



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2024년09월04일

(11) 등록번호 10-2701443

(24) 등록일자 2024년08월28일

- (51) 국제특허분류(Int. Cl.)  
A01K 67/027 (2024.01) C07K 14/705 (2006.01)  
C12N 15/86 (2006.01)
  - (52) CPC특허분류  
A01K 67/0278 (2024.01)  
C07K 14/705 (2013.01)
  - (21) 출원번호 10-2020-7002408
  - (22) 출원일자(국제) 2018년06월27일  
심사청구일자 2021년06월25일
  - (85) 번역문제출일자 2020년01월23일
  - (65) 공개번호 10-2020-0022012
  - (43) 공개일자 2020년03월02일
  - (86) 국제출원번호 PCT/US2018/039864
  - (87) 국제공개번호 WO 2019/006034  
국제공개일자 2019년01월03일
  - (30) 우선권주장  
62/525,524 2017년06월27일 미국(US)
  - (56) 선행기술조사문헌  
W02017087780 A1\*  
(뒷면에 계속)
- 전체 청구항 수 : 총 19 항

- (73) 특허권자  
리제너론 파마슈티칼스 인코포레이티드  
미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리버 로드 777
- (72) 발명자  
무지카 알렉산더 오.  
미국 뉴욕 10591 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내  
구사로바 빅토리아  
미국 뉴욕 10591 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내  
(뒷면에 계속)
- (74) 대리인  
특허법인와이에스장

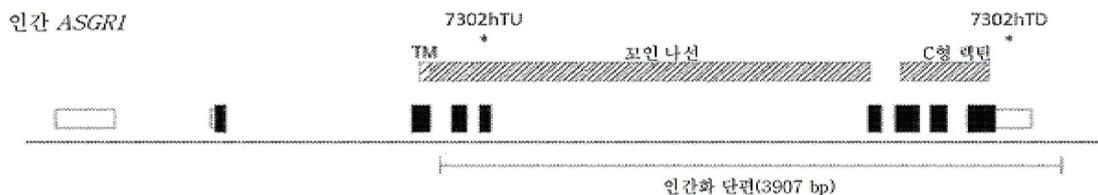
심사관 : 박영관

(54) 발명의 명칭 인간화 ASGR1 유전자좌를 포함하는 비인간 동물

(57) 요약

인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물 세포와 비인간 동물 및 이러한 비인간 동물 세포와 비인간 동물을 사용하는 방법이 제공된다. 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물 세포 또는 비인간 동물은 인간 ASGR1 단백질 또는 *Asgr1* 단백질을 발현하고, 이의 단편은 인간 ASGR1로부터 유래된다. 인간화 *Asgr1* 유전자좌를 포함하는 이러한 비인간 동물을 사용해 인간 ASGR1 매개에 의해 간에 전달된 치료 분자 또는 치료 복합체의 생체 내 효능을 평가하고, 인간 ASGR1 매개 메커니즘을 통해 작용하는 치료 분자 또는 치료 복합체의 효능을 평가하는 방법이 제공된다.

대표도



(52) CPC특허분류

**C12N 15/86** (2013.01)  
*A01K 2207/15* (2013.01)  
*A01K 2217/072* (2013.01)  
*A01K 2227/105* (2013.01)  
*A01K 2267/01* (2013.01)  
*C07K 2317/31* (2013.01)  
*C07K 2319/41* (2013.01)  
*C12N 2750/14122* (2013.01)  
*C12N 2750/14143* (2013.01)

(72) 발명자

**왕 쉥**

미국 뉴욕 10591 테리타운 올드 소우 밀 리버 로드  
777 리제너론 파마슈티칼스 인코포레이티드 내

**키랏소스 크리스토스**

미국 뉴욕 10591 테리타운 올드 소우 밀 리버 로드  
777 리제너론 파마슈티칼스 인코포레이티드 내

**포토키 데라**

미국 뉴욕 10591 테리타운 올드 소우 밀 리버 로드  
777 리제너론 파마슈티칼스 인코포레이티드 내

**시그나 캐서린**

미국 뉴욕 10591 테리타운 올드 소우 밀 리버 로드  
777 리제너론 파마슈티칼스 인코포레이티드 내

**마틴 조엘**

미국 뉴욕 10591 테리타운 올드 소우 밀 리버 로드  
777 리제너론 파마슈티칼스 인코포레이티드 내

(56) 선행기술조사문헌

WO2017100467 A2  
US20170164588 A1  
J Mol Biol. vol.300 no.4 pp.857-865 2000.  
J Biol Chem vol.260 no.4 pp.1979-82 1985.  
PLoS One vol.5 no.9 pp.e12934 2010.  
Gene. vol.148 no.2 pp.237-244 1994.

\*는 심사관에 의하여 인용된 문헌

**명세서**

**청구범위**

**청구항 1**

변형된 *Asgr1* 단백질을 암호화하는 유전적으로 변형된 내인성 *Asgr1* 유전자좌를 포함하는 설치류로서,  
 변형된 *Asgr1* 단백질은 세포질 도메인, 막관통 도메인, 및 세포외 도메인을 포함하고,  
 세포외 도메인은 꼬인 나선 도메인 및 C형 렉틴 도메인을 포함하고,

(i) 꼬인 나선 도메인의 전부는 결실되고 이종상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화되거나,

(ii) C형 렉틴 도메인의 전부는 결실되고 이종상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화되거나,

(iii) 꼬인 나선 도메인의 전부 및 C형 렉틴 도메인의 전부는 결실되고 이종상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 각각 암호화되거나, 또는

(iv) 세포외 도메인의 전부는 결실되고 이종상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화되고,

설치류는 변형된 *Asgr1* 단백질을 발현하며, 그리고

설치류는 랫트 또는 마우스인, 설치류.

**청구항 2**

제1 항에 있어서,

(a) C형 렉틴 도메인은 인간 *ASGR1* C형 렉틴 도메인이거나, 또는

(b) C형 렉틴 도메인은 서열 번호 28로 제시되는 서열을 포함하는 것을 특징으로 하는 설치류.

**청구항 3**

제1 항에 있어서,

(a) 꼬인 나선 도메인은 인간 *ASGR1* 꼬인 나선 도메인이거나, 또는

(b) 꼬인 나선 도메인은 서열 번호 27로 제시되는 서열을 포함하는 것을 특징으로 하는 설치류.

**청구항 4**

제1 항에 있어서,

(a) C형 렉틴 도메인은 인간 *ASGR1* C형 렉틴 도메인이고, 꼬인 나선 도메인은 인간 *ASGR1* 꼬인 나선 도메인이거나, 또는

(b) C형 렉틴 도메인은 서열 번호 28로 제시되는 서열을 포함하고, 꼬인 나선 도메인은 서열 번호 27로 제시되는 서열을 포함하는 것을 특징으로 하는 설치류.

**청구항 5**

제1 항에 있어서,

(a) 이종상동성 인간 *ASGR1* 서열은 인간 *ASGR1* 유전자의 엑손 3~8을 포함하거나,

(b) 이종상동성 인간 *ASGR1* 서열은 서열 번호 31로 제시되는 서열을 포함하는 *ASGR1* 단백질 분절을 암호화하거나, 또는

(c) 이종상동성 인간 *ASGR1* 서열은 인간 *ASGR1* 유전자의 엑손 3~8을 포함하고, 이종상동성 인간 *ASGR1* 서열은

서열 번호 31로 제시되는 서열을 포함하는 ASGR1 단백질 분절을 암호화하는 것을 특징으로 하는 설치류.

**청구항 6**

제1 항에 있어서,

- (a) 세포질 도메인의 전부 또는 일부는 내인성 설치류 *Asgr1* 서열에 의해 암호화되거나;
- (b) 막관통 도메인의 전부 또는 일부는 내인성 설치류 *Asgr1* 서열에 의해 암호화되거나; 또는
- (c) 세포질 도메인 및 막관통 도메인 모두의 전부 또는 일부는 내인성 설치류 *Asgr1* 서열에 의해 암호화된 것을 특징으로 하는 설치류.

**청구항 7**

제1 항에 있어서,

설치류는 유전적으로 변형된 내인성 *Asgr1* 유전자좌에 대한 이형접합체인 것을 특징으로 하는 설치류.

**청구항 8**

제1 항에 있어서,

설치류는 유전적으로 변형된 내인성 *Asgr1* 유전자좌에 대한 동형접합체인 것을 특징으로 하는 설치류.

**청구항 9**

제1 항에 있어서, 설치류는 마우스이고,

- (a) 세포질 도메인의 전부 또는 일부는 내인성 마우스 *Asgr1* 서열에 의해 암호화되거나;
- (b) 세포질 도메인의 전부 또는 일부는 내인성 마우스 *Asgr1* 서열에 의해 암호화되고, 세포질 도메인은 서열 번호 29로 제시되는 서열을 포함하거나;
- (c) 막관통 도메인의 전부 또는 일부는 내인성 마우스 *Asgr1* 서열에 의해 암호화되거나; 또는
- (d) 막관통 도메인의 전부 또는 일부는 내인성 마우스 *Asgr1* 서열에 의해 암호화되고, 막관통 도메인은 서열 번호 30으로 제시되는 서열을 포함하는 것을 특징으로 하는 설치류.

**청구항 10**

제9 항에 있어서,

- (a) 세포질 도메인 및 막관통 도메인 모두의 전부 또는 일부는 내인성 마우스 *Asgr1* 서열에 의해 암호화되거나; 또는
- (b) 세포질 도메인 및 막관통 도메인 모두의 전부 또는 일부는 내인성 마우스 *Asgr1* 서열에 의해 암호화되고, 세포질 도메인은 서열 번호 29로 제시되는 서열을 포함하고, 막관통 도메인은 서열 번호 30으로 제시되는 서열을 포함하는 것을 특징으로 하는 설치류.

**청구항 11**

제1 항에 있어서,

- (a) 꼬인 나선 도메인 및 C형 렉틴 도메인 모두의 전부는 결실되고 이종상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화되고, 세포질 도메인 및 막관통 도메인 모두의 전부 또는 일부는 내인성 마우스 *Asgr1* 서열에 의해 암호화되거나;
- (b) C형 렉틴 도메인은 인간 *ASGR1* C형 렉틴 도메인이고, 꼬인 나선 도메인은 인간 *ASGR1* 꼬인 나선 도메인이고, 세포질 도메인은 마우스 *Asgr1* 세포질 도메인이고, 막관통 도메인은 마우스 *Asgr1* 막관통 도메인이거나; 또는
- (c) 꼬인 나선 도메인 및 C형 렉틴 도메인 모두의 전부는 결실되고 이종상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화되고, C형 렉틴 도메인은 인간 *ASGR1* C형 렉틴 도메인이고, 꼬인 나선

도메인은 인간 ASGR1 꼬인 나선 도메인이며; 그리고 세포질 도메인 및 막관통 도메인 모두의 전부 또는 일부는 내인성 마우스 *Asgr1* 서열에 의해 암호화되고, 세포질 도메인은 마우스 *Asgr1* 세포질 도메인이고, 막관통 도메인은 마우스 *Asgr1* 막관통 도메인인 것을 특징으로 하는 설치류.

**청구항 12**

제11 항에 있어서,

(a) C형 렉틴 도메인은 서열 번호 28로 제시되는 서열을 포함하고, 꼬인 나선 도메인은 서열 번호 27로 제시되는 서열을 포함하고, 세포질 도메인은 서열 번호 29로 제시되는 서열을 포함하고, 막관통 도메인은 서열 번호 30으로 제시되는 서열을 포함하거나;

(b) 변형된 *Asgr1* 단백질은 서열 번호 3으로 제시되는 서열을 포함하거나; 또는

(c) C형 렉틴 도메인은 서열 번호 28로 제시되는 서열을 포함하고, 꼬인 나선 도메인은 서열 번호 27로 제시되는 서열을 포함하고, 세포질 도메인은 서열 번호 29로 제시되는 서열을 포함하고, 막관통 도메인은 서열 번호 30으로 제시되는 서열을 포함하며, 변형된 *Asgr1* 단백질은 서열 번호 3으로 제시되는 서열을 포함하는 것을 특징으로 하는 설치류.

**청구항 13**

인간-ASGR1-매개 내재화를 통해 치료 복합체의 간으로의 생체 내 전달을 평가하는 방법으로서,

(a) 제1 항 내지 제12 항 중 어느 한 항의 설치류에게 치료 복합체를 투여하는 단계로서, 치료 복합체는 치료 분자 및 인간 ASGR1과 특이적으로 결합하는 항원-결합 단백질 또는 리간드를 포함하는 단계; 및

(b) 설치류의 간으로의 치료 분자 전달을 평가하는 단계를 포함하며,

여기서 설치류는 랫트 또는 마우스인, 방법.

**청구항 14**

제13 항에 있어서,

(i) 치료 분자는 리소솜 대체 단백질 또는 효소이거나, 리소솜 대체 단백질 또는 효소를 암호화하는 핵산이고, 단계 (b)는 설치류의 간에서 리소솜 대체 단백질 또는 효소의 존재 또는 활성을 평가하는 단계를 포함하거나; 또는

(ii) 치료 분자는 치료 분비 단백질을 암호화하는 핵산이고, 단계 (b)는 설치류에서 치료 분비 단백질의 혈청 수준 또는 활성을 평가하는 단계를 포함하는, 방법.

**청구항 15**

인간 ASGR1을 통한 간으로의 내재화에 대해 간 세포 표면 단백질 또는 가용성 단백질을 표적화하는 치료 분자의 생체 내 효능을 평가하는 방법으로서,

(a) 제1 항 내지 제12 항 중 어느 한 항의 설치류에게 치료 분자를 투여하는 단계로서, 치료 분자는 간 세포 표면 단백질 또는 가용성 단백질에 특이적으로 결합하고 인간 ASGR1에 특이적으로 결합하는 이중특이적 항원-결합 단백질을 포함하는 단계; 및

(b) 설치류의 간에서 간 세포 표면 단백질의 세포 표면 수준 또는 활성을 평가하거나 설치류의 간에서 가용성 단백질의 발현 또는 활성을 평가하는 단계를 포함하며, 여기서 설치류는 랫트 또는 마우스인, 방법.

**청구항 16**

변형된 *Asgr1* 단백질을 암호화하는 유전적으로 변형된 내인성 *Asgr1* 유전자좌를 포함하는 설치류 세포로서,

변형된 *Asgr1* 단백질은 세포질 도메인, 막관통 도메인, 및 세포외 도메인을 포함하고,

세포외 도메인은 꼬인 나선 도메인 및 C형 렉틴 도메인을 포함하며,

(i) 꼬인 나선 도메인의 전부 또는 일부는 결실되고 이중상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자

좌의 분절에 의해 암호화되거나,

(ii) C형 렉틴 도메인의 전부 또는 일부는 결실되고 이종상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자 좌의 분절에 의해 암호화되거나,

(iii) 꼬인 나선 도메인의 전부 또는 일부 및 C형 렉틴 도메인의 전부는 결실되고 이종상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 각각 암호화되거나, 또는

(iv) 세포의 도메인의 전부 또는 일부는 결실되고 이종상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화되며,

여기서 설치류는 랫트 또는 마우스인 것을 특징으로 하는 설치류 세포.

**청구항 17**

제16 항에 있어서,

- (a) 설치류 세포는 간세포이거나;
- (b) 설치류 세포는 다능성 또는 전능성 세포이거나;
- (c) 설치류 세포는 ES 세포이거나; 또는
- (d) 설치류 세포는 생식세포인 것을 특징으로 하는 설치류 세포.

**청구항 18**

제16 항 또는 제17 항의 설치류 세포를 포함하는 조성물.

