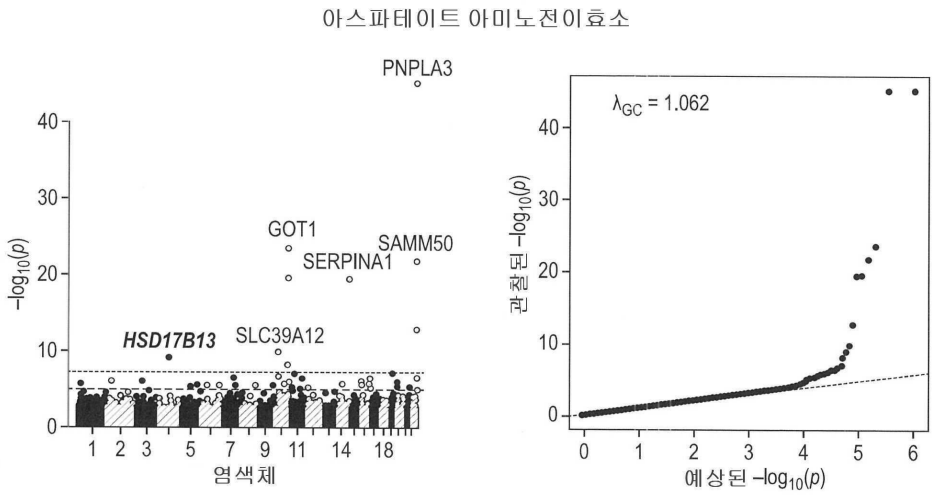
도면

도면1a



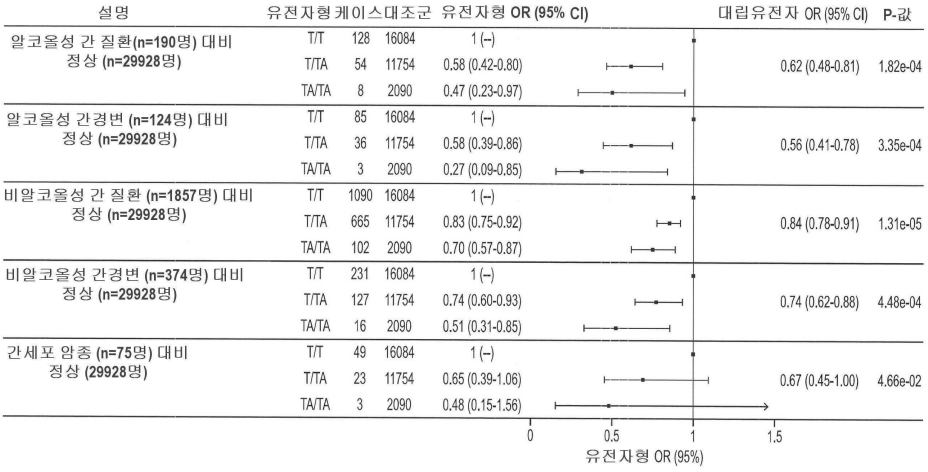
도 1A

도면1b



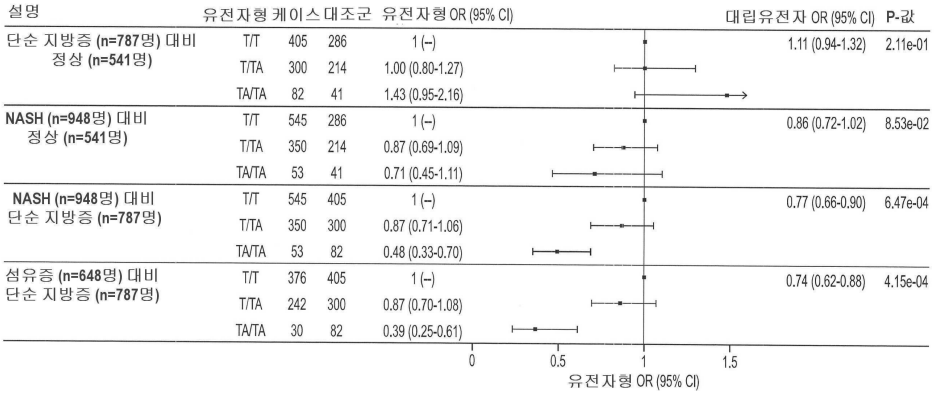
도 1B

도면2a



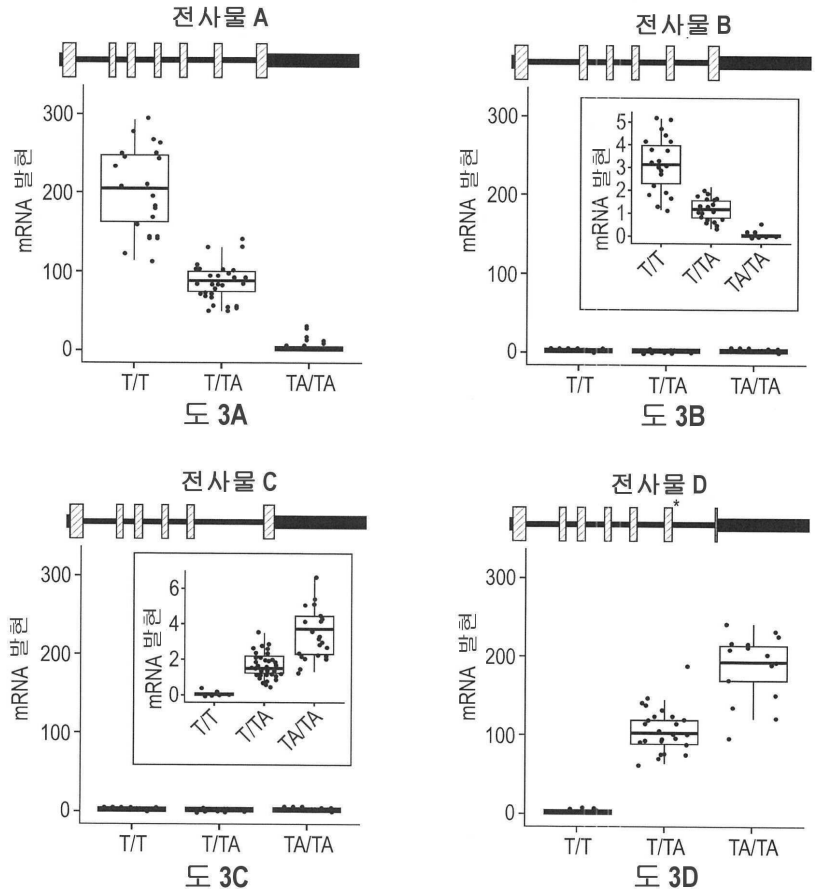
도 2A

도면2b



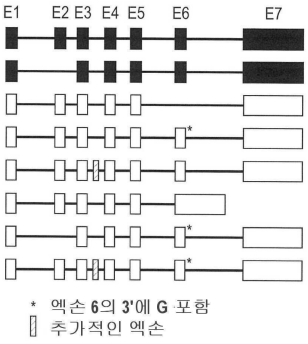
도 2B

도면3



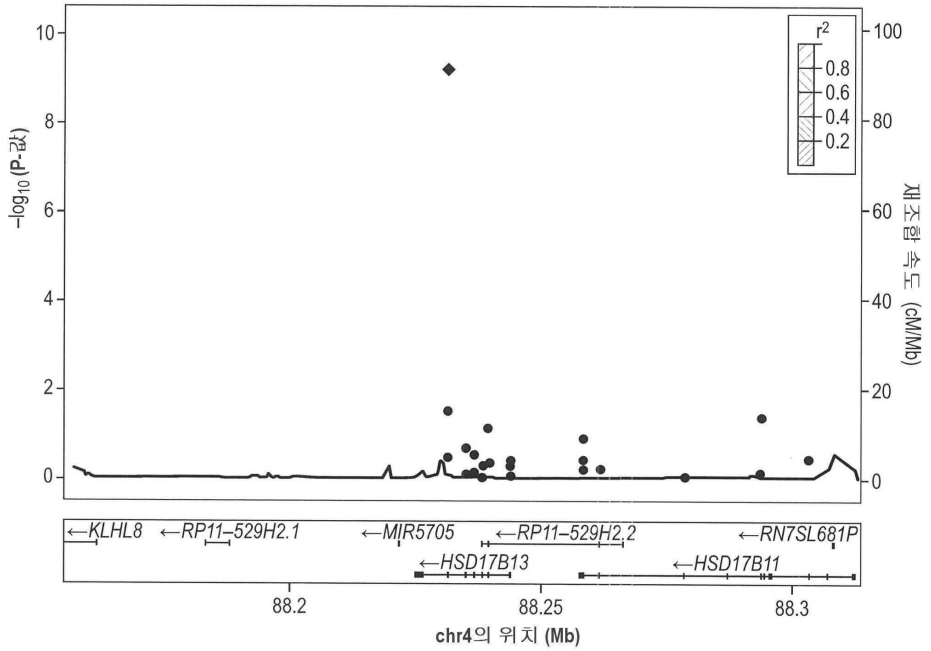
도면4

전사물	유형	설명	발현 FPKM WT	Het	Hom
A	공지		188.4	88	3.2
B	공지	엑손 2 건너뜀	10.8	1.2	0
C	신규	엑손 6 건너뜀	0.2	1.8	4.1
D	신규	엑손 6의 3'에 G 포함	5.7	103	185.6
E	신규	엑손 3과 4 사이에 추가적인 엑손	0.9	0.5	0
F	신규	엑손 6으로부터 인트론 6으로의 초과번역	0.3	2.3	4.6
G	신규	엑손 2 건너뜀; 엑손 6의 3'에 G 포함	0.3	1.4	2.8
H	신규	E; 엑손 6의 3'에 G 포함	0	0.6	0.8



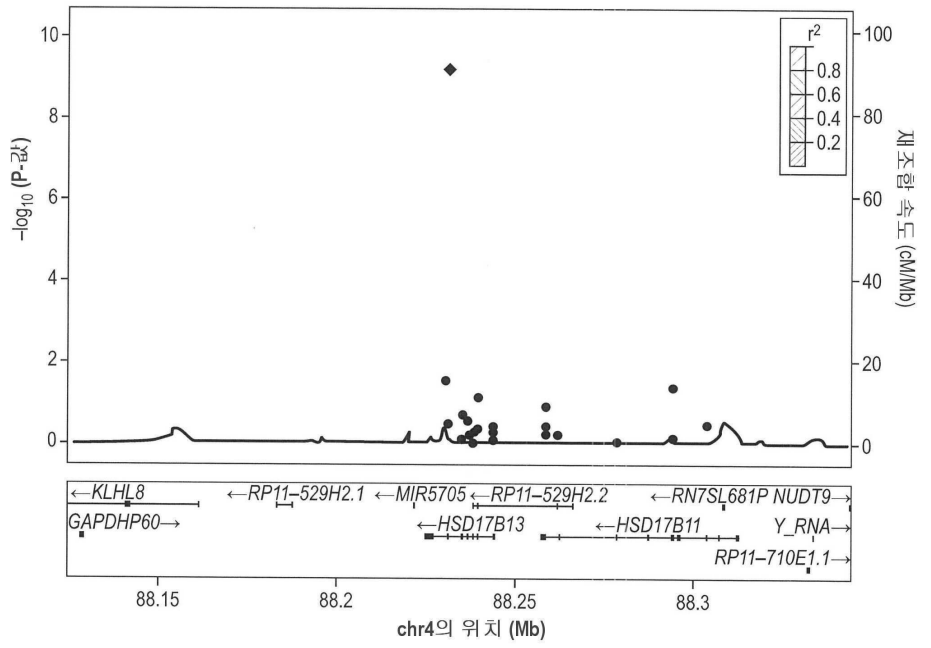
도 4

도면5a



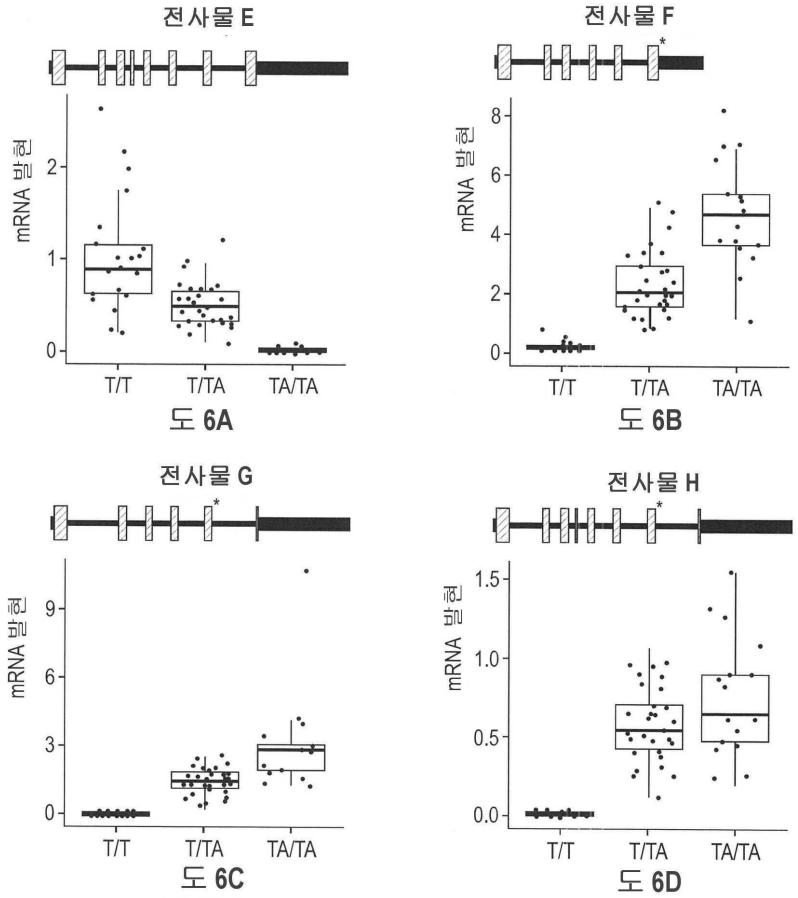
도 5A

도면5b

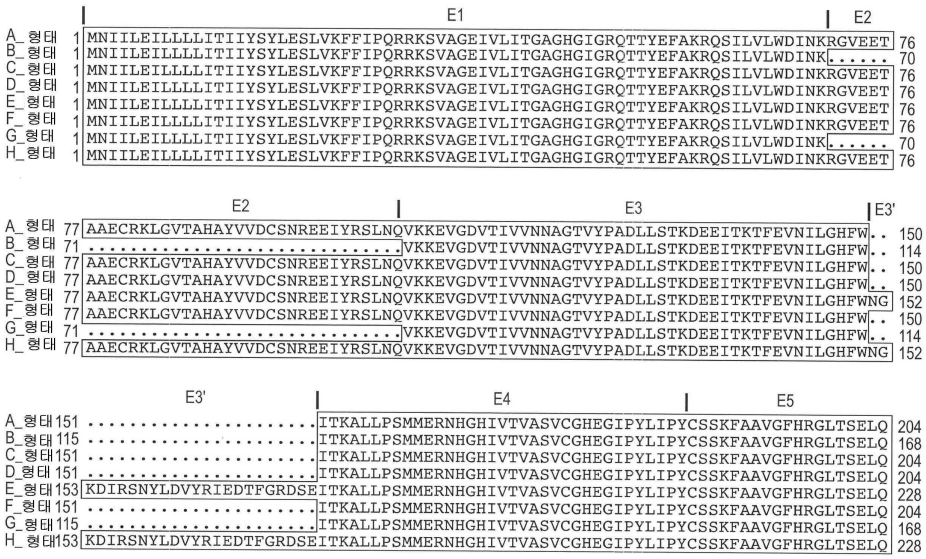


도 5B

도면6

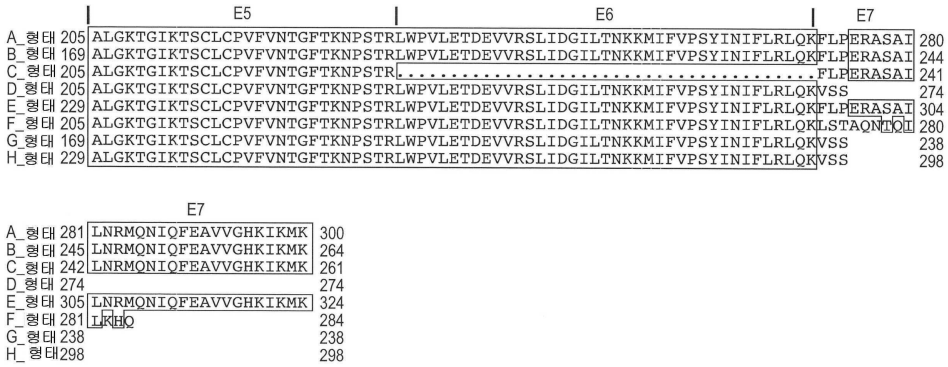

전사물 E
mRNA 발현
T/T
T/TA
TA/TA
도 6A
전사물 F
mRNA 발현
T/T
T/TA
TA/TA
도 6B
전사물 G
mRNA 발현
T/T
T/TA
TA/TA
도 6C
전사물 H
mRNA 발현
T/T
T/TA
TA/TA
도 6D

도면7a


E1
E2
A_형태 1 MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSILVLWDINKRGVEET 76
B_형태 1 MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSILVLWDINK..... 70
C_형태 1 MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSILVLWDINKRGVEET 76
D_형태 1 MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSILVLWDINKRGVEET 76
E_형태 1 MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSILVLWDINKRGVEET 76
F_형태 1 MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSILVLWDINKRGVEET 76
G_형태 1 MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSILVLWDINK..... 70
H_형태 1 MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSILVLWDINKRGVEET 76
E2
E3
E3'
A_형태 77 AAECRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDVTIVVNNAGTVYPADLLSTKDEEITKTFEVNILGHFW.. 150
B_형태 71 .....VKKEVGDVTIVVNNAGTVYPADLLSTKDEEITKTFEVNILGHFW.. 114
C_형태 77 AAECRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDVTIVVNNAGTVYPADLLSTKDEEITKTFEVNILGHFW.. 150
D_형태 77 AAECRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDVTIVVNNAGTVYPADLLSTKDEEITKTFEVNILGHFW.. 150
E_형태 77 AAECRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDVTIVVNNAGTVYPADLLSTKDEEITKTFEVNILGHFWNG 152
F_형태 77 AAECRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDVTIVVNNAGTVYPADLLSTKDEEITKTFEVNILGHFW.. 150
G_형태 71 .....VKKEVGDVTIVVNNAGTVYPADLLSTKDEEITKTFEVNILGHFW.. 114
H_형태 77 AAECRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDVTIVVNNAGTVYPADLLSTKDEEITKTFEVNILGHFWNG 152
E3'
E4
E5
A_형태 151 .....ITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYCSSKFAAVGFHRGLTSELQ 204
B_형태 115 .....ITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYCSSKFAAVGFHRGLTSELQ 168
C_형태 151 .....ITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYCSSKFAAVGFHRGLTSELQ 204
D_형태 151 .....ITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYCSSKFAAVGFHRGLTSELQ 204
E_형태 153 KDIRSNYLDVYRIEDTFGRDSEITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYCSSKFAAVGFHRGLTSELQ 228
F_형태 151 .....ITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYCSSKFAAVGFHRGLTSELQ 204
G_형태 115 .....ITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYCSSKFAAVGFHRGLTSELQ 168
H_형태 153 KDIRSNYLDVYRIEDTFGRDSEITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYCSSKFAAVGFHRGLTSELQ 228

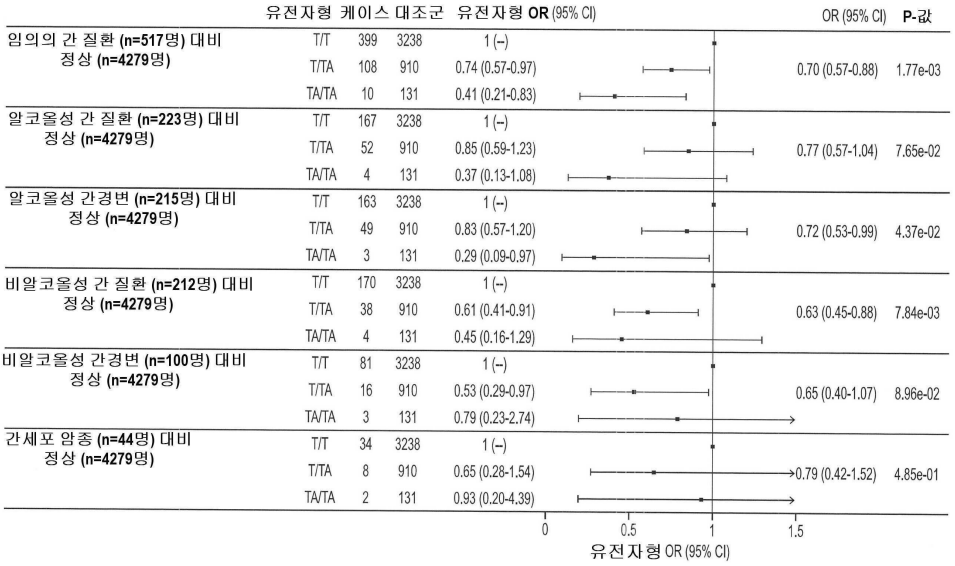
도 7A

도면7b



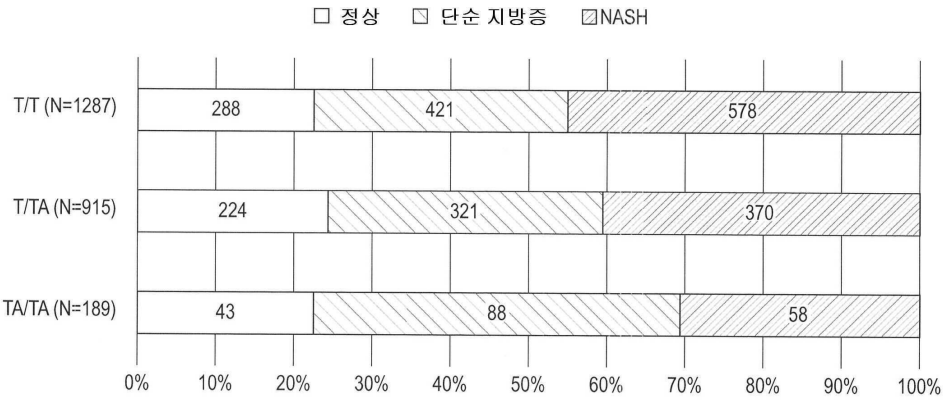
도 7B

도면8



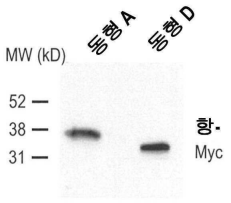
도 8

도면9



도 9

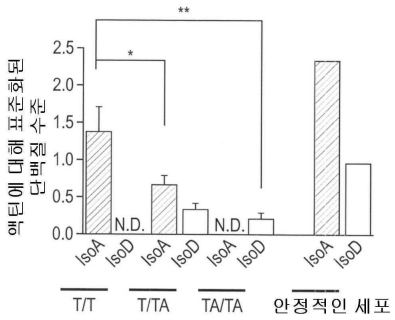
도면10ac



도 10A

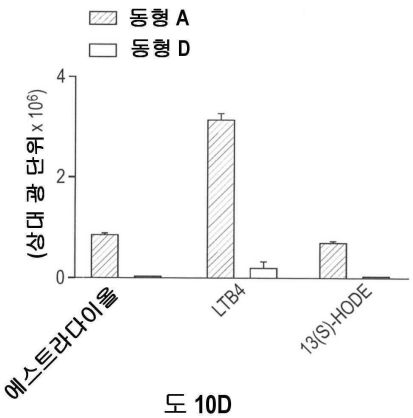


도 10B

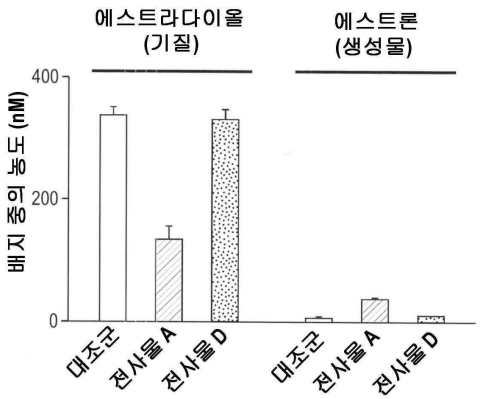


도 10C

도면10de

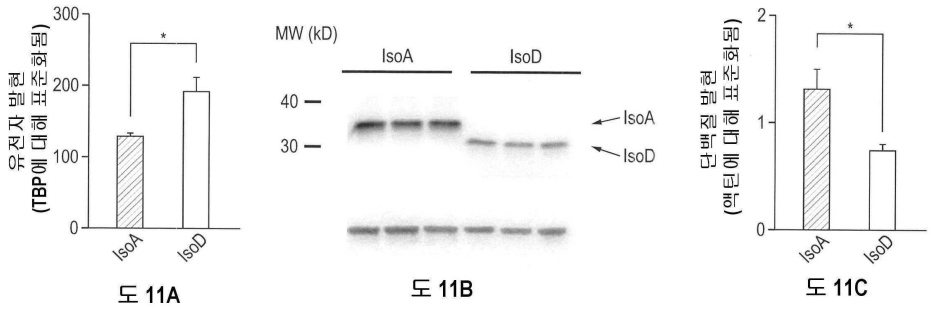


도 10D



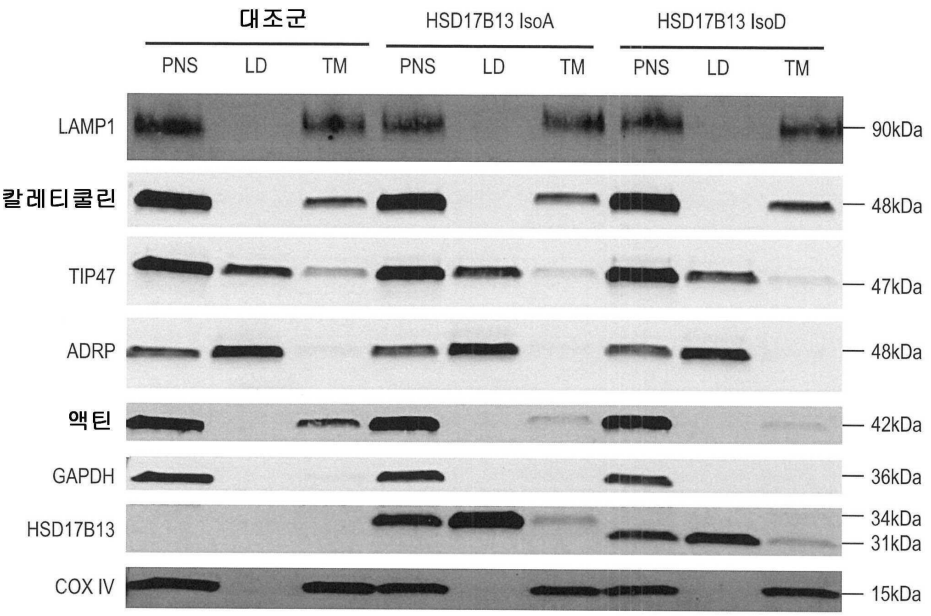
도 10E

도면11



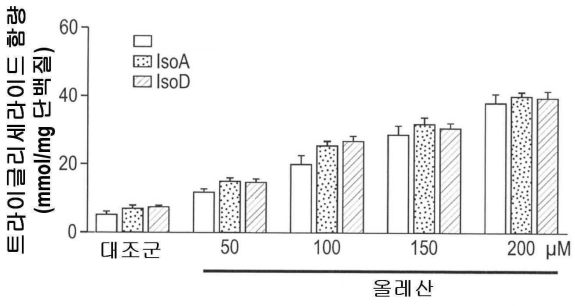
도 11A 도 11B 도 11C

도면12

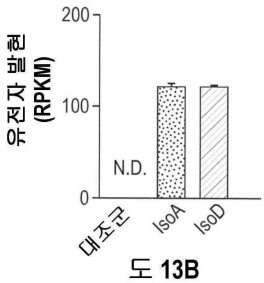


도 12

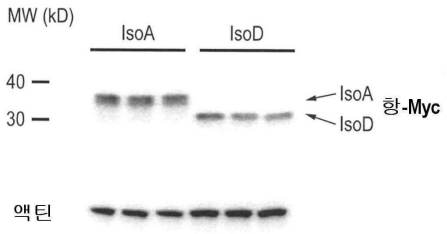
도면13



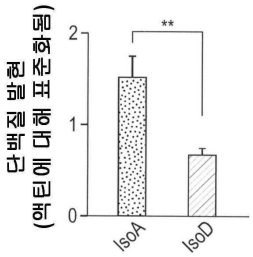
도 13A



도 13B

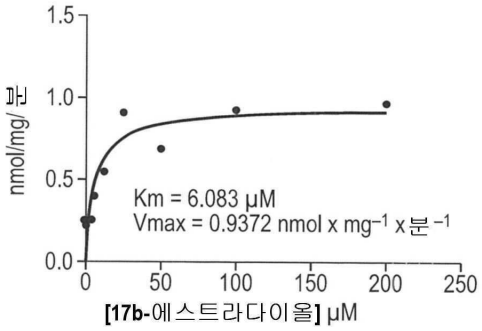


도 13C



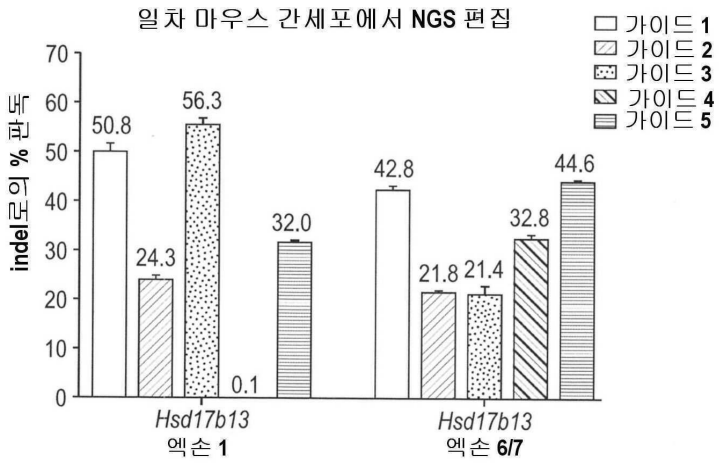
도 13D

도면14



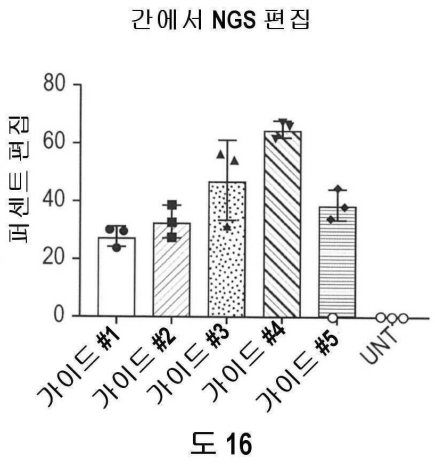
도 14

도면15



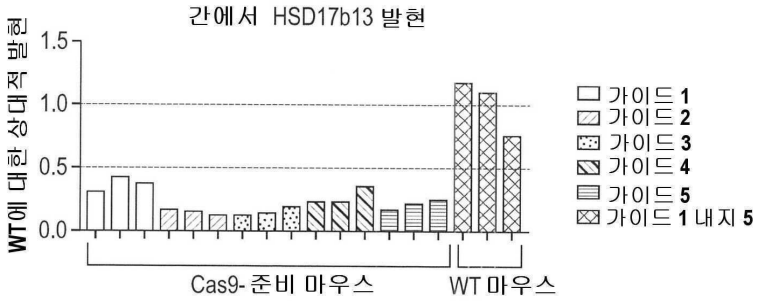
도 15

도면16

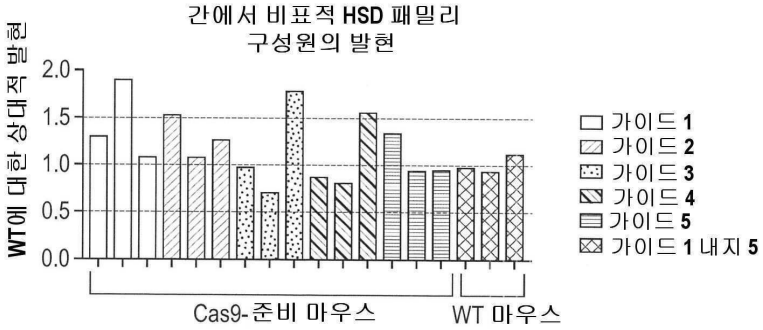


도 16

도면17



도 17A



도 17B

서 열 목 록

SEQUENCE LISTING

<110> Regeneron Pharmaceuticals, Inc.

<120> HSD17B13 Variants And Uses Thereof

<130> 057766/507242

<150> US 62/449,335

<151> 2017-01-23

<150> US 62/472,972

<151> 2017-03-17

<150> US 62/581,918

<151> 2017-11-06

<160> 1652

<170> PatentIn version 3.5

<210> 1

<211> 19118

<212> DNA

<213> Homo sapiens

<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

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

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 tcccccttac 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 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

tttttttttt tttttttttt 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 tttttttttt tttttttttt 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 ttttatttttt 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 sapiens

<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

gacccttac 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

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 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

tttttttgag 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 ttttttttt 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
ttttttgag 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 atttttttaa 19020
aagtcacaaa aacaattcac taaaaataaa tatgtcattg tgctttaaaa aaataacctc 19080
ttgtagttat aaaataaaac gtttgacttc taaactctg 19119

<210> 3
<211> 600
<212> DNA
<213> Homo sapiens
<220><221> misc_feature
<222> (1)..(499)
<223> Upstream of TSS
<400> 3
cactgcacca cgaactcatg gactgaagca atcctcctgc ctcagcctcc tgggtagctg 60
ggactacaga cacatgccac catatccagc taatttttt ctatagtttt ttttttttt 120
tttgagacag ggtcttacta tgttgcccag actggtctcg aactcctggg ctcaagcaat 180
cctctgcctc agcctcccaa agtgctggga ttacagatgt gagccactgc acctggcccc 240
tagaattgtt tctagaggtg aaacttcaag gtgaaatata gtacataact gcttttcaga 300

taaacaagtc cagagagcac actctcttgt gctcttggca tcacttggca tcacttcata 360
tttgaggtgt ttcaaaccca ttagaacacg tgaacaaggc ctgcttccaa agctggcttc 420
catctggtag tcccattaac aactgggcac acccctccc tagagctctg tgtagacagt 480
acctcctccc taggactaca caaggactga accagaagga agaggacaga gcaaagccat 540
gaacatcatc ctagaaatcc ttctgcttct gatcaccatc atctactcct acttggagtc 600
<210> 4
<211> 2397
<212> DNA
<213> Homo sapiens
<220><221> misc_feature
<222> (1)..(275)

<223> Exon 1

<220>

<221> misc_feature

<222> (276)..(383)

<223> Exon 2

<220><221> misc_feature

<222> (384)..(515)

<223> Exon 3

<220><221> misc_feature

<222> (516)..(622)

<223> Exon 4

<220><221> misc_feature

<222> (623)..(760)

<223> Exon 5

<220><221> misc_feature

<222> (761)..(877)

<223> Exon 6v1

<220><221> misc_feature

<222> (878)..(2397)

<223> Exon 7

<400> 4

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

ttttgtgaa 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> 5

<211> 2289

<212> DNA

<213> Homo sapiens

<220><221> misc_feature

<222> (1)..(275)

<223> Exon 1

<220><221> misc_feature

<222> (276)..(407)

<223> Exon 3

<220><221> misc_feature

<222> (408)..(514)

<223> Exon 4

<220><221> misc_feature

<222

> (515)..(652)

<223> Exon 5

<220><221> misc_feature

<222> (653)..(769)

<223> Exon 6v1

<220><221> misc_feature

<222> (770)..(2289)

<223> Exon 7

<400> 5

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 tcttcctttt 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 tgtttttttg 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 tttttttttt 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> 6

<211> 2280

<212> DNA

<213> Homo sapiens

<220><221> misc_feature

<222> (1)..(275)

<223> Exon 1

<220><221> misc_feature

<222> (276)..(383)

<223> Exon 2

<220><221> misc_feature

<222> (384)..(515)

<223> Exon 3

<220><221> misc_feature

<222> (516)..(622)

<223> Exon 4

<220><221> misc_feature

<222> (623)..(760)

<223> Exon 5

<220><221> misc_feature

<222> (761)..(2280)

<223> Exon 7

<400> 6

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> 7

<211> 2398

<212> DNA

<213> Homo sapiens

<220><221> misc_feature

<222> (1)..(275)

<223> Exon 1

<220><221> misc_feature

<222> (276)..(383)

<223> Exon 2

<220><221> misc_feature

<222> (384)..(515)

<223> Exon 3

<220><221> misc_feature

<222> (516)..(622)

<223> Exon 4

<220><221> misc_feature

<222> (623)..(760)

<223> Exon 5

<220><221> misc_feature

<222> (761)..(878)

<223

> Exon 6v2 - includes additional residue 878 at 3' end

<220><221> misc_feature

<222> (879)..(2398)

<223> Exon 7

<400> 7

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 ttttattttt 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 ttttttttaaa agtcacaaaa acaattcact aaaaataaat atgtcattgt 2340

gctttaaaaa aataacctct tgtagttata aaataaaacg tttgacttct aaactctg 2398

<210> 8

<211> 2469

<212> DNA

<213> Homo sapiens

<220><221> misc_feature

<222> (1)..(275)

<223> Exon 1

<220><221> misc_feature

<222> (276)..(383)

<223> Exon 2

<220><221> misc_feature

<222> (384)..(515)

<223> Exon 3

<220><221> misc_feature

<222> (516)..(587)

<223> Exon 3'

<220><221> misc_feature

<222> (588)..(694)

<223> Exon 4

<220><221> misc_feature

<222> (695)..(832)

<223> Exon 5

<220><221> misc_feature

<222> (833)..(949)

<223> Exon 6v1

<220><221> misc_feature

<222> (950)..(2469)

<223> Exon 7

<400> 8

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 ttttttttttt 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> 9

<211> 1715

<212> DNA

<213> Homo sapiens

<220><221> misc_feature

<222> (1)..(275)

<223> Exon 1

<220><221> misc_feature

<222> (276)..(383)

<223> Exon 2

<220><221> misc_feature

<222> (384)..(515)

<223> Exon 3

<220><221> misc_feature

<222> (516)..(622)

<223> Exon 4

<220><221> misc_feature

<222> (623)..(760)

<223> Exon 5

<220><221> misc_feature

<222> (761)..(1715)

<223

> Exon 6v3

<220><221> misc_feature

<222> (878)..(1715)

<223> Read-through from exon 6 into intron 6

<220><221> misc_feature

<222> (1716)..(1715)

<223> Read-through from exon 6 into intron 6

<400> 9

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> 10

<211> 2290

<212> DNA

<213> Homo sapiens

<220><221> misc_feature

<222> (1)..(275)

<223> Exon 1

<220><221> misc_feature

<222> (276)..(407)

<223> Exon 3

<220><221> misc_feature

<222> (408)..(514)

<223> Exon 4

<220><221> misc_feature

<222> (515)..(652)

<223> Exon 5

<220><221> misc_feature

<222> (653)..(770)

<223> Exon 6v2 - includes additional residue 770 at 3' end

<220><221> misc_feature

<222> (771)..(2290)

<223> Exon 7

<400> 10

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 ctttttttttt ttaatagaaa aaataatccg actcccacta catcaagact 2100

aatcttgttt tgtgtgtttt tcacatgtat tatagaatgc ttttgcatgg actatcctct 2160

tgtttttatt aaaaacaaat gatttttttа aaagtcacaa aaacaattca ctaaaaataa 2220

atatgtcatt gtgctttaaa aaaataacct cttgtagtta taaaataaaa cgtttgactt 2280

ctaaactctg 2290

<210> 11

<211> 2470

<212> DNA

<213> Homo sapiens

<220><221> misc_feature

<222> (1)..(275)

<223> Exon 1

<220><221> misc_feature

<222> (276)..(383)

<223> Exon 2

<220><221> misc_feature

<222> (384)..(515)

<223> Exon 3

<220><221> misc_feature

<222> (516)..(587)

<223> Exon 3'

<220><221> misc_feature

<222> (588)..(694)

<223> Exon 4

<220><221> misc_feature

<222> (695)..(832)

<223

> Exon 5

<220><221> misc_feature

<222> (833)..(950)

<223> Exon 6v2 - includes additional residue 950 at 3' end

<220><221> misc_feature

<222> (951)..(2470)

<223> Exon 7

<400> 11

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 ttgtttttttt 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> 12

<211> 300

<212> PRT

<213> Homo sapiens

<220><221> MISC_FEATURE

<222> (1)..(70)

<223> Region Encoded by Exon 1

<220><221> MISC_FEATURE

<222> (71)..(106)

<223> Region Encoded by Exon 2

<220><221> MISC_FEATURE

<222> (107)..(150)

<223> Region Encoded by Exon 3

<220><221> MISC_FEATURE

<222> (151)..(185)

<223>

Region Encoded by Exon 4

<220><221> MISC_FEATURE

<222> (186)..(232)

<223> Region Encoded by Exon 5

<220><221> MISC_FEATURE

<222> (233)..(271)

<223> Region Encoded by Exon 6v1

<220><221> MISC_FEATURE

<222> (272)..(300)

<223> Region Encoded by Exon 7

<400> 12

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> 13

<211> 264

<212> PRT

<213> Homo sapiens

<220><221> MISC_FEATURE

<222> (1)..(70)

<223> Region Encoded by Exon 1

<220><221> MISC_FEATURE

<222> (71)..(114)

<223> Region Encoded by Exon 3

<220><221> MISC_FEATURE

<222> (115)..(149)

<223> Region Encoded by Exon 4

<220><221> MISC_FEATURE

<222> (150)..(196)

<223> Region Encoded by Exon 5

<220><221> MISC_FEATURE

<222> (197)..(235)

<223> Region Encoded by Exon 6v1

<220><221> MISC_FEATURE

<222> (236)..(264)

<223> Region Encoded by Exon 7

<400> 13

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> 14

<211> 261

<212> PRT

<213> Homo sapiens

<220><221> MISC_FEATURE

<222> (1)..(70)

<223> Region Encoded by Exon 1

<220><221> MISC_FEATURE

<222> (71)..(106)

<223> Region Encoded by Exon 2

<220><221> MISC_FEATURE

<222> (107)..(150)

<223> Region Encoded by Exon 3

<220><

221> MISC_FEATURE

<222> (151)..(185)

<223> Region Encoded by Exon 4

<220><221> MISC_FEATURE

<222> (186)..(232)

<223> Region Encoded by Exon 5

<220><221> MISC_FEATURE

<222> (233)..(261)

<223> Region Encoded by Exon 7

<400> 14

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> 15
<211> 274
<212> PRT
<213> Homo sapiens
<220><221> MISC_FEATURE
<222> (1)..(70)
<223> Region Encoded by Exon 1
<220><221> MISC_FEATURE
<222> (71)..(106)

<223> Region Encoded by Exon 2

<220><221> MISC_FEATURE

<222> (107)..(150)

<223> Region Encoded by Exon 3

<220><221> MISC_FEATURE

<222> (151)..(185)

<223> Region Encoded by Exon 4

<220><221> MISC_FEATURE

<222> (186)..(232)

<223> Region Encoded by Exon 5

<220><221> MISC_FEATURE

<222> (233)..(271)

<223> Region Encoded by Exon 6v2

<220><221> MISC_FEATURE

<222> (272)..(274)

<223> Region Encoded by Exon 7

<400> 15

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> 16

<211> 324

<212> PRT

<213> Homo sapiens

<220><221> MISC_FEATURE

<222> (1)..(70)

<223> Region Encoded by Exon 1

<220><221> MISC_FEATURE

<222> (71)..(106)

<223> Region Encoded by Exon 2

<220><221> MISC_FEATURE

<222> (107)..(150)

<223> Region Encoded by Exon 3

<220><221> MISC_FEATURE

<222> (151)..(174)

<223> Region Encoded by Exon 3'

<220><221> MISC_FEATURE

<222> (175)..(209)

<223> Region Encoded by Exon 4

<220><221> MISC_FEATURE

<222> (210)..(256)

<223> Region Encoded by Exon 5

<220><221> MISC_FEATURE

<222> (257)..(295)

<223> Region Encoded by Exon 6v1

<220><221> MISC_FEATURE

<222> (296)..(324)

<223> Region Encoded by Exon 7

<400> 16

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> 17
<211> 284
<212> PRT
<213> Homo sapiens

<220><221> MISC_FEATURE

<222> (1)..(70)

<223> Region Encoded by Exon 1

<220><221> MISC_FEATURE

<222> (71)..(106)

<223> Region Encoded by Exon 2

<220><221> MISC_FEATURE

<222> (107)..(150)

<223> Region Encoded by Exon 3

<220><221> MISC_FEATURE

<222> (151)..(185)

<223> Region Encoded by Exon 4

<220><221> MISC_FEATURE

<222> (186)..(232)

<223> Region Encoded by Exon 5

<220><221> MISC_FEATURE

<222> (233)..(284)

<223> Region Encoded by Exon 6v3

<220><221> MISC_FEATURE

<222> (272)..(284)

<223> Region Encoded by Read-Through Into Intron 6

<400> 17

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> 18

<211> 238

<212> PRT

<213> Homo sapiens

<220><221> MISC_FEATURE

<222> (1)..(70)

<223> Region Encoded by Exon 1

<220><221> MISC_FEATURE

<222> (71)..(114)

<223> Region Encoded by Exon 3

<220><221> MISC_FEATURE

<222> (115)..(149)

<223> Region Encoded by Exon 4

<220><221> MISC_FEATURE

<222> (150)..(196)

<223> Region Encoded by Exon 5

<220><221> MISC_FEATURE

<222> (197)..(235)

<223> Region Encoded by Exon 6v2

<220><221> MISC_FEATURE

<222> (236)..(238)

<223> Region Encoded by Exon 7

<400> 18

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> 19

<211> 298

<212> PRT

<213> Homo sapiens

<220

><221> MISC_FEATURE

<222> (1)..(70)

<223> Region Encoded by Exon 1

<220><221> MISC_FEATURE

<222> (71)..(106)

<223> Region Encoded by Exon 2

<220><221> MISC_FEATURE

<222> (107)..(150)

<223> Region Encoded by Exon 3

<220><221> MISC_FEATURE

<222> (151)..(174)

<223> Region Encoded by Exon 3'

<220><221> MISC_FEATURE

<222> (175)..(209)

<223> Region Encoded by Exon 4

<220><221> MISC_FEATURE

<222> (210)..(256)

<223> Region Encoded by Exon 5

<220><221> MISC_FEATURE

<222> (257)..(295)

<223> Region Encoded by Exon 6v2

<220><221> MISC_FEATURE

<222> (296)..(298)