**청구항 19**

제1 항 내지 제12 항 중 어느 한 항의 설치류를 만드는 방법에서 사용하기 위한 핵산으로서,

- (a) 서열 번호 18, 서열 번호 19, 서열 번호 20, 서열 번호 23, 및 이들의 조합으로 이루어진 군으로부터 선택 되는 뉴클레오티드 서열;
- (b) 서열 번호 24로 제시되는 뉴클레오티드 서열;
- (c) 서열 번호 18, 서열 번호 19, 서열 번호 20, 서열 번호 23, 및 이들의 조합으로 이루어진 군으로부터 선택 되는 뉴클레오티드 서열로서, 서열 번호 3으로 제시되는 아미노산 서열을 암호화하는 것을 특징으로 하는 뉴클레오티드 서열; 또는
- (d) 서열 번호 24로 제시되는 서열로서, 서열 번호 3으로 제시되는 아미노산 서열을 암호화하는 것을 특징으로 하는 뉴클레오티드 서열이며, 여기서 설치류는 랫트 또는 마우스인, 핵산.

**청구항 20**

삭제

**청구항 21**

삭제

**청구항 22**

삭제

**청구항 23**

삭제

**청구항 24**

삭제

- 청구항 25
- 삭제
- 청구항 26
- 삭제
- 청구항 27
- 삭제
- 청구항 28
- 삭제
- 청구항 29
- 삭제
- 청구항 30
- 삭제
- 청구항 31
- 삭제
- 청구항 32
- 삭제
- 청구항 33
- 삭제
- 청구항 34
- 삭제
- 청구항 35
- 삭제
- 청구항 36
- 삭제
- 청구항 37
- 삭제
- 청구항 38
- 삭제
- 청구항 39
- 삭제
- 청구항 40
- 삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

**발명의 설명**

**배경 기술**

[0001] 서열 목록에 대한 참조

[0002] EFS 웹을 통해 텍스트 파일로 제출됨

[0003] 10362W001\_ST25.txt 파일에 기록된 서열 목록은 50킬로바이트이고, 2018년 6월 26일에 생성되었으며, 본원에 참조로서 통합된다.

[0004] 배경 기술

[0005] 많은 질병은 관련 조직이나 기관에 대해 치료 분자의 잘못된 표적화로 부적절하게 치료된다. 대상체에게 생물학적 활성제를 전달하는 것은 표적 세포 또는 조직에 성분을 도달시키는 것의 어려움으로 인해 종종 제약을 받는다.

[0006] 특이적 표적 세포에 대한 치료 분자의 생체 내 전달을 용이하게 하고 이를 증진시키기 위해 세포 유형 특이적 내재화 이펙터, 예컨대, 표적 세포에 의해 간접적으로 또는 직접적으로 내재화되는 인간 ASGR1과 같은 세포 유형 특이적 수용체 등이 사용될 수 있다. 유사하게, 이러한 세포 유형 특이적 내재화 이펙터는 생체 내에서 표적 세포 표면 수용체 또는 표적 가용성 단백질의 내재화를 용이하게 하기 위한 치료 목적에 적합할 수 있다. 그러나, 이러한 전달 메커니즘 및 치료 메커니즘의 생체 내 효능을 평가하기 위한 적절한 모델에 대한 필요성이 남아 있다.

**발명의 내용**

[0007] 인간화 *Asgr1* 유전자좌를 포함하고 인간화 *Asgr1* 유전자좌로부터 인간화 또는 키메라 ASGR1 단백질을 발현하는 비인간 동물, 뿐만 아니라 이러한 동물, 및 이러한 동물을 만드는데 유용한 뉴클레오티드(예를 들어, 표적화 벡터, 계놈 등)로부터 유도된 이러한 비인간 동물(예를 들어, 설치류, 예컨대, 랫트 또는 마우스), 세포 및 조직을 사용하는 방법이 제시되어 있다.

[0008] 일 양태에서, 변형된 *Asgr1* 단백질을 암호화하는 유전적으로 변형된 내인성 *Asgr1* 유전자좌를 포함하는 비인간 동물을 제공하되, 변형된 *Asgr1* 단백질은 세포질 도메인, 막관통 도메인, 및 세포외 도메인을 포함하고, 세포외 도메인의 전부 또는 일부가 결실되고 이종상동성 인간 ASGR1 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화된다. 일부 이러한 비인간 동물에서, 세포외 도메인은 꼬인 나선 도메인 및 C형 렉틴 도메인을 포함하며, C형 렉틴 도메인의 전부 또는 일부는 결실되고 이종상동성 인간 ASGR1 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화된다. 임의로, C형 렉틴 도메인은 인간 ASGR1 C형 렉틴 도메인이다. 임의로, C형 렉틴 도메인은 서열 번호 28로 제시되는 서열을 포함한다.

[0009] 일부 이러한 동물에서, 세포외 도메인은 꼬인 나선 도메인 및 C형 렉틴 도메인을 포함하며, 꼬인 나선 도메인의

전부 또는 일부는 결실되고 이중상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화된다. 임의로, *코인* 나선 도메인은 인간 *ASGR1* *코인* 나선 도메인이다. 임의로, *코인* 나선 도메인은 서열 번호 27로 제시되는 서열을 포함한다.

- [0010] 일부 이러한 동물에서, *코인* 나선 도메인 및 C형 렉틴 도메인 모두의 전부 또는 일부가 결실되고 이중상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화된다. 임의로, C형 렉틴 도메인은 인간 *ASGR1* C형 렉틴 도메인이고, *코인* 나선 도메인은 인간 *ASGR1* *코인* 나선 도메인이다. 임의로, C형 렉틴 도메인은 서열 번호 28로 제시되는 서열을 포함하고, *코인* 나선 도메인은 서열 번호 27로 제시되는 서열을 포함한다.
- [0011] 일부 이러한 동물에서, 이중상동성 인간 *ASGR1* 서열은 인간 *ASGR1* 유전자의 엑손 3~8을 포함한다. 임의로, 이중상동성 인간 *ASGR1* 서열은 서열 번호 31로 제시되는 서열을 포함하는 *ASGR1* 단백질 분절을 암호화한다.
- [0012] 일부 이러한 동물에서, 세포질 도메인의 전부 또는 일부는 내인성 비인간 동물 *Asgr1* 서열에 의해 암호화된다. 일부 이러한 동물에서, 막관통 도메인의 전부 또는 일부는 내인성 비인간 동물 *Asgr1* 서열에 의해 암호화된다. 임의로, 세포질 도메인 및 막관통 도메인 모두의 전부 또는 일부는 내인성 비인간 동물 *Asgr1* 서열에 의해 암호화된다.
- [0013] 또 다른 양태에서, 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물 세포 또는 조직(예를 들어, 비인간 간세포, 비인간 배아 줄기(ES) 세포 또는 생식 세포와 같은 다능성 세포), 및 이를 포함하는 시험관 내 조성물이 제공된다. 일 양태에서, 변형된 *Asgr1* 단백질을 암호화하는 유전적으로 변형된 내인성 *Asgr1* 유전자좌를 포함하는 비인간 동물 세포를 제공하되, 변형된 *Asgr1* 단백질은 세포질 도메인, 막관통 도메인, 및 세포외 도메인을 포함하고, 세포외 도메인의 전부 또는 일부가 결실되고 이중상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화된다. 일부 이러한 비인간 동물 세포에서, 세포외 도메인은 *코인* 나선 도메인 및 C형 렉틴 도메인을 포함하며, C형 렉틴 도메인의 전부 또는 일부는 결실되고 이중상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화된다. 임의로, C형 렉틴 도메인은 인간 *ASGR1* C형 렉틴 도메인이다. 임의로, C형 렉틴 도메인은 서열 번호 28로 제시되는 서열을 포함한다.
- [0014] 일부 이러한 비인간 동물 세포에서, 세포외 도메인은 *코인* 나선 도메인 및 C형 렉틴 도메인을 포함하며, *코인* 나선 도메인의 전부 또는 일부는 결실되고 이중상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화된다. 임의로, *코인* 나선 도메인은 인간 *ASGR1* *코인* 나선 도메인이다. 임의로, *코인* 나선 도메인은 서열 번호 27로 제시되는 서열을 포함한다.
- [0015] 일부 이러한 비인간 동물 세포에서, *코인* 나선 도메인 및 C형 렉틴 도메인 모두의 전부 또는 일부가 결실되고 이중상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화된다. 임의로, C형 렉틴 도메인은 인간 *ASGR1* C형 렉틴 도메인이고, *코인* 나선 도메인은 인간 *ASGR1* *코인* 나선 도메인이다. 임의로, C형 렉틴 도메인은 서열 번호 28로 제시되는 서열을 포함하고, *코인* 나선 도메인은 서열 번호 27로 제시되는 서열을 포함한다.
- [0016] 일부 이러한 비인간 동물 세포에서, 이중상동성 인간 *ASGR1* 서열은 인간 *ASGR1* 유전자의 엑손 3~8을 포함한다. 임의로, 이중상동성 인간 *ASGR1* 서열은 서열 번호 31로 제시되는 서열을 포함하는 *ASGR1* 단백질 분절을 암호화한다.
- [0017] 일부 이러한 비인간 동물 세포에서, 세포질 도메인의 전부 또는 일부는 내인성 비인간 동물 *Asgr1* 서열에 의해 암호화된다. 일부 이러한 동물 세포에서, 막관통 도메인의 전부 또는 일부는 내인성 비인간 동물 *Asgr1* 서열에 의해 암호화된다. 임의로, 세포질 도메인 및 막관통 도메인 모두의 전부 또는 일부는 내인성 비인간 동물 *Asgr1* 서열에 의해 암호화된다.
- [0018] 일부 이러한 동물 및/또는 동물 세포는 유전적으로 변형된 내인성 *Asgr1* 유전자좌에 대해서 이형접합체이다. 일부 이러한 동물 및/또는 동물 세포는 유전적으로 변형된 내인성 *Asgr1* 유전자좌에 대해서 동형접합체이다.
- [0019] 일부 이러한 비인간 동물 및/또는 동물 세포는 포유류 및/또는 포유류 세포이다. 임의로, 포유동물은 설치류이다. 임의로, 설치류는 랫트 또는 마우스이다. 임의로, 설치류는 마우스이다. 임의로, 세포질 도메인의 전부 또는 일부는 내인성 마우스 *Asgr1* 서열에 의해 암호화된다. 임의로, 세포질 도메인은 서열 번호 29로 제시되는 서열을 포함한다. 임의로, 막관통 도메인의 전부 또는 일부는 내인성 마우스 *Asgr1* 서열에 의해 암호화된다. 임의로, 막관통 도메인은 서열 번호 30으로 제시되는 서열을 포함한다. 임의로, 세포질 도메인 및 막관통 도메인 모두의 전부 또는 일부는 내인성 마우스 *Asgr1* 서열에 의해 암호화된다. 임의로, 세포질 도메인은 서열 번호 29로 제시되는 서열을 포함하고, 막관통 도메인은 서열 번호 30으로 제시되는 서열을 포함한다. 임의로, 세포외 도메

인은 *코인* 나선 도메인 및 C형 렉틴 도메인을 포함하되, *코인* 나선 도메인 및 C형 렉틴 도메인 모두의 전부 또는 일부가 결실되고 이중상동성 인간 *ASGR1* 서열로 대체된 내인성 *Asgr1* 유전자좌의 분절에 의해 암호화되고, 세포질 도메인 및 막관통 도메인의 전부 또는 일부가 내인성 마우스 *Asgr1* 서열에 의해 암호화된다. 임의로, C형 렉틴 도메인은 인간 *ASGR1* C형 렉틴 도메인이고, *코인* 나선 도메인은 인간 *ASGR1* *코인* 나선 도메인이고, 세포질 도메인은 마우스 *Asgr1* 세포질 도메인이고, 막관통 도메인은 마우스 *Asgr1* 막관통 도메인이다. 임의로, C형 렉틴 도메인은 서열 번호 28로 제시되는 서열을 포함하고, *코인* 나선 도메인은 서열 번호 27로 제시되는 서열을 포함하고, 세포질 도메인은 서열 번호 29로 제시되는 서열을 포함하고, 막관통 도메인은 서열 번호 30으로 제시되는 서열을 포함한다. 임의로, 변형된 *Asgr1* 단백질은 서열 번호 3으로 제시되는 서열을 포함한다.

[0020] 또 다른 양태에서, *생체 내* 인간-*ASGR1*-매개 내재화를 통해 간으로 치료 복합체의 전달을 평가하는 방법이 제공된다. 일부 이러한 방법은 (a) 상기 비인간 동물 중 임의의 동물에게 치료 복합체를 투여하는 단계를 포함하되, 치료 복합체는 인간 *ASGR1*에 특이적으로 결합하는 치료 분자 및 항원 결합 단백질 또는 리간드를 포함하고; (b) 비인간 동물의 간으로 치료 분자의 전달을 평가하는 단계를 포함한다. 임의로, 치료 분자는 리소솜 대체 단백질 또는 효소 또는 리소솜 대체 단백질 또는 효소를 암호화하는 핵산이고, 단계 (b)는 비인간 동물의 간에서 리소솜 대체 단백질 또는 효소의 존재 또는 활성을 평가하는 단계를 포함한다. 임의로, 치료 분자는 치료 분비 단백질을 암호화하는 핵산이고, 단계 (b)는 비인간 동물에서 치료 분비 단백질의 혈청 수준 또는 활성을 평가하는 단계를 포함한다.

[0021] 또 다른 양태에서, *생체 내* 인간 *ASGR1*을 통한 내재화를 위해 간 세포 표면 단백질 또는 간 내 가용성 단백질을 표적화하는 치료 분자의 효능을 평가하는 방법이 제공된다. 일부 이러한 방법은: (a) 상기 비인간 동물들 중 임의의 동물에 치료 분자를 투여하는 단계를 포함하되, 치료 분자는 간 세포 표면 단백질 또는 가용성 단백질에 특이적으로 결합하고 인간 *ASGR1*에 특이적으로 결합하는 이중 특이적 항원 결합 단백질을 포함하고; (b) 비인간 동물의 간에서 간 세포 표면 단백질의 세포 표면 수준 또는 활성을 평가하거나, 비인간 동물의 간 내 가용성 단백질의 발현 또는 활성을 평가하는 단계를 포함한다.

[0022] 또 다른 양태에서, 핵산(예를 들어 본원에 기술된 바와 같이 인간화 *Asgr1* 유전자좌), 인간화 또는 키메라 *ASGR1* 단백질을 발현하는 내인성 비인간 동물 *Asgr1* 유전자좌 및/또는 본원에 기술된 바와 같이 인간 *Asgr1* 유전자좌를 포함하는 비인간 동물 게놈이 제공된다. 일부 이러한 핵산은 서열 번호 18, 서열 번호 19, 서열 번호 20, 서열 번호 22, 서열 번호 23, 및 이들의 조합을 포함하는 군으로부터 선택되는 서열을 포함한다. 일부 이러한 핵산은 서열 번호 3으로 제시되는 아미노산 서열을 암호화하는 서열을 포함하는데, 예를 들어 서열 번호 24로 제시되는 서열을 포함한다.

**도면의 간단한 설명**

[0023] 특허 또는 출원 파일은 컬러로 작업된 적어도 하나의 도면을 포함한다. 컬러 도면이 있는 본 특허 또는 특허 출원 공개의 사본은 요청 시 필요한 수수료를 납부하면 특허청에서 제공될 것이다.