<223> Region Encoded by Exon 7

<400> 19

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> 20

<211> 20

<212>

DNA

<213> Homo sapiens

<400> 20

tgaagttttt cattcctcag 20

<210> 21

<211> 20

<212> DNA

<213> Homo sapiens

<400> 21

cttcaccaac gactccaagt 20

<210> 22

<211> 20

<212> DNA

<213> Homo sapiens

<400> 22

ctactcctac ttggagtcgt 20

<210> 23

<211> 20

<212> DNA

<213> Homo sapiens

<400> 23

ctccaagtag gagtagatga 20

<210> 24

<211> 20

<212> DNA

<213> Homo sapiens

<400> 24

caccatcatc tactcctact 20

<210> 25

<211> 20

<212> DNA

<213> Homo sapiens

<400> 25

tgatggtgat cagaagcaga 20

<210> 26

<211> 20

<212> DNA

<213> Homo sapiens

<400> 26

tcagaagcag aaggatttct 20

<210> 27

<211> 20

<212> DNA

<213> Homo sapiens

<400> 27

gatttctagg atgatgttca 20

<210> 28

<211> 20

<212> DNA

<213> Homo sapiens

<400> 28

ttgctctgtc ctcttccttc 20

<210> 29

<211> 20

<212> DNA

<213> Homo sapiens

<400> 29

aggactgaac cagaaggaag 20

<210> 30

<211> 20

<212> DNA

<213> Homo sapiens

<400> 30

tacacaagga ctgaaccaga 20

<210> 31

<211> 20

<212> DNA

<213> Homo sapiens

<400> 31

ttcagtcctt gtgtagtcct 20

<210> 32

<211> 20

<212> DNA

<213> Homo sapiens

<400> 32

tcagtccttg tgtagtccta 20

<210> 33

<211> 20

<212> DNA

<213> Homo sapiens

<400> 33

gtccttgtgt agtcctaggg 20

<210> 34

<211> 20

<212> DNA

<213> Homo sapiens

<400> 34

cttgtgtagt cctagggagg 20

<210> 35

<211> 20

<212> DNA

<213> Homo sapiens

<400> 35

ctcctcccta ggactacaca 20

<210> 36

<211> 20

<212> DNA

<213> Homo sapiens

<400> 36

gtagacagta cctcctccct 20

<210> 37

<211> 20

<212> DNA

<213> Homo sapiens

<400> 37

tactgtctac acagagctct 20

<210> 38

<211> 20

<212> DNA

<213> Homo sapiens

<400> 38

actgtctaca cagagctcta 20

<210> 39

<211> 20

<212> DNA

<213> Homo sapiens

<400> 39

tctacacaga gctctaggga 20

<210> 40

<211> 20

<212> DNA

<213> Homo sapiens

<400> 40

ctacacagag ctctagggaa 20

<210> 41

<211> 20

<212> DNA

<213> Homo sapiens

<400> 41

tacacagagc tctagggaag 20

<210> 42

<211> 20

<212> DNA

<213> Homo sapiens

<400> 42

tgtcaggtta gttagatgaa 20

<210> 43

<211> 20

<212> DNA

<213> Homo sapiens

<400> 43

gtgtcaggtt agttagatga 20

<210> 44

<211> 20

<212> DNA

<213> Homo sapiens

<400> 44

cctgacacat atacagacta 20

<210> 45

<211> 20

<212> DNA

<213> Homo sapiens

<400> 45

ctgacacata tacagactaa 20

<210> 46

<211> 20

<212> DNA

<213> Homo sapiens

<400> 46

ccttagtctg tatatgtgtc 20

<210> 47

<211> 20

<212> DNA

<213> Homo sapiens

<400> 47

catatacaga ctaagggacc 20

<210> 48

<211> 20

<212> DNA

<213> Homo sapiens

<400> 48

atatacagac taagggacca 20

<210> 49

<211> 20

<212> DNA

<213> Homo sapiens

<400> 49

tcaaagtttg ataaattccc 20

<210> 50

<211> 20

<212> DNA

<213> Homo sapiens

<400> 50

aaaatacaaa gataagtaga 20

<210> 51

<211> 20

<212> DNA

<213> Homo sapiens

<400> 51

actctgtgac tttaaaaagt 20

<210> 52

<211> 20

<212> DNA

<213> Homo sapiens

<400> 52

ggttctgtgg gatattaata 20

<210> 53

<211> 20

<212> DNA

<213> Homo sapiens

<400> 53

acagagcata ttggttctgt 20

<210> 54

<211> 20

<212> DNA

<213> Homo sapiens

<400> 54

gacagagcat attggttctg 20

<210> 55

<211> 20

<212> DNA

<213> Homo sapiens

<400> 55

tgcaaaacga cagagcatat 20

<210> 56

<211> 20

<212> DNA

<213> Homo sapiens

<400> 56

gagctgggca tggaataggc 20

<210> 57

<211> 20

<212> DNA

<213> Homo sapiens

<400> 57

actggagctg ggcatggaat 20

<210> 58

<211> 20

<212> DNA

<213> Homo sapiens

<400> 58

ctcattactg gagctgggca 20

<210> 59

<211> 20

<212> DNA

<213> Homo sapiens

<400> 59

ttgttctcat tactggagct 20

<210> 60

<211> 20

<212> DNA

<213> Homo sapiens

<400> 60

attgttctca ttactggagc 20

<210> 61

<211> 20

<212> DNA

<213> Homo sapiens

<400> 61

ggggagattg ttctcattac 20

<210> 62

<211> 20

<212> DNA

<213> Homo sapiens

<400> 62

gaggagaaaa tctgtggctg 20

<210> 63

<211> 20

<212> DNA

<213> Homo sapiens

<400> 63

agaggagaaa atctgtggct 20

<210> 64

<211> 20

<212> DNA

<213> Homo sapiens

<400> 64

cagaggagaa aatctgtggc 20

<210> 65

<211> 20

<212> DNA

<213> Homo sapiens

<400> 65

tcctcagagg agaaaatctg 20

<210> 66

<211> 20

<212> DNA

<213> Homo sapiens

<400> 66

ggggtgtgcc cagttgttaa 20

<210> 67

<211> 20

<212> DNA

<213> Homo sapiens

<400> 67

gggtgtgccc agttgttaat 20

<210> 68

<211> 20

<212> DNA

<213> Homo sapiens

<400> 68

tggtagtccc attaacaact 20

<210> 69

<211> 20

<212> DNA

<213> Homo sapiens

<400> 69

ctggtagtcc cattaacaac 20

<210> 70

<211> 20

<212> DNA

<213> Homo sapiens

<400> 70

ttgttaatgg gactaccaga 20

<210> 71

<211> 20

<212> DNA

<213> Homo sapiens

<400> 71

taccagatgg aagccagctt 20

<210> 72

<211> 20

<212> DNA

<213> Homo sapiens

<400> 72

ttccaaagct ggcttccatc 20

<210> 73

<211> 20

<212> DNA

<213> Homo sapiens

<400> 73

tggaagccag ctttggaagc 20

<210> 74

<211> 20

<212> DNA

<213> Homo sapiens

<400> 74

acaaggcctg cttccaaagc 20

<210> 75

<211> 20

<212> DNA

<213> Homo sapiens

<400> 75

gccttgttca cgtgttctaa 20

<210> 76

<211> 20

<212> DNA

<213> Homo sapiens

<400> 76

ccttgttcac gtgttctaat 20

<210> 77

<211> 20

<212> DNA

<213> Homo sapiens

<400> 77

cccattagaa cacgtgaaca 20

<210> 78

<211> 20

<212> DNA

<213> Homo sapiens

<400> 78

ttggcatcac ttcatatttg 20

<210> 79

<211> 20

<212> DNA

<213> Homo sapiens

<400> 79

cttgtgctct tggcatcact 20

<210> 80

<211> 20

<212> DNA

<213> Homo sapiens

<400> 80

agcacactct cttgtgctct 20

<210> 81

<211> 20

<212> DNA

<213> Homo sapiens

<400> 81

gcacaagaga gtgtgctctc 20

<210> 82

<211> 20

<212> DNA

<213> Homo sapiens

<400> 82

gcttaatctc acacatagaa 20

<210> 83

<211> 20

<212> DNA

<213> Homo sapiens

<400> 83

cttaatctca cacatagaaa 20

<210> 84

<211> 20

<212> DNA

<213> Homo sapiens

<400> 84

ttaatctcac acatagaaag 20

<210> 85

<211> 20

<212> DNA

<213> Homo sapiens

<400> 85

aggagtgctg gtttatcaac 20

<210> 86

<211> 20

<212> DNA

<213> Homo sapiens

<400> 86

ttctttgaca gcaggagtgc 20

<210> 87

<211> 20

<212> DNA

<213> Homo sapiens

<400> 87

actctggttt ctttgacagc 20

<210> 88

<211> 20

<212> DNA

<213> Homo sapiens

<400> 88

accagagttg agaaaacccc 20

<210> 89

<211> 20

<212> DNA

<213> Homo sapiens

<400> 89

tccaggggtt ttctcaactc 20

<210> 90

<211> 20

<212> DNA

<213> Homo sapiens

<400> 90

cagttattaa atgaatccag 20

<210> 91

<211> 20

<212> DNA

<213> Homo sapiens

<400> 91

gcagttatta aatgaatcca 20

<210> 92

<211> 20

<212> DNA

<213> Homo sapiens

<400> 92

ggcagttatt aaatgaatcc 20

<210> 93

<211> 20

<212> DNA

<213> Homo sapiens

<400> 93

tggatggtaa cagctacatc 20

<210> 94

<211> 20

<212> DNA

<213> Homo sapiens

<400> 94

gctgttacca tccacatcct 20

<210> 95

<211> 20

<212> DNA

<213> Homo sapiens

<400> 95

tcaagaacca aggatgtgga 20

<210> 96

<211> 20

<212> DNA

<213> Homo sapiens

<400> 96

tccttcaaga accaaggatg 20

<210> 97

<211> 20

<212> DNA

<213> Homo sapiens

<400> 97

tgagtgtcct tcaagaacca 20

<210> 98

<211> 20

<212> DNA

<213> Homo sapiens

<400> 98

ttttatttta taactacaag 20

<210> 99

<211> 20

<212> DNA

<213> Homo sapiens

<400> 99

ttgtttttaa taaaaacaag 20

<210> 100

<211> 20

<212> DNA

<213> Homo sapiens

<400> 100

tattatagaa tgcttttgca 20

<210> 101

<211> 20

<212> DNA

<213> Homo sapiens

<400> 101

caagattagt cttgatgtag 20

<210> 102

<211> 20

<212> DNA

<213> Homo sapiens

<400> 102

aagattagtc ttgatgtagt 20

<210> 103

<211> 20

<212> DNA

<213> Homo sapiens

<400> 103

agtcttgatg tagtgggagt 20

<210> 104

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 104

tttttctatt aaaaaaaaaa 20

<210> 105

<211> 20

<212> DNA

<213> Homo sapiens

<400> 105

tctattaaaa aaaaaaaggc 20

<210> 106

<211> 20

<212> DNA

<213> Homo sapiens

<400> 106

ctattaaaaa aaaaaaggct 20

<210> 107

<211> 20

<212> DNA

<213> Homo sapiens

<400> 107

aaaaaaaaaa aggctgggca 20

<210> 108

<211> 20

<212> DNA

<213> Homo sapiens

<400> 108

aaaaaaaagg ctgggcacgg 20

<210> 109

<211> 20

<212> DNA

<213> Homo sapiens

<400> 109

cacccgtaat cccagcactt 20

<210> 110

<211> 20

<212> DNA

<213> Homo sapiens

<400> 110

acccgtaatc ccagcacttt 20

<210> 111

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 111

cgtaatccca gcactttggg 20

<210> 112

<211> 20

<212> DNA

<213> Homo sapiens

<400> 112

tcccaaagtg ctgggattac 20

<210> 113

<211> 20

<212> DNA

<213> Homo sapiens

<400> 113

ctcccaaagt gctgggatta 20

<210> 114

<211> 20

<212> DNA

<213> Homo sapiens

<400> 114

cccagcactt tgggaggccg 20

<210> 115

<211> 20

<212> DNA

<213> Homo sapiens

<400> 115

cctcggcctc ccaaagtgct 20

<210> 116

<211> 20

<212> DNA

<213> Homo sapiens

<400> 116

gcactttggg aggccgaggc 20

<210> 117

<211> 20

<212> DNA

<213> Homo sapiens

<400> 117

ctttgggagg ccgaggcagg 20

<210> 118

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 118

gccgaggcag gtggatcacg 20

<210> 119

<211> 20

<212> DNA

<213> Homo sapiens

<400> 119

acctcgtgat ccacctgcct 20

<210> 120

<211> 20

<212> DNA

<213> Homo sapiens

<400> 120

ggcaggtgga tcacgaggtc 20

<210> 121

<211> 20

<212> DNA

<213> Homo sapiens

<400> 121

tcaggagatc gagaccatct 20

<210> 122

<211> 20

<212> DNA

<213> Homo sapiens

<400> 122

cgagaccatc ttggctaaca 20

<210> 123

<211> 20

<212> DNA

<213> Homo sapiens

<400> 123

tttcaccatg ttagccaaga 20

<210> 124

<211> 20

<212> DNA

<213> Homo sapiens

<400> 124

ttgtattttt tgtagagacg 20

<210> 125

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 125

tttgtatttt ttgtagagac 20

<210> 126

<211> 20

<212> DNA

<213> Homo sapiens

<400> 126

ttttgtattt tttgtagaga 20

<210> 127

<211> 20

<212> DNA

<213> Homo sapiens

<400> 127

aaaaaataca aaaaattagc 20

<210> 128

<211> 20

<212> DNA

<213> Homo sapiens

<400> 128

aaaaatacaa aaaattagcc 20

<210> 129

<211> 20

<212> DNA

<213> Homo sapiens

<400> 129

tacaaaaaat tagccgggtg 20

<210> 130

<211> 20

<212> DNA

<213> Homo sapiens

<400> 130

aaaaaattag ccgggtgtgg 20

<210> 131

<211> 20

<212> DNA

<213> Homo sapiens

<400> 131

aaattagccg ggtgtggtgg 20

<210> 132

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 132

aattagccgg gtgtggtggc 20

<210> 133

<211> 20

<212> DNA

<213> Homo sapiens

<400> 133

caggcgcccg ccaccacacc 20

<210> 134

<211> 20

<212> DNA

<213> Homo sapiens

<400> 134

gcctgtagtc ccagctactc 20

<210> 135

<211> 20

<212> DNA

<213> Homo sapiens

<400> 135

tgtagtccca gctactcagg 20

<210> 136

<211> 20

<212> DNA

<213> Homo sapiens

<400> 136

tcctgagtag ctgggactac 20

<210> 137

<211> 20

<212> DNA

<213> Homo sapiens

<400> 137

cccagctact caggaggctg 20

<210> 138

<211> 20

<212> DNA

<213> Homo sapiens

<400> 138

cctcagcctc ctgagtagct 20

<210> 139

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 139

gcctcagcct cctgagtagc 20

<210> 140

<211> 20

<212> DNA

<213> Homo sapiens

<400> 140

aggaggctga ggcaggagaa 20

<210> 141

<211> 20

<212> DNA

<213> Homo sapiens

<400> 141

gcaggagaat ggcgtgaacc 20

<210> 142

<211> 20

<212> DNA

<213> Homo sapiens

<400> 142

caggagaatg gcgtgaaccc 20

<210> 143

<211> 20

<212> DNA

<213> Homo sapiens

<400> 143

gagaatggcg tgaacccggg 20

<210> 144

<211> 20

<212> DNA

<213> Homo sapiens

<400> 144

aatggcgtga acccgggagg 20

<210> 145

<211> 20

<212> DNA

<213> Homo sapiens

<400> 145

cactgcaagc tccacctccc 20

<210> 146

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 146

tcactgcaag ctccacctcc 20

<210> 147

<211> 20

<212> DNA

<213> Homo sapiens

<400> 147

cataccactg cactccagcc 20

<210> 148

<211> 20

<212> DNA

<213> Homo sapiens

<400> 148

ataccactgc actccagcct 20

<210> 149

<211> 20

<212> DNA

<213> Homo sapiens

<400> 149

tcgcccaggc tggagtgcag 20

<210> 150

<211> 20

<212> DNA

<213> Homo sapiens

<400> 150

tctcactctt tcgcccaggc 20

<210> 151

<211> 20

<212> DNA

<213> Homo sapiens

<400> 151

ggagtctcac tctttcgccc 20

<210> 152

<211> 20

<212> DNA

<213> Homo sapiens

<400> 152

tgttttttgt ttttttgaga 20

<210> 153

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 153

aggaagaaag aaaggttttt 20

<210> 154

<211> 20

<212> DNA

<213> Homo sapiens

<400> 154

agaagaaaag gaagaaagaa 20

<210> 155

<211> 20

<212> DNA

<213> Homo sapiens

<400> 155

ctttcttcct tttcttctct 20

<210> 156

<211> 20

<212> DNA

<213> Homo sapiens

<400> 156

tttcttcctt ttcttctctt 20

<210> 157

<211> 20

<212> DNA

<213> Homo sapiens

<400> 157

aatggaccca agagaagaaa 20

<210> 158

<211> 20

<212> DNA

<213> Homo sapiens

<400> 158

ggctattaca taagaaacaa 20

<210> 159

<211> 20

<212> DNA

<213> Homo sapiens

<400> 159

cacaggaaaa ggaactgtac 20

<210> 160

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 160

attaaagcta acacaggaaa 20

<210> 161

<211> 20

<212> DNA

<213> Homo sapiens

<400> 161

tcaaaaatta aagctaacac 20

<210> 162

<211> 20

<212> DNA

<213> Homo sapiens

<400> 162

taaaattgtc taaacatctc 20

<210> 163

<211> 20

<212> DNA

<213> Homo sapiens

<400> 163

agagatgttt agacaatttt 20

<210> 164

<211> 20

<212> DNA

<213> Homo sapiens

<400> 164

tctaaacatc tctgggacca 20

<210> 165

<211> 20

<212> DNA

<213> Homo sapiens

<400> 165

tttatgcttt catatatcct 20

<210> 166

<211> 20

<212> DNA

<213> Homo sapiens

<400> 166

agcataaatt acaaagaaaa 20

<210> 167

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 167

tacaaagaaa aaggttatca 20

<210> 168

<211> 20

<212> DNA

<213> Homo sapiens

<400> 168

acaaagaaaa aggttatcat 20

<210> 169

<211> 20

<212> DNA

<213> Homo sapiens

<400> 169

caaagaaaaa ggttatcatg 20

<210> 170

<211> 20

<212> DNA

<213> Homo sapiens

<400> 170

tctgagattt aaaatagagt 20

<210> 171

<211> 20

<212> DNA

<213> Homo sapiens

<400> 171

cttataagat acattatgaa 20

<210> 172

<211> 20

<212> DNA

<213> Homo sapiens

<400> 172

tatcttataa gactataaaa 20

<210> 173

<211> 20

<212> DNA

<213> Homo sapiens

<400> 173

atcttataag actataaaaa 20

<210> 174

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 174

ttataagact ataaaaaggg 20

<210> 175

<211> 20

<212> DNA

<213> Homo sapiens

<400> 175

taaaaaggga ggaaatatag 20

<210> 176

<211> 20

<212> DNA

<213> Homo sapiens

<400> 176

aaaaagggag gaaatataga 20

<210> 177

<211> 20

<212> DNA

<213> Homo sapiens

<400> 177

aaatatagag ggtccacttt 20

<210> 178

<211> 20

<212> DNA

<213> Homo sapiens

<400> 178

tatagagggt ccacttttgg 20

<210> 179

<211> 20

<212> DNA

<213> Homo sapiens

<400> 179

actctgaagt ccaccaaaag 20

<210> 180

<211> 20

<212> DNA

<213> Homo sapiens

<400> 180

agaatagagt tgcaccgttt 20

<210> 181

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 181

aaaacggtgc aactctattc 20

<210> 182

<211> 20

<212> DNA

<213> Homo sapiens

<400> 182

ccgttttggg ctaatgaaaa 20

<210> 183

<211> 20

<212> DNA

<213> Homo sapiens

<400> 183

cctttttcat tagcccaaaa 20

<210> 184

<211> 20

<212> DNA

<213> Homo sapiens

<400> 184

tgggctaatg aaaaaggaag 20

<210> 185

<211> 20

<212> DNA

<213> Homo sapiens

<400> 185

taatgaaaaa ggaagaggct 20

<210> 186

<211> 20

<212> DNA

<213> Homo sapiens

<400> 186

aatgaaaaag gaagaggcta 20

<210> 187

<211> 20

<212> DNA

<213> Homo sapiens

<400> 187

ctgaatctta aaatatgtcc 20

<210> 188

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 188

caggcagctt tatctcaacc 20

<210> 189

<211> 20

<212> DNA

<213> Homo sapiens

<400> 189

ctaagagatc aagtttcagc 20

<210> 190

<211> 20

<212> DNA

<213> Homo sapiens

<400> 190

gtgttcttgt tgatattctg 20

<210> 191

<211> 20

<212> DNA

<213> Homo sapiens

<400> 191

cttgttgata ttctgtggca 20

<210> 192

<211> 20

<212> DNA

<213> Homo sapiens

<400> 192

tctgtggcat ggctacagat 20

<210> 193

<211> 20

<212> DNA

<213> Homo sapiens

<400> 193

agaacttatt tacacaggga 20

<210> 194

<211> 20

<212> DNA

<213> Homo sapiens

<400> 194

aaagagaact tatttacaca 20

<210> 195

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 195

caaagagaac ttatttacac 20

<210> 196

<211> 20

<212> DNA

<213> Homo sapiens

<400> 196

ttctctttgt atttactttt 20

<210> 197

<211> 20

<212> DNA

<213> Homo sapiens

<400> 197

tctctttgta tttactttta 20

<210> 198

<211> 20

<212> DNA

<213> Homo sapiens

<400> 198

ctttgtattt acttttaggg 20

<210> 199

<211> 20

<212> DNA

<213> Homo sapiens

<400> 199

agcttttgtc cacctttaaa 20

<210> 200

<211> 20

<212> DNA

<213> Homo sapiens

<400> 200

tttatttttc catttaaagg 20

<210> 201

<211> 20

<212> DNA

<213> Homo sapiens

<400> 201

tattttattt ttccatttaa 20

<210> 202

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 202

cttacataaa catacttaaa 20

<210> 203

<211> 20

<212> DNA

<213> Homo sapiens

<400> 203

taagcacaga agtttttaag 20

<210> 204

<211> 20

<212> DNA

<213> Homo sapiens

<400> 204

aagtttttaa gaggcatgaa 20

<210> 205

<211> 20

<212> DNA

<213> Homo sapiens

<400> 205

atatttacgt agtttttcat 20

<210> 206

<211> 20

<212> DNA

<213> Homo sapiens

<400> 206

cgtaaatatt cttgagaaac 20

<210> 207

<211> 20

<212> DNA

<213> Homo sapiens

<400> 207

ttcttgagaa acaggaagac 20

<210> 208

<211> 20

<212> DNA

<213> Homo sapiens

<400> 208

taatattaaa aacattggtt 20

<210> 209

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 209

ccaatgtttt taatattatc 20

<210> 210

<211> 20

<212> DNA

<213> Homo sapiens

<400> 210

cctgataata ttaaaaacat 20

<210> 211

<211> 20

<212> DNA

<213> Homo sapiens

<400> 211

cattatcatg catacatctc 20

<210> 212

<211> 20

<212> DNA

<213> Homo sapiens

<400> 212

atcatgcata catctctggc 20

<210> 213

<211> 20

<212> DNA

<213> Homo sapiens

<400> 213

ttcatttcat tttgattttg 20

<210> 214

<211> 20

<212> DNA

<213> Homo sapiens

<400> 214

attcaatttg aagcagtggt 20

<210> 215

<211> 20

<212> DNA

<213> Homo sapiens

<400> 215

gaatattcaa tttgaagcag 20

<210> 216

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 216

catacgattt aaaatcgctg 20

<210> 217

<211> 20

<212> DNA

<213> Homo sapiens

<400> 217

aaaatcgctg aggcgcgttc 20

<210> 218

<211> 20

<212> DNA

<213> Homo sapiens

<400> 218

ttttttttttc ttttttgtac 20

<210> 219

<211> 20

<212> DNA

<213> Homo sapiens

<400> 219

ctgttgtcaa agattttaaa 20

<210> 220

<211> 20

<212> DNA

<213> Homo sapiens

<400> 220

tgacaacaga gttctgtttt 20

<210> 221

<211> 20

<212> DNA

<213> Homo sapiens

<400> 221

agaatacgct gagagttatc 20

<210> 222

<211> 20

<212> DNA

<213> Homo sapiens

<400> 222

gcaagagaag aaaagaacgg 20

<210> 223

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 223

gttgcaagag aagaaaagaa 20

<210> 224

<211> 20

<212> DNA

<213> Homo sapiens

<400> 224

atgcacacgt aaaagagagg 20

<210> 225

<211> 20

<212> DNA

<213> Homo sapiens

<400> 225

aagatgcaca cgtaaaagag 20

<210> 226

<211> 20

<212> DNA

<213> Homo sapiens

<400> 226

atcatgcata catctctggc 20

<210> 227

<211> 20

<212> DNA

<213> Homo sapiens

<400> 227

ttcatttcat tttgattttg 20

<210> 228

<211> 20

<212> DNA

<213> Homo sapiens

<400> 228

attcaatttg aagcagtggt 20

<210> 229

<211> 20

<212> DNA

<213> Homo sapiens

<400> 229

gaatattcaa tttgaagcag 20

<210> 230

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 230

catacgattt aaaatcgctg 20

<210> 231

<211> 20

<212> DNA

<213> Homo sapiens

<400> 231

aaaatcgctg aggcgcgttc 20

<210> 232

<211> 20

<212> DNA

<213> Homo sapiens

<400> 232

ttttttttttc ttttttgtac 20

<210> 233

<211> 20

<212> DNA

<213> Homo sapiens

<400> 233

ctgttgtcaa agattttaaa 20

<210> 234

<211> 20

<212> DNA

<213> Homo sapiens

<400> 234

tgacaacaga gttctgtttt 20

<210> 235

<211> 20

<212> DNA

<213> Homo sapiens

<400> 235

agaatacgct gagagttatc 20

<210> 236

<211> 20

<212> DNA

<213> Homo sapiens

<400> 236

gcaagagaag aaaagaacgg 20

<210> 237

<211> 20

<212> DNA

<213> Homo sapiens

<400

> 237

gttgcaagag aagaaaagaa 20

<210> 238

<211> 20

<212> DNA

<213> Homo sapiens

<400> 238

atgcacacgt aaaagagagg 20

<210> 239

<211> 20

<212> DNA

<213> Homo sapiens

<400> 239

aagatgcaca cgtaaaagag 20

<210> 240

<211> 300

<212> PRT

<213> Homo sapiens

<400> 240

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> 241

<211> 264

<212> PRT

<213> Homo sapiens

<400> 241

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> 242

<211> 300

<212> PRT

<213> Homo sapiens

<400> 242

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> 243

<211> 264

<212> PRT

<213> Homo sapiens

<400> 243

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> 244

<211> 2397

<212> DNA

<213> Homo sapiens

<400> 244

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> 245

<211> 2289

<212> DNA

<213> Homo sapiens

<400> 245

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 tcttcctttt 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 tgtttttttg 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 tttttttttt 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> 246

<211> 1714

<212> DNA

<213> Homo sapiens

<220><221> misc_feature

<222> (1)..(275)

<223> Exon 1

<220><221> misc_feature

<222> (276)..(383)

<223> Exon 2

<220><221> misc_feature

<222> (384)..(515)

<223> Exon 3

<220><221> misc_feature

<222> (516)..(622)

<223> Exon 4

<220><221> misc_feature

<222> (623)..(760)

<223> Exon 5

<220><221> misc_feature

<222> (761)..(1714)

<223> Exon 6v4

<220><221> misc_feature

<222> (878)..(1714)

<223> Read-through from exon 6 into intron 6

<400> 246

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> 247

<211> 271

<212> PRT

<213> Homo sapiens

<220><221> MISC_FEATURE

<222> (1)..(70)

<223> Region Encoded by Exon 1

<220><221> MISC_FEATURE

<222> (71)..(106)

<223> Region Encoded by Exon 2

<220><221> MISC_FEATURE

<222> (107)..(150)

<223> Region Encoded by Exon 3

<220><221> MISC_FEATURE

<222> (151)..(185)

<223> Region Encoded by Exon 4

<220><221> MISC_FEATURE

<222> (186)..(232)

<223> Region Encoded by Exon 5

<220><221> MISC_FEATURE

<222> (233)..(271)

<223> Region Encoded by Exon 6v4

<400> 247

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> 248
<211> 23
<212> DNA
<213> Artificial Sequence
<220><223> Synthetic
<220><221> misc_feature

<222>

(2)..(21)

<223> n = A, T, C, or G

<400> 248

gnnnnnnnnn nnnnnnnnnn ngg 23

<210> 249

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<220><221> misc_feature

<222> (1)..(21)

<223> n = A, T, C, or G

<400> 249

nnnnnnnnnn nnnnnnnnnn ngg 23

<210> 250

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<220><221> misc_feature

<222> (3)..(23)

<223> n = A, T, C, or G

<400> 250

ggnnnnnnnn nnnnnnnnnn nnngg 25

<210> 251

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 251

atgaacatca tcctagaaat ccttc 25

<210> 252

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 252

atcatgcata catctctggc tggag 25

<210> 253

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 253

atcagaactt caggccttgg 20

<210> 254

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 254

gcaaagccat gaacatcatc c 21

<210> 255

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 255

tcttgatgta gtgggagtcg gatt 24

<210> 256

<211> 82

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 256

guuggaacca uucaaaacag cauagcaagu uaaaauaagg cuaguccguu aucaacuuga 60

aaaaguggca ccgagucggu gc 82

<210> 257

<211> 76

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 257

guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc cguuaucaac uugaaaaagu 60

ggcaccgagu cggugc 76

<210> 258

<211> 86

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 258

guuuaagagc uaugcuggaa acagcauagc aaguuuaaau aaggcuaguc cguuaucaac 60

uugaaaaagu ggcaccgagu cggugc 86

<210> 259

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 259

ggcagaccgt tctcatcacg 20

<210> 260

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 260

ctttaccagt gactccaggt 20

<210> 261

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 261

gtcacagatt tccttctccg 20

<210> 262

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 262

agatgatgac gcccaccaga 20

<210> 263

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 263

ggagaaggaa atctgtgacc 20

<210> 264

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 264

tgcgaggaac ttacttttcc 20

<210> 265

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 265

agagaaatat tgatatagga 20

<210> 266

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 266

tatcaatatt tctctgatcc 20

<210> 267

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 267

atcgctttta aggcacgctc 20

<210> 268

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 268

tatacgactg atcgctttta 20

<210> 269

<211> 21947

<212> DNA

<213> Mus musculus

<400> 269

gcagcagccc ataatacctg acagggctct ctgtggagct ctgagcagag gtgtctgttg 60

tgagaacaga gccatgaacc tcatcctgga atttctcctt ctggtgggcg tcatcatcta 120

ctcctacctg gagtcactgg taaagttctt cattccccgg agaaggaaat ctgtgaccgg 180

gcagaccgtt ctcatcacgg gggccggaca cggaataggc aggctgactg catatgaatt 240

tgcaaagcag aaaagcagac tggttctatg ggatatcaat aaggtaataa taggtgtttc 300

taacaatctt ttaaagtcac aaaataacaa ggcatatgtt tctctagaga cgtaacttag 360

tgctggcttt tctctgtgtt ttccttttttc aaagttttaa taaactcctg ggtcctttct 420

ggatccctgc gacaggtgaa caagatggaa ccaacgagct ggagaactaa gcctgggggg 480

tttcattttg tttgactggt tggttggttg gttggttggt tctactgtgg tgtacaactt 540

gataacaaga gacgcaggag ccagagagct tcccagctgc tgaaattagg caagccacat 600

aacttacctc aacttcaaat ttaaaaaaca aaaaacaaac aaaaaactct accattggga 660

taaaatctca gagggagaca aaatctacag gtaatgattg ccagcaacca aattggccaa 720

tatgcagttt attacagcta ggggtgactt ttatactttt agattttcat ttattctgtt 780

attcatttgt tcctttctgt gatagactga agctgggtct tagtacctac cactgagcca 840

cagtacctta gcacgtgcca ggcaagcact ctatcaccga gcaaccgctc aggtttttac 900

acatttgagg gattgattct caaaacaaac aaacaaaacc acatgaagaa ctaacagcag 960

aaacccatgg tacttccact cacactccag caatcctaaa atacttacat ctgacctttt 1020

gcagaagaag aaaaaatact tgccgggtcc tgagcaagaa gggtaaagaa taatacttaa 1080

gtgcttagca tggcctggtg cacgcagcat cacatatgga aatgtctaca ccagtgcctg 1140

gctttcacta gtacatttaa gaggttgttg atctctgttt atttataaat gctcttaact 1200

accttaagtc tgaaagtgtt gacgtcctgt gtcactttct cctaaacatg cacattccca 1260

aaaggtcgtg aaaaagatgt gccacaacaa actgtctctg ggattgtaat ctttgaagaa 1320

ctggactggg gatgtagccc agcggtacaa caattatcta gaatatacaa actagatatt 1380

agatccagtt gtggctgggt gtgagcctca gcactacaag gaaaccacaa accgctgctg 1440

tcatgtaccg gatacagtgg ctcagactag agagactggg gttttggtgt ttgtttgttt 1500

gtttgtttgt ttgtttgttg agaagcagag tttccaaagc tttccttggg gttttgctcg 1560

gcttcaggta accctctgtt acgtttttgc catctgcagt taacaccact tttcccagcc 1620

ttcagaacgc tcttgaacat aaggagctaa atctgctttg gttccttttg agcagtgaag 1680

aatcccgaga ggcaccatgc atctttgctt agcagataca ggcgttggtg ggcaagatga 1740

gctaacacgg tacttggaga tgattctgag attatcaaat cctgcacagg tcccaggcag 1800

tcttcttgga gctgagaaaa ctgagtcagc tgttctcagt ccctcgtaaa aaatataaac 1860

gagaatgaga atacatcatg aagatggtcc acgggtcttc agatcacctc ccaagtgtgc 1920

atggtgggaa gacatcaata cttagtagac tacctttaat ttgaaaattt tcagacacag 1980

aaagagtaaa actagaaaga ggtcacatta aagatgactt ccacttttt tttcctgtgc 2040

ctgaaaactg agtaactcaa ccttgtaaat ttctctttga attcttgaac ttcatgatga 2100

aacatttaag acatcaaatg attgactgaa aagctgcaag tgtaactcca gtttccctgt 2160

tcactgagag gccagctagc tttccatcaa agaaatatgg ctccttttcc ttaaaaggca 2220

agaagcaagg ccacagagaa gcctcctcca cccccagctt taaaaacagt ggcacagaga 2280

tgtgaaacac aactgttagg aagggaattc agcagtgggg agggaggggc tgcaggaagg 2340

taacagggga taaaataata agagcataca tgtgtatgca aggaatgaaa agcaaattta 2400

agaagaacgc aatgacacaa tagcaaaact gtggacccag ccaaggtgcc cacccagaga 2460

catgggttgg aaaatgtggt ttcacacaca ggacatcgtt cagtcttttg gaagaatgac 2520

atcattctgt ttgctggaca gtggaaaaaa ctggcagtca caactttagg tgaaaaggga 2580

gccagactca caaagggaag gatcagtggg gggcgggggg ggggggaggt cactctgctc 2640

tggtcccagc attggaccag gaaccaaggc aagagggttg ctgcaagtta gaacccagcc 2700

ctggctacag atgaaattat agcaagagac aaaagagggt aggggaaggc agcaagcatg 2760

ggctctgtca tgcctcgtgt gtgggaatgt catcatggaa catactagtt tgtacgttaa 2820

taaaaaggac acagggatta gcaatgggca gaggggttga caccccagaa ggtgtttctt 2880

gtgcaaaagt gacgaagaga tacattcttc catgctccag attaattata aggctggata 2940

gttcgcaaat tattagcatc tatcactcat acatgatgct aggtttcacc gtaagctttc 3000

atagatgtgt atattctggt catattcact gttcacccat cattatctct ggtcccctcc 3060

cactcttgct gacccccttt ctaagtgccc ccttgtgctt tcatgtgtgt gtgctaatga 3120

ctttagatag gattgctaac ttgagcctag acgagggctt gttacaggag tagggatggc 3180

ttgctggtgg cgaagaaagt aaccctgata cactgtttct tcagtcacat cttttgctta 3240

tgctactgaa gaaagtaaca ctgataaatt gtttcttctt cagtcaaatc tttgcttgag 3300

gctctggtta ctgcctctct ctacagctct gacgtctgta tgaggctagt gaggagggca 3360

gctgggccat ctgacagtaa gcagccttca ttagtggaaa taaagccaac attagtggaa 3420

aactctccta cccagaccac agctctattc cacagcccca cgaggaccta tgctctatgc 3480

ggcagtccat cccaagtcta tgtgaagact atgttggttt cgtaagtgac accttaaccc 3540

tccttaaagt ctgttcttta agacttctat tgttgctttg tggtttagcc aagataagga 3600

aatgcaaagg gtcggagggg aggcaagcag tgtaagccaa acataccatc cactcaggca 3660

aggcccaaat gatggctaca ctattgccat cccagcttgc acaagaaaga gcttcctggg 3720

tgtcctcaca gagaagtggg gacaaccatc agaggaaaaa gtgtgatttt taaaaagaat 3780

ttgaacccca taaaatcaat aggtgtcaaa acttggaaag agtccaaatt ttcctgaaat 3840

gtaggttaaa aaaaaaaaaa aagccaaggg tagtgactca cacctgtaag tgcttgagga 3900

ggttgaggca ggaggattgg cctgagttga agacccgctt agggttcaca tagtaagcaa 3960

gggatagaga atgacacctg gtctaaaaca ccaagagaga acacacagac acacagacac 4020

acacacacac accacattag gacaatctta ggtaagaggc ttctggattt aggctaaaaa 4080

ttagctgtct ccatgtcttg gaggacttgt gtgcaattct cactttcgtg agaattcagc 4140

tgttggctct tggccttttt ccctccagag cttcagaaga ataaatactc agttaagact 4200
gtcacgtgta gtaagttaca gtttcctcgt gcaagctttt cctggataag attttctcct 4260
cgcatccgcg ggggtgaatc cacgttgtga taagccatct agtgtcactc tagagggctg 4320
cttagcactc tactgccaac aatgggtgat gcccggcaac tgtgtggact agtcttcctt 4380
ctgtcctccg tgctcagaca tcttccctga cactgtgagg agcagcatct tcgataactt 4440
ataggaagcc cttcatgcta ccacagcatg acgctagaca aatcacctac aggcactccc 4500
aggacaagat gcgaagctgc aaaagcctta acccgaaggg cttgtttctc tttgattccg 4560

ctgctgcttt ggccagctaa ttagttaatg tagttagatg agttcacata acgaggctta 4620
ttggacaagg atccgaaagt attttacaga gcctgatctt ttctcaaatt tgcctccccc 4680
ctccttcacc cctcccgccc cccaccccgc ccagcgtggt gttgaggaaa ccgcggacaa 4740
atgcaggaaa ctgggggccg tcgtgcacgt gtttgtggtg gactgcagca accgggccga 4800
gatttacaac tctgtggatc aggtgagaac ccgggtgcag gtgtcgcatc agctgacact 4860
gagctgggaa acgaaaaggt gggctaggga aggttgactt ccggtagacg gcagccgccc 4920
tgtcaagttg ctgtgtgacc gaaggagtga atattaatat ccactcagct ctggcctctg 4980

tcactgtttc atattctgag tcaaatctgc tgtgtctaaa aagtttggcc cagaggaaac 5040
gaaagatgat ttatcaagag aaaatttgat ctcttttcta acttttttaa aattagactt 5100
tgttctgctt tttctttaac ttattccctc tgagatttta ttggtttgcg tataatctag 5160
tcttttatga aaaacagagt taatgacttt ttaaacattg tatttgtttt atttactgca 5220
tcgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgcgcgc gcgcatggtg 5280
ggggtgtatg tagaatttag atgacaaatt ttgaaagctg gtctctcctt ccatgatggg 5340
gattctgggg gttgttaggt atggcaaaaa cagaagagga ggaggaggag gaggaggagg 5400

aggaggagga ggaggtggtg gtggtggtgg tggtggtgaa gaagaagaag gagcaggagg 5460
aggggaggag gggaggagta agaggagaaa gaggaaaggg aaggggaggg gggaggggga 5520
gaagaaaatg cctttaccca ctgagccgtc tccccacccc acaataccta ctctttcaaa 5580
aagtgaacat caagagctta ccaaaaaaaa aaaaacaaaa aacaaaaaac aaaaaaacaa 5640
aaaacaaacc ctccgctaag ctactgcctt ctttaatctg atttaatttt taaaaatccc 5700
attagcctca ctatagataa gcaaatagaa gcttggaagg tttggcccag acaaagacca 5760
cgtggccagt gggtagcaaa accaagatta attacttaac actttaaaaa tatcttaaca 5820

gaacttaagt attatttacc agcacaaaca ttaattcttg tttatttaag ttcataagtt 5880

ttcaccctac ctatcacata ctgagcacag aatttagggt tttaaccaaa acaggttagc 5940

cctacctcat gacaggtttc ccttcctgtc actgcttggg ctggttcttt cacgcctgaa 6000

ctctgagacc tcgtgcactg gaataacaaa agcctagatg gcatagtgac caagctagct 6060

gtcaccaaag ccagtgagcc cagacagaaa tgcatcgtaa agctattctt tattcttgga 6120

agtcagatcg atatcttgga tatctgtctt ccatgtcacg ctgtgatgtc tgtgtgtctc 6180

tcattatcat cccttggaca acttcctcct tttgtgttat tgggctcagg catttcttgg 6240

agttgagttg gcctagttta tgctggtatc aatcttagac tcttgggatg gctaggggta 6300

gggtgtggag tggttagctc agaacacatg gtgtggagtt gctgagctat aaatattgtg 6360

tatgctctgg tttcttttta catggtggag tatatatatc cttataatat agaacataaa 6420

aatgcctcta cccatctgtc tcctatcaac atgagctttg tgagaatgac cggtcctgtg 6480

tgtgcatttt aaaaggtaaa gagagaagta ggtgatgtcg agatcgtggt aaacaacgcc 6540

ggggcgatat atccagcaga ccttcttagt gccaaggacg aggagatcac caagaccttt 6600

gaggtcaata tcctcggaca tttttgggta agtcagaaac atttctgggt gtgtgtgtct 6660

gacatctctt cggtacagga aactcattga ggtcacaata tttgtggagg gaggtgtttt 6720

tggttttgtt tcttaagatt tatttattta tttatttttat gtatgtgagt acactgtagc 6780

tgtcttcaga cacaccagaa gagggcatcc gatcccgtta cagatggttg tgagccacca 6840

tgaggttgtt gggaattgaa ctcagtacct ctgggagtgc tattaaccac tgagccatct 6900

ctccagcctt gtttttgtt ttgcaagtca ggaatcttat aatgtgtttg aatagaagag 6960

ccacaggtct ctggcaacag ggttcagagt tcaaggagac acaggacacc caaaagagag 7020

agccatttgg tttattttgt caaacaaaat ctaacatagg gcctgatgaa taccagatct 7080

acaaaaagta aattaataat aaaaaattta ttaattaaaa atatttgtaa aggtttactg 7140

atatttcctc tgtacatctt ttcactattt tagaggaaat gtatgctgca tacattggtc 7200

tgcatcttcc tttcataggt gtttagagct gaaaggacat agaatactta gatttgtgca 7260

gaagttgggg acacttagag ttaggctctg agataagtag acattttgac ataacattat 7320

tatctacaaa gttcctgcct gagaacttca caattgagtt accaataggg gggggaaaca 7380

atctcaaaat gttttactgt agctcacaaa tttgcattga tctatattcc cagctaatct 7440

cagtcataag cagcctatgg gcggtgggat gagcatgcct attagttctg gagtgtagag 7500

agaatgaatc caattcctta ctattgggag tttgtgagaa actgaagtac ccttctcagc 7560

cctttggagg gagcatttgt atgaattaat accgtgggac ctcaaaatta ctatagcacc 7620

tggagcaaac ctggatttga agaacatcct tcatcctcat gacccagcta acatgacagt 7680

cacacaacac agggacccta tatactagtt atctaataat taataggaaa ttataactat 7740

gtcacaaaat atatcaaaag gtagaggttt agagaaccag gagccctcat tacctccata 7800

cagatctgca ggcagggctg ataggggggat tatcttggca acctgggctc taggaattta 7860

cacagaattt gaagagtggg tctgtaaaat tctcataaaa tgttctgata tttcccactg 7920

tacagtgagg cttttgtttc ccataatgta tcctatactg tatactgtta caagagagag 7980

atttgggtct gctccctgtg cacggtggtt gtgggccagg atagagcttc caccgctttg 8040

acgtgagtaa agatccccat tcttcaaatt ggagccaaga acttgtaaaa cggtaaacca 8100

tcctccagtt cccgcactgg tggatctctc ccttgacaat attgcttagg gctcagcagc 8160

accaggcttg cctggcatgc acaaagtcct ctggttaatt cctagcacgg cggaaaagta 8220

acaatgttgc ttctgtggca aatcccggaa ggagctactt aattcttcct ggagggatga 8280

aggggcatgc tcgcccgcta gcactggtcc tcagcacaat ctaggggact cagctctgag 8340

tgctgtatgc acgcttgcac aggtcctctg catgatctaa gggacactct gagtactgcc 8400

tggaaccatc acagttactc tctctttctg tcactgtgtg gaaggaacct cagggtctta 8460

caagtgtgta atttccaatg cctcctcctc cgctatggtg gaggcattgg aaatgaaaga 8520

gcctggcctg gaatccacct aataaccgtg tcattgtttc tgttcttttt acgtggacct 8580

taagatcata aaagcactcc ttccatcgat gctgagaaga aactctggcc acattgtcac 8640

agtggcttcg gtgtgcggcc atggagtgat tccttatctc atcccttatt ggtaggtgtc 8700

tcttgctcac agcgctgtgc atttctgctt attagtttgg ggattattta tagggattat 8760

tataattagt atatttatgt gttttataaa tattatatac tataatttta taaatgtatt 8820

catatattat ataactaaat aattataggg attaattata ggaaattatt agggattatt 8880

agttttcagg cctcaatgcc cacagcctag taaagtgccc agaagctaac acaagcacag 8940

aggtatagcg tggtctatgt gtgtgcatgt gtgagagaag agaacggagc agccctgcct 9000

tcccaggaga tgcacacact gcacaccatg tcccagggac ccagacacac tcgcagtctg 9060

cagcatatcc ctgagaagtc tgctattgcc atcttgtttc cttactgcta tacgatgatc 9120

tcagaaacac ttttaccaag atctgtggat gtagcagatg caggtgttat ttctttatag 9180

gcatattcaa agctaaactt gaccatggac taagcaatct ctcacactat acatggtcac 9240

agatgagtta aatttcagtg tagttcaggt ttccattaca cgcacacact cacacacata 9300

cacatacaca cacaatacac acacatgcac acacatacac cacacacacc atacacaaca 9360

catacacaca caacacacat tcacatatgc acacacatat acacacacac acatcacaca 9420

tgcacacagg agaggaagac agagacagca ataacttcta acactgatat tttttgtcac 9480

ctggacaaaa actacaatta aaacttatat agctccatat atgcatatgc taatttattg 9540

aattctccaa gacagcctta tattttgtta tctatatttt ataatcagaa aaatgaaaga 9600

acaaaaggtt aaaggttgca catgcaaaca caaccctacc aatgccatct ggagcccctt 9660
tgctccaaat actcgttgtg acattaggac acacatgaga ctgccagtgt caagggagaa 9720
aacatcagct ctttaatcaa gagaagagaa atgagctcag acccacaagt ttgtcgatgt 9780
tggtttcaca gtgttttaag tgtgtttttc tgactcacat ttttggagat aaaaactaat 9840
aaaacctgat taagaagcac ttggaaggtt gcctggacca tagagaaggg tcacacatac 9900
agtgagaaac attaagccta tctgtaaagt ggtccttttg ctatctttgc tagttagaag 9960
ttagaaacag gggctttggg gttagcaatg gttgaagaca tgcatgtgtc aatccctgag 10020

ttcaatcccc agcacctaaa agaaaaaagg aaaaagaaac aggtattttg catgctcaga 10080
attctgcttc ctacataaga aaataaactc gagcggcagg tgacaggtct ggacacatga 10140
agctgccttc tctcaggcag ctgagactgt ccccaccacc accaccacca ccaccaccac 10200
cacacacaca cacacacaca cacacacaca cacacacaca cacacacaca caccctttgct 10260
gatggctgat tttctcacag ctccagcaag tttgctgctg tgggcttcca ccgagcactg 10320
accgcagaac tggacacctt ggggaaaacc ggtatccaaa cctcgtgtct ctgccctgtg 10380
ttcgtgaata ctggcttcac caaaaacccg agtacaaggt gagatggaag acctttaagc 10440

gggagccccc ccctccaggc aatggagtct tgagtaggtc ataggtactt aacaagcatg 10500
tcacgatcat ttaaataagt ccagtagcat gatctcaacc cctaggctga cactgaggtg 10560
cccatgtcag atgtggctct aaggctggat tgaagactcc attgcaaaat taggaagggg 10620
agagacagag gcaggataat ctaaaaagaa ttgatttgtg ctgctgtgtg atggtttgct 10680
agtccacatc acaggcatca caatcgcacg ggagtctgcg gaaaccacag actatctcat 10740
gggattcagt tccaacaccc acacagcctt gcctcggcat gctaaccttg gtgtgacatc 10800
taaaatccat ttttatgctt tgatttttcc ctcctcttca ttaatgggca aggctccaac 10860

actggacatg ggttatacgc atttctcagt tatacacggg tgcaactttt gaatcacgaa 10920
acgccccagt gctgtaacag aacactcctt gtacaaatct taaattccca taaaattgca 10980
ggagaagcaa atcttagtac gagatgcctc agaacaactc agctctgtgg atacccacac 11040
acaacgggaa gtgtactcag acatagaata ttctttttac atacagcagg cattggagat 11100
ccggtgtttc cttcctggcc attttgtccc ctggctcctc agtccctccc catcttgtac 11160
cttgttgtga tttaattatt ttgacctggt gaggtaacag acagacctac tggacccaag 11220
agtacaagaa catgataaca agttatcagt gtgggatgtg ggaaggggag agggtgatgt 11280

tcactcaggg aagcctggaa cattgtaaat cctaattcta aatttggtga agagacagga 11340

taatttcctt catttgttaa ttaaaaacaa tataaatcaa gaggctgggg ttttacctca 11400

acatcagggt atgtgtgaaa gttctggatt taattccagc atccaaagga agaaataaaa 11460

tataaacgga gataaactct ttcccccttt ccttcaggtt atggcctgta ttagagccgg 11520

aagaagttgc aaggagtctg atcaatggaa tacttaccaa caagaaaatg atcttcgttc 11580

catcctatat caatatttct ctgatcctgg aaaagtaagt tcctcgcaga accctgaaac 11640

actcaaacac aaatgttaga aatgttgcta tggaaactcc ccatggctgc agtgaaagtc 11700

acgttttcca tgttgcctta cccatgagtc tttgattccc tgataccctt tctggaatac 11760

atctaccatc ttggcccaaa ttttgttcaa gtctatttaa cagaaagaaa gaaaaactac 11820

agggtatggg cacaccataa aggctctgtg tgtgtgttca tgtggtgtat gtgggggata 11880

gagatgtggt acttgtggag acagaggcaa agctctcagg aaccatccac ctaggtttct 11940

gatttgggtt tatttctctt ctctccttcc cctcccctcc cctcccctcc cttcccctcc 12000

cttcccctcc cctcccctcc tctcttcctt ttcctttcct ttcctttcct tttcttttct 12060

tttcttctct tctcttctct tctcttctct tctcttctct tctcttctct tctcttcctc 12120

ttccttccct ccctccctcc ctctctttcc tctctctctt tttttctgat acaatctctt 12180

attgagacct ggactcacca gcttggctag gatagatatg catcaccagg gatatcccag 12240

gccctctatt cccacagtac tggggttaca agtgctgaca gctatgccca gctgtgtgtg 12300

tgttctgagg attgaactca ggccctcctg ttgtcgtggc aatctcccta aaccctcggg 12360

tgctcagttt tgcagttctt tgtgacagtg tgcatctcga atgtctcata acatgctttc 12420

acatgtttct atttcacttc atttataacc caagagcctt atgatagaat cacttctttt 12480

catctcaggc ttgtgagccc ccaaagaggg caaaatccca caaattgaat ttcctcaaca 12540

cagctcaaaa caggtctcat taaaagttga aaaaaaattc caaaattggc atgtatatga 12600

ccaaagtgca aaaatggttg actgctctgc ctgggatagg accagttacc tgtctcaaag 12660

cccttgccac actgtgggtg atttgttacg ctggcaacag ggatttttgt tgctccttgc 12720

atttaactcc ttttggagag agacagttct gtgatcagaa gcaatcacta gtctctgggc 12780

ttgttgttca atggcctgag caagctctct ccttaccacc tcacagcagg gtccccacct 12840

cccagtctct tctcccacct tctgccctct atccagggaa gtaaatgttg ctttagtaaa 12900

caataccaat gtgatctgaa atttggcttt ctctgtagaa atcgcacctc ctgcctcttt 12960

cacttcactc acacccttcc ctattctgtc aactcagacc taaccctccg acatgctgag 13020

aggtattcag cccctgtaaa cccttccttc cacatcttgt tatagtcagc tgatggtttt 13080

ctgtattcat ttccagttgg ccatgtttaa ccgctgggca gcgcaataca ctaagaggca 13140

gtagcttttg aaaaggctct ttagtttttc tttctctaaa ataataggtc tggacacagg 13200

cagaggggggg aaggggggtgt atgtggtgga ctctcagttc ttggccccag aatacactgc 13260

cgtttttttgg aagtgagctt ggtgtcggtt ttaggttgtc attggtagaa gacgtcatcc 13320

aattttgctg ttctgagaga acaacgctat attatatata gataactgta gatggggtgg 13380

gggaggggag actaaacaat aagcaaagag tagaggaaaa ctggacttga gtaagtatgc 13440

tcttggcgct ttttaaattg acaacaaaaa ggaacactgg catcaaaact tcaaactgaa 13500

atttctcatt ttatttttaa tgtaggtttc ttcctgagcg tgccttaaaa gcgatcagtc 13560

gtatacagaa cattcaattt gaagcaattg tgggccacaa aaccaagatg aagtagtgca 13620

tgcagagacg tgtggacacc aatgatgtga agccaagttt agaggggaca cacagctttc 13680

tttcacatgt tttaagtgtt ccacatgttt aaaatgtagg cttgacccta gcagccatcg 13740

agtgcataag cgtggtcaac tgtccttcta gtttcctata cttatagatg tctcagctcc 13800

tggtagtggg ttctggcaat ggatatgtaa aggagggaaa aggcaatcca tgtgttttta 13860

taaaataaat tactaaatag aaatgtgggc aaaagggcaa gataataaag tcttgggcaa 13920

atgttggcac gctgtagtga tttctctttc gatggagaga ttggtggtgt tttaaatttg 13980

ctctttgtgg acaggatgaa aactctacct ggtttccttg acatatgaat taggaaggag 14040

tatttttgtt ccacattttt tttctcagtg tgaggttttc ttcacatcat gtgaatggac 14100

aatgaggtca ttttatttac ctaacttggg gaagcaaata ttgacacaga tttgcaactg 14160

tgaaatctca acaccttgtt ctttgcttta gaaagacctt ggacacattt gtttttcttt 14220

ttattctatt aggcaataac tgaaaatacc caaccgtata catagtaaca gtgacatgac 14280

ctcccacaca gatcccatta tgaagcagga aaccttcctg ggttttcctt tggcttctgt 14340

attttgctgc aggggactga tcccacagac taagacaggt tacataggca gtctgtgttg 14400

acctcctgtc tatcaacttg cagagttact tccacatttt gctttgatac ctttaaagtg 14460

aagctcagag ggatgcctca gtatgtgaca atgatctctt ttgcccacgt ctgcatattt 14520

tgagtgtaaa tatcttccct acaaggtcca gctccatttc taaatatttt tctatgataa 14580

tcttcacagt cacagttctc actcctgcta accattgcag ccgcttaggc taccccagga 14640

agcacagatc cctaagtaag accccatcta cacaaatcat ttagttagag acacatcata 14700

gaattcctag gtccctgcga gcacggggtc atgatttatt gatgagtatt tcagaatggt 14760

ggaagtattt gtttattgtc tactcagttc atggaaacaa ccttttggt aagcagacag 14820

tatctcctct ttttttaat ggtagaatat gatattgcat atacataagc acattttgaa 14880

catagctgaa tataagtatc actgtttcat ccttcttctt cttcttcttc ttcttcttct 14940

tcttcttctt cttcttcttc ttcttccgga gcatataatt ttattgctct tctttttaa 15000

agaatagaag aatatctttt atactaggtc catgggctat ctagtgtctg gtttttgttc 15060

acccaagcag tgctagatac aagctccatc ttgtgtagta ggctttaagt caaaacagat 15120
ttggttggtt actctcacaa cattgggcta ccattttcct accatatctt ttaggcacaa 15180
caccaatgtt gatccatggg tgtgtggctt tgtttgttgt tcttttccta ggagcatgca 15240
gaatgacttc ctacacctaa gatactagta gataggagtg aagagtctat gtagtcacca 15300
gattgacttc tagatgttca gtgagttgag ttgcctttag caataggacc ttgttgacag 15360
gtggtgtaga gaaacatatg gccatggtaa cagcttgagt tctttgcaga ttcccatgta 15420
aacatttgga caacatctca gtaaaatgta attgagtttg aatactggaa gcttcttttg 15480