**도 1a**는 인간 *ASGR1* 유전자좌의 개략도를 도시한다. 8개의 엑손은 흑색 상자로 표시되고, 미번역 영역(UTR)은 백색 상자로 표시되고, 막관통, *코인* 나선, 및 C형 렉틴 도메인은 도면의 상단에 표시되어 있다. 별표는 대립유전자의 획득 검정을 위한 상류(7302hTU) 및 하류(7302hTD) 프라이머의 위치를 나타낸다. 인간화를 위한 마우스 *Asgr1* 유전자좌에 삽입될 단편이 하기에 도시되어 있다.

**도 1b**는 마우스 *Asgr1* 유전자좌의 개략도를 도시한다. 8개의 엑손은 흑색 상자로 표시되고, UTR은 백색 상자로 표시되고, 막관통, *코인* 나선, 및 C형 렉틴 도메인은 도면의 상단에 표시되어 있다. 별표는 대립유전자의 상실 검정을 위한 상류(7302mTU) 및 하류(7302mTD) 프라이머의 위치를 나타낸다. 삭제되고 인간 *ASGR1* 유전자좌로부터 상응하는 단편으로 교체될 단편이 하기에 도시되어 있다.

**도 2a**는 히그로마이신 저항성 자가 결실 카세트를 포함하는 인간화 *Asgr1* 대립유전자(7302 대립유전자)를 생성하기 위한 대형 표적화 벡터의 개략도를 도시한다. 마우스 엑손은 회색 상자로 표시되고, 인간 엑손은 검정색 박스로 표시되고, UTR은 흰색 박스로 표시되어 있다. 상이한 영역(마우스/인간, 인간/카세트, 및 카세트/마우스) 사이의 경계는 도면의 하단에서 각각, A, B, 및 C로 표시된 선으로 표시되어 있다.

**도 2b**는 도 2a에서 인간화 *Asgr1* 대립유전자의 카세트 결실 버전의 개략도를 도시한다. 마우스 엑손은 회색 상자로 표시되고, 인간 엑손은 검정색 박스로 표시되고, UTR은 흰색 박스로 표시되어 있다. 상이한 영역 사이의 경계(5' 마우스/인간 경계, 3' 인간/마우스 경계)는 도면의 하단에서 각각, A 및 D로 표시된 선으로 표시되어

있다.

**도 3**은 인간 ASGR1 단백질(hASGR1; 서열 번호 1), 마우스 *Asgr1* 단백질(mAsgr1; 서열 번호 2), 및 인간화 마우스 *Asgr1* 단백질(7302 humIn; 서열 번호 3)의 정렬을 도시한다. 밑줄 친 잔기는 도입된 인간 엑손에 의해 암호화된 잔기이다. 상자 안 잔기는 막관통 도메인을 구성한다. 점선은 C형 렉틴 도메인을 나타낸다. 두꺼운 실선은 꼬인 나선 영역을 나타낸다.

**도 4**는 인간화 *Asgr1* 마우스(*Asgr1*<sup>hu/hu</sup>)가 그들의 야생형(*Asgr1*<sup>+/+</sup>) 한배새끼(littermate)와 유사한 혈장 지질 프로파일(총 콜레스테롤, 트리글리세리드, LDL-C, 및 HDL-C를 포함함)을 갖는 것을 도시한다.

**도 5**는 인간화 *Asgr1* 마우스(*Asgr1*<sup>hu/hu</sup>)가 그들의 야생형 한배새끼와 유사한 체중 및 혈당을 갖는 것을 도시한다.

**도 6**은 인간화 *Asgr1* 마우스(*Asgr1*<sup>hu/hu</sup>)에서 인간화 *Asgr1* 단백질이 마우스 *Asgr1*과 유사하게 간 막에 공동 국재화하는 것을 도시한다. 트랜스페린 수용체(TRFR) 및 글리세르알데히드-3-인산 탈수소효소(GAPDH)를, 각각, 간의 막 및 세포핵 분획에 대한 로딩 대조군으로서 사용하였다. 그룹 당 4개씩(N=4) 도시되어 있다.

**도 7**은 인간 ASGR1 단백질(서열 번호 1), 필리핀 원숭이(마카카 파스키쿨라리스(*Macaca fascicularis*))(cyno)(서열 번호 39), 마우스 *Asgr1* 단백질(서열 번호 2), 및 랫트 *Asgr1* 단백질(서열 번호 41)의 정렬을 도시한다. 밑줄 친 잔기는 도입된 인간 엑손에 의해 암호화된 잔기이다. 인간 및 사이노 ASGR1 단백질 사이에 97.6% 서열 동일성(세포의 도메인에서는 98.3% 서열 동일성), 인간 및 마우스 ASGR1 단백질 사이에 77% 서열 동일성, 및 인간 및 랫트 ASGR1 단백질 사이에 78.4% 서열 동일성이 존재한다.

**도 8a-8c**는 간 세포가 인간 ASGR1을 발현하도록 유전자 이식에 의해 변형된 C57BL/6 마우스(**i-iv**)로부터 채취되거나,  $1 \times 10^{11}$ 개의 야생형 scAAV2-CMV-eGFP(**i, v**), 식염수(**ii, vi**),  $1 \times 10^{11}$ 개의 scAAV2-N587myc-CMV-eGFP 바이러스 벡터 단독(**iii, vii**), 또는 이중 특이적 항-myc-ASGR1 항체와 함께 scAAV2-N587myc-CMV-eGFP 바이러스 벡터(**iv, viii**)를 정맥 내 주입한 후 10일차에 야생형 C57BL/6 마우스(**v-viii**)로부터 채취된 간(A), 비장(B), 또는 신장(C) 시료의 면역 형광 현미경 이미지를 제공한다.

**도 9a-9r**은 간 세포가 인간 ASGR1을 발현하도록 유전자 이식에 의해 변형된 C57BL/6 마우스(**도 9d-9f, 9j-9l, 9p-9r**)로부터 채취되거나,  $2.18 \times 10^{11}$ 개의 야생형 ssAAV2-CAGG-eGFP(**도 9b, 9c, 9e, 9f**), 식염수(**도 9a, 9d**),  $2.18 \times 10^{11}$ 개의 ssAAV2-N587myc-CAGG-eGFP 바이러스 벡터 단독(**도 9g-9i, 9j-9l**), 또는 이중 특이적 항-myc-ASGR1 항체와 함께 ssAAV2-N587myc-CAGG-eGFP 바이러스 벡터(**도 9m-9o, 9p-9r**)를 정맥 내 주입한 후 4주차에 야생형 C57BL/6 마우스(**도 9a-9c, 9g-9i, 9m-9o**)로부터 채취된 간 시료의 면역 형광 현미경 이미지를 제공한다. 각각의 이미지는 하나의 마우스를 나타낸다.

**발명을 실시하기 위한 구체적인 내용**

- [0024] 정의
- [0025] 본원에서 상호호환적으로 사용되는, 용어 "단백질", "폴리펩티드" 및 "펩티드"는 암호화 및 비암호화된 아미노산 및 화학적으로 또는 생화학적으로 변형된 또는 유도된 아미노산을 포함하는, 임의 길이의 아미노산의 중합체 형태를 포함한다. 상기 용어는 또한 변형된 펩티드 골격을 갖는 폴리펩티드와 같은, 변형된 중합체를 포함한다. 용어 "도메인"은 특정 기능 또는 구조를 갖는 단백질 또는 폴리펩티드의 임의의 부분을 지칭한다.
- [0026] 단백질은 "N-말단" 및 "C-말단"을 갖는다고 언급된다. 용어 "N-말단"은 자유 아민기(-NH<sub>2</sub>)를 갖는 아미노산으로 종결된, 단백질 또는 폴리펩티드의 시작과 관련된다. 용어 "C-말단"은 자유 카르복실기(-COOH)를 갖는 아미노산으로 종결된, 아미노산 사슬(단백질 또는 폴리펩티드)의 말단과 관련된다.
- [0027] 본원에서 상호호환적으로 사용되는, 용어 "핵산" 및 "폴리뉴클레오티드"는 리보뉴클레오티드, 데옥시리보뉴클레오티드, 또는 유사체 또는 이의 변형된 버전을 포함하는, 임의 길이의 뉴클레오티드의 중합체 형태를 포함한다. 이들은 단일-, 이중 및 다중-가닥 DNA 또는 RNA, 게놈 DNA, cDNA, DNA-RNA 하이브리드, 및 퓨린 염기, 피리미딘 염기, 또는 다른 천연, 화학적으로 변형된, 생화학적으로 변형된, 비-천연 또는 유도된 뉴클레오티드 염기를 포함하는 중합체를 포함한다.
- [0028] 핵산은 하나의 모노뉴클레오티드 오탄당 고리의 5' 포스페이트가 포스포디에스테르 결합을 통해 한 방향으로 그

이웃의 3' 산소에 부착되는 방식으로 모노뉴클레오티드를 반응시켜 올리고뉴클레오티드를 만들기 때문에 핵산은 "5' 말단" 및 "3' 말단"을 갖는다고 언급된다. 올리고뉴클레오티드의 말단은 그의 5' 포스페이트가 모노뉴클레오티드 오탄당 고리의 3' 산소에 연결되어 있지 않은 경우 "5' 말단"으로 지칭된다. 올리고뉴클레오티드의 말단은 그의 3' 산소가 또 다른 모노뉴클레오티드 오탄당 고리의 5' 포스페이트에 연결되어 있지 않은 경우 "3' 말단"으로 지칭된다. 핵산 서열은, 더 큰 올리고뉴클레오티드의 내부에 있더라도, 또한 5' 및 3' 말단을 갖는 것으로 언급될 수 있다. 선형 또는 원형 DNA 분자에서, 불연속 요소는 "하류" 또는 3' 요소의 "상류" 또는 5'으로 지칭된다.

- [0029] 용어 "게놈적으로 통합된"은 뉴클레오티드 서열이 세포의 게놈 내로 통합되고 이의 자손에 의해 유전될 수 있도록 세포 내에 도입된 핵산을 지칭한다. 임의의 프로토콜은 세포의 게놈 내로의 핵산의 안정적인 혼입을 위해 사용될 수 있다.
- [0030] 용어 "표적화 벡터"는 상동성 재조합, 비-상동성-말단-결합-매개 연결, 또는 세포의 게놈에서 표적 위치로의 임의의 다른 재조합 수단에 의해 도입될 수 있는 재조합 핵산을 지칭한다.
- [0031] 용어 "바이러스 벡터"는 바이러스 기원의 적어도 하나의 요소를 포함하고, 바이러스 벡터 입자로 패키징하기에 충분하거나 패키징이 허용되는 요소를 포함하는 재조합 핵산을 지칭한다. 벡터 및/또는 입자는 생체 외 또는 생체 내에서 DNA, RNA, 또는 다른 핵산을 세포 내로 전달하기 위한 목적으로 사용될 수 있다. 수많은 형태의 바이러스 벡터가 알려져 있다.
- [0032] 용어 "야생형"은 정상(돌연변이, 질환, 변형 등과 대조됨) 상태 또는 맥락에서 발견되는 것과 동일한 구조 및/또는 활성을 가지는 엔티티를 포함한다. 야생형 유전자 및 폴리펩티드는 흔히 다수의 상이한 형태(예: 대립 유전자)로 존재한다.
- [0033] 용어 "내인성"은 세포 또는 비인간 동물 내에서 자연적으로 발생하는 핵산 서열을 지칭한다. 예를 들어, 비인간 동물의 내인성 *Asgr1* 서열은 비인간 동물에서의 *Asgr1* 유전자좌에서 자연적으로 발생하는 천연 *Asgr1* 서열을 지칭한다.
- [0034] "외인성" 분자 또는 서열은 그 형태의 세포에는 정상적으로 존재하지 않은 분자 또는 서열을 포함한다. 정상적인 존재는 세포의 특정한 발달 단계 및 환경 조건에 대한 존재를 포함한다. 외인성 분자 또는 서열은, 예를 들어, 내인성 서열의 인간화 버전과 같이, 세포 내에서 상응하는 내인성 서열의 돌연변이된 버전을 포함할 수 있거나, 세포 내의 내인성 서열에 상응하는 서열이지만, 상이한 형태(즉, 염색체 내에 있지 않음)로 포함할 수 있다. 대조적으로, 내인성 분자 또는 서열은 특정 환경 조건 하에서 특정 발달 단계에서 특정 세포에 그 형태로 정상적으로 존재하는 분자 또는 서열을 포함한다.
- [0035] 핵산 또는 단백질의 문맥에서 사용될 때, 용어 "이종"은 핵산 또는 단백질이 동일한 분자 내에서 자연적으로 함께 발생하지 않은 적어도 두 부분을 포함한다는 것을 나타낸다. 예를 들어, 용어 "이종"은, 핵산의 일부 또는 단백질의 일부와 관련하여 사용될 때, 핵산 또는 단백질이 천연에서 서로 동일한 관계(예를 들어, 함께 결합됨)로 발견되지 않은 두 개 이상의 하위 서열을 포함하는 것을 나타낸다. 일례로, 핵산 벡터의 "이종" 영역은 천연에서 다른 분자와 연관되어 발견되지 않은 또 다른 핵산 분자 내에 또는 그것에 부착된 핵산의 분절이다. 예를 들어, 핵산 벡터의 이종 영역은 천연에서 암호화 서열과 관련하여 발견되지 않은 서열이 측면에 위치한 암호화 서열을 포함할 수 있다. 마찬가지로, 단백질의 "이종" 영역은 천연에서 다른 펩티드 분자(예를 들어, 융합 단백질, 또는 태그를 갖는 단백질)와 관련하여 발견되지 않은 또 다른 펩티드 분자 내에 또는 그것에 부착된 아미노산의 분절이다. 유사하게, 핵산 또는 단백질은 이종 라벨 또는 이종 분비 또는 국소화 서열을 포함할 수 있다.
- [0036] "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 핵산은 자연적으로 발생하는 핵산 서열에 비교하여, 박테리아 세포, 효모 세포, 인간 세포, 비인간 세포, 포유류 세포, 설치류 세포, 마우스 세포, 랫트 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, "Codon Usage Database"에서 입수 가능하다. 이런 표들은 많은 방법으로 조정될 수 있다. Nakamura 외 (2000) *Nucleic Acids Research* 28:292를 참조하고, 그 전체는 모든 목적을 위해 참조로서 본원에 통합된다.

특정 숙주에서의 발현을 위해 특정 서열의 코돈 최적화를 위한 컴퓨터 알고리즘 또한 활용될 수 있다 (예컨대, Gene Forge 참조).