gtgaggaaag ataccctgtt agaaacaagt ctcctttgtt atagggtgat ttcatttata 15540
tgtgtaattg cctttgtatg tgtatatttt tctatcaatc tgtctatgta tcatctatct 15600
aatctctctg tcatctaata tatgtatcat ctgtcagtta tccacctatc acacatatat 15660
agctattgtc tatcaatcac cttttaatct acctttttcat tctatcaaac tattaatata 15720
tcaatcactt atgtatttgt gattgatcta tcttttaacc atctaattat cacatatctt 15780
tcttctattg atgatctatt ttcatctgtt tctctgtctg cctctcattt ccttttgaat 15840
agctctagct atcatatgca aagggaagac attttcacaa aaataagcag tcttgagggt 15900

cctcatgatg ttttcagtcc ctcagtgtct cattccaaaa gcccaaagac accatagaat 15960
tcagacagaa tgccatgtaa caaagcccca cactcaatac ccagttacaa ccaccttagc 16020
agaacactaa agagattaaa ctatccttgg ctgccactaa aatcaggtag tcttcatggt 16080
aagtgtggat ctagcttttg gggaagtcct gtattctttg ctaacccact ttcatttcaa 16140
gattgaacaa aacacaggaa gatggtgccc tgtgaggcct tgacactcac atagcctcag 16200
tccagttagc tccctgtcca tgtgttcaca acagatgctt gtgataaaga aagccccaca 16260
ggatgggtga ggtctgagtt tctgaccaca gatccttcct gttatttttg atgttatttt 16320

tatgtttgtg tggttatctt cttttgggtt tgttgaaaga ggatttcttt ttctttttct 16380
aaggtgtagt ttcccttctt gtgttggcat tttccatcta ttatccttta tagggctgga 16440
tttgtggaaa gatattgtgt aaatttggtt tgtcatggaa tatcttggtc tctctgtcta 16500
tggtaattga gagttttgct gggtatagta gcctaggctg gcatttgtgt tctcttagga 16560
tctgtatgac atctgcctag gatcttctag ctttcatagt ctctggtaag aagtctggtg 16620
gaattctgat aggtctgcct tttacttgac cttttcccct tactgctttt aatattcttt 16680
ctttgttttg tgcatttgat gttttgatta ttatgtgaca ggaggaattt cttttctggt 16740