- [0037] 용어 "유전자좌"는 유전자(또는 유의 서열), DNA 서열, 폴리펩티드 암호화 서열의 특정한 위치, 또는 유기체 게놈의 염색체 상에서의 위치를 지칭한다. 예를 들어, "Asgr1 유전자좌"는 *Asgr1* 유전자, *Asgr1* DNA 서열, *Asgr1* 암호화 서열의 특정한 위치, 또는 이러한 서열이 어디에 상주하는지 확인된 유기체 게놈의 염색체 상에서의 *Asgr1* 위치를 지칭할 수 있다. "Asgr1 유전자좌"는, 예를 들어, 인핸서, 프로모터, 5' 및/또는 3' 미번역 영역(UTR), 또는 이들의 조합을 포함하는 *Asgr1* 유전자의 조절 요소를 포함할 수 있다.
- [0038] 용어 "유전자"는 생성물(예컨대, RNA 생성물 및/또는 폴리펩티드 생성물)을 암호화하고 비암호화 인트론이 끼어 있는 암호화 영역 및 5' 및 3' 양 단부에서 암호화 영역에 인접하여 위치한 서열을 포함하여서 유전자가 전체 길이 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, 절열 서열(insulating sequence), 및 매트릭스 부착 영역을 포함하는 다른 비암호화 서열을 포함한다. 이들 서열은 유전자의 암호화 영역에 가깝거나(예컨대, 10 kb 이내) 먼 부위에 있을 수 있고, 유전자의 전사 및 번역의 수준 또는 속도에 영향을 준다.
- [0039] 용어 "대립유전자"는 유전자의 변이체 형태를 나타낸다. 일부 유전자는 염색체 상의 동일한 위치, 또는 유전자좌에 위치하는 다양한 상이한 형태를 가진다. 이배체 유기체는 각각의 유전자좌에서 2개의 대립유전자를 가진다. 대립유전자의 각각의 쌍은 특이적 유전자좌의 유전자형을 나타낸다. 유전자형은 특정 유전자좌에 두 개의 동일한 대립유전자가 존재한다면 동형접합체로서 기술되고, 두 대립유전자가 상이하다면 이형접합체로서 기술된다.
- [0040] "프로모터"는 일반적으로 RNA 중합효소 II를 지시할 수 있는 TATA 박스를 포함하여 특정 폴리뉴클레오티드 서열에 대해 적절한 전사 개시 부위에서 RNA 합성을 개시할 수 있는 DNA의 조절 영역이다. 프로모터는 전사 개시 속도에 영향을 미치는 다른 영역을 추가적으로 포함할 수 있다. 본원에 개시된 프로모터 서열은 작동 가능하게 연결된 폴리뉴클레오티드의 전사를 조절한다. 프로모터는 본원에 개시된 세포 유형 중 하나 이상(예컨대, 진핵 세포, 비인간 포유류 세포, 인간 세포, 설치류 세포, 다능성 세포, 단일 세포 단계 배아, 분화 세포, 또는 이들의 조합)에서 활성화될 수 있다. 프로모터는, 예를 들어, 구성적으로 활성화된 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 예는, 예를 들어, 그 전체가 모든 목적을 위해 참조로서 본원에 통합되는 WO 2013/176772에서 찾을 수 있다.
- [0041] "작동가능한 연결" 또는 "작동가능하게 연결된"은 두 구성요소가 정상적으로 기능하고 구성요소들 중 적어도 하나가 다른 구성요소 중 적어도 하나에 가해지는 기능을 매개할 수 있는 가능성을 허용하도록 둘 이상의 구성요소(예컨대, 프로모터 및 또 다른 서열 요소)의 병치(juxtaposition)를 포함한다. 예를 들어, 프로모터는, 프로모터가 하나 이상의 전사 조절 인자의 존재 또는 부재에 대한 반응으로 암호화 서열의 전사 수준을 제어하는 경우, 암호화 서열에 작동 가능하게 연결될 수 있다. 작동 가능한 연결은 서로 연속적이거나 트랜스로 작용하는 이러한 서열을 포함할 수 있다(예컨대, 조절 서열은 암호화 서열의 전사를 조절하기 위한 거리에서 작용할 수 있다).
- [0042] 용어 "변이체"는 집단에서 가장 우세한 서열과 상이한 뉴클레오티드 서열(예컨대, 한 뉴클레오티드 차이) 또는 집단에서 가장 우세한 서열과 상이한 단백질 서열(예컨대, 한 아미노산 차이)을 나타낸다.
- [0043] 단백질을 지칭할 때, 용어 "단편"은 전체 길이 단백질보다 더 짧거나 더 적은 아미노산을 갖는 단백질을 의미한다. 핵산을 지칭할 때, 용어 "단편"은 전체 길이 핵산보다 더 짧거나 더 적은 뉴클레오티드를 갖는 핵산을 의미한다. 단편은, 예를 들어, N 말단 단편(즉, 단백질의 C 말단의 일부 제거), C 말단 단편(즉, 단백질의 N 말단의 일부 제거), 또는 내부 단편일 수 있다.
- [0044] 두 폴리뉴클레오티드 또는 폴리펩티드 서열의 맥락에서 "서열 동일성" 또는 "동일성"은 명시된 비교 창(window)에 걸쳐 최대 상응성을 위해 정렬되었을 때 동일한 두 서열에서의 잔기를 참조한다. 서열 동일성의 백분율이 단백질을 참조하여 사용될 때, 동일하지 않은 잔기 위치는 보존적 아미노산 치환에 의해 종종 상이하하며, 여기서 아미노산 잔기는 유사한 화학적 특성(예컨대, 전하 또는 소수성)을 가지는 다른 아미노산 잔기로 치환되며, 따라서 분자의 기능적 특성은 변화하지 않는다. 서열이 보존적 치환에서 상이할 때, 백분율 서열 동일성은 치환의 보존적 성질에 대해 보정하도록 상향으로 조정될 수 있다. 이러한 보존적 치환에 의해 상이한 서열은 "서열 유

사성" 또는 "유사성"을 가지는 것으로 언급될 수 있다. 이러한 조정을 만들기 위한 수단은 잘 알려져 있다. 전형적으로, 이는 보존적 치환을 완전한 미스매치(mismatch)보다는 부분적인 것으로서 스코어링(scoring)함으로써, 백분율 서열 동일성을 증가시킨다. 따라서, 예를 들어, 동일한 아미노산은 1의 점수로 주어지고 비보존적 치환은 0의 점수로 주어지고, 보존적 치환은 0과 1사이의 점수로 주어진다. 보존적 치환의 스코어링은, 예컨대 프로그램 PC/유전자(Intelligenetics, Mountain View, California)에서 구현되는 바와 같이 계산된다.

[0045] "서열 동일성의 백분율"은 비교 창에 걸쳐 두 개의 최적 정렬된 서열을 비교함으로써 측정된 값(완벽하게 매치된 잔기들의 가장 큰 수)을 포함하되, 비교 창에서 폴리뉴클레오티드 서열의 일부는 두 서열의 최적의 정렬을 위해 (첨가 또는 결실을 포함하지 않는) 참조 서열과 비교하여 첨가 또는 결실(즉, 갭)을 포함할 수 있다. 백분율은 두 서열에서 동일한 핵산 염기 또는 아미노산 잔기가 발생하는 위치의 수를 결정하고, 일치된 위치의 수를 비교 창에서 총 위치의 수로 나눈 후, 서열 동일성의 백분율을 산출하기 위해 그 결과에 100을 곱함으로써 산출된다. 달리 명시되지 않는 한(예를 들어, 더 짧은 서열은 연결된 이중성 서열을 포함함), 비교 창은 비교되는 두 서열 중 더 짧은 서열의 전체 길이이다.

[0046] 달리 언급하지 않은 한, 서열 동일성/유사성 값은 다음의 파라미터를 사용하여 GAP 버전 10을 사용하여 얻어진 값을 포함한다: 50의 GAP 무게 및 3의 길이 무게, 및 nwsgapdna.cmp 스코어링 매트릭스를 사용하는 뉴클레오티드 서열에 대한 % 동일성 및 % 유사성; 8의 GAP 무게 및 2의 길이 무게, 및 BLOSUM62 스코어링 매트릭스를 사용하는 아미노산 서열에 대한 % 동일성 및 % 유사성; 또는 이것들의 임의의 동등한 프로그램. "동등한 프로그램"은, 문제에서 임의의 두 서열에 대해, GAP 버전 10에 의해 생성된 상응하는 정렬과 비교할 때 동일한 뉴클레오티드 또는 아미노산 잔기 매치 및 동일한 백분율 서열 동일성을 가지는 정렬을 생성하는 임의의 서열 비교 프로그램을 포함한다.

[0047] 용어 "보존적 아미노산 치환"은 유사한 크기, 전하, 또는 극성의 상이한 아미노산을 가지는 서열에 정상적으로 존재하는 아미노산의 치환을 나타낸다. 보존적 치환의 예시는 이소류신, 발린, 또는 류신과 같은 비극성(소수성) 잔기로 또 다른 비극성 잔기의 치환을 포함한다. 마찬가지로, 보존적 치환의 예로는, 아르기닌과 리신 사이, 글루타민과 아스파라긴 사이, 또는 글리신과 세린 사이의 치환과 같이 한 극성(친수성) 잔기로 또 다른 극성 잔기의 치환을 들 수 있다. 추가적으로, 리신, 아르기닌, 또는 히스티딘과 같은 염기성 잔기로 또 다른 염기성 잔기의 치환, 또는 아스파르트산 또는 글루탐산과 같은 한 산성 잔기로 또 다른 산성 잔기의 치환은 보존적 치환의 추가적인 예시이다. 비보존적 치환의 예시는, 시스테인, 글루타민, 글루탐산 또는 리신과 같은 극성(친수성) 잔기를 이소류신, 발린, 류신, 알라닌, 또는 메티오닌과 같은 비극성(소수성) 아미노산 잔기로의 치환 및/또는 비극성 잔기를 극성 잔기로의 치환을 포함한다. 일반적인 아미노산 분류가 아래에 요약되어 있다.

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
이소류신	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

[0048]

[0049]

"상동성" 서열(예컨대, 핵산 서열)은 공지된 참조 서열과 동일하거나 실질적으로 동일한 참조 서열과 동일하거나 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, 이종상동성 서열(orthologous sequence) 및 동종상동성 서열(paralogous sequences)을 포함할 수 있다. 상동성 유전자는, 예를 들어, 통상적으로 공통 조상 DNA 서열로부터, 종분화 사건(이종상동성 유전자) 또는 유전자 중복 사건(동종상동성 유전자) 어느 한 쪽을 통해 내려온다. "이종상동성" 유전자는 종분화에 의해 공통 조상 유전자로부터 진화된 상이한 종에서의 유전자를 포함한다. 이종상동성은 전형적으로 진화 과정에서 동일한 기능을 유지한다. "동종상동성" 유전자는 유전자는 게놈 내에서의 중복에 의해 관련된 유전자를 포함한다. 직렬상동체는 진화 과정에서 새로운 기능을 진화시킬 수 있다.

[0050]

용어 "시험관 내"는 인공 환경 및 인공 환경(예컨대, 시험관) 내에서 발생하는 공정 또는 반응을 포함한다. 용어 "생체 내"는 천연 환경 및 천연 환경(예를 들어, 세포 또는 유기체 또는 신체) 내에서 발생하는 공정 또는 반응을 포함한다. 용어 "생체 외"는 개인의 신체로부터 제거된 세포 및 이러한 세포 내에서 발생하는 공정 또는 반응을 포함한다.

[0051]

용어 "리포터 유전자"는 이종 프로모터 및/또는 인핸서 요소에 작동 가능하게 연결된 리포터 유전자 서열을 포함하는 구조체가 프로모터 및/또는 인핸서 요소의 활성화를 위해 필요한 (또는 이를 함유하도록 제조될 수 있는) 인자를 포함하여 세포 내로 도입될 때, 용이하고 정량적으로 분석되는 유전자 산물(통상적으로 효소)을 암호화하는 서열을 갖는 핵산을 나타낸다. 리포터 유전자의 예시는 베타-갈락토시다아제(lacZ)를 암호화하는 유전자, 박테리아 클로람페니콜 아세틸트랜스퍼라아제(cat) 유전자, 반딧불이 루시페라아제 유전자, 베타-글루쿠로니다아제(GUS)를 암호화하는 유전자, 및 형광 단백질을 암호화하는 유전자를 포함하지만, 이에 한정되지 않는다. "리포터 단백질"은 리포터 유전자에 의해 암호화된 단백질을 나타낸다.

[0052]

본원에서 사용되는 용어 "형광 리포터 단백질"은 형광에 기초하여 검출될 수 있는 리포터 단백질을 의미하되, 형광은 리포터 단백질 직접적으로, 형광형성 기재 상의 리포터 단백질의 활성일 수 있거나, 형광 태그된 화합물에 대한 결합에 대한 친화성을 갖는 단백질을 의미한다. 형광 단백질의 예는 녹색 형광 단백질(예컨대, GFP, GFP-2, tagGFP, turboGFP, eGFP, 에메랄드(Emerald), 아자미 그린(Azami Green), 단량체성(Monomeric) 아자미 그린, CopGFP, AceGFP, ZsGreen1), 황색 형광 단백질(예컨대, YFP, eYFP, 시트린(Citrine), 비너스(Venus), YPet, PhiYFP, 및 ZsYellow1), 청색 형광 단백질(예컨대, BFP, eBFP, eBFP2, 아주라이트(Azurite), mKalamal,

GFPuv, 사파이어(Sapphire), 및 T-사파이어), 시안 형광 단백질(예컨대, CFP, eCFP, 세룰리안(Cerulean), CyPet, AmCyan1, 및 미도리시-시안(Midoriishi-Cyan)), 적색 형광 단백질(예컨대, RFP, mKate, mKate2, mPlum, DsRed 단량체, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-단량체, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, 및 Jred), 오렌지색 형광 단백질(예컨대, mOrange, mKO, 쿠사비라-오렌지(Kusabira-Orange), 단량체성 쿠사비라-오렌지, mTangerine, tdTomato), 및 세포 내에서 존재가 유동 세포 측정법에 의해 검출될 수 있는 임의의 다른 적합한 형광 단백질을 포함한다.

[0053] 용어 "재조합"은 두 폴리뉴클레오티드 사이의 유전 정보의 교환 중 임의의 공정을 포함하며, 임의의 메커니즘에 의해 발생할 수 있다. 이중 가닥 절단(DSG)에 대한 반응으로의 재조합은 주로 2개의 보존된 DNA 수선 경로를 통해 발생한다: 비상동성 말단융합(NHEJ) 및 상동재조합(HR) Kasperek & Humphrey (2011) *Seminars in Cell & Dev. Biol.* 22:886-897를 참조하고, 그 전체는 모든 목적을 위해 참조로서 본원에 통합된다. 마찬가지로, 외인성 공여 핵산에 의해 매개되는 표적 핵산의 수선은 두 폴리뉴클레오티드 사이의 유전 정보의 교환 중 임의의 공정을 포함할 수 있다.

[0054] NHEJ는 상동성 주형을 필요로 하지 않고 절단 말단을 서로 또는 외인성 서열에 직접 연결함으로써 핵산에서 이중-가닥 절단의 수선을 포함한다. NHEJ에 의한 비연속 서열의 연결은 종종 이중 가닥 절단부의 부근에 결실, 삽입, 또는 전좌를 초래할 수 있다. 예를 들어, NHEJ는 또한 외인성 공여 핵산의 말단과 절단 말단의 직접적인 연결(즉, NHEJ-기반된 포착)을 통해 외인성 공여 핵산의 표적화된 통합을 야기할 수 있다. 이러한 NHEJ-매개된 표적화된 통합은 상동성 지향된 수선(HDR) 경로가 (예를 들면, 비-분할 세포, 1차 세포, 및 상동성-기반된 DNA 수선을 저조하게 수행하는 세포에서) 쉽게 사용 가능하지 않은 경우 외인성 공여 핵산의 삽입에 바람직할 수 있다. 또한, 상동성-지향된 복구와 대조적으로, 절단 부위의 측부에 있는 서열 동일성의 큰 영역에 관한 지식은 필요없고, 이는 게놈 서열의 제한된 지식이 있는 게놈을 갖는 유기체에 표적화된 삽입을 시도하는 경우 유익할 수 있다. 통합은 외인성 공여 핵산과 절단된 게놈 서열 사이 평활 말단의 연결을 통해, 또는 절단된 게놈 서열에서 핵산 분해효소 시약에 의해 생성된 것과 양립가능한 돌출부에 의해 측부에 위치되는 외인성 공여 핵산을 사용하는(즉, 5' 또는 3' 돌출부를 갖는) 점성 말단의 연결을 통해 진행할 수 있다. 예를 들어, US 2011/020722, WO 2014/033644, WO 2014/089290, 및 Maresca 외 (2013) *Genome Res.* 23(3):539-546을 참조하고, 그 전체는 모든 목적을 위해 참조로서 본원에 통합된다. 평활 말단이 연결되면, 표적 및/또는 공여자 절단이 단편 융합에 필요한 미세상동성의 영역을 생성하기 위해 필요할 수 있으며, 이는 표적 서열에서 원치 않는 변형을 생성할 수 있다.