caagtctctt tggagttctg taggcttctt gtatgtttat gggcatctct ttctttagat 16800

tagggaagtt ttcttctata actttgttaa agatatttac tggcccttta agttggggat 16860

cttcactctc ttctatacct attatcctta ggtttggtct tctcattgtg tcctggattt 16920

ccaggatgtt ttgggttagg agctttttgc tttttgcatt ttctttgact gttgtgtcaa 16980

tgttttctat ggtatcttct gcacctgaga ttctctcttc tacctcttgt atactgctgg 17040

tgatgcttgc atctataact actgatctct ttcctaggtt ttctatctcc agggttgtcc 17100

ctttgtgatt tctttattgt ttctatttcc attttcagat cctggatagt tttgttcaac 17160

tccttcacct atttggttgt gttttcctgt aattctttaa gggagttttg tgtttcctct 17220

ttaagggttt ctagacattt acctgtgttt ctttaaagga gttatttatg tccttaaagt 17280

cctctatcat catcatgaga tgtgatttta actcagggtc ttgcttttct ggtgtgttgg 17340

gatatccagg gctcactgtg ttgggagaac tgggttctga tgatgccaag tagccttggt 17400

ttctgttgct tatgttcttg ctcttacctc ttgccacctg gttatctctg gtgttaactg 17460

gtctggctgt ctctgactgt ggcttgtccc ttctgcaagt ctgcattgtt agtacttctg 17520

ggagaccagt tctctcagga ggaatttggg tatggagatc tgtggcacag ggtcagctct 17580

ggagtgtaga cagaaaccag aaggatcctg tgcctggttg ttccttgatt cctgtgtcct 17640

gatggctctg ggaaggtttc gcttgggcca cgagtttgaa cagaagtggt gctcttacct 17700

gtgctcgggg gtgtgtcagc actcctggca aaccagctct ctcctggccg tatttgggta 17760

tgtagcggtg tggcacagga tcagctccgg gcacagacgc aaaccagaaa gcttctgtcc 17820

caggccgctc ctggttcctg tgtcctgaga gttccaggag ggtccttctg agcagcagtg 17880

gtggttttac ctgagttcaa aggcttgtcc acactcctgg gaggcaagct ctctcccagc 17940

actatttggg tatggagccc tgtggcagag gatcagctct gggccaaaca taaaccggaa 18000

ggctgacatc tttcaaaagg agtctttttt tctttctttt tttttggggg gggggggggtt 18060

gatttttttt tttttcgaga cagggtgtct ctgtatagcc ctggctgtcc tggaactcac 18120

tctgtagacc aggctggtct cgaactcaga aatctgcctg cctctgcctc ccgcatgctg 18180

ggattaaagg tgtgcgccac cacgacccgc tcaaaaggac ttaatgttct ttatttaaca 18240

atgttatata ttttctagca acctatcttc tctccctctc tccctttctg cttctctctt 18300

tcttcctctt tctttttctg caccttttctc tctctgcctc tctctcttcc tttctctctg 18360

tctctgtctc tctctctctc tctttctctc tctctctctc tctctctctc tctctctcgt 18420

gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtattaaa ggctgcatta 18480

caccaggcaa aggctgtgtt acaccaggta aaggctacat tacacaggta aaggttgtgt 18540

tatgccaagc aaaggagttt ctttgaactt taatctcatt gcacccacaa gagagtaagc 18600

atacctcagg ctaggtgtgt tcactggaat acgtgaagta aaggttgaaa cttctgttta 18660

tgactaattc tttatcttaa aactttgttt tttacagtat atacgatatg tgattatatt 18720

atgacacgat atatactaag taaatattag gtacacactg aagtttagct gcaatgctca 18780

gtcagaccca cagtcctttg cccactcctg ttctctctct ctctctctcc ctctctgtct 18840

ctctctctct ctcgctctgt actctatagc tgtctttgag caaagatcac aggaaccaaa 18900

aaactgagag gtccagagca tgagcctagt gcacatttta gtattcctac atggttatag 18960

gtgcactgtc tcctgtccat ggtgtctacc accctaggaa aagctgcagg gtctagagat 19020

cttagaaata gaataagcca ggctaaatac acagaccact gggtaaatca gtgcagtggt 19080

cactgttgtg ggctagtcac cagagaagac tgttcaagcc aatggaaagt cattactagc 19140

tgactggtga ctacactgga tgttcaggat cccagtgtag cactgagcct tcctcaggtt 19200

gagcttttaa gcacaaaaac catgttcttg gttgacgtac ttcagttaac aagaacagtt 19260

agctggaaga ggaactacag aagctaaaag gcaacattag tacattatga gactttacca 19320

aaacaatgaa ccctaatgaa ttatgccttt tagttttgac aggtagggct gtctaagtgc 19380

tgagttttac agcctgaaca gtacttccat catggagtca gatatgctaa gatctgggac 19440

cctgttacat tccacactgt tcttagtgaa tgagtcaaat catggattca tcctgtttta 19500

gtttaagatg aagaatgaga attaaagtgc tgacacaaaa tcattagcca gggagaccac 19560

gagaatggag actggtactg actacttcct caaccatgtc tatattttct ctaattattc 19620

cttaccaatt agtatagaaa caataggttt ctttcagagc catgtatagt ttcctttatg 19680

tcatttgaag tctaggcctc tccaattttg tagggagaat ctacctgcaa gggaatctaa 19740

ctgttgtttt aattcagaaa ctgaaacttc aaatacctgt aggtcctgat caacctatcc 19800

tcttagtttg gaaagcctca gaagaggcat atttagtcaa accattaata tatgctcacc 19860

aagtagtctc agaggtttgg taatattgct gttgtccaga tgaattaaac catcgggcag 19920

aataacacac cttggaatgt ggggaagtgc gtaggttaaa aattttggat attagacagg 19980

ttttagcccc ttctaagtcc tctagcatta atctgggctg cccccagtca cagtgagttt 20040

tgtcagtcaa tagattggtt agtccccatt tctctgtcct acagcagatg ggtgtctaag 20100

cagtttcatc aaccctggat ggttttgact gggcatgatt acctctgcct cccagggact 20160

ctttctgtgg catagagtct gatgctccag gccttttgg ttttacaaag accagaatct 20220

tttaaagagt tttgggggcc tagtgagatg gttaagaaca ctgactgttc ttccaaaggt 20280

cctgagttca attcccagcg accacatggt gggtcacaac catccgtaat gagatctgac 20340

accctcttct ggtgtgtctg aagacagcta cagtgtagtg tacttagata taataataaa 20400

taaatctaaa agaattttct gtaggaagct tcggttttaa gttccaattt ttggaatgga 20460

aactcctctg ttttaagctg tggtacaggg tcagctctgg ggttcagaca gaaaccagaa 20520

ggatcctgtc cccaactgtt ctttggttcc agtgtcctga tggctctggg catgtccctc 20580

ttgggccagg aatttgagca gaaatggtgg tcttacctgt ccacactcac aggcatgtcc 20640

acactcctgg gagaccagtt ctctcctggt ggtatttggg tatggagcgc tgtggcacag 20700

gatcaagctc tgggtacagg cggccagtat ctcctctcgt ttgatcctga attgttaaat 20760

tctttattat caaataaacc attcttggcc aatcttgtgg ccacactgaa tgccagtagt 20820

aaggagccta tgaacatctg acatcactgt gcaacaataa aagacggcac catgttcaga 20880

cataaccaat ggaattggat tcatagatgt acacatgtgg agtgaagatc aatagtccag 20940

agccagaaga ggaccagagg ccccagcaac attgctccca gacctgctaa cactttggtt 21000

tcctccatac actgctgaac aagagagtac ttactccaga aagctatgag ctgggtgatg 21060

aggagagggc tcaggtttta aaagcacttc tctttctttc agaggacctg gattcagttc 21120

caagcaccca catggtggtt ctcaacagcc tgttactcca atcccagggg atttgacacc 21180

ctcttcagac tttctcaaac actaggcact cgtggggtac gcatacgtca ttcaggcaag 21240

ccctcataca cataaaatag aataagcaca tattttaaat gtgctactat tactgatgtt 21300

ttaagccagt taaattatgt gattttctta acagcaataa aatattgtag ataatgtcat 21360

catgctagct gcatcactct gactgtataa acccaagaga atttcactcc atctccttct 21420

gccaaaatat ttacaacact taagaaataa gggctgattt ttatttttt taacataaaa 21480

tatatgagac agttgtatcc tgtcagctaa aaatgttttc tctggcctgt gagttggctc 21540

agtagatgaa gacacttgct gccaagccta acaacctgaa ttcaactccc agaacccatg 21600

taatgggaag acagatggga ttcagagaaa ttgtcctctg atcattgtac ccatgccttg 21660

gtatatgagt atctgtgatg gtttgtatat gctcagtcca gggagtggca ctactagaag 21720

gtgtggtcct gttggagtag atgtatcact gtgggtgtgg gctttaataa cctcatccta 21780

gctgcctgga agccagtatt ctgctagcag ccttcatatg aagatacaga actctcagct 21840

gtgtttgcac catgcctgca tgaacgctgc catgctccca ccttgatgat aatggactga 21900

acctctgaac ctgtaagcca gccccgatta aatgttgtca cttataa 21947

<210> 270

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 270

ugucagguua guuagaugaa guuuuagagc uaugcu 36

<210> 271

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 271

gugucagguu aguuagauga guuuuagagc uaugcu 36

<210> 272

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 272

ccugacacau auacagacua guuuuagagc uaugcu 36

<210> 273

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 273

cugacacaua uacagacuaa guuuuagagc uaugcu 36

<210> 274

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223

> Synthetic

<400> 274

ccuuagucug uauauguguc guuuuagagc uaugcu 36

<210> 275

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 275

cauauacaga cuaagggacc guuuuagagc uaugcu 36

<210> 276

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 276

auauacagac uaagggacca guuuuagagc uaugcu 36

<210> 277

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 277

ucaaaguuug auaaaauuccc guuuuagagc uaugcu 36

<210> 278

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 278

aaaauacaaa gauaaguaga guuuuagagc uaugcu 36

<210> 279

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 279

acucugugac uuuaaaaagu guuuuagagc uaugcu 36

<210> 280

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 280

gguucugugg gauauuaaua guuuuagagc uaugcu 36

<210> 281

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 281

acagagcaua uugguucugu guuuuagagc uaugcu 36

<210> 282

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 282

gacagagcau auugguucug guuuuagagc uaugcu 36

<210> 283

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 283

ugcaaaacga cagagcauau guuuuagagc uaugcu 36

<210> 284

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 284

gagcugggca uggaauaggc guuuuagagc uaugcu 36

<210> 285

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 285

acuggagcug ggcauggaau guuuuagagc uaugcu 36

<210> 286

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 286

cucauuacug gagcugggca guuuuagagc uaugcu 36

<210> 287

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 287

uuguucucau uacuggagcu guuuuagagc uaugcu 36

<210> 288

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 288

auuguucuca uuacuggagc guuuuagagc uaugcu 36

<210> 289

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 289

ggggagauug uucucauuac guuuuagagc uaugcu 36

<210> 290

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 290

gaggagaaaa ucuguggcug guuuuagagc uaugcu 36

<210> 291

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 291

agaggagaaa aucuguggcu guuuuagagc uaugcu 36

<210> 292

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 292

cagaggagaa aaucuguggc guuuuagagc uaugcu 36

<210> 293

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 293

uccucagagg agaaaaucug guuuuagagc uaugcu 36

<210> 294

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 294

ugaaguuuuu cauuccucag guuuuagagc uaugcu 36

<210> 295

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 295

cuucaccaac gacuccaagu guuuuagagc uaugcu 36

<210> 296

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 296

cuacuccuac uuggagucgu guuuuagagc uaugcu 36

<210> 297

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 297

cuccaaguag gaguagauga guuuuagagc uaugcu 36

<210> 298

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 298

caccaucauc uacuccuacu guuuuagagc uaugcu 36

<210> 299

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 299

ugauggugau cagaagcaga guuuuagagc uaugcu 36

<210> 300

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 300

ucagaagcag aaggauuucu guuuuagagc uaugcu 36

<210> 301

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 301

gauuucuagg augauguuca guuuuagagc uaugcu 36

<210> 302

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 302

uugcucuguc cucuuccuuc guuuuagagc uaugcu 36

<210> 303

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 303

aggacugaac cagaaggaag guuuuagagc uaugcu 36

<210> 304

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 304

uacacaagga cugaaccaga guuuuagagc uaugcu 36

<210> 305

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 305

uucaguccuu guguaguccu guuuuagagc uaugcu 36

<210> 306

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 306

ucaguccuug uguaguccua guuuuagagc uaugcu 36

<210> 307

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 307

guccuugugu aguccuaggg guuuuagagc uaugcu 36

<210> 308

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 308

cuuguguagu ccuagggagg guuuuagagc uaugcu 36

<210> 309

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 309

cuccucccua ggacuacaca guuuuagagc uaugcu 36

<210> 310

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 310

guagacagua ccuccucccu guuuuagagc uaugcu 36

<210> 311

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 311

uacugucuac acagagcucu guuuuagagc uaugcu 36

<210> 312

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 312

acugucuaca cagagcucua guuuuagagc uaugcu 36

<210> 313

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 313

ucuacacaga gcucuaggga guuuuagagc uaugcu 36

<210> 314

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 314

cuacacagag cucuagggaa guuuuagagc uaugcu 36

<210> 315

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 315

uacacagagc ucuagggaag guuuuagagc uaugcu 36

<210> 316

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 316

ggggugugcc caguuguuaa guuuuagagc uaugcu 36

<210> 317

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 317

gggugugccc aguuguuaau guuuuagagc uaugcu 36

<210> 318

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 318

ugguaguccc auuaacaacu guuuuagagc uaugcu 36

<210> 319

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 319

cugguagucc cauuaacaac guuuuagagc uaugcu 36

<210> 320

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 320

uuguuaaugg gacuaccaga guuuuagagc uaugcu 36

<210> 321

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 321

uaccagaugg aagccagcuu guuuuagagc uaugcu 36

<210> 322

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 322

uuccaaagcu ggcuuccauc guuuuagagc uaugcu 36

<210> 323

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 323

uggaagccag cuuuggaagc guuuuagagc uaugcu 36

<210> 324

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 324

acaaggccug cuuccaaagc guuuuagagc uaugcu 36

<210> 325

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 325

gccuuguuca cguguucuaa guuuuagagc uaugcu 36

<210> 326

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 326

ccuuguucac guguucuaau guuuuagagc uaugcu 36

<210> 327

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 327

cccauuagaa cacgugaaca guuuuagagc uaugcu 36

<210> 328

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 328

uuggcaucac uucauauuug guuuuagagc uaugcu 36

<210> 329

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 329

cuugugcucu uggcaucacu guuuuagagc uaugcu 36

<210> 330

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 330

agcacacucu cuugugcucu guuuuagagc uaugcu 36

<210> 331

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 331

gcacaagaga gugugcucuc guuuuagagc uaugcu 36

<210> 332

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 332

gcuuaaucuc acacauagaa guuuuagagc uaugcu 36

<210> 333

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 333

cuuaaucuca cacauagaaa guuuuagagc uaugcu 36

<210> 334

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 334

uuaaucucac acauagaaag guuuuagagc uaugcu 36

<210> 335

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 335

aggagugcug guuuaucaac guuuuagagc uaugcu 36

<210> 336

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 336

uucuuugaca gcaggagugc guuuuagagc uaugcu 36

<210> 337

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 337

acucugguuu cuuugacagc guuuuagagc uaugcu 36

<210> 338

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 338

accagaguug agaaaacccc guuuuagagc uaugcu 36

<210> 339

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 339

uccaggguu uucucaacuc guuuuagagc uaugcu 36

<210> 340

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 340

caguuauuaa augaauccag guuuuagagc uaugcu 36

<210> 341

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 341

gcaguuauua aaugaaucca guuuuagagc uaugcu 36

<210> 342

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 342

ggcaguuauu aaaugaaucc guuuuagagc uaugcu 36

<210> 343

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 343

uggaugguaa cagcuacauc guuuuagagc uaugcu 36

<210> 344

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 344

gcuguuacca uccacauccu guuuuagagc uaugcu 36

<210> 345

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 345

ucaagaacca aggaugugga guuuuagagc uaugcu 36

<210> 346

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 346

uccuucaaga accaaggaug guuuuagagc uaugcu 36

<210> 347

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 347

ugaguguccu ucaagaacca guuuuagagc uaugcu 36

<210> 348

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 348

uuuuauuuua uaacuacaag guuuuagagc uaugcu 36

<210> 349

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 349

uuguuuuuaa uaaaaacaag guuuuagagc uaugcu 36

<210> 350

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 350

uauuauagaa ugcuuuugca guuuuagagc uaugcu 36

<210> 351

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 351

caagauuagu cuugauguag guuuuagagc uaugcu 36

<210> 352

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 352

aagauuaguc uugauguagu guuuuagagc uaugcu 36

<210> 353

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 353

agucuugaug uagugggagu guuuuagagc uaugcu 36

<210> 354

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 354

uuuuucuauu aaaaaaaaaa guuuuagagc uaugcu 36

<210> 355

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 355

ucuauuaaaa aaaaaaaggc guuuuagagc uaugcu 36

<210> 356

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 356

cuauuaaaaa aaaaaaggcu guuuuagagc uaugcu 36

<210> 357

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 357

aaaaaaaaaa aggcugggca guuuuagagc uaugcu 36

<210> 358

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 358

aaaaaaaagg cugggcacgg guuuuagagc uaugcu 36

<210> 359

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 359

cacccguaau cccagcacuu guuuuagagc uaugcu 36

<210> 360

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 360

acccguaauc ccagcacuuu guuuuagagc uaugcu 36

<210> 361

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 361

cguaauccca gcacuuuggg guuuuagagc uaugcu 36

<210> 362

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 362

ucccaaagug cugggauuac guuuuagagc uaugcu 36

<210> 363

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 363

cucccaaagu gcugggauua guuuuagagc uaugcu 36

<210> 364

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 364

cccagcacuu ugggaggccg guuuuagagc uaugcu 36

<210> 365

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 365

ccucggccuc ccaaagugcu guuuuagagc uaugcu 36

<210> 366

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 366

gcacuuuggg aggccgaggc guuuuagagc uaugcu 36

<210> 367

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 367

cuuugggagg ccgaggcagg guuuuagagc uaugcu 36

<210> 368

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 368

gccgaggcag guggaucacg guuuuagagc uaugcu 36

<210> 369

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 369

accucgugau ccaccugccu guuuuagagc uaugcu 36

<210> 370

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 370

ggcaggugga ucacgagguc guuuuagagc uaugcu 36

<210> 371

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 371

ucaggagauc gagaccaucu guuuuagagc uaugcu 36

<210> 372

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 372

cgagaccauc uuggcuaaca guuuuagagc uaugcu 36

<210> 373

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 373

uuucaccaug uuagccaaga guuuuagagc uaugcu 36

<210> 374

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 374

uuguauuuuu uguagagacg guuuuagagc uaugcu 36

<210> 375

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 375

uuuguauuuu uuguagagac guuuuagagc uaugcu 36

<210> 376

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 376

uuuuguauuu uuuguagaga guuuuagagc uaugcu 36

<210> 377

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 377

aaaaaauaca aaaaauuagc guuuuagagc uaugcu 36

<210> 378

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 378

aaaaauacaa aaaauuagcc guuuuagagc uaugcu 36

<210> 379

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 379

uacaaaaaau uagccgggug guuuuagagc uaugcu 36

<210> 380

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 380

aaaaaauuag ccgggugugg guuuuagagc uaugcu 36

<210> 381

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 381

aaauuagccg gguguggugg guuuuagagc uaugcu 36

<210> 382

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 382

aauuagccgg gugugguggc guuuuagagc uaugcu 36

<210> 383

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 383

caggcgcccg ccaccacacc guuuuagagc uaugcu 36

<210> 384

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 384

gccuguaguc ccagcuacuc guuuuagagc uaugcu 36

<210> 385

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 385

uguaguccca gcuacucagg guuuuagagc uaugcu 36

<210> 386

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 386

uccugaguag cugggacuac guuuuagagc uaugcu 36

<210> 387

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 387

cccagcuacu caggaggcug guuuuagagc uaugcu 36

<210> 388

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 388

ccucagccuc cugaguagcu guuuuagagc uaugcu 36

<210> 389

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 389

gccucagccu ccugaguagc guuuuagagc uaugcu 36

<210> 390

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 390

aggaggcuga ggcaggagaa guuuuagagc uaugcu 36

<210> 391

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 391

gcaggagaau ggcgugaacc guuuuagagc uaugcu 36

<210> 392

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 392

caggagaaug gcgugaaccc guuuuagagc uaugcu 36

<210> 393

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 393

gagaauggcg ugaacccggg guuuuagagc uaugcu 36

<210> 394

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 394

aauggcguga acccgggagg guuuuagagc uaugcu 36

<210> 395

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 395

cacugcaagc uccaccuccc guuuuagagc uaugcu 36

<210> 396

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 396

ucacugcaag cuccaccucc guuuuagagc uaugcu 36

<210> 397

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 397

cauaccacug cacuccagcc guuuuagagc uaugcu 36

<210> 398

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 398

auaccacugc acuccagccu guuuuagagc uaugcu 36

<210> 399

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 399

ucgcccaggc uggagugcag guuuuagagc uaugcu 36

<210> 400

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 400

ucucacucuu ucgcccaggc guuuuagagc uaugcu 36

<210> 401

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 401

ggagucucac ucuuucgccc guuuuagagc uaugcu 36

<210> 402

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 402

uguuuuuugu uuuuuugaga guuuuagagc uaugcu 36

<210> 403

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 403

aggaagaaag aaagguuuuu guuuuagagc uaugcu 36

<210> 404

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 404

agaagaaaag gaagaaagaa guuuuagagc uaugcu 36

<210> 405

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 405

cuuucuuccu uuucuucucu guuuuagagc uaugcu 36

<210> 406

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 406

uuucuuccuu uucuucucuu guuuuagagc uaugcu 36

<210> 407

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 407

aauggaccca agagaagaaa guuuuagagc uaugcu 36

<210> 408

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 408

ggcuauuaca uaagaaacaa guuuuagagc uaugcu 36

<210> 409

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 409

cacaggaaaa ggaacuguac guuuuagagc uaugcu 36

<210> 410

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 410

auuaaagcua acacaggaaa guuuuagagc uaugcu 36

<210> 411

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 411

ucaaaaauua aagcuaacac guuuuagagc uaugcu 36

<210> 412

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 412

uaaaauuguc uaaacaucuc guuuuagagc uaugcu 36

<210> 413

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 413

agagauguuu agacaauuuu guuuuagagc uaugcu 36

<210> 414

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 414

ucuaaacauc ucugggacca guuuuagagc uaugcu 36

<210> 415

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 415

uuuaugcuuu cauauauccu guuuuagagc uaugcu 36

<210> 416

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 416

agcauaaauu acaaagaaaa guuuuagagc uaugcu 36

<210> 417

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 417

uacaaagaaa aagguuauca guuuuagagc uaugcu 36

<210> 418

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 418

acaaagaaaa agguuaucau guuuuagagc uaugcu 36

<210> 419

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 419

caaagaaaaa gguuaucaug guuuuagagc uaugcu 36

<210> 420

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 420

ucugagauuu aaaauagagu guuuuagagc uaugcu 36

<210> 421

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 421

cuuauaagau acauuaugaa guuuuagagc uaugcu 36

<210> 422

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 422

uaucuuauaa gacuauaaaa guuuuagagc uaugcu 36

<210> 423

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 423

aucuuauaag acuauaaaaa guuuuagagc uaugcu 36

<210> 424

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 424

uuauaagacu auaaaaaggg guuuuagagc uaugcu 36

<210> 425

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 425

uaaaaaggga ggaaauauag guuuuagagc uaugcu 36

<210> 426

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 426

aaaaagggag gaaauauaga guuuuagagc uaugcu 36

<210> 427

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 427

aaauauagag gguccacuuu guuuuagagc uaugcu 36

<210> 428

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 428

uauagagggu ccacuuuugg guuuuagagc uaugcu 36

<210> 429

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 429

acucugaagu ccaccaaaag guuuuagagc uaugcu 36

<210> 430

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 430

agaauagagu ugcaccguuu guuuuagagc uaugcu 36

<210> 431

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 431

aaaacggugc aacucuauuc guuuuagagc uaugcu 36

<210> 432

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 432

ccguuuuggg cuaaugaaaa guuuuagagc uaugcu 36

<210> 433

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 433

ccuuuuucau uagcccaaaa guuuuagagc uaugcu 36

<210> 434

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 434

ugggcuaaug aaaaaggaag guuuuagagc uaugcu 36

<210> 435

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 435

uaaugaaaaa ggaagaggcu guuuuagagc uaugcu 36

<210> 436

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 436

aaugaaaaag gaagaggcua guuuuagagc uaugcu 36

<210> 437

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 437

cugaaucuua aaauaugucc guuuuagagc uaugcu 36

<210> 438

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 438

caggcagcuu uaucucaacc guuuuagagc uaugcu 36

<210> 439

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 439

cuaagagauc aaguuucagc guuuuagagc uaugcu 36

<210> 440

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 440

guguucuugu ugauauucug guuuuagagc uaugcu 36

<210> 441

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 441

cuuguugaua uucuguggca guuuuagagc uaugcu 36

<210> 442

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 442

ucuguggcau ggcuacagau guuuuagagc uaugcu 36

<210> 443

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 443

agaacuuauu uacacaggga guuuuagagc uaugcu 36

<210> 444

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 444

aaagagaacu uauuuacaca guuuuagagc uaugcu 36

<210> 445

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 445

caaagagaac uuauuuacac guuuuagagc uaugcu 36

<210> 446

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 446

uucucuuugu auuuacuuuu guuuuagagc uaugcu 36

<210> 447

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 447

ucucuuugua uuuacuuuua guuuuagagc uaugcu 36

<210> 448

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 448

cuuuguauuu acuuuuaggg guuuuagagc uaugcu 36

<210> 449

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 449

agcuuuuguc caccuuuaaa guuuuagagc uaugcu 36

<210> 450

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 450

uuuauuuuuc cauuuaaagg guuuuagagc uaugcu 36

<210> 451

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 451

uauuuuauuu uuccauuuaa guuuuagagc uaugcu 36

<210> 452

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 452

cuuacauaaa cauacuuaaa guuuuagagc uaugcu 36

<210> 453

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 453

uaagcacaga aguuuuuaag guuuuagagc uaugcu 36

<210> 454

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 454

aaguuuuuaa gaggcaugaa guuuuagagc uaugcu 36

<210> 455

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 455

auauuuacgu aguuuuucau guuuuagagc uaugcu 36

<210> 456

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 456

cguaaauauu cuugagaaac guuuuagagc uaugcu 36

<210> 457

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 457

uucuugagaa acaggaagac guuuuagagc uaugcu 36

<210> 458

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 458

uaauauuuaaa aacauugguu guuuuagagc uaugcu 36

<210> 459

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 459

ccaauguuuu uaauauuuauc guuuuagagc uaugcu 36

<210> 460

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 460

ccugauaaua uuaaaaacau guuuuagagc uaugcu 36

<210> 461

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 461

cauuaucaug cauacaucuc guuuuagagc uaugcu 36

<210> 462

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 462

aucaugcaua caucucuggc guuuuagagc uaugcu 36

<210> 463

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 463

uucauuucau uuugauuuug guuuuagagc uaugcu 36

<210> 464

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 464

auucaauuug aagcaguggu guuuuagagc uaugcu 36

<210> 465

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 465

gaauauucaa uuugaagcag guuuuagagc uaugcu 36

<210> 466

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 466

cauacgauuu aaaaucgcug guuuuagagc uaugcu 36

<210> 467

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 467

aaaaucgcug aggcgcguuc guuuuagagc uaugcu 36

<210> 468

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 468

uuuuuuuuuc uuuuuuguac guuuuagagc uaugcu 36

<210> 469

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 469

cuguugucaa agauuuuaaa guuuuagagc uaugcu 36

<210> 470

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 470

ugacaacaga guucuguuuu guuuuagagc uaugcu 36

<210> 471

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 471

agaauacgcu gagaguuauc guuuuagagc uaugcu 36

<210> 472

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 472

gcaagagaag aaaagaacgg guuuuagagc uaugcu 36

<210> 473

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 473

guugcaagag aagaaaagaa guuuuagagc uaugcu 36

<210> 474

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 474

augcacacgu aaaagagagg guuuuagagc uaugcu 36

<210> 475

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 475

aagaugcaca cguaaaagag guuuuagagc uaugcu 36

<210> 476

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 476

aucaugcaua caucucuggc guuuuagagc uaugcu 36

<210> 477

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 477

uucauuucau uuugauuuug guuuuagagc uaugcu 36

<210> 478

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 478

auucaauuug aagcaguggu guuuuagagc uaugcu 36

<210> 479

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 479

gaauauucaa uuugaagcag guuuuagagc uaugcu 36

<210> 480

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 480

cauacgauuu aaaaucgcug guuuuagagc uaugcu 36

<210> 481

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 481

aaaaucgcug aggcgcguuc guuuuagagc uaugcu 36

<210> 482

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 482

uuuuuuuuuc uuuuuuguac guuuuagagc uaugcu 36

<210> 483

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 483

cuguugucaa agauuuuaaa guuuuagagc uaugcu 36

<210> 484

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 484

ugacaacaga guucuguuuu guuuuagagc uaugcu 36

<210> 485

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 485

agaauacgcu gagaguuauc guuuuagagc uaugcu 36

<210> 486

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 486

gcaagagaag aaaagaacgg guuuuagagc uaugcu 36

<210> 487

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 487

guugcaagag aagaaaagaa guuuuagagc uaugcu 36

<210> 488

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 488

augcacacgu aaaagagagg guuuuagagc uaugcu 36

<210> 489

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 489

aagaugcaca cguaaaagag guuuuagagc uaugcu 36

<210> 490

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 490

ggcagaccgu ucucaucacg guuuuagagc uaugcu 36

<210> 491

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 491

cuuuaccagu gacuccaggu guuuuagagc uaugcu 36

<210> 492

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 492

gucacagauu uccuucuccg guuuuagagc uaugcu 36

<210> 493

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 493

agaugaugac gcccaccaga guuuuagagc uaugcu 36

<210> 494

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 494

ggagaaggaa aucugugacc guuuuagagc uaugcu 36

<210> 495

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 495

ugcgaggaac uuacuuuucc guuuuagagc uaugcu 36

<210> 496

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 496

agagaaauau ugauauagga guuuuagagc uaugcu 36

<210> 497

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 497

uaucaauauu ucucugaucc guuuuagagc uaugcu 36

<210> 498

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 498

aucgcuuuua aggcacgcuc guuuuagagc uaugcu 36

<210> 499

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 499

uauacgacug aucgcuuuua guuuuagagc uaugcu 36

<210> 500

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 500

ugucagguua guuagaugaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 501

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223

> Synthetic

<400> 501

gugucagguu aguuagauga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 502

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 502

ccugacacau auacagacua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 503

<211> 97

<212> RNA

<213> Artificial Sequence

<220><

223> Synthetic

<400> 503

cugacacaua uacagacuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 504

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 504

ccuuagucug uauauguguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 505

<211> 97

<212> RNA

<213> Artificial Sequence

<220

><223> Synthetic

<400> 505

cauauacaga cuaagggacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 506

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 506

auauacagac uaagggacca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 507

<211> 97

<212> RNA

<213> Artificial Sequence

<

220><223> Synthetic

<400> 507

ucaaaguuug auaaauuccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 508

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 508

aaaauacaaa gauaaguaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 509

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 509

acucugugac uuuaaaaagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 510

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 510

gguucugugg gauauuaaua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 511

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 511

acagagcaua uugguucugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 512

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 512

gacagagcau auugguucug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 513

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 513

ugcaaaacga cagagcauau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 514

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 514

gagcugggca uggaauaggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 515

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 515

acuggagcug ggcauggaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 516

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 516

cucauuacug gagcugggca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 517

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 517

uuguucucau uacuggagcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 518

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 518

auuguucuca uuacuggagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 519

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 519

ggggagauug uucucauuac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 520

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 520

gaggagaaaa ucuguggcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 521

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 521

agaggagaaa aucuguggcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 522

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 522

cagaggagaa aaucuguggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 523

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 523

uccucagagg agaaaaucug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 524

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 524

ugaaguuuuu cauuccucag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 525

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 525

cuucaccaac gacuccaagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 526

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 526

cuacuccuac uuggagucgu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 527

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 527

cuccaaguag gaguagauga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 528

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 528

caccaucauc uacuccuacu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 529

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 529

ugauggugau cagaagcaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 530

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 530

ucagaagcag aaggauuucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 531

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 531

gauuucuagg augauguuca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 532

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 532

uugcucuguc cucuuccuuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 533

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 533

aggacugaac cagaaggaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 534

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 534

uacacaagga cugaaccaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 535

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 535

uucaguccuu guguaguccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 536

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 536

ucaguccuug uguaguccua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 537

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 537

guccuugugu aguccuaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 538

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 538

cuuguguagu ccuagggagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 539

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 539

cuccucccua ggacuacaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 540

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 540

guagacagua ccuccucccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 541

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 541

uacugucuac acagagcucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 542

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 542

acugucuaca cagagcucua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 543

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 543

ucuacacaga gcucuaggga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 544

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 544

cuacacagag cucuagggaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 545

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 545

uacacagagc ucuagggaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 546

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 546

ggggugugcc caguuguuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 547

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 547

gggugugccc aguuguuaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 548

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 548

ugguaguccc auuaacaacu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 549

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 549

cugguagucc cauuaacaac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 550

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 550

uuguuaaugg gacuaccaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 551

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 551

uaccagaugg aagccagcuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 552

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 552

uuccaaagcu ggcuuccauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 553

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 553

uggaagccag cuuuggaagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 554

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 554

acaaggccug cuuccaaagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 555

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 555

gccuuguuca cguguucuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 556

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 556

ccuuguucac guguucuaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 557

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 557

cccauuagaa cacgugaaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 558

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 558

uuggcaucac uucauauuug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 559

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 559

cuugugcucu uggcaucacu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 560

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 560

agcacacucu cuugugcucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 561

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 561

gcacaagaga gugugcucuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 562

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 562

gcuuaaucuc acacauagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 563

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 563

cuuaaucuca cacauagaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 564

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 564

uuaaucucac acauagaaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 565

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 565

aggagugcug guuuaucaac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 566

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 566

uucuuugaca gcaggagugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 567

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 567

acucugguuu cuuugacagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 568

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 568

accagaguug agaaaacccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 569

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 569

uccagggguu uucucaacuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 570

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 570

caguuauuaa augaauccag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 571

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 571

gcaguuauua aaugaaucca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 572

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 572

ggcaguuauu aaaugaaucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 573

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 573

uggaugguaa cagcuacauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 574

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 574

gcuguuacca uccacauccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 575

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 575

ucaagaacca aggaugugga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 576

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 576

uccuucaaga accaaggaug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 577

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 577

ugaguguccu ucaagaacca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 578

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 578

uuuuauuuua uaacuacaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 579

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 579

uuguuuuuaa uaaaaacaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 580

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 580

uauuauagaa ugcuuuugca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 581

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 581

caagauuagu cuugauguag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 582

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 582

aagauuaguc uugauguagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 583

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 583

agucuugaug uagugggagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 584

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 584

uuuuucuauu aaaaaaaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 585

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 585

ucuauuaaaa aaaaaaaggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 586

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 586

cuauuaaaaa aaaaaaggcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 587

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 587

aaaaaaaaaa aggcugggca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 588

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 588

aaaaaaaagg cugggcacgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 589

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 589

cacccguaau cccagcacuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 590

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 590

acccguaauc ccagcacuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 591

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 591

cguaauccca gcacuuuggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 592

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 592

ucccaaagug cugggauuac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 593

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 593

cucccaaagu gcugggauua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 594

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 594

cccagcacuu ugggaggccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 595

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 595

ccucggccuc ccaaagugcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 596

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 596

gcacuuuggg aggccgaggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 597

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 597

cuuugggagg ccgaggcagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 598

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 598

gccgaggcag guggaucacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 599

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 599

accucgugau ccaccugccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 600

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 600

ggcaggugga ucacgagguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 601

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 601

ucaggagauc gagaccaucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 602

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 602

cgagaccauc uuggcuaaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 603

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 603

uuucaccaug uuagccaaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 604

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 604

uuguauuuuu uguagagacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 605

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 605

uuuguauuuu uuguagagac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 606

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 606

uuuuguauuu uuuguagaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 607

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 607

aaaaaauaca aaaaauuagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 608

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 608

aaaaauacaa aaaauuagcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 609

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 609

uacaaaaaau uagccgggug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 610

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 610

aaaaaauuag ccgggugugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 611

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 611

aaauuagccg gguguggugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 612

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 612

aauuagccgg gugugguggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 613

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 613

caggcgcccg ccaccacacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 614

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 614

gccuguaguc ccagcuacuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 615

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 615

uguaguccca gcuacucagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 616

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 616

uccugaguag cugggacuac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 617

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 617

cccagcuacu caggaggcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 618

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 618

ccucagccuc cugaguagcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 619

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 619

gccucagccu ccugaguagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 620

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 620

aggaggcuga ggcaggagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 621

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 621

gcaggagaau ggcgugaacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 622

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 622

caggagaaug gcgugaaccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 623

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 623

gagaauggcg ugaacccggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 624

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 624

aauggcguga acccgggagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 625

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 625

cacugcaagc uccaccuccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 626

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 626

ucacugcaag cuccaccucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 627

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 627

cauaccacug cacuccagcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 628

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 628

auaccacugc acuccagccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 629

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 629

ucgcccaggc uggagugcag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 630

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 630

ucucacucuu ucgcccaggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 631

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 631

ggagucucac ucuuucgccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 632

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 632

uguuuuuugu uuuuuugaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 633

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 633

aggaagaaag aaagguuuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 634

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 634

agaagaaaag gaagaaagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 635

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 635

cuuucuuccu uuucuucucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 636

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 636

uuucuuccuu uucuucucuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 637

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 637

aauggaccca agagaagaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 638

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 638

ggcuauuaca uaagaaacaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 639

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 639

cacaggaaaa ggaacuguac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 640

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 640

auuaaagcua acacaggaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 641

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 641

ucaaaaauua aagcuaacac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 642

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 642

uaaaauuguc uaaacaucuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 643

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 643

agagauguuu agacaauuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 644

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 644

ucuaaacauc ucugggacca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 645

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 645

uuuaugcuuu cauauauccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 646

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 646

agcauaaauu acaaagaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 647

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 647

uacaaagaaa aagguuauca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 648

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 648

acaaagaaaa agguuaucau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 649

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 649

caaagaaaaa gguuaucaug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 650

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 650

ucugagauuu aaaauagagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 651

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 651

cuuauaagau acauuaugaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 652

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 652

uaucuuauaa gacuauaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 653

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 653

aucuuauaag acuauaaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 654

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 654

uuauaagacu auaaaaaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 655

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 655

uaaaaaggga ggaaauauag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 656

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 656

aaaaagggag gaaauauaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 657

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 657

aaauauagag gguccacuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 658

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 658

uauagagggu ccacuuuugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 659

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 659

acucugaagu ccaccaaaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 660

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 660

agaauagagu ugcaccguuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 661

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 661

aaaacggugc aacucuauuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 662

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 662

ccguuuuggg cuaaugaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 663

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 663

ccuuuuucau uagcccaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 664

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 664

ugggcuaaug aaaaaggaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 665

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 665

uaaugaaaaa ggaagaggcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 666

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 666

aaugaaaaag gaagaggcua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 667

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 667

cugaaucuua aaauaugucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 668

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 668

caggcagcuu uaucucaacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 669

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 669

cuaagagauc aaguuucagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 670

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 670

guguucuugu ugauauucug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 671

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 671

cuuguugaua uucuguggca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 672

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 672

ucuguggcau ggcuacagau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 673

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 673

agaacuuauu uacacaggga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 674

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 674

aaagagaacu uauuuacaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 675

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 675

caaagagaac uuauuuacac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 676

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 676

uucucuuugu auuuacuuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 677

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 677

ucucuuugua uuuacuuuua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 678

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 678

cuuuguauuu acuuuuaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 679

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 679

agcuuuuguc caccuuuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 680

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 680

uuuauuuuuc cauuuaaagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 681

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 681

uauuuuauuu uuccauuuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 682

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 682

cuuacauaaa cauacuuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 683

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 683

uaagcacaga aguuuuuaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 684

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 684

aaguuuuuaa gaggcaugaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 685

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 685

auauuuacgu aguuuuucau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 686

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 686

cguaaauauu cuugagaaac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 687

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 687

uucuugagaa acaggaagac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 688

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 688

uaauauuaaa aacauugguu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 689

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 689

ccaauguuuu uaauauuuauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 690

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 690

ccugauaaua uuaaaaacau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 691

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 691

cauuaucaug cauacaucuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 692

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 692

aucaugcaua caucucuggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 693

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 693

uucauuucau uuugauuuug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 694

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 694

auucaauuug aagcaguggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 695

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 695

gaauauucaa uuugaagcag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 696

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 696

cauacgauuu aaaaucgcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 697

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 697

aaaaucgcug aggcgcguuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 698

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 698

uuuuuuuuuc uuuuuuguac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 699

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 699

cuguugucaa agauuuuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 700

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 700

ugacaacaga guucuguuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 701

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 701

agaauacgcu gagaguuauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 702

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 702

gcaagagaag aaaagaacgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 703

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 703

guugcaagag aagaaaagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 704

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 704

augcacacgu aaaagagagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 705

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 705

aagaugcaca cguaaaagag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 706

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 706

aucaugcaua caucucuggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 707

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 707

uucauuucau uuugauuuug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 708

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 708

auucaauuug aagcaguggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 709

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 709

gaauauucaa uuugaagcag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 710

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 710

cauacgauuu aaaaucgcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 711

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 711

aaaaucgcug aggcgcguuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 712

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 712

uuuuuuuuuc uuuuuuguac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 713

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 713

cuguugucaa agauuuuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 714

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 714

ugacaacaga guucuguuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 715

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 715

agaauacgcu gagaguuauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 716

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 716

gcaagagaag aaaagaacgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 717

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 717

guugcaagag aagaaaagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 718

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 718

augcacacgu aaaagagagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 719

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 719

aagaugcaca cguaaaagag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 720

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 720

ggcagaccgu ucucaucacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 721

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 721

cuuuaccagu gacuccaggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 722

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 722

gucacagauu uccuucuccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 723

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 723

agaugaugac gcccaccaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 724

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 724

ggagaaggaa aucugugacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 725

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 725

ugcgaggaac uuacuuuucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 726

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 726

agagaaauau ugauauagga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 727

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 727

uaucaauauu ucucugaucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 728

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 728

aucgcuuuua aggcacgcuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 729

<211> 97

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 729

uauacgacug aucgcuuuua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcu 97

<210> 730

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 730

ugucagguua guuagaugaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 731

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 731

gugucagguu aguuagauga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 732

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 732

ccugacacau auacagacua guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 733

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 733

cugacacaua uacagacuaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 734

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 734

ccuuagucug uauauguguc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 735

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 735

cauauacaga cuaagggacc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 736

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 736

auauacagac uaagggacca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 737

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 737

ucaaaguuug auaaauuccc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 738

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 738

aaaauacaaa gauaaguaga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 739

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 739

acucugugac uuuaaaaagu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 740

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 740

gguucugugg gauauuaaua guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 741

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 741

acagagcaua uugguucugu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 742

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 742

gacagagcau auugguucug guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 743

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 743

ugcaaaacga cagagcauau guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 744

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 744

gagcugggca uggaauaggc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 745

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 745

acuggagcug ggcauggaau guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 746

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 746

cucauuacug gagcugggca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 747

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 747

uuguucucau uacuggagcu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 748

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 748

auuguucuca uuacuggagc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 749

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 749

ggggagauug uucucauuac guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 750

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 750

gaggagaaaa ucuguggcug guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 751

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 751

agaggagaaa aucuguggcu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 752

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 752

cagaggagaa aaucuguggc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 753

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 753

uccucagagg agaaaaucug guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 754

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 754

ugaaguuuuu cauuccucag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 755

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 755

cuucaccaac gacuccaagu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 756

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 756

cuacuccuac uuggagucgu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 757

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 757

cuccaaguag gaguagauga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 758

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 758

caccaucauc uacuccuacu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 759

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 759

ugauggugau cagaagcaga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 760

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 760

ucagaagcag aaggauuucu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 761

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 761

gauuucuagg augauguuca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 762

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 762

uugcucuguc cucuuccuuc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 763

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 763

aggacugaac cagaaggaag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 764

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 764

uacacaagga cugaaccaga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 765

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 765

uucaguccuu guguaguccu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 766

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 766

ucaguccuug uguaguccua guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 767