[0055] 재조합은 또한 상동성 지향된 수선(HDR) 또는 상동재조합(HR)을 통해 발생할 수도 있다. HDR 또는 HR은 뉴클레오티드 서열 상동성을 요구할 수 있는 핵산 수선의 형태를 포함하고, "표적" 분자 (즉, 이중-가닥 절단을 경험 하였던 것)의 수선용 주형으로서 "공여체" 분자를 사용하고, 공여체부터 표적으로의 유전 정보의 전달을 유발시킨다. 임의의 특정 이론에 구속되고자 하는 바람 없이, 상기 전달은 절단된 표적과 공여체 사이에 형성되는 이중이중나선 DNA의 미스매치 교정, 및/또는 합성-의존적 가닥 어닐링(공여체는 표적의 일부가 될 유전 정보를 합성하는데 사용됨) 및/또는 관련 공정을 포함할 수 있다. 일부 경우에서, 공여체 폴리뉴클레오티드, 공여체 폴리뉴클레오티드의 일부, 공여체 폴리뉴클레오티드의 사본, 또는 공여체 폴리뉴클레오티드의 사본의 일부는 표적 DNA에 통합한다. Wang 외 (2013) *Cell* 153:910-918; Mandalos 외 (2012) *PLOS ONE* 7:e45768:1-9; and Wang 외 (2013) *Nat Biotechnol.* 31:530-532를 참조하고, 그 전체는 모든 목적을 위해 참조로서 본원에 통합된다.

[0056] 용어 "항원-결합 단백질"은 항원에 결합하는 임의의 단백질을 포함한다. 항원-결합 단백질의 예시로는 항체, 항체의 항원-결합 단편, 다중특이적 항체(예컨대, 이중-특이적 항체), scFV, 비스-scFV, 디아바디, 트리아바디, 테트라바디, V-NAR, VHH, VL, F(ab), F(ab)<sub>2</sub>, DVD(이중 가변 도메인 항원-결합 단백질), SVD(단일 가변 도메인 항원-결합 단백질), 이중특이적 T-세포 인게이저(BiTE), 또는 다비스바디(미국 특허 번호 8,586,713, 모든 목적을 위하여 그 전체가 본 명세서에서 참고로 본 명세서에서 편입됨)를 포함한다.

[0057] 항원-결합 단백질과 관련하여, 용어 "다중-특이적" 또는 "이중-특이적"은 동일한 항원 또는 상이한 항원 상의 다른 에피토프를 단백질이 인식하는 것을 의미한다. 다중-특이적 항원-결합 단백질은 단일한 다기능성 폴리펩티드일 수 있거나, 상호 간에 공유적으로 또는 비 공유적으로 결합되는 2개 이상의 폴리펩티드의 다량체 복합체일 수 있다. 예를 들어, 항체 또는 그의 단편은 (예, 화학적 커플링, 유전 융합, 비공유 결합 또는 그 외에 의하여) 단백질 또는 그의 단편과 같은, 하나 이상의 다른 분자 엔티티와 기능적으로 연결되어서 제2 결합 특이성을 갖는 이중특이적 또는 다중특이적 항원-결합 분자를 생산할 수 있다.

[0058] 용어 "항원"은, 그 기체에 결합 특이성을 가진 항체의 생산을 이끌어낼 수 있는, 전체 분자 또는 분자 이내 도

메인이든, 기제를 지칭한다. 용어 항원은 또한, 야생형 숙주 유기체가 자가-인식의 덕택으로 항체 생산을 유도하지 못하지만, 면역학적 내성을 파괴시키기 위해 적절한 유전적 공학기술로 숙주 동물에서 그와 같은 반응을 유도할 수 있는, 기제를 포함한다.

[0059] 용어 "에피토프"는 항원-결합 단백질 (예컨대, 항체)가 결합하는 항원에서 부위를 지칭한다. 에피토프는 하나 이상의 단백질의 3차 접힘에 의해 병치된 연속 아미노산 또는 비인접 아미노산으로부터 형성될 수 있다. 연속하는 아미노산으로부터 형성된 (또한 선형 에피토프로서 공지된) 에피토프는 통상적으로 변성 용매에 노출시 유지되는 반면, 3차 접힘에 의해 형성된 (또한 형태적 에피토프로서 공지된) 에피토프는 통상적으로 변성 용매로 처리시 손실된다. 에피토프는 특유의 공간적 형태에서 전형적으로 적어도 3개, 및 더 일반적으로, 적어도 5개 또는 8-10개 아미노산을 포함한다. 에피토프의 공간적 형태의 결정 방법은, 예를 들어, x-선 결정학 및 2-차원 핵자기 공명을 포함한다. 예를 들어, Epitope Mapping Protocols, in Methods in Molecular Biology, Vol. 66, Glenn E. Morris, Ed. (1996)를 참조하고, 그 전체는 모든 목적을 위해 참조로서 본원에 통합된다.

[0060] 본원에 기술된 바와 같은 항체 파라토프는 일반적으로 이종 에피토프를 특이적으로 인식하는 상보성 결정 영역(CDR)(예를 들어, 중쇄 및/또는 경쇄 가변 도메인의 CDR3 영역)을 최소한으로 포함한다.

[0061] 용어 "항체"는 4개의 폴리펩티드 사슬, 즉 이황화 결합에 의해 상호 연결된 2개의 중(H)쇄 및 2개의 경(L)쇄를 포함하는 면역글로불린 분자를 포함한다. 각각의 중쇄는 중쇄 가변 도메인 및 중쇄 불변 영역(C<sub>H</sub>)을 포함한다. 중쇄 불변 영역은 세 개의 도메인을 포함한다: C<sub>H1</sub>, C<sub>H2</sub> 및 C<sub>H3</sub>. 각각의 경쇄는 경쇄 가변 도메인 및 경쇄 불변 영역(C<sub>L</sub>)을 포함한다. 중쇄 및 경쇄 가변 도메인은 프레임워크 영역(FR)으로 불리는 보다 보존적인 영역이 사이 사이에 배치되는 상보성 결정 영역(CDR)으로 불리는 추가변성 영역으로 다시 나뉘질 수 있다. 각각의 중쇄 및 경쇄 가변 도메인은 아미노 말단으로부터 카르복시 말단 방향으로 다음의 순서로 배열된 3개의 CDR과 4개의 FR을 포함한다: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4 (중쇄 CDR은 HCDR1, HCDR2 및 HCDR3로서 약칭될 수 있고; 경쇄 CDR는 LCDR1, LCDR2 및 LCDR3로서 약칭될 수 있음). 용어 "높은 친화도" 항체는 이의 표적 에피토프에 대해 약 10<sup>-9</sup> M 또는 그 이하(예: 1X10<sup>-9</sup> M, 1X10<sup>-10</sup> M, 1X10<sup>-11</sup> M, 또는 약 1X10<sup>-12</sup> M)의 K<sub>D</sub>를 갖는 항체를 지칭한다. 일 구현예에서, K<sub>D</sub>는 표면 플라즈몬 공명, 예를 들어, BIACORE제에 의해 측정되고; 다른 구현예에서, K<sub>D</sub>는 ELISA에 의해 측정된다.

[0062] 용어 "이중특이적 항체"는 2개 이상의 에피토프에 선택적으로 결합할 수 있는 항체를 포함한다. 이중특이적 항체는 일반적으로 상이한 2개의 중쇄를 포함하며, 각각의 중쇄는 2개의 상이한 분자 (예, 2개의 상이한 항원 상의) 또는 (예를 들어, 동일한 항원 상의) 동일한 분자 상에서 상이한 에피토프에 특이적으로 결합한다. 이중특이적 항체가 2개의 상이한 에피토프 (제1 에피토프 및 제2 에피토프)를 선택적으로 결합할 수 있다면, 제1 에피토프에 대한 제1 중쇄의 친화성은 일반적으로 제2 에피토프에 대한 제1 중쇄의 친화성보다 적어도 1에서 2 또는 3 또는 4배 낮은 세기 순, 및 그 반대일 것이다. 이중특이적 항체에 의해 인식되는 에피토프는 동일하거나 상이한 표적(예, 동일하거나 상이한 단백질)에 존재할 수 있다. 예를 들어, 동일한 항원의 상이한 에피토프를 인식하는 중쇄들을 조합함으로써, 이중특이적 항체가 생성될 수 있다. 예를 들어, 동일한 항원의 상이한 에피토프를 인식하는 중쇄 가변 서열을 암호화하는 핵산 서열은 상이한 중쇄 불변 영역을 암호화하는 핵산 서열에 융합될 수 있으며, 이러한 서열들은 면역글로불린 경쇄를 발현하는 세포에서 발현될 수 있다. 전형적인 이중특이적 항체는 각각 3개의 중쇄 CDR, 그 다음에 (N-말단에서 C-말단으로) CH1 도메인, 힌지, CH2 도메인 및 CH3 도메인이 뒤따르는, 2개의 중쇄, 및 항체-결합 특이성을 부여하지는 않지만 각각의 중쇄와 결합할 수 있는 것, 또는 각각의 중쇄와 결합할 수 있고 중쇄 항원-결합 영역에 의해 결합된 하나 이상의 에피토프를 결합할 수 있는 것, 또는 각각의 중쇄와 결합할 수 있고 중쇄의 하나 또는 둘 모두를 하나 또는 둘 모두의 에피토프에 결합시킬 수 있는 것인, 면역글로불린 경쇄를 갖는다.

[0063] 용어 "중쇄" 또는 "면역글로불린 중쇄"는 임의의 유기체의 면역글로불린 중쇄 불변 영역 서열을 포함하여, 면역글로불린 중쇄 서열을 포함한다. 중쇄 가변 도메인은 달리 명시되지 않는 한, 3개의 중쇄 CDR 및 4개의 FR 영역을 포함한다. 중쇄의 단편에는 CDR, CDR 및 FR 및 이들의 조합이 포함된다. 일반적인 중쇄는, (N 말단에서 C 말단까지) 가변 도메인에 이어서, C<sub>H1</sub> 도메인, 힌지, C<sub>H2</sub> 도메인 및 C<sub>H3</sub> 도메인을 갖는다. 중쇄의 기능성 단편은 에피토프를 특이적으로 인식할 수 있고 (예, 마이크로몰, 나노몰 또는 피코몰 범위의 K<sub>D</sub>로 에피토프를 인식함), 세포로부터 발현 및 분비가 가능하고, 적어도 하나의 CDR을 포함하는, 단편을 포함한다. 중쇄 가변 도메인은, 생식선에 존재하는 V<sub>H</sub>, D<sub>H</sub>, 및 J<sub>H</sub> 분절의 레파토리로부터 유래된 V<sub>H</sub>, D<sub>H</sub>, 및 J<sub>H</sub> 분절을 일반적으로 포함하는 가변

영역 유전자 서열에 의해 암호화된다. 다양한 유기체에 대한 V, D 및 J 중쇄 분절에 대한 서열, 위치 및 명명법은 IMGT 데이터베이스에서 찾을 수 있으며, 인터넷을 통해 "www.imgt.org" URL에 접속할 수 있다.

- [0064] 용어 "경쇄"는 임의의 유기체의 면역글로불린 경쇄 서열을 포함하며, 달리 명시되지 않은 한 인간 카파( $\kappa$ ) 및 람다( $\lambda$ ) 경쇄 및 VpreB, 및 대용(surrogate) 경쇄를 포함한다. 경쇄 가변 (VL) 도메인은 달리 명시되지 않은 한, 일반적으로 3개의 경쇄 CDR 및 4개의 프레임워크(FR) 영역을 포함한다. 일반적으로, 전장 경쇄는 아미노 말단에서 카르복실 말단까지, FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4를 포함하는 가변 도메인, 및 경쇄 불변 영역 아미노산 서열을 포함한다. 경쇄 가변 도메인은, 생식선에 존재하는 V 및 J 유전자 분절의 레파토리로부터 유래된 경쇄  $V_L$  및  $J_L$  유전자 분절을 일반적으로 포함하는, 경쇄 가변 영역 뉴클레오티드 서열에 의해 암호화된다. 다양한 유기체에 대한 V 및 J 유전자 분절에 대한 서열, 위치 및 명명법은 IMGT 데이터베이스에서 찾을 수 있으며, 인터넷을 통해 "www.imgt.org" URL에 접속할 수 있다. 경쇄는 예를 들어 이들이 나타나는 에피토프 결합 단백질에 의해 선택적으로 결합된 제1 에피토프 또는 제2 에피토프 중 어느 하나에 선택적으로 결합하지 않는 것을 포함한다. 경쇄는 또한 이들이 나타나는 에피토프 결합 단백질에 의해 선택적으로 결합된 1개 이상의 에피토프에 결합하고 인식하거나, 중쇄가 이것과 결합하고 인식하는 것을 돕는 것을 포함한다.
- [0065] 본원에서 사용된 바와 같이 용어 "상보성 결정 영역", 또는 "CDR"은 보통 (즉, 야생형 동물에서) 면역글로불린 분자 (예, 항체 또는 T 세포 수용체)의 경쇄 또는 중쇄의 가변 영역의 2개의 프레임워크 영역 사이에 나타나는, 유기체의 면역글로불린 유전자의 핵산 서열에 의해 암호화된 아미노산 서열을 포함한다. CDR은 예를 들어, 생식 계열 서열 또는 재배열 서열에 의해 암호화될 수 있으며, 예를 들어, 미처리 B 세포 또는 성숙한 B 세포 또는 T 세포에 의해 암호화될 수 있다. CDR은 체세포 돌연변이 (예를 들어, 동물의 생식선에서 암호화되는 서열에 따라 달라짐), 인간화, 및/또는 아미노산 치환, 부가, 또는 결실로 변형될 수 있다. 일부 경우에(예: CDR3의 경우), CDR은, (예컨대, 재배열되지 않은 핵산 서열에서는) 서로 인접하지 않지만 B 세포 핵산 서열 내에서는 인접하는 2개 이상의 서열(예: 생식계열 서열)에 의해, 예를 들어 서열의 접합 또는 연결(예를 들어, 중쇄 CDR3을 형성하기 위한 V-D-J 재조합)의 결과로서 암호화될 수 있다.
- [0066] 항원-결합 단백질의 그 표적 항원에 대한 특이적 결합은 적어도  $10^6$ ,  $10^7$ ,  $10^8$ ,  $10^9$  또는  $10^{10} M^{-1}$ 의 친화력을 갖는 결합을 포함한다. 특이적 결합은 크기가 검출 가능하고, 적어도 하나의 관련되지 않은 표적에 발생하는 비특이적 결합과 구별될 수 있다. 특이적 결합은 특정 작용기 또는 특정 공간적 끼워맞춤(예를 들어, 잠금 및 키 유형) 사이의 결합 형성의 결과일 수 있는 반면에, 비특이적 결합은 일반적으로 반데르발스 힘의 결과이다. 그러나, 특이적 결합은 항원-결합 단백질이 하나 및 단 하나의 표적에 결합하는 것을 반드시 의미하지는 않는다.
- [0067] 하나 이상의 인용된 요소를 "포함하는" 또는 "포괄하는" 조성물 또는 방법은 특이적으로 인용되지 않은 다른 요소를 포함할 수 있다. 예를 들어, 단백질을 "포함하는" 또는 "포괄하는" 조성물은 단백질을 단독으로 또는 다른 성분과 조합으로 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. 따라서, 용어 "본질적으로 ~으로 구성되는"은 본 발명의 청구범위에서 사용될 때 "포함하는"과 동등한 것으로 해석되는 것으로 의도된 것이 아니다.
- [0068] "임의의" 또는 "임의로"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.
- [0069] 값의 범위의 지정은 그 범위 이내 또는 상기를 한정하는 모든 정수, 그리고 그 범위 이내 정수에 의해 정의된 모든 하위범위를 포함한다.
- [0070] 문맥으로부터 달리 분명하지 않는 한, 용어 "약"은 언급된 값의 측정의 표준 오차 한계 (예를 들면, SEM) 이내 값을 포괄한다.
- [0071] 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.
- [0072] 용어 "또는"은 특정 목록의 임의의 한 구성원을 나타내며 또한 그 목록의 구성원들의 임의의 조합을 포함한다.
- [0073] 본문 중 단수 형태("a", "an", 및 "the")는 문맥상 명시적으로 달리 언급하지 않는 한 복수의 참조를 포함한다. 예를 들어, 용어 "단백질" 또는 "적어도 하나의 단백질"은, 이의 혼합물을 포함하는, 복수의 단백질을 포함할 수 있다

[0074] 통계적으로 유의한 의미는  $p \leq 0.05$ 이다.