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 767

guccuugugu aguccuaggg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 768

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 768

cuuguguagu ccuagggagg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 769

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 769

cuccucccua ggacuacaca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 770

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 770

guagacagua ccuccucccu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 771

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 771

uacugucuac acagagcucu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 772

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 772

acugucuaca cagagcucua guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 773

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 773

ucuacacaga gcucuaggga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 774

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 774

cuacacagag cucuagggaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 775

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 775

uacacagagc ucuagggaag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 776

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 776

ggggugugcc caguuguuaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 777

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 777

gggugugccc aguuguuaau guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 778

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 778

ugguaguccc auuaacaacu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 779

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 779

cugguagucc cauuaacaac guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 780

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 780

uuguuaaugg gacuaccaga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 781

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 781

uaccagaugg aagccagcuu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 782

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 782

uuccaaagcu ggcuuccauc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 783

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 783

uggaagccag cuuuggaagc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 784

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 784

acaaggccug cuuccaaagc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 785

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 785

gccuuguuca cguguucuaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 786

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 786

ccuuguucac guguucuaau guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 787

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 787

cccauuagaa cacgugaaca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 788

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 788

uuggcaucac uucauauuug guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 789

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 789

cuugugcucu uggcaucacu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 790

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 790

agcacacucu cuugugcucu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 791

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 791

gcacaagaga gugugcucuc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 792

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 792

gcuuaaucuc acacauagaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 793

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 793

cuuaaucuca cacauagaaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 794

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 794

uuaaucucac acauagaaag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 795

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 795

aggagugcug guuuaucaac guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 796

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 796

uucuuugaca gcaggagugc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 797

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 797

acucugguuu cuuugacagc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 798

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 798

accagaguug agaaaacccc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 799

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 799

uccagggguu uucucaacuc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 800

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 800

caguuauuaa augaauccag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 801

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 801

gcaguuauua aaugaaucca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 802

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 802

ggcaguuauu aaaugaaucc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 803

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 803

uggaugguaa cagcuacauc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 804

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 804

gcuguuacca uccacauccu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 805

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 805

ucaagaacca aggaugugga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 806

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 806

uccuucaaga accaaggaug guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 807

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 807

ugaguguccu ucaagaacca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 808

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 808

uuuuauuuua uaacuacaag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 809

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 809

uuguuuuuaa uaaaaacaag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 810

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 810

uauuauagaa ugcuuuugca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 811

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 811

caagauuagu cuugauguag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 812

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 812

aagauuaguc uugauguagu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 813

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 813

agucuugaug uagugggagu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 814

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 814

uuuuucuauu aaaaaaaaaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 815

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 815

ucuauuaaaa aaaaaaaggc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 816

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 816

cuauuaaaaa aaaaaaggcu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 817

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 817

aaaaaaaaaa aggcugggca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 818

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 818

aaaaaaaagg cugggcacgg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 819

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 819

cacccguaau cccagcacuu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 820

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 820

acccguaauc ccagcacuuu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 821

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 821

cguaauccca gcacuuuggg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 822

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 822

ucccaaagug cugggauuac guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 823

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 823

cucccaaagu gcugggauua guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 824

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 824

cccagcacuu ugggaggccg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 825

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 825

ccucggccuc ccaaagugcu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 826

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 826

gcacuuuggg aggccgaggc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 827

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 827

cuuugggagg ccgaggcagg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 828

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 828

gccgaggcag guggaucacg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 829

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 829

accucgugau ccaccugccu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 830

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 830

ggcaggugga ucacgagguc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 831

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 831

ucaggagauc gagaccaucu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 832

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 832

cgagaccauc uuggcuaaca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 833

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 833

uuucaccaug uuagccaaga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 834

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 834

uuguauuuuu uguagagacg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 835

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 835

uuuguauuuu uuguagagac guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 836

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 836

uuuuguauuu uuuguagaga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 837

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 837

aaaaaauaca aaaaauuagc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 838

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 838

aaaaauacaa aaaauuagcc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 839

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 839

uacaaaaaau uagccgggug guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 840

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 840

aaaaaauuag ccgggugugg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 841

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 841

aaauuagccg gguguggugg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 842

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 842

aauuagccgg gugugguggc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 843

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 843

caggcgcccg ccaccacacc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 844

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 844

gccuguaguc ccagcuacuc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 845

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 845

uguaguccca gcuacucagg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 846

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 846

uccugaguag cugggacuac guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 847

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 847

cccagcuacu caggaggcug guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 848

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 848

ccucagccuc cugaguagcu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 849

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 849

gccucagccu ccugaguagc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 850

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 850

aggaggcuga ggcaggagaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 851

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 851

gcaggagaau ggcgugaacc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 852

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 852

caggagaaug gcgugaaccc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 853

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 853

gagaauggcg ugaacccggg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 854

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 854

aauggcguga acccgggagg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 855

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 855

cacugcaagc uccaccuccc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 856

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 856

ucacugcaag cuccaccucc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 857

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 857

cauaccacug cacuccagcc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 858

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 858

auaccacugc acuccagccu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 859

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 859

ucgcccaggc uggagugcag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 860

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 860

ucucacucuu ucgcccaggc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 861

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 861

ggagucucac ucuuucgccc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 862

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 862

uguuuuuugu uuuuuugaga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 863

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 863

aggaagaaag aaagguuuuu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 864

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 864

agaagaaaag gaagaaagaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 865

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 865

cuuucuuccu uuucuucucu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 866

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 866

uuucuuccuu uucuucucuu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 867

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 867

aauggaccca agagaagaaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 868

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 868

ggcuauuaca uaagaaacaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 869

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 869

cacaggaaaa ggaacuguac guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 870

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 870

auuaaagcua acacaggaaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 871

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 871

ucaaaaauua aagcuaacac guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 872

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 872

uaaaauuguc uaaacaucuc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 873

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 873

agagauguuu agacaauuuu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 874

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 874

ucuaaacauc ucugggacca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 875

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 875

uuuaugcuuu cauauauccu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 876

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 876

agcauaaauu acaaagaaaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 877

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 877

uacaaagaaa aagguuauca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 878

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 878

acaaagaaaa agguuaucau guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 879

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 879

caaagaaaaa gguuaucaug guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 880

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 880

ucugagauuu aaaauagagu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 881

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 881

cuuauaagau acauuaugaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 882

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 882

uaucuuauaa gacuauaaaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 883

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 883

aucuuauaag acuauaaaaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 884

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 884

uuauaagacu auaaaaaggg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 885

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 885

uaaaaaggga ggaaauauag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 886

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 886

aaaaagggag gaaauauaga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 887

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 887

aaauauagag gguccacuuu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 888

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 888

uauagagggu ccacuuuugg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 889

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 889

acucugaagu ccaccaaaag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 890

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 890

agaauagagu ugcaccguuu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 891

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 891

aaaacggugc aacucuauuc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 892

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 892

ccguuuuggg cuaaugaaaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 893

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 893

ccuuuuucau uagcccaaaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 894

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 894

ugggcuaaug aaaaaggaag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 895

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 895

uaaugaaaaa ggaagaggcu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 896

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 896

aaugaaaaag gaagaggcua guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 897

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 897

cugaaucuua aaauaugucc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 898

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 898

caggcagcuu uaucucaacc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 899

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 899

cuaagagauc aaguuucagc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 900

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 900

guguucuugu ugauauucug guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 901

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 901

cuuguugaua uucuguggca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 902

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 902

ucuguggcau ggcuacagau guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 903

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 903

agaacuuauu uacacaggga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 904

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 904

aaagagaacu uauuuacaca guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 905

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 905

caaagagaac uuauuuacac guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 906

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 906

uucucuuugu auuuacuuuu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 907

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 907

ucucuuugua uuuacuuuua guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 908

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 908

cuuuguauuu acuuuuaggg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 909

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 909

agcuuuuguc caccuuuaaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 910

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 910

uuuauuuuuc cauuuaaagg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 911

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 911

uauuuuauuu uuccauuuaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 912

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 912

cuuacauaaa cauacuuaaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 913

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 913

uaagcacaga aguuuuuaag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 914

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 914

aaguuuuuaa gaggcaugaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 915

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 915

auauuuacgu aguuuuucau guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 916

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 916

cguaaauauu cuugagaaac guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 917

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 917

uucuugagaa acaggaagac guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 918

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 918

uaauauuaaa aacauugguu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 919

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 919

ccaauguuuu uaauauuauc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 920

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 920

ccugauaaua uuaaaaacau guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 921

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 921

cauuaucaug cauacaucuc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 922

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 922

aucaugcaua caucucuggc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 923

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 923

uucauuucau uuugauuuug guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 924

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 924

auucaauuug aagcaguggu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 925

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 925

gaauauucaa uuugaagcag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 926

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 926

cauacgauuu aaaaucgcug guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 927

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 927

aaaaucgcug aggcgcguuc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 928

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 928

uuuuuuuuuc uuuuuuguac guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 929

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 929

cuguugucaa agauuuuaaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 930

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 930

ugacaacaga guucuguuuu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 931

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 931

agaauacgcu gagaguuauc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 932

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 932

gcaagagaag aaaagaacgg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 933

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 933

guugcaagag aagaaaagaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 934

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 934

augcacacgu aaaagagagg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 935

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 935

aagaugcaca cguaaaagag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 936

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 936

aucaugcaua caucucuggc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 937

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 937

uucauuucau uuugauuuug guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 938

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 938

auucaauuug aagcaguggu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 939

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 939

gaauauucaa uuugaagcag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 940

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 940

cauacgauuu aaaaucgcug guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 941

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 941

aaaaucgcug aggcgcguuc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 942

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 942

uuuuuuuuuc uuuuuuguac guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 943

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 943

cuguugucaa agauuuuaaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 944

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 944

ugacaacaga guucuguuuu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 945

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 945

agaauacgcu gagaguuauc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 946

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 946

gcaagagaag aaaagaacgg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 947

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 947

guugcaagag aagaaaagaa guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 948

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 948

augcacacgu aaaagagagg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 949

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 949

aagaugcaca cguaaaagag guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 950

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 950

ggcagaccgu ucucaucacg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 951

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 951

cuuuaccagu gacuccaggu guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 952

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 952

gucacagauu uccuucuccg guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 953

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 953

agaugaugac gcccaccaga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 954

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 954

ggagaaggaa aucugugacc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 955

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 955

ugcgaggaac uuacuuuucc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 956

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 956

agagaaauau ugauauagga guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 957

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 957

uaucaauauu ucucugaucc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 958

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 958

aucgcuuuua aggcacgcuc guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 959

<211> 102

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 959

uauacgacug aucgcuuuua guuggaacca uucaaaacag cauagcaagu uaaaauaagg 60

cuaguccguu aucaacuuga aaaaguggca ccgagucggu gc 102

<210> 960

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 960

ugucagguua guuagaugaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 961

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 961

gugucagguu aguuagauga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 962

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 962

ccugacacau auacagacua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 963

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 963

cugacacaua uacagacuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 964

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 964

ccuuagucug uauauguguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 965

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 965

cauauacaga cuaagggacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 966

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 966

auauacagac uaagggacca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 967

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 967

ucaaaguuug auaaauuccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 968

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 968

aaaauacaaa gauaaguaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 969

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 969

acucugugac uuuaaaaagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 970

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 970

gguucugugg gauauuaaua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 971

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 971

acagagcaua uugguucugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 972

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 972

gacagagcau auugguucug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 973

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 973

ugcaaaacga cagagcauau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 974

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 974

gagcugggca uggaauaggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 975

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 975

acuggagcug ggcauggaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 976

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 976

cucauuacug gagcugggca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 977

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 977

uuguucucau uacuggagcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 978

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 978

auuguucuca uuacuggagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 979

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 979

ggggagauug uucucauuac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 980

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 980

gaggagaaaa ucuguggcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 981

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 981

agaggagaaa aucuguggcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 982

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 982

cagaggagaa aaucuguggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 983

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 983

uccucagagg agaaaaucug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 984

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 984

ugaaguuuuu cauuccucag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 985

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 985

cuucaccaac gacuccaagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 986

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 986

cuacuccuac uuggagucgu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 987

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 987

cuccaaguag gaguagauga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 988

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 988

caccaucauc uacuccuacu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 989

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 989

ugauggugau cagaagcaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 990

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 990

ucagaagcag aaggauuucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 991

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 991

gauuucuagg augauguuca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 992

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 992

uugcucuguc cucuuccuuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 993

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 993

aggacugaac cagaaggaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 994

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 994

uacacaagga cugaaccaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 995

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 995

uucaguccuu guguaguccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 996

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 996

ucaguccuug uguaguccua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 997

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 997

guccuugugu aguccuaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 998

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 998

cuuguguagu ccuagggagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 999

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 999

cuccucccua ggacuacaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1000

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1000

guagacagua ccuccucccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1001

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1001

uacugucuac acagagcucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1002

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1002

acugucuaca cagagcucua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1003

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1003

ucuacacaga gcucuaggga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1004

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1004

cuacacagag cucuagggaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1005

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1005

uacacagagc ucuagggaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1006

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1006

ggggugugcc caguuguuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1007

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1007

gggugugccc aguuguuaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1008

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1008

ugguaguccc auuaacaacu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1009

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1009

cugguagucc cauuaacaac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1010

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1010

uuguuaaugg gacuaccaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1011

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1011

uaccagaugg aagccagcuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1012

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1012

uuccaaagcu ggcuuccauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1013

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1013

uggaagccag cuuuggaagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1014

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1014

acaaggccug cuuccaaagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1015

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1015

gccuuguuca cguguucuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1016

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1016

ccuuguucac guguucuaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1017

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1017

cccauuagaa cacgugaaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1018

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1018

uuggcaucac uucauauuug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1019

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1019

cuugugcucu uggcaucacu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1020

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1020

agcacacucu cuugugcucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1021

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1021

gcacaagaga gugugcucuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1022

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1022

gcuuaaucuc acacauagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1023

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1023

cuuaaucuca cacauagaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1024

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1024

uuaaucucac acauagaaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1025

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1025

aggagugcug guuuaucaac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1026

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1026

uucuuugaca gcaggagugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1027

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1027

acucugguuu cuuugacagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1028

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1028

accagaguug agaaaacccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1029

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1029

uccagggguu uucucaacuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1030

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1030

caguuauuaa augaauccag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1031

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1031

gcaguuauua aaugaaucca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1032

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1032

ggcaguuauu aaaugaaucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1033

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1033

uggaugguaa cagcuacauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1034

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1034

gcuguuacca uccacauccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1035

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1035

ucaagaacca aggaugugga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1036

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1036

uccuucaaga accaaggaug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1037

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1037

ugaguguccu ucaagaacca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1038

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1038

uuuuauuuua uaacuacaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1039

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1039

uuguuuuuaa uaaaaacaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1040

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1040

uauuauagaa ugcuuuugca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1041

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1041

caagauuagu cuugauguag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1042

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1042

aagauuaguc uugauguagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1043

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1043

agucuugaug uagugggagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1044

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1044

uuuuucuauu aaaaaaaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1045

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1045

ucuauuaaaa aaaaaaaggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1046

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1046

cuauuaaaaa aaaaaaggcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1047

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1047

aaaaaaaaaa aggcugggca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1048

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1048

aaaaaaaagg cugggcacgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1049

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1049

cacccguaau cccagcacuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1050

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1050

acccguaauc ccagcacuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1051

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1051

cguaauccca gcacuuuggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1052

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1052

ucccaaagug cugggauuac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1053

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1053

cucccaaagu gcugggauua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1054

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1054

cccagcacuu ugggaggccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1055

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1055

ccucggccuc ccaaagugcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1056

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1056

gcacuuuggg aggccgaggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1057

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1057

cuuugggagg ccgaggcagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1058

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1058

gccgaggcag guggaucacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1059

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1059

accucgugau ccaccugccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1060

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1060

ggcaggugga ucacgagguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1061

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1061

ucaggagauc gagaccaucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1062

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1062

cgagaccauc uuggcuaaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1063

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1063

uuucaccaug uuagccaaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1064

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1064

uuguauuuuu uguagagacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1065

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1065

uuuguauuuu uuguagagac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1066

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1066

uuuuguauuu uuuguagaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1067

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1067

aaaaaauaca aaaaauuagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1068

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1068

aaaaauacaa aaaauuagcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1069

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1069

uacaaaaaau uagccgggug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1070

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1070

aaaaaauuag ccgggugugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1071

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1071

aaauuagccg gguguggugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1072

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1072

aauuagccgg gugugguggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1073

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1073

caggcgcccg ccaccacacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1074

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1074

gccuguaguc ccagcuacuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1075

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1075

uguaguccca gcuacucagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1076

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1076

uccugaguag cugggacuac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1077

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1077

cccagcuacu caggaggcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1078

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1078

ccucagccuc cugaguagcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1079

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1079

gccucagccu ccugaguagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1080

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1080

aggaggcuga ggcaggagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1081

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1081

gcaggagaau ggcgugaacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1082

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1082

caggagaaug gcgugaaccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1083

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1083

gagaauggcg ugaacccggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1084

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1084

aauggcguga acccgggagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1085

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1085

cacugcaagc uccaccuccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1086

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1086

ucacugcaag cuccaccucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1087

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1087

cauaccacug cacuccagcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1088

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1088

auaccacugc acuccagccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1089

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1089

ucgcccaggc uggagugcag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1090

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1090

ucucacucuu ucgcccaggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1091

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1091

ggagucucac ucuuucgccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1092

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1092

uguuuuuugu uuuuuugaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1093

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1093

aggaagaaag aaagguuuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1094

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1094

agaagaaaag gaagaaagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1095

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1095

cuuucuuccu uuucuucucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1096

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1096

uuucuuccuu uucuucucuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1097

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1097

aauggaccca agagaagaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1098

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1098

ggcuauuaca uaagaaacaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1099

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1099

cacaggaaaa ggaacuguac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1100

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1100

auuaaagcua acacaggaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1101

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1101

ucaaaaauua aagcuaacac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1102

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1102

uaaaauuguc uaaacaucuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1103

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1103

agagauguuu agacaauuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1104

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1104

ucuaaacauc ucugggacca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1105

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1105

uuuaugcuuu cauauauccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1106

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1106

agcauaaauu acaaagaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1107

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1107

uacaaagaaa aagguuauca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1108

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1108

acaaagaaaa agguuaucau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1109

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1109

caaagaaaaa gguuaucaug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1110

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1110

ucugagauuu aaaauagagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1111

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1111

cuuauaagau acauuaugaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1112

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1112

uaucuuauaa gacuauaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1113

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1113

aucuuauaag acuauaaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1114

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1114

uuauaagacu auaaaaaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1115

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1115

uaaaaaggga ggaaauauag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1116

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1116

aaaaagggag gaaauauaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1117

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1117

aaauauagag gguccacuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1118

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1118

uauagagggu ccacuuuugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1119

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1119

acucugaagu ccaccaaaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1120

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1120

agaauagagu ugcaccguuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1121

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1121

aaaacggugc aacucuauuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1122

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1122

ccguuuuggg cuaaugaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1123

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1123

ccuuuuucau uagcccaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1124

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1124

ugggcuaaug aaaaaggaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1125

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1125

uaaugaaaaa ggaagaggcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1126

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1126

aaugaaaaag gaagaggcua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1127

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1127

cugaaucuua aaauaugucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1128

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1128

caggcagcuu uaucucaacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1129

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1129

cuaagagauc aaguuucagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1130

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1130

guguucuugu ugauauucug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1131

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1131

cuuguugaua uucuguggca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1132

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1132

ucuguggcau ggcuacagau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1133

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1133

agaacuuauu uacacaggga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1134

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1134

aaagagaacu uauuuacaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1135

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1135

caaagagaac uuauuuacac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1136

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1136

uucucuuugu auuuacuuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1137

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1137

ucucuuugua uuuacuuuua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1138

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1138

cuuuguauuu acuuuuaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1139

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1139

agcuuuuguc caccuuuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1140

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1140

uuuauuuuuc cauuuaaagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1141

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1141

uauuuuauuu uuccauuuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1142

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1142

cuuacauaaa cauacuuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1143

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1143

uaagcacaga aguuuuuaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1144

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1144

aaguuuuuaa gaggcaugaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1145

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1145

auauuuacgu aguuuuucau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1146

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1146

cguaaauauu cuugagaaac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1147

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1147

uucuugagaa acaggaagac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1148

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1148

uaauauuuaa aacauugguu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1149

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1149

ccaauguuuu uaauauuauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1150

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1150

ccugauaaua uuaaaaacau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1151

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1151

cauuaucaug cauacaucuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1152

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1152

aucaugcaua caucucuggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1153

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1153

uucauuucau uuugauuuug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1154

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1154

auucaauuug aagcaguggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1155

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1155

gaauauucaa uuugaagcag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1156

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1156

cauacgauuu aaaaucgcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1157

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1157

aaaaucgcug aggcgcguuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1158

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1158

uuuuuuuuuc uuuuuuguac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1159

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1159

cuguugucaa agauuuuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1160

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1160

ugacaacaga guucuguuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1161

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1161

agaauacgcu gagaguuauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1162

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1162

gcaagagaag aaaagaacgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1163

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1163

guugcaagag aagaaaagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1164

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1164

augcacacgu aaaagagagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1165

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1165

aagaugcaca cguaaaagag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1166

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1166

aucaugcaua caucucuggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1167

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1167

uucauuucau uuugauuuug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1168

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1168

auucaauuug aagcaguggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1169

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1169

gaauauucaa uuugaagcag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1170

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1170

cauacgauuu aaaaucgcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1171

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1171

aaaaucgcug aggcgcguuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1172

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1172

uuuuuuuuuc uuuuuuguac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1173

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1173

cuguugucaa agauuuuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1174

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1174

ugacaacaga guucuguuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1175

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1175

agaauacgcu gagaguuauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1176

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1176

gcaagagaag aaaagaacgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1177

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1177

guugcaagag aagaaaagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1178

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1178

augcacacgu aaaagagagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1179

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1179

aagaugcaca cguaaaagag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1180

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1180

ggcagaccgu ucucaucacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1181

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1181

cuuuaccagu gacuccaggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1182

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1182

gucacagauu uccuucuccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1183

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1183

agaugaugac gcccaccaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1184

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1184

ggagaaggaa aucugugacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1185

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1185

ugcgaggaac uuacuuuucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1186

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1186

agagaaauau ugauauagga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1187

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1187

uaucaauauu ucucugaucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1188

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1188

aucgcuuuua aggcacgcuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1189

<211> 96

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1189

uauacgacug aucgcuuuua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 1190

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1190

ugucagguua guuagaugaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1191

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1191

gugucagguu aguuagauga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1192

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1192

ccugacacau auacagacua guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1193

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1193

cugacacaua uacagacuaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1194

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1194

ccuuagucug uauauguguc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1195

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1195

cauauacaga cuaagggacc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1196

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1196

auauacagac uaagggacca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1197

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1197

ucaaaguuug auaaauuccc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1198

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1198

aaaauacaaa gauaaguaga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1199

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1199

acucugugac uuuaaaaagu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1200

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1200

gguucugugg gauauuaaua guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1201

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1201

acagagcaua uugguucugu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1202

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1202

gacagagcau auugguucug guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1203

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1203

ugcaaaacga cagagcauau guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1204

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1204

gagcugggca uggaauaggc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1205

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1205

acuggagcug ggcauggaau guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1206

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1206

cucauuacug gagcugggca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1207

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1207

uuguucucau uacuggagcu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1208

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1208

auuguucuca uuacuggagc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1209

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1209

ggggagauug uucucauuac guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1210

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1210

gaggagaaaa ucuguggcug guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1211

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1211

agaggagaaa aucuguggcu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1212

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1212

cagaggagaa aaucuguggc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1213

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1213

uccucagagg agaaaaucug guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1214

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1214

ugaaguuuuu cauuccucag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1215

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1215

cuucaccaac gacuccaagu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1216

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1216

cuacuccuac uuggagucgu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1217

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1217

cuccaaguag gaguagauga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1218

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1218

caccaucauc uacuccuacu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1219

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1219

ugauggugau cagaagcaga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1220

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1220

ucagaagcag aaggauuucu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1221

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1221

gauuucuagg augauguuca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1222

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1222

uugcucuguc cucuuccuuc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1223

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1223

aggacugaac cagaaggaag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1224

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1224

uacacaagga cugaaccaga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1225

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1225

uucaguccuu guguaguccu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1226

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1226

ucaguccuug uguaguccua guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1227

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1227

guccuugugu aguccuaggg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1228

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1228

cuuguguagu ccuagggagg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1229

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1229

cuccucccua ggacuacaca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1230

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1230

guagacagua ccuccucccu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1231

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1231

uacugucuac acagagcucu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1232

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1232

acugucuaca cagagcucua guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1233

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1233

ucuacacaga gcucuaggga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1234

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1234

cuacacagag cucuagggaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1235

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1235

uacacagagc ucuagggaag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1236

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1236

ggggugugcc caguuguuaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1237

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1237

gggugugccc aguuguuaau guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1238

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1238

ugguaguccc auuaacaacu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1239

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1239

cugguagucc cauuaacaac guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1240

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1240

uuguuaaugg gacuaccaga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1241

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1241

uaccagaugg aagccagcuu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1242

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1242

uuccaaagcu ggcuuccauc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1243

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1243

uggaagccag cuuuggaagc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1244

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1244

acaaggccug cuuccaaagc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1245

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1245

gccuuguuca cguguucuaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1246

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1246

ccuuguucac guguucuaau guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1247

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1247

cccauuagaa cacgugaaca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1248

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1248

uuggcaucac uucauauuug guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1249