[0075] **발명의 상세한 설명**

[0076] **I. 개요**

[0077] 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물 세포와 비인간 동물 및 이러한 비인간 동물 세포와 비인간 동물을 사용하는 방법이 본원에 개시된다. 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물 세포 또는 비인간 동물은 인간 ASGR1 단백질 또는 인간 ASGR1 단백질의 하나 이상의 단편을 포함하는 키메라 *Asgr1* 단백질(예컨대, 인간 ASGR1 세포의 도메인의 전부 또는 일부)을 발현한다. 고효능 엔도시토시스 간 특이적 수용체로서, 인간 ASGR1은 항체, 작은 분자(항체-약물 접합체의 일부로서) 및 DNA와 같은 치료제의 간 특이적 전달을 위해 사용될 수 있다. 그러나, 인간 ASGR1과 특이적으로 결합하는 항원-결합 단백질 또는 이중특이적 항원-결합 단백질은 인간 ASGR1과 비인간 동물 *Asgr1* 사이의 서열 차이로 인해 종종 마우스 *Asgr1*과 같은 이중상동성 비인간 동물 *Asgr1* 단백질에 결합하지 않을 것이다. 예를 들어, 인간 ASGR1 세포의 도메인에 대해 생성된 항체는 마우스 *Asgr1* 동원체에 결합하지 않는다(데이터는 도시되지 않음). 이로 인해 인간 ASGR1 매개 전달 메커니즘 또는 치료 메커니즘의 생체 내 효능은 변형되지 않는 내인성(즉, 천연) *Asgr1* 유전자좌들을 가진 야생형 비인간 동물에서 효율적으로 평가될 수 없다. 인간화 *Asgr1* 비인간 동물(예컨대, 인간화 *Asgr1* 마우스)은 다수의 상이한 접근법을 이용하여 인간 ASGR1 매개 내재화를 통해 상이한 치료제의 간 특이적 전달을 검증하기 위해 사용될 수 있다. 예를 들어, 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물은 간으로 치료 분자 또는 치료 복합체의 인간 ASGR1 매개 전달의 생체 내 효능을 평가하기 위해 사용될 수 있다. 유사하게, 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물은 인간 ASGR1 매개 메커니즘을 통해 작용하는 치료 분자 또는 치료 복합체의 효능을 평가하기 위해 사용될 수 있다.

[0078] **II. 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물**

[0079] 본원에 개시된 세포 및 비인간 동물은 인간화 *Asgr1* 유전자좌를 포함한다. 인간화 *Asgr1* 유전자좌를 포함하는 세포 또는 비인간 동물은 인간 ASGR1 단백질 또는 천연 *Asgr1* 단백질의 하나 이상의 단편(예컨대, 세포의 도메인의 전부 또는 일부)이 인간 ASGR1으로부터 유래된 상응하는 단편으로 대체된 부분적으로 인간화된 키메라 *Asgr1* 단백질을 발현한다.

[0080] **A. 아시알로당단백질 수용체 1(ASGR1)**

[0081] 본원에 기술된 세포 및 비인간 동물은 인간화 *Asgr1* 유전자좌를 포함한다. 아시알로당단백질 수용체 1(C형 렉틴 도메인 패밀리에 4 구성원 H1, 간 렉틴 H1, HL-1, ASGP-R 1, ASGPR 1, ASGR1)은 *ASGR1 (CLECAH1)* 유전자에 의해 암호화되고, 아시알로당단백질 수용체(ASGPR 또는 ASGR)의 주요 서브유닛이다. ASGPR은 간 세포의 세포 표면에서 주로 발현되는 이중 올리고머 단백질이며, 대략  $1 \sim 5 \times 10^5$ 개의 결합 부위/세포를 가지고, 탈시알틸화된 당단백질을 내재화하고 분해하여 순환에서 이들을 제거하는 역할을 한다. ASGPR은 간세포의 골막관 표면에서 주로 발현되는 잘 특성화된 C형 간 렉틴이다. ASGPR은 수용체 매개 세포 내 유입을 통해 간 실질 세포에 의한 갈락토오스-말단 및 N-아세틸갈락토사민-말단의 선택적 결합 및 내재화를 담당한다. 유전자 *ASGR1* 및 *ASGR2*에 의해 암호화된, 2개의 단백질, 아시알로당단백질 수용체 1 및 2(ASGR1 및 ASGR2)를 함한다. 두 서브유닛은 모두 II형이고, N-말단 세포질 도메인, 단일 막관통 도메인, 및 C-말단 세포외 탄수화물 인식 도메인(CRD)을 광범위하게 포함하는 단일 통과 단백질이다. ASGR1은 N-말단 세포질 도메인(~40개 아미노산), 단일-통과 막관통 도메인(~20개 아미노산), 세포외 꼬인 나선 줄기(올리고머화) 영역(~80개 아미노산), 및 기능적 C형(칼슘-의존성) 탄수화물 인식 도메인(C형 렉틴 도메인)(~140개 아미노산)을 함유한다. CRD는 말단 갈락토오스 또는 N-아세틸갈락토사민(GalNac) 모티프를 갖는 당단백질과 결합한다. CRD는 단량체 상태에서 탈시알틸화 당단백질에 대해 낮은 친화성을 갖는다.

[0082] ASGR1 및 ASGR2를 암호화하는 유전자(각각 *ASGR1* 및 *ASGR2*)는 염색체 17의 짧은 아암(arm) 상에서, 대략 58.6 킬로베이스(kb) 떨어져 위치한다. 유전자는 진화적으로 관련되어 있지만 그들의 구조 조직에서는 상당히 상이하다: *ASGR1*은 8개의 엑손을 포함하며, 대략 6 kb 길이이고, *ASGR2*는 9개의 엑손을 함유하며, 대략 13.5 kb 길이이다.

[0083] 인간 *ASGR1*에 대한 서열을 암호화하는 예시에는 NCBI 수탁번호 NM\_001671(서열 번호 5)이 할당된다. 미우스 *Asgr1*에 대한 서열을 암호화하는 예시에는 NCBI 수탁번호 NM\_009714(서열 번호 4)가 할당된다. 예시적인 인간 ASGR1 단백질에는 UniProt 수탁 번호 P07306(서열 번호 1)이 할당된다. 예시적인 마우스 *Asgr1* 단백질에는 UniProt 수탁 번호 P34927(서열 번호 2)이 할당된다. 인간화 꼬인 나선 및 C형 렉틴 도메인을 가지는 예시적인

마우스 *Asgr1* 단백질은 서열 번호 3으로 제시된다. 예시적인 랫트 *Asgr1* 단백질에는 UniProt 수탁 번호 P02706 이 할당된다. 예시적인 오랑우탄 *Asgr1* 단백질에는 UniProt 수탁 번호 Q5RBQ8이 할당된다.

[0084] **B. 인간화 *Asgr1* 유전자좌들**

[0085] 인간화 *Asgr1* 유전자좌는 상응하는 이종상동성 인간 *ASGR1* 서열로 대체된 전체 *Asgr1* 유전자가 있는 *Asgr1* 유전자좌가 될 수 있거나, 상응하는 이종상동성 인간 *ASGR1* 서열(즉, 인간화)로 대체된 *Asgr1* 유전자의 일부만 있는 *Asgr1* 유전자좌가 될 수 있다. 임의로, 상응하는 이종상동성 인간 *ASGR1* 서열은 비인간 동물에서 코돈 사용에 기초하여 코돈 최적화된 것으로 변형된다. 대체된(즉, 인간화된) 영역은 엑손과 같은 암호화 영역, 인트론, 비번역 영역, 또는 조절 영역(예컨대, 프로모터, 인핸서, 또는 전사 억제제-결합 요소), 또는 이들의 임의의 조합과 같은 비암호화 영역을 포함할 수 있다. 일 실시예로서, 인간 *ASGR1* 유전자의 1, 2, 3, 4, 5, 6, 7 또는 8개의 모든 엑손에 상응하는 엑손은 인간화될 수 있다. 예를 들어, 인간 *ASGR1* 유전자의 엑손 3-8에 상응하는 엑손은 인간화될 수 있다. 대안적으로, 항-인간-*ASGR1* 항원-결합 단백질에 의해 인식되는 에피토프를 암호화하는 *Asgr1*의 영역은 인간화될 수 있다. 또 다른 실시예로서, N-말단 세포질 도메인, 막관통 도메인, 꼬인 나선 도메인, 또는 C형 렉틴 도메인 중 하나 이상 또는 전부가 인간화될 수 있다. 예를 들어, 꼬인 나선 도메인을 암호화하는 *Asgr1* 유전자좌의 영역 중 전부 또는 일부가 인간화될 수 있고, C형 렉틴 도메인을 암호화하는 *Asgr1* 유전자좌의 영역 중 전부 또는 일부가 인간화될 수 있고/있거나, 세포질 도메인을 암호화하는 *Asgr1* 유전자좌의 영역 중 전부 또는 일부가 인간화될 수 있다. 일 실시예로, 꼬인 나선 도메인을 암호화하는 *Asgr1* 유전자좌의 영역 중 전부 또는 일부만이 인간화될 수 있고, C형 렉틴 도메인을 암호화하는 *Asgr1* 유전자좌의 영역 중 전부 또는 일부만이 인간화될 수 있거나, 세포외 영역(즉, 꼬인 나선 도메인 및 C형 렉틴 도메인)을 암호화하는 *Asgr1* 유전자좌의 영역 중 전부 또는 일부만. 예를 들어, 꼬인 나선 도메인 및 C형 렉틴 도메인을 암호화하는 *Asgr1* 유전자좌의 영역은 키메라 *Asgr1* 단백질이 내인성 N-말단 세포질 도메인, 내인성 막관통 도메인, 인간화 꼬인 나선 도메인, 및 인간화 C형 렉틴 도메인과 함께 생성되도록 인간화될 수 있다. 마찬가지로, 인간 *ASGR1* 유전자의 1, 2, 3, 4, 5, 6, 또는 7개의 모든 인트론에 상응하는 인트론은 인간화될 수 있다. 조절 서열을 포함하는 5' 미번역 영역도 인간화될 수 있다. 예를 들어, 5' 미번역 영역(UTR), 3' UTR, 또는 5' UTR 및 3' UTR 둘 모두는 인간화될 수 있거나 5' UTR, 3' UTR, 또는 5' UTR 및 3' UTR 둘 모두가 내인성 상태로 남을 수 있다. 구체적인 일 실시예에서, 3' UTR은 인간화되지만, 5' UTR은 내인성 상태로 남아있다. 이종상동성 서열에 의한 대체의 정도에 따라, 프로모터와 같은 조절 서열은 내인성 또는 대체용 인간 이종상동성 서열에 의해 공급될 수 있다. 예를 들어, 인간화 *Asgr1* 유전자좌는 내인성 비인간 동물 *Asgr1* 프로모터를 포함할 수 있다.

[0086] 인간화 *Asgr1* 유전자좌에 의해 암호화되는 *Asgr1* 단백질은 인간 *ASGR1* 단백질로부터 유래된 하나 이상의 도메인을 포함할 수 있다. 예를 들어, *Asgr1* 단백질은 인간 *ASGR1* 꼬인 나선 도메인, 인간 *ASGR1* C형 렉틴 도메인, 인간 *ASGR1* 막관통 도메인, 및 인간 *ASGR1* 세포질 도메인 중 하나 이상 또는 전부를 포함할 수 있다. 일 실시예로서, *Asgr1* 단백질은 인간 *ASGR1* 꼬인 나선 도메인만, 인간 *ASGR1* C형 렉틴 도메인만, 또는 인간 *ASGR1* 세포외 도메인(즉, 꼬인 나선 도메인 및 C형 렉틴 도메인)만을 포함할 수 있다. 임의로, 인간화 *Asgr1* 유전자좌에 의해 암호화되는 *Asgr1* 단백질은 내인성 비인간 동물 *Asgr1* 단백질로부터 유래된 하나 이상의 도메인을 또한 포함할 수 있다. 일 실시예로서, 인간화 *Asgr1* 유전자좌에 의해 암호화되는 *Asgr1* 단백질은 인간 *ASGR1* 단백질로부터 유래된 꼬인 나선 도메인, 인간 *ASGR1* 단백질로부터 유래된 C형 렉틴 도메인, 내인성(즉, 천연) 비인간 동물 *Asgr1* 단백질로부터 유래된 N-말단 세포질 도메인, 및 내인성(즉, 천연) 비인간 동물 *Asgr1* 단백질로부터 유래된 막관통 도메인을 포함할 수 있다. 인간 *ASGR1* 단백질로부터 유래된 도메인은 완전한 인간화 서열(즉, 도메인이 이종상동성 인간 *ASGR1* 서열을 암호화하는 전체 서열)에 의해 암호화될 수 있거나, 부분적 인간화 서열(즉, 도메인이 이종상동성 인간 *ASGR1* 서열을 암호화하는 서열의 일부, 및 암호화된 도메인이 인간 *ASGR1* 단백질에 있는 그 도메인과 동일하도록 도메인이 이종상동성 인간 *ASGR1* 서열과 동일한 아미노산을 암호화하는 나머지 내인성(즉, 천연) 서열 암호화에 의해 암호화될 수 있다.