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1249

cuugugcucu uggcaucacu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1250

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1250

agcacacucu cuugugcucu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1251

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1251

gcacaagaga gugugcucuc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1252

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1252

gcuuaaucuc acacauagaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1253

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1253

cuuaaucuca cacauagaaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1254

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1254

uuaaucucac acauagaaag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1255

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1255

aggagugcug guuuaucaac guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1256

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1256

uucuuugaca gcaggagugc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1257

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1257

acucugguuu cuuugacagc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1258

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1258

accagaguug agaaaacccc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1259

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1259

uccagggguu uucucaacuc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1260

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1260

caguuauuaa augaauccag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1261

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1261

gcaguuauua aaugaaucca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1262

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1262

ggcaguuauu aaaugaaucc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1263

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1263

uggaugguaa cagcuacauc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1264

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1264

gcuguuacca uccacauccu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1265

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1265

ucaagaacca aggaugugga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1266

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1266

uccuucaaga accaaggaug guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1267

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1267

ugaguguccu ucaagaacca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1268

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1268

uuuuauuuua uaacuacaag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1269

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1269

uuguuuuuaa uaaaaacaag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1270

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1270

uauuauagaa ugcuuuugca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1271

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1271

caagauuagu cuugauguag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1272

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1272

aagauuaguc uugauguagu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1273

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1273

agucuugaug uagugggagu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1274

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1274

uuuuucuauu aaaaaaaaaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1275

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1275

ucuauuaaaa aaaaaaaggc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1276

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1276

cuauuaaaaa aaaaaaggcu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1277

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1277

aaaaaaaaaa aggcugggca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1278

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1278

aaaaaaaagg cugggcacgg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1279

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1279

cacccguaau cccagcacuu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1280

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1280

acccguaauc ccagcacuuu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1281

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1281

cguaauccca gcacuuuggg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1282

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1282

ucccaaagug cugggauuac guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1283

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1283

cucccaaagu gcugggauua guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1284

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1284

cccagcacuu ugggaggccg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1285

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1285

ccucggccuc ccaaagugcu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1286

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1286

gcacuuuggg aggccgaggc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1287

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1287

cuuugggagg ccgaggcagg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1288

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1288

gccgaggcag guggaucacg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1289

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1289

accucgugau ccaccugccu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1290

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1290

ggcaggugga ucacgagguc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1291

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1291

ucaggagauc gagaccaucu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1292

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1292

cgagaccauc uuggcuaaca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1293

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1293

uuucaccaug uuagccaaga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1294

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1294

uuguauuuuu uguagagacg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1295

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1295

uuuguauuuu uuguagagac guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1296

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1296

uuuuguauuu uuuguagaga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1297

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1297

aaaaaauaca aaaaauuagc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1298

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1298

aaaaauacaa aaaauuagcc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1299

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1299

uacaaaaaau uagccgggug guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1300

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1300

aaaaaauuag ccgggugugg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1301

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1301

aaauuagccg gguguggugg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1302

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1302

aauuagccgg gugugguggc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1303

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1303

caggcgcccg ccaccacacc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1304

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1304

gccuguaguc ccagcuacuc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1305

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1305

uguaguccca gcuacucagg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1306

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1306

uccugaguag cugggacuac guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1307

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1307

cccagcuacu caggaggcug guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1308

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1308

ccucagccuc cugaguagcu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1309

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1309

gccucagccu ccugaguagc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1310

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1310

aggaggcuga ggcaggagaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1311

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1311

gcaggagaau ggcgugaacc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1312

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1312

caggagaaug gcgugaaccc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1313

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1313

gagaauggcg ugaacccggg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1314

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1314

aauggcguga acccgggagg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1315

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1315

cacugcaagc uccaccuccc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1316

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1316

ucacugcaag cuccaccucc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1317

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1317

cauaccacug cacuccagcc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1318

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1318

auaccacugc acuccagccu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1319

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1319

ucgcccaggc uggagugcag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1320

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1320

ucucacucuu ucgcccaggc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1321

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1321

ggagucucac ucuuucgccc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1322

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1322

uguuuuuugu uuuuuugaga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1323

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1323

aggaagaaag aaagguuuuu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1324

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1324

agaagaaaag gaagaaagaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1325

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1325

cuuucuuccu uuucuucucu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1326

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1326

uuucuuccuu uucuucucuu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1327

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1327

aauggaccca agagaagaaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1328

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1328

ggcuauuaca uaagaaacaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1329

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1329

cacaggaaaa ggaacuguac guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1330

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1330

auuaaagcua acacaggaaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1331

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1331

ucaaaaauua aagcuaacac guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1332

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1332

uaaaauuguc uaaacaucuc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1333

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1333

agagauguuu agacaauuuu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1334

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1334

ucuaaacauc ucugggacca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1335

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1335

uuuaugcuuu cauauauccu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1336

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1336

agcauaaauu acaaagaaaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1337

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1337

uacaaagaaa aagguuauca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1338

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1338

acaaagaaaa agguuaucau guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1339

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1339

caaagaaaaa gguuaucaug guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1340

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1340

ucugagauuu aaaauagagu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1341

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1341

cuuauaagau acauuaugaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1342

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1342

uaucuuauaa gacuauaaaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1343

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1343

aucuuauaag acuauaaaaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1344

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1344

uuauaagacu auaaaaaggg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1345

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1345

uaaaaaggga ggaaauauag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1346

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1346

aaaaagggag gaaauauaga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1347

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1347

aaauauagag gguccacuuu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1348

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1348

uauagagggu ccacuuuugg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1349

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1349

acucugaagu ccaccaaaag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1350

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1350

agaauagagu ugcaccguuu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1351

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1351

aaaacggugc aacucuauuc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1352

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1352

ccguuuuggg cuaaugaaaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1353

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1353

ccuuuuucau uagcccaaaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1354

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1354

ugggcuaaug aaaaaggaag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1355

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1355

uaaugaaaaa ggaagaggcu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1356

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1356

aaugaaaaag gaagaggcua guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1357

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1357

cugaaucuua aaauaugucc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1358

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1358

caggcagcuu uaucucaacc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1359

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1359

cuaagagauc aaguuucagc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1360

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1360

guguucuugu ugauauucug guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1361

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1361

cuuguugaua uucuguggca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1362

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1362

ucuguggcau ggcuacagau guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1363

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1363

agaacuuauu uacacaggga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1364

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1364

aaagagaacu uauuuacaca guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1365

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1365

caaagagaac uuauuuacac guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1366

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1366

uucucuuugu auuuacuuuu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1367

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1367

ucucuuugua uuuacuuuua guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1368

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1368

cuuuguauuu acuuuuaggg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1369

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1369

agcuuuuguc caccuuuaaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1370

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1370

uuuauuuuuc cauuuaaagg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1371

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1371

uauuuuauuu uuccauuuaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1372

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1372

cuuacauaaa cauacuuaaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1373

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1373

uaagcacaga aguuuuuaag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1374

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1374

aaguuuuuaa gaggcaugaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1375

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1375

auauuuacgu aguuuuucau guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1376

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1376

cguaaauauu cuugagaaac guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1377

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1377

uucuugagaa acaggaagac guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1378

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1378

uaauauuaaa aacauugguu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1379

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1379

ccaauguuuu uaauauuauc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1380

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1380

ccugauaaua uuaaaaacau guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1381

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1381

cauuaucaug cauacaucuc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1382

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1382

aucaugcaua caucucuggc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1383

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1383

uucauuucau uuugauuuug guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1384

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1384

auucaauuug aagcaguggu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1385

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1385

gaauauucaa uuugaagcag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1386

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1386

cauacgauuu aaaaucgcug guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1387

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1387

aaaaucgcug aggcgcguuc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1388

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1388

uuuuuuuuuc uuuuuuguac guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1389

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1389

cuguugucaa agauuuuaaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1390

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1390

ugacaacaga guucuguuuu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1391

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1391

agaauacgcu gagaguuauc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1392

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1392

gcaagagaag aaaagaacgg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1393

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1393

guugcaagag aagaaaagaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1394

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1394

augcacacgu aaaagagagg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1395

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1395

aagaugcaca cguaaaagag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1396

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1396

aucaugcaua caucucuggc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1397

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1397

uucauuucau uuugauuuug guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1398

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1398

auucaauuug aagcaguggu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1399

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1399

gaauauucaa uuugaagcag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1400

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1400

cauacgauuu aaaaucgcug guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1401

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1401

aaaaucgcug aggcgcguuc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1402

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1402

uuuuuuuuuc uuuuuuguac guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1403

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1403

cuguugucaa agauuuuaaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1404

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1404

ugacaacaga guucuguuuu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1405

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1405

agaauacgcu gagaguuauc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1406

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1406

gcaagagaag aaaagaacgg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1407

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1407

guugcaagag aagaaaagaa guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1408

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1408

augcacacgu aaaagagagg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1409

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1409

aagaugcaca cguaaaagag guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1410

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1410

ggcagaccgu ucucaucacg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1411

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1411

cuuuaccagu gacuccaggu guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1412

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1412

gucacagauu uccuucuccg guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1413

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1413

agaugaugac gcccaccaga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1414

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1414

ggagaaggaa aucugugacc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1415

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1415

ugcgaggaac uuacuuuucc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1416

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1416

agagaaauau ugauauagga guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1417

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1417

uaucaauauu ucucugaucc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1418

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1418

aucgcuuuua aggcacgcuc guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1419

<211> 106

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1419

uauacgacug aucgcuuuua guuuaagagc uaugcuggaa acagcauagc aaguuuaaau 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 1420

<211> 77

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1420

guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc cguuaucaac uugaaaaagu 60

ggcaccgagu cggugcu 77

<210> 1421

<211> 16

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1421

guuuuagagc uaugcu 16

<210> 1422

<211> 67

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1422

agcauagcaa guuaaaauaa ggcuaguccg uuaucaacuu gaaaaagugg caccgagucg 60

gugcuuu 67

<210> 1423

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1423

ugucagguua guuagaugaa 20

<210> 1424

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1424

gugucagguu aguuagauga 20

<210> 1425

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1425

ccugacacau auacagacua 20

<210> 1426

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1426

cugacacaua uacagacuaa 20

<210> 1427

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1427

ccuuagucug uauauguguc 20

<210> 1428

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1428

cauauacaga cuaagggacc 20

<210> 1429

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1429

auauacagac uaagggacca 20

<210> 1430

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1430

ucaaaguuug auaaauuccc 20

<210> 1431

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1431

aaaauacaaa gauaaguaga 20

<210> 1432

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1432

acucugugac uuuaaaaagu 20

<210> 1433

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1433

gguucugugg gauauuaaua 20

<210> 1434

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1434

acagagcaua uugguucugu 20

<210> 1435

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1435

gacagagcau auugguucug 20

<210> 1436

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1436

ugcaaaacga cagagcauau 20

<210> 1437

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1437

gagcugggca uggaauaggc 20

<210> 1438

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1438

acuggagcug ggcauggaau 20

<210> 1439

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1439

cucauuacug gagcugggca 20

<210> 1440

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1440

uuguucucau uacuggagcu 20

<210> 1441

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1441

auuguucuca uuacuggagc 20

<210> 1442

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1442

ggggagauug uucucauuac 20

<210> 1443

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1443

gaggagaaaa ucuguggcug 20

<210> 1444

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1444

agaggagaaa aucuguggcu 20

<210> 1445

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1445

cagaggagaa aaucuguggc 20

<210> 1446

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1446

uccucagagg agaaaaucug 20

<210> 1447

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1447

ugaaguuuuu cauuccucag 20

<210> 1448

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1448

cuucaccaac gacuccaagu 20

<210> 1449

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1449

cuacuccuac uuggagucgu 20

<210> 1450

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1450

cuccaaguag gaguagauga 20

<210> 1451

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1451

caccaucauc uacuccuacu 20

<210> 1452

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1452

ugauggugau cagaagcaga 20

<210> 1453

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1453

ucagaagcag aaggauuucu 20

<210> 1454

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1454

gauuucuagg augauguuca 20

<210> 1455

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1455

uugcucuguc cucuuccuuc 20

<210> 1456

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1456

aggacugaac cagaaggaag 20

<210> 1457

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1457

uacacaagga cugaaccaga 20

<210> 1458

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1458

uucaguccuu guguaguccu 20

<210> 1459

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1459

ucaguccuug uguaguccua 20

<210> 1460

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1460

guccuugugu aguccuaggg 20

<210> 1461

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1461

cuuguguagu ccuagggagg 20

<210> 1462

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1462

cuccucccua ggacuacaca 20

<210> 1463

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1463

guagacagua ccuccucccu 20

<210> 1464

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1464

uacugucuac acagagcucu 20

<210> 1465

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1465

acugucuaca cagagcucua 20

<210> 1466

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1466

ucuacacaga gcucuaggga 20

<210> 1467

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1467

cuacacagag cucuagggaa 20

<210> 1468

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1468

uacacagagc ucuagggaag 20

<210> 1469

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1469

ggggugugcc caguuguuaa 20

<210> 1470

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1470

gggugugccc aguuguuaau 20

<210> 1471

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1471

ugguaguccc auuaacaacu 20

<210> 1472

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1472

cugguagucc cauuaacaac 20

<210> 1473

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1473

uuguuaaugg gacuaccaga 20

<210> 1474

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1474

uaccagaugg aagccagcuu 20

<210> 1475

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1475

uuccaaagcu ggcuuccauc 20

<210> 1476

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1476

uggaagccag cuuuggaagc 20

<210> 1477

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1477

acaaggccug cuuccaaagc 20

<210> 1478

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1478

gccuuguuca cguguucuaa 20

<210> 1479

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1479

ccuuguucac guguucuaau 20

<210> 1480

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1480

cccauuagaa cacgugaaca 20

<210> 1481

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1481

uuggcaucac uucauauuug 20

<210> 1482

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1482

cuugugcucu uggcaucacu 20

<210> 1483

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1483

agcacacucu cuugugcucu 20

<210> 1484

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1484

gcacaagaga gugugcucuc 20

<210> 1485

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1485

gcuuaaucuc acacauagaa 20

<210> 1486

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1486

cuuaaucuca cacauagaaa 20

<210> 1487

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1487

uuaaucucac acauagaaag 20

<210> 1488

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1488

aggagugcug guuuaucaac 20

<210> 1489

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1489

uucuuugaca gcaggagugc 20

<210> 1490

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1490

acucugguuu cuuugacagc 20

<210> 1491

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1491

accagaguug agaaaacccc 20

<210> 1492

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1492

uccaggggguu uucucaacuc 20

<210> 1493

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1493

caguuauuaa augaauccag 20

<210> 1494

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1494

gcaguuauua aaugaaucca 20

<210> 1495

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1495

ggcaguuauu aaaugaaucc 20

<210> 1496

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1496

uggaugguaa cagcuacauc 20

<210> 1497

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1497

gcuguuacca uccacauccu 20

<210> 1498

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1498

ucaagaacca aggaugugga 20

<210> 1499

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1499

uccuucaaga accaaggaug 20

<210> 1500

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1500

ugaguguccu ucaagaacca 20

<210> 1501

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1501

uuuuauuuua uaacuacaag 20

<210> 1502

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1502

uuguuuuuaa uaaaaacaag 20

<210> 1503

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1503

uauuauagaa ugcuuuugca 20

<210> 1504

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1504

caagauuagu cuugauguag 20

<210> 1505

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1505

aagauuaguc uugauguagu 20

<210> 1506

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1506

agucuugaug uagugggagu 20

<210> 1507

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1507

uuuuucuauu aaaaaaaaaa 20

<210> 1508

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1508

ucuauuaaaa aaaaaaggc 20

<210> 1509

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1509

cuauuaaaaa aaaaaggcu 20

<210> 1510

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1510

aaaaaaaaaa aggcugggca 20

<210> 1511

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1511

aaaaaaaagg cugggcacgg 20

<210> 1512

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1512

cacccguaau cccagcacuu 20

<210> 1513

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1513

acccguaauc ccagcacuuu 20

<210> 1514

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1514

cguaauccca gcacuuuggg 20

<210> 1515

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1515

ucccaaagug cugggauuac 20

<210> 1516

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1516

cucccaaagu gcugggauua 20

<210> 1517

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1517

cccagcacuu ugggaggccg 20

<210> 1518

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1518

ccucggccuc ccaaagugcu 20

<210> 1519

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1519

gcacuuuggg aggccgaggc 20

<210> 1520

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1520

cuuugggagg ccgaggcagg 20

<210> 1521

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1521

gccgaggcag guggaucacg 20

<210> 1522

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1522

accucgugau ccaccugccu 20

<210> 1523

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1523

ggcaggugga ucacgagguc 20

<210> 1524

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1524

ucaggagauc gagaccaucu 20

<210> 1525

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1525

cgagaccauc uuggcuaaca 20

<210> 1526

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1526

uuucaccaug uuagccaaga 20

<210> 1527

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1527

uuguauuuuu uguagagacg 20

<210> 1528

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1528

uuuguauuuu uuguagagac 20

<210> 1529

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1529

uuuuguauuu uuuguagaga 20

<210> 1530

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1530

aaaaaauaca aaaaauuagc 20

<210> 1531

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1531

aaaaauacaa aaaauuagcc 20

<210> 1532

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1532

uacaaaaaau uagccgggug 20

<210> 1533

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1533

aaaaaauuag ccgggugugg 20

<210> 1534

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1534

aaauuagccg gguguggugg 20

<210> 1535

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1535

aauuagccgg gugugguggc 20

<210> 1536

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1536

caggcgcccg ccaccacacc 20

<210> 1537

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1537

gccuguaguc ccagcuacuc 20

<210> 1538

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1538

uguaguccca gcuacucagg 20

<210> 1539

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1539

uccugaguag cugggacuac 20

<210> 1540

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1540

cccagcuacu caggaggcug 20

<210> 1541

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1541

ccucagccuc cugaguagcu 20

<210> 1542

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1542

gccucagccu ccugaguagc 20

<210> 1543

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1543

aggaggcuga ggcaggagaa 20

<210> 1544

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1544

gcaggagaau ggcgugaacc 20

<210> 1545

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1545

caggagaaug gcgugaaccc 20

<210> 1546

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1546

gagaauggcg ugaacccggg 20

<210> 1547

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1547

aauggcguga acccgggagg 20

<210> 1548

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1548

cacugcaagc uccaccuccc 20

<210> 1549

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1549

ucacugcaag cuccaccucc 20

<210> 1550

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1550

cauaccacug cacuccagcc 20

<210> 1551

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1551

auaccacugc acuccagccu 20

<210> 1552

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1552

ucgcccaggc uggagugcag 20

<210> 1553

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1553

ucucacucuu ucgcccaggc 20

<210> 1554

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1554

ggagucucac ucuuucgccc 20

<210> 1555

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1555

uguuuuuugu uuuuuugaga 20

<210> 1556

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1556

aggaagaaag aaagguuuuu 20

<210> 1557

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1557

agaagaaaag gaagaaagaa 20

<210> 1558

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1558

cuuucuuccu uuucuucucu 20

<210> 1559

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1559

uuucuuccuu uucuucucuu 20

<210> 1560

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1560

aauggaccca agagaagaaa 20

<210> 1561

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1561

ggcuauuaca uaagaaacaa 20

<210> 1562

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1562

cacaggaaaa ggaacuguac 20

<210> 1563

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1563

auuaaagcua acacaggaaa 20

<210> 1564

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1564

ucaaaaauua aagcuaacac 20

<210> 1565

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1565

uaaaauuguc uaaacaucuc 20

<210> 1566

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1566

agagauguuu agacaauuuu 20

<210> 1567

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1567

ucuaaacauc ucugggacca 20

<210> 1568

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1568

uuuaugcuuu cauauauccu 20

<210> 1569

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1569

agcauaaauu acaaagaaaa 20

<210> 1570

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1570

uacaaagaaa aagguuauca 20

<210> 1571

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1571

acaaagaaaa agguuaucau 20

<210> 1572

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1572

caaagaaaaa gguuaucaug 20

<210> 1573

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1573

ucugagauuu aaaauagagu 20

<210> 1574

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1574

cuuauaagau acauuaugaa 20

<210> 1575

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1575

uaucuuauaa gacuauaaaa 20

<210> 1576

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1576

aucuuauaag acuauaaaaa 20

<210> 1577

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1577

uuauaagacu auaaaaaggg 20

<210> 1578

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1578

uaaaaaggga ggaaauauag 20

<210> 1579

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1579

aaaaagggag gaaauauaga 20

<210> 1580

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1580

aaauauagag gguccacuuu 20

<210> 1581

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1581

uauagagggu ccacuuuugg 20

<210> 1582

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1582

acucugaagu ccaccaaaag 20

<210> 1583

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1583

agaauagagu ugcaccguuu 20

<210> 1584

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1584

aaaacggugc aacucuauuc 20

<210> 1585

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1585

ccguuuuggg cuaaugaaaa 20

<210> 1586

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1586

ccuuuuucau uagcccaaaa 20

<210> 1587

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1587

ugggcuaaug aaaaaggaag 20

<210> 1588

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1588

uaaugaaaaa ggaagaggcu 20

<210> 1589

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1589

aaugaaaaag gaagaggcua 20

<210> 1590

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1590

cugaaucuua aaauaugucc 20

<210> 1591

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1591

caggcagcuu uaucucaacc 20

<210> 1592

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1592

cuaagagauc aaguuucagc 20

<210> 1593

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1593

guguucuugu ugauauucug 20

<210> 1594

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1594

cuuguugaua uucuguggca 20

<210> 1595

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1595

ucuguggcau ggcuacagau 20

<210> 1596

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1596

agaacuuauu uacacaggga 20

<210> 1597

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1597

aaagagaacu uauuuacaca 20

<210> 1598

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1598

caaagagaac uuauuuacac 20

<210> 1599

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1599

uucucuuugu auuuacuuuu 20

<210> 1600

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1600

ucucuuugua uuuacuuuua 20

<210> 1601

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1601

cuuuguauuu acuuuuaggg 20

<210> 1602

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1602

agcuuuuguc caccuuuaaa 20

<210> 1603

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1603

uuuauuuuuc cauuuaaagg 20

<210> 1604

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1604

uauuuuauuu uuccauuuaa 20

<210> 1605

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1605

cuuacauaaa cauacuuaaa 20

<210> 1606

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1606

uaagcacaga aguuuuuaag 20

<210> 1607

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1607

aaguuuuuaa gaggcaugaa 20

<210> 1608

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1608

auauuuacgu aguuuuucau 20

<210> 1609

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1609

cguaaauauu cuugagaaac 20

<210> 1610

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1610

uucuugagaa acaggaagac 20

<210> 1611

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1611

uaauauuaaa aacauugguu 20

<210> 1612

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1612

ccaauguuuu uaauauuauc 20

<210> 1613

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1613

ccugauaaua uuaaaaacau 20

<210> 1614

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1614

cauuaucaug cauacaucuc 20

<210> 1615

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1615

aucaugcaua caucucuggc 20

<210> 1616

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1616

uucauuucau uuugauuuug 20

<210> 1617

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1617

auucaauuug aagcaguggu 20

<210> 1618

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1618

gaauauucaa uuugaagcag 20

<210> 1619

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1619

cauacgauuu aaaaucgcug 20

<210> 1620

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1620

aaaaucgcug aggcgcguuc 20

<210> 1621

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1621

uuuuuuuuuc uuuuuuguac 20

<210> 1622

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1622

cuguugucaa agauuuuaaa 20

<210> 1623

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1623

ugacaacaga guucuguuuu 20

<210> 1624

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1624

agaauacgcu gagaguuauc 20

<210> 1625

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1625

gcaagagaag aaaagaacgg 20

<210> 1626

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1626

guugcaagag aagaaaagaa 20

<210> 1627

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1627

augcacacgu aaaagagagg 20

<210> 1628

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1628

aagaugcaca cguaaaagag 20

<210> 1629

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1629

aucaugcaua caucucuggc 20

<210> 1630

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1630

uucauuucau uuugauuuug 20

<210> 1631

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1631

auucaauuug aagcaguggu 20

<210> 1632

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1632

gaauauucaa uuugaagcag 20

<210> 1633

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1633

cauacgauuu aaaaucgcug 20

<210> 1634

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1634

aaaaucgcug aggcgcguuc 20

<210> 1635

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1635

uuuuuuuuuc uuuuuuguac 20

<210> 1636

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1636

cuguugucaa agauuuuaaa 20

<210> 1637

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1637

ugacaacaga guucuguuuu 20

<210> 1638

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1638

agaauacgcu gagaguuauc 20

<210> 1639

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1639

gcaagagaag aaaagaacgg 20

<210> 1640

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1640

guugcaagag aagaaaagaa 20

<210> 1641

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1641

augcacacgu aaaagagagg 20

<210> 1642

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1642

aagaugcaca cguaaaagag 20

<210> 1643

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1643

ggcagaccgu ucucaucacg 20

<210> 1644

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1644

cuuuaccagu gacuccaggu 20

<210> 1645

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1645

gucacagauu uccuucuccg 20

<210> 1646

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1646

agaugaugac gcccaccaga 20

<210> 1647

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1647

ggagaaggaa aucugugacc 20

<210> 1648

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1648

ugcgaggaac uuacuuuucc 20

<210> 1649

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1649

agagaaauau ugauauagga 20

<210> 1650

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1650

uaucaauauu ucucugaucc 20

<210> 1651

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1651

aucgcuuuua aggcacgcuc 20

<210> 1652

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 1652

uauacgacug aucgcuuuua 20