[0087] 일 실시예로서, 인간화 *Asgr1* 유전자좌에 의해 암호화되는 *Asgr1* 단백질은 인간 *ASGR1* 꼬인 나선 도메인을 포함할 수 있다. 임의로, 인간 *ASGR1* 꼬인 나선 도메인은 서열 번호 27과 적어도 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 서열을 포함하거나, 이로 필수적으로 구성되거나, 이로 구성되며, *Asgr1* 단백질은 천연 *Asgr1*의 활성을 유지한다(즉, 수용체-매개 세포질 함입을 통해 갈락토오스-말단 및 N-아세틸갈락토사민-말단 당 단백질을 선택적으로 결합하고 내재화하는 능력을 유지함). 또 다른 실시예로서, 인간화 *Asgr1* 유전자좌에 의해 암호화되는 *Asgr1* 단백질은 인간 *ASGR1* C형 렉틴 도메인을 포함할 수 있다. 임의로, 인간 *ASGR1* C형 렉틴 도메인은 서열 번호 28과 적어도 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 서열을 포함하거나, 이로

필수적으로 구성되거나, 이로 구성되며, Asgr1 단백질은 천연 Asgr1의 활성을 유지한다. 예를 들어, 인간 ASGR1로부터 유래된 인간 ASGR1은 서열 번호 31과 적어도 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 서열을 포함하거나, 이로 필수적으로 구성되거나, 이로 구성되며, Asgr1 단백질은 천연 Asgr1의 활성을 유지한다. 또 다른 실시예로서, 인간화 *Asgr1* 유전자좌에 의해 암호화되는 Asgr1 단백질은 내인성 비인간 동물 Asgr1 세포질 도메인(예컨대, 마우스 Asgr1 세포질 도메인)을 포함할 수 있다. 임의로, 비인간 동물 Asgr1 세포질 도메인은 서열 번호 29과 적어도 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 서열을 포함하거나, 이로 필수적으로 구성되거나, 이로 구성되며, Asgr1 단백질은 천연 Asgr1의 활성을 유지한다. 또 다른 실시예로서, 인간화 *Asgr1* 유전자좌에 의해 암호화되는 Asgr1 단백질은 내인성 비인간 동물 Asgr1 막관통 도메인(예컨대, 마우스 Asgr1 막관통 도메인)을 포함할 수 있다. 임의로, 비인간 동물 Asgr1 막관통 도메인은 서열 번호 30과 적어도 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 서열을 포함하거나, 이로 필수적으로 구성되거나, 이로 구성되며, Asgr1 단백질은 천연 Asgr1의 활성을 유지한다. 예를 들어, 인간화 *Asgr1* 유전자좌에 의해 암호화되는 Asgr1 단백질은 서열 번호 3과 적어도 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 서열을 포함하거나, 이로 필수적으로 구성되거나, 이로 구성되며, Asgr1 단백질은 천연 Asgr1의 활성을 유지한다.

[0088] 임의로, 인간화 *Asgr1* 유전자좌는 다른 요소를 포함할 수 있다. 이러한 요소의 실시예는 선택 카세트, 리포터 유전자, 재조합효소 인식 부위, 또는 다른 요소를 포함할 수 있다. 대안적으로, 인간화 *Asgr1* 유전자좌는 다른 요소들을 결합할 수 있다(예컨대, 선택 마커 또는 선택 카세트가 결합될 수 있음). 적절한 리포터 유전자 및 리포터 단백질의 실시예는 본원의 다른 곳에 개시된다. 적절한 선택 마커의 실시예는 네오마이신 포스포트랜스퍼라아제(neo<sup>r</sup>), 히그로마이신 B 포스포트랜스퍼라아제(hyg<sup>r</sup>), 퓨로마이신-N-아세틸트랜스퍼라아제(puro<sup>r</sup>), 블라스티딘 S 디아미나아제(bsr<sup>r</sup>), 크산틴/구아닌 포스포리보실 트랜스퍼라아제(gpt), 및 단순 포진 바이러스 티미딘 키나아제(HSV-k)를 포함한다. 재조합 효소의 실시예는 Cre, Flp, 및 Dre 재조합효소를 포함한다. Cre 재조합효소 유전자의 일 실시예는 Crei이며, Cre 재조합효소를 암호화하는 두 개의 엑손이 인트론에 의해 분리되어 원핵 세포에서 그의 발현을 방지한다. 이러한 재조합 효소는 핵으로 국소화를 용이하게 하기 위해 핵 국소화 신호를 추가로 포함할 수 있다(예컨대, NLS-Crei). 재조합효소 인식 부위는 부위-특이적 재조합효소에 의해 인식되고 재조합 사건을 위한 기체로서의 역할을 할 수 있는 뉴클레오티드 서열을 포함한다. 재조합효소 인식 부위의 실시예로는 FRT, FRT11, FRT71, attp, att, rox, 및 loxP, lox511, lox2272, lox66, lox71, loxM2, 및 lox5171 과 같은 lox 부위를 포함한다.

[0089] 리포터 유전자 또는 선택 카세트와 같은 다른 요소는 재조합효소 인식 부위에 의해 측부에 위치한 자가-결실 카세트일 수 있다. 예를 들어, US 8,697,851 및 US 2013/0312129를 참조하며, 이들 각각은 그 전체가 모든 목적을 위해 참조로서 본원에 통합된다. 실시예로서, 자가-결실 카세트는, 마우스 *Prm1* 프로모터에 작동 가능하게 연결되는 (인트론에 의하여 분리된 Cre 재조합효소를 암호화하는 두 엑손을 포함하는) Crei 유전자, 및 인간 유비퀴틴 프로모터에 작동 가능하게 연결되는 네오마이신 저항성 유전자를 포함할 수 있다. *Prm1* 프로모터를 이용함으로써, 자가-결실 카세트는 F0 동물의 수컷 생식 세포에서 특이적으로 결실될 수 있다. 선택 마커를 암호화하는 폴리뉴클레오티드는 표적화되는 세포에서 활성인 프로모터에 작동 가능하게 연결될 수 있다. 프로모터의 실시예는 본원의 다른 곳에서 기술된다. 또 다른 특이적 실시예로서, 자가-결실 선택 카세트는 하나 이상의 프로모터(예를 들어, 인간 유비퀴틴 및 EM7 프로모터 둘 다)에 작동 가능하게 연결된 히그로마이신 저항성 유전자 암호화 서열, 이어서 폴리아데닐화 신호, 이어서 하나 이상의 프로모터(예컨대, mPrm1 프로모터)에 작동 가능하게 연결된 Crei 암호화 서열, 이어서 또 다른 폴리아데닐화 신호를 포함할 수 있되, 전체 카세트는 loxP 부위에 의해 측부에 위치된다.

[0090] 인간화 *Asgr1* 유전자좌는 또한 조건부 대립유전자일 수 있다. 예를 들어, 조건부 대립유전자는, 모든 목적을 위해 그 전체가 참조로 본원에 통합된 US 2011/0104799에 기술된 바와 같이, 다기능성 대립유전자일 수 있다. 예를 들어, 조건부 대립유전자는 하기를 포함할 수 있다: (a) 표적 유전자의 전사에 관한, 센스 배향의 구동 서열; (b) 센스 또는 안티센스 배향의 약물 선택 카세트(DSC); (c) 안티센스 배향의 관심 뉴클레오티드 서열(NSI); 및 (d) 역방향 배향의 역전 모듈에 의한 조건부(엑손-분할 인트론, 및 가역적 유전자-트랩-유사 모듈을 이용하는, COIN) **참고: 예컨대**, US 2011/0104799. 조건부 대립유전자는 제1 재조합효소에 노출시 재조합하는 재조합가능 유닛을 추가로 포함하여, (i) 구동 서열 및 DSC가 결핍되며; 그리고 (ii) 센스 배향으로의 NSI, 및 안티센스 배향으로의 COIN을 함유하는 조건부 대립유전자를 형성할 수 있다. **참고: 예컨대**, US 2011/0104799.

[0091] 하나의 예시적인 인간화 *Asgr1* 유전자좌(예컨대, 인간화 마우스 *Asgr1* 유전자좌)는 암호화 엑손 3-8이 상응하는 인간 서열로 대체되는 것이다. 이러한 엑손은 Asgr1의 꼬인 나선과 C형 렉틴 도메인을 암호화한다. 임의로, 인간화 서열은 종결 코돈 및 3' UTR을 통해, 그리고 임의로 3' UTR의 바로 하류에 있는 서열 내에 존재할 수

있다. 임의로, 암호화 엑손 3의 인트론 상류의 일부는 또한 인간화된다. 참조: 도 2a 및 2b 및 서열 번호 21과 24.

**[0092] C. 인간화 *Asgr1* 유전자좌를 포함하는 비인간 세포 및 비인간 동물**

**[0093]** 본원의 다른 곳에 기술된 바와 같이 인간화 유전자좌 *Asgr1*을 포함하는 비인간 동물 세포 및 비인간 동물이 제공된다. 세포 또는 비인간 동물은 인간화 *Asgr1* 유전자좌에 대해서 이형접합체 또는 동형접합체일 수 있다. 이 배체 유기체는 각각의 유전자좌에서 2개의 대립유전자를 가진다. 대립유전자의 각각의 쌍은 특이적 유전자좌의 유전자형을 나타낸다. 유전자형은 특정 유전자좌에 두 개의 동일한 대립유전자가 존재한다면 동형접합체로서 기술되고, 두 대립유전자가 상이하다면 이형접합체로서 기술된다.

**[0094]** 본원에서 제공되는 비인간 동물 세포는, 예를 들어, *Asgr1* 유전자좌 또는 인간 *ASGR1* 유전자좌에 대해 상동성 또는 이종상동성인 게놈 유전자좌를 포함하는 임의의 비인간 세포일 수 있다. 세포는, 예를 들어, 진균 세포(예컨대, 효모), 식물 세포, 동물 세포, 포유류 세포, 비인간 포유류 세포 및 인간 세포를 포함하는, 진핵 세포일 수 있다. 용어 "동물"은 포유류, 어류 및 조류를 포함한다. 포유류 세포는, 예를 들어, 비인간 포유류 세포, 설치류 세포, 랫트 세포, 마우스 세포, 또는 햄스터 세포일 수 있다. 다른 비인간 포유동물은, 예를 들어, 비인간 영장류, 원숭이, 유인원, 오랑우탄, 고양이, 개, 토끼, 말, 황소, 사슴, 들소, 가축(예컨대, 소, 거세소 등과 같은 소 종; 양, 염소 등과 같은 양 종; 및 돼지 및 멧돼지와 같은 돼지 종)을 포함한다. 조류로는, 예를 들어 닭, 칠면조, 타조, 거위, 오리 등을 포함한다. 가축 및 농업용 동물 또한 포함된다. 용어 "비인간"은 인간을 배제한다.

**[0095]** 세포는 또한 임의 유형의 미분화 또는 분화된 상태일 수 있다. 예를 들어, 세포는 전능성 세포, 다능성 세포(예컨대, 인간 다능성 세포 또는 마우스 배아 줄기(ES) 세포 또는 랫트 ES 세포와 같은 비인간 다능성 세포), 또는 비다능성 세포일 수 있다. 전능성 세포는 임의의 세포 유형을 생기게 할 수 있는 미분화된 세포를 포함하고, 다능성 세포는 1 초과 분화된 세포 유형으로 발생하는 능력을 갖는 미분화된 세포를 포함한다. 이러한 다능성 및/또는 전능성 세포는, 예를 들어, ES 세포 또는 유도 만능 줄기(iPS) 세포와 같은 ES-유사 세포일 수 있다. ES 세포는 배아에 도입시 발생 배아의 임의의 조직에 기여할 수 있는 배아 유래 전능성 또는 다능성 세포를 포함한다. ES 세포는 배반포의 내부 세포 덩어리로부터 유래될 수 있고 3개 척추동물 배엽층 (내배엽, 외배엽, 및 중배엽)의 어느 하나의 세포로 분화할 수 있다.

**[0096]** 본원에서 제공된 세포는 또한 생식 세포(예컨대, 정자 또는 난모세포)일 수도 있다. 세포는 체세포 분열을 할 수 있는 세포 또는 체세포 분열 비활성 세포, 감수분열을 할 수 있는 세포 또는 감수분열 비활성 세포일 수 있다. 유사하게, 세포는 일차 체세포 또는 일차 체세포가 아닌 세포일 수도 있다. 체세포는 배우자, 생식 세포, 생식 모세포 또는 미분화 줄기세포가 아닌 임의의 세포를 포함한다. 예를 들어, 세포는 간모세포 또는 간세포와 같은 간 세포일 수 있다.

**[0097]** 본원에 제공되는 적합한 세포는 또한 일차 세포를 포함한다. 일차 세포는 유기체, 기관 또는 조직으로부터 직접 단리된 세포 또는 세포의 배양물을 포함한다. 일차 세포는 형질 전환되지 않거나 불멸이 아닌 세포를 포함한다. 이들은 조직 배양에서 이전에 통과되지 않았거나, 조직 배양에서 이전에 통과되었지만 조직 배양에서 무기한으로 통과될 수 없는 유기체, 기관 또는 조직으로부터 수득된 임의의 세포를 포함한다. 이러한 세포는 통상적인 기술에 의해 단리될 수 있고, 예를 들어, 간세포를 포함할 수 있다.

**[0098]** 본원에서 제공되는 다른 적합한 세포는 불멸화 세포를 포함한다. 불멸화 세포는 정상적으로는 무기한으로 증식하지 않지만 돌연변이 또는 변형으로 인해 정상적인 세포 노화를 회피하고 대신 분열을 계속할 수 있는 다세포 유기체로부터의 세포를 포함한다. 이러한 돌연변이 또는 변형은 자연적으로 발생하거나 의도적으로 유도될 수 있다. 불멸화 세포주의 특정 실시예는 HepG2 인간 간암 세포주이다. 불멸화 세포의 수많은 유형이 잘 알려져 있다. 불멸화 또는 일차 세포는 배양을 위해 또는 재조합 유전자 또는 단백질을 발현하기 위해 통상적으로 사용되는 세포를 포함한다.

**[0099]** 본원에서 제공된 세포는 또한 1세포기 배아(즉, 수정된 난모세포 또는 접합체)를 포함한다. 이러한 1세포기 배아는 임의의 유전적 배경(예컨대, BALB/c, C57BL/6, 129, 또는 마우스에서 이들의 조합)으로부터 유래될 수 있고, 신선하거나 냉동될 수 있고, 천연 교배 또는 시험관 내 수정으로부터 유도될 수 있다.

**[0100]** 본원에서 제공된 세포는 정상적인, 건강한 세포일 수 있거나, 질환 또는 돌연변이-보유 세포일 수 있다.

**[0101]** 본원에 기술된 바와 같이 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물은 본원의 다른 곳에 기술된 방법에 의해 만들어질 수 있다. 용어 "동물"은 포유류, 어류 및 조류를 포함한다. 비인간 포유동물은, 예를 들어, 비인간

영장류, 원숭이, 유인원, 오랑우탄, 고양이, 개, 말, 황소, 사슴, 들소, 양, 토끼, 설치류(예컨대, 마우스, 랫트, 햄스터, 및 기니피그), 및 가축(예컨대, 소, 거세소와 같은 소 종; 양, 염소와 같은 양 종; 및 돼지 및 멧돼지와 같은 돼지 종)을 포함한다. 조류로는, 예를 들어 닭, 칠면조, 타조, 거위, 및 오리를 포함한다. 가축 및 농업용 동물 또한 포함된다. 용어 "비인간 동물"은 인간을 배제한다. 바람직한 비인간 동물은, 예를 들어, 마우스 및 랫트와 같은 설치류를 포함한다.

[0102] 비인간 동물은 임의의 유전적 배경으로부터 유래될 수 있다. 예를 들어, 적합한 마우스는 129 균주, C57BL/6 균주, 129 및 C57BL/6 혼합, BALB/c 균주, 또는 스위스 웹스터 균주로부터 유래될 수 있다. 129 균주의 실시예는 129P1, 129P2, 129P3, 129X1, 129S1(예컨대, 129S1/SV, 129S1/SvIm), 129S2, 129S4, 129S5, 129S9/SvEvH, 129S6(129/SvEvTac), 129S7, 129S8, 129T1, 및 129T2를 포함한다. 예를 들어, Festing 외 (1999) *Mammalian Genome* 10:836을 참조하고, 그 전체는 모든 목적을 위해 참조로서 본원에 통합된다. C57BL 균주의 실시예는 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 포함한다. 적합한 마우스는 또한 전술한 129 균주 및 전술한 C57BL/6 균주의 혼합(예컨대, 50% 129 및 50% C57BL/6)으로부터 유래될 수 있다. 마찬가지로, 적합한 마우스는 전술한 129 균주의 혼합 또는 전술한 BL/6 균주의 혼합(예컨대, 129S6(129/SvEvTac) 균주)으로부터 유래될 수 있다.

[0103] 유사하게, 랫트는, 예를 들어, ACI 랫트 계통, 검은 아구티(Dark Agouti, DA) 랫트 계통, 위스타 랫트 계통, LEA 랫트 계통, 스프래그 다울리(SD) 랫트 계통, 또는 피셔 F344나 피셔 F6과 같은 피셔 랫트 계통을 포함하는 임의의 랫트 계통으로부터 유래될 수 있다. 랫트는 위에 인용된 하나 이상의 계통의 혼합체로부터 유래된 계통으로부터 수득될 수도 있다. 예를 들어, 적절한 랫트는 DA 균주 또는 ACI 균주로부터 유래될 수 있다. ACI 랫트 계통은 흰 배와 다리 및  $RTI^{av1}$  일배체형을 가진 검은 아구티를 가지는 것을 특징으로 한다. 이러한 계통은 Harlan Laboratories를 포함하는 다양한 공급원으로부터 이용이 가능하다. 검은 아구티(DA) 랫트 계통은 아구티 코트(coat) 및  $RTI^{av1}$  일배체형을 가지는 것을 특징으로 한다. 이러한 랫트는 Charles River 및 Harlan Laboratories를 포함하는 다양한 공급원으로부터 이용이 가능하다. 일부 적합한 랫트는 동종번식된 쥐 균주로부터 유래될 수 있다. 예를 들어, US 2014/0235933를 참조하며, 그 전체는 모든 목적을 위해 참조로서 본원에 통합된다.

[0104] **III. 간으로 치료 복합체 및 인간-ASGR1-매개 메커니즘을 통해 작용하는 치료 분자의 인간-ASGR1-매개 전달의 생체 내 효능을 평가하기 위해 인간화 Asgr1 유전자좌를 포함하는 비인간 동물을 사용하는 방법**

[0105] 간으로 치료 분자 또는 치료 복합체의 인간-ASGR1-매개 전달 또는 인간-ASGR1-매개 메커니즘의 생체 내 효능을 평가하기 위해 본원에서 다른 곳에 기술된 바와 같이 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물을 사용하기 위한 다양한 방법이 제공된다. 비인간 동물은 인간화 *Asgr1* 유전자좌를 포함하기 때문에, 비인간 동물은 비인간 *Asgr1* 유전자좌를 가진 비인간 동물보다 인간 ASGR1 또는 인간 ASGR1 매개 치료 메커니즘에 의해 매개된 전달을 보다 정확하게 반영할 것이다. 일 실시예로서, 방법은 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물에 치료 복합체를 투여하는 단계를 포함하여, 생체 내 인간-ASGR1 매개 내재화를 통해 간으로 치료 복합체의 전달을 평가할 수 있되, 치료 복합체는 인간 ASGR1과 특이적으로 결합하는 치료 분자 및 항원-결합 단백질 또는 리간드를 포함한 다음, 비인간 동물의 간으로 치료 분자의 전달을 평가하는 단계를 포함한다. 또 다른 실시예로서, 방법은 인간 ASGR1을 통해 표적 간 세포 표면 단백질을 내재화하도록 설계된 치료 복합체와 같은, 인간-ASGR1-매개 내재화를 통해 작용하도록 설계된 치료 분자 또는 복합체의 생체 내 효능일 수 있다.

[0106] **A. 인간-ASGR1-매개 내재화를 통해 간으로 치료 분자의 전달의 생체 내 효능을 평가하기 위한 방법**

[0107] 인간-ASGR1-매개 내재화를 통해 간으로 치료 분자 전달의 생체 내 효능을 평가하기 위해 본원에서 다른 곳에 기술된 바와 같이 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물을 사용하기 위한 다양한 방법이 제공된다. 예를 들어, 이러한 방법은 (a) 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물에게 치료 복합체를 투여하는 단계를 포함하되, 치료 복합체는 인간 ASGR1에 특이적으로 결합하는 치료 분자 및 항원 결합 단백질 또는 리간드를 포함하고; (b) 비인간 동물의 간으로 치료 분자의 전달을 평가하는 단계를 포함할 수 있다.

[0108] 치료 분자는 임의의 질환 또는 장애의 치료 또는 예방에 사용되는 생물학적 또는 화학적 제제일 수 있다. 예를 들어, 치료 분자는 치료 핵산(예컨대, CRISPR/Cas 가이드 RNA, 짧은 헤어핀 RNA(shRNA), 또는 작은 간섭 RNA(siRNA)), 또는 치료 단백질(예컨대, Cas9 단백질, 대체 효소, 분비된 치료적 단백질 등과 같은 Cas 단백질)을 암호화하는 핵산일 수 있다. 대안적으로, 치료 분자는 치료 단백질, 치료 항체 또는 항원-결합 단백질,

또는 임의의 다른 치료 거대 분자 또는 작은 분자일 수 있다.

- [0109] 치료 분자 및 인간 ASGR1 항원-결합 단백질 또는 리간드는 임의의 수단에 의해 함께 복합될 수 있다. 예를 들어, 치료 분자 및 인간 ASGR1 항원-결합 단백질 또는 리간드는 직접적인 공유 결합을 통해 커플링될 수 있거나 펩티드 링커 또는 화학적 링커와 같은 링커를 통해 커플링될 수 있다. 치료 분자 및 인간 ASGR1 항원-결합 단백질 또는 리간드는 또한 특정 기능 그룹 또는 특정 공간 맞춤(예컨대, 잠금 및 키 유형) 사이에서 결합의 형성을 통해 함께 복합체화할 수 있다. 특정 실시예로서, 인간 ASGR1 항원 결합 단백질은 치료 분자와 또한 특이적으로 결합하는 이중 특이적 항원-결합 단백질일 수 있다.
- [0110] 치료 복합체의 투여단계는 본원의 다른 곳에 그리고 임의의 투여 경로에 의해 더 자세히 개시된 바와 같은 임의의 수단에 의해 이루어질 수 있다. 치료 복합체 및 분자를 전달하는 수단 및 투여 경로는 본원의 다른 곳에서 더욱 자세히 개시된다.
- [0111] 인간 ASGR1과 특이적으로 결합하는 인간 ASGR1 항원-결합 단백질 또는 인간 ASGR1 리간드가 사용될 수 있다. 인간 ASGR1 단백질은 간에서 특정 당단백질의 세포 내 유입을 매개하는 막관통 단백질이기 때문에, 인간 ASGR1과 함께 복합체가 되는 분자는 인간 ASGR1과 함께 내재화될 수 있다. 적합한 항원-결합 단백질의 실시예는 수용체-융합 분자, 트랩 분자, 수용체-Fc 융합 분자, 항체, Fab 단편, F(ab')<sub>2</sub> 단편, Fd 단편, Fv 단편, 단쇄 Fv (scFv) 분자, dAb 단편, 단리된 상보성 결정 영역 (CDR), CDR3 펩티드, 제약적 FR3-CDR3-FR4 펩티드, 도메인-특이적 항체, 단일 도메인 항체, 도메인-결실된 항체, 키메라 항체, CDR-그래프트된 항체, 이중체, 삼중체, 사중체, 미니바디, 나노바디, 1가 나노바디, 2가 나노바디, 작은 모듈의 면역의약품(SMIP), 카멜리드 항체 (VHH 중쇄 동종이량체 항체), 상어 가변성 IgNAR 도메인 등을 포함한다. 하나의 특정 실시예에서, 항원-결합 단백질은 인간 ASGR1 및 치료 분자(예컨대, 대체 단백질 또는 효소 또는 AAV와 같은 전달 비허클)에 결합하는 이중 특이적 항원이다. 대안적으로, 인간 ASGR1과 특이적으로 상호작용하는 리간드 또는 리간드의 일부(예컨대, ASGR1에 대한 아시알로오소뮤코이드(ASOR) 또는 베타-GalNAc 또는 이들의 수용체-결합 부분)가 사용될 수 있다.
- [0112] 비인간 동물의 간으로 치료 분자의 전달을 임의의 알려진 수단에 의해 평가할 수 있다. 일 실시예로서, 치료 분자의 존재는 간에서 평가될 수 있다. 예를 들어, 치료 분자가 치료 단백질인 경우, 치료 단백질의 존재는 단백질을 검출하기 위해 알려진 분석법을 사용하여 비인간 동물의 간에서 평가될 수 있다. 유사하게, 치료 분자가 치료 단백질을 암호화하는 핵산인 경우, 핵산의 발현(예컨대, mRNA 발현 또는 단백질 발현)은 알려진 분석법을 사용하여 비인간 동물의 간에서 평가될 수 있다. 암호화된 치료 단백질이 분비 단백질인 경우, 치료 단백질의 혈청 수준을 측정하거나 그 의도된 표적 세포 유형, 조직 유형 또는 기관에서 분비된 치료 단백질의 활성을 알려진 분석법에 의해 평가할 수 있다. 비인간 동물의 간에서 치료 분자의 활성은 치료 분자의 의도된 기능에 따라 알려진 분석법에 의해 평가될 수 있다. 예를 들어, CRISPR/Cas와 같은 게놈 편집 체계가 도입되고 있는 경우, 알려진 분석법이 특정 표적 게놈 유전자좌에서 게놈 편집을 평가하기 위해 사용될 수 있다.
- [0113] 특정 실시예에서, 치료 복합체는 바이러스 벡터 조성물(예를 들어, 간 특이적 바이러스 벡터 조성물)을 포함할 수 있다. 이러한 바이러스 벡터 조성물은 간을 표적화하기 위해 인간 ASGR1에 지향되도록 설계된, 폐지된 천연 향성(tropism)을 가질 수 있다. 예를 들어, 이러한 변형된 바이러스 벡터 복합체는: (i) 치료 핵산 또는 단백질을 암호화하는 핵산을 포함하는 변형된 바이러스 벡터, 여기서 변형된 바이러스 벡터는 천연의 향성을 폐지하거나 감소시키고, 이중 에피토프를 포함함; 및 (ii) (1) 이중 에피토프와 특이적으로 결합하는 항원-결합 단백질(예를 들어, 항체) 파라토프; 및 (2) 인간 ASGR1과 특이적으로 결합하는 표적화 리간드를 포함하는 재표적화 모이어티를 포함할 수 있다.
- [0114] 이를 달성하는 방법의 한 실시예로서, SpyCatcher-SpyTag와 같은 단백질 태그 부착 시스템이 바이러스 표면에 항체를 공유 결합하는 데 사용될 수 있다. 대안적으로, 이중 특이적 항원-결합 단백질이 사용될 수 있다(예를 들어, 항체의 하나의 아암이 바이러스와 결합하고 다른 아암이 인간 ASGR1에 결합하는 이중특이적 항체). 특정 실시예에서, 재지향화 모이어티는 제1 및 제2 항원-결합 도메인을 포함하는 이중특이적 항원-결합 단백질(예를 들어 이중특이적 항체)이되, 제1 항원-결합 도메인은 재조합 인간 바이러스 캡시드 단백질에 삽입되고/되거나 이에 의해 디스플레이되는 이중 에피토프에 특이적으로 결합하는 파라토프를 포함하고, 제2 항원 결합 도메인은 인간 ASGR1에 특이적으로 결합한다.
- [0115] 적절한 이중 에피토프의 실시예는 Myc 태그이다. 예를 들어, Myc 태그는 AAV2의 N587 이후에 삽입될 수 있다. 이러한 삽입은 AAV2의 천연 리간드-결합 활성을 폐지하고 또한 항-Myc 항체에 의한 변형된 AAV의 인지를 가능하게 한다. 그런 다음 Myc 및 인간 ASGR1에 특이적으로 결합하는 이중특이적 항체의 사용은 AAV(예컨대, AAV2 N587 Myc)를 인간 ASGR1(또는 인간화 Asgr1)을 발현하는 간 세포로 재표적화할 수 있다.

- [0116] 또 다른 실시예에서, 치료 복합체는 리소좀 대체 효소 또는 단백질 또는 리소좀 대체 효소 또는 단백질을 암호화하는 핵산을 포함한다. 리소좀 저장 질환은, 스펅고지질, 뮤코다당류, 당단백질, 글리코젠, 및 올리고당을 포함하는 리소좀에서의 상이한 기체의 분해에 영향을 미치는 희귀 질환의 등급으로서, 질환 세포 내에 축적될 수 있어서 세포 사멸을 초래할 수 있다. 리소좀 저장 질환에 의해 영향을 받는 장기는 간을 포함한다. 이 질환의 병인은 보통 단백질 기능의 상실로 인한, 리소좀에서의 불완전한 분해 산물의 축적(buildup)에 기인한다. 리소좀 축적 질환은 일반적으로 그것의 정상적인 기능이 리소좀 함유물을 분해시키거나 분해를 조정하는 것인 단백질에서의 기능 상실 또는 변이체를 감소시킴으로써 발생한다. 리소좀 저장 질환의 실시예는 모든 목적을 위해 그 전체가 참조로 본원에 통합된 WO 2017/100467에 제공된다. 예를 들어, 가장 일반적인 리소좀 저장 질환 중 하나는 폼페병이다. 폼페병은 결함 있는 리소좀 효소 알파-글루코시다아제 (GAA)에 의해 야기되며, 이로 인해 리소좀 글리코젠에 대한 결함 있는 가공이 초래된다. 리소좀 글리코젠의 축적은 골격, 심장 및 간 조직에서 주로 일어난다.
- [0117] 리소좀 저장 질환의 치료를 위한 하나의 옵션은 효소 또는 단백질 대체 요법이다. 대체 효소 또는 단백질은 인간 ASGR1에 특이적으로 결합하는 항원 결합 단백질 또는 리간드와 함께 치료적 복합체에 연관될 때 특정 표적 세포의 리소좀에 효과적으로 전달될 수 있다. WO 2017/100467을 참조하고, 그 전체는 모든 목적을 위해 참조로서 본원에 통합된다. 이러한 치료 복합체는 대상체에게 투여될 수 있고, 치료 복합체는 대상체의 표적화된 세포의 리소좀으로 진입할 수 있으며 리소좀 저장 질환과 연관된 효소 활성을 대체하는 효소 활성을 제공할 수 있다. 인간 ASGR1 단백질은 간에서 노출된 말단 갈락토오스 또는 N-아세틸갈락토사민 잔기를 갖는 당단백질의 세포 내 유입 및 리소좀 분해를 매개하는 막관통 단백질이다. 결과적으로, 인간 ASGR1과 결합하고 함께 내재화되는 단백질은 리소좀으로 표적화될 것이다. 이와 같이, 인간 ASGR1에 치료 분자 또는 복합체를 지향하는 전달 방법은 치료제를 리소좀(예컨대, 간에서 리소좀)에 표적화하는 데 사용될 수 있다.
- [0118] 따라서, 본원에서 제공된 방법들 중 일부는 생체 내 인간 ASGR1을 통해 리소좀 대체 단백질 또는 효소를 포함하는 치료 복합체의 전달을 평가하는 방법이다. 예를 들어, 이러한 방법은 (a) 본원의 다른 곳에서 기술된 비인간 동물에 치료 분자 또는 치료 복합체를 투여하는 단계, 여기서 치료 분자 또는 치료 복합체는 인간 ASGR1과 특이적으로 결합하는 리소좀 대체 단백질 또는 효소(또는 리소좀 대체 단백질 또는 효소를 암호화하는 핵산) 및 항원-결합 단백질을 포함하고, 여기서 ASGR1은 세포 결합 및 리소좀 구획으로의 흡수를 매개함; 및 (b) 비인간 동물의 간에서 리소좀 대체 단백질 또는 효소의 존재 또는 활성을 평가하는 단계를 포함할 수 있다. 비인간 동물의 간의 대체 단백질 또는 효소의 활성은 특정 대체 단백질 또는 효소를 알려진 분석법에 의해 또는 알려진 분석에 의해 리소좀에서 상이한 영향을 받는 기체들의 분해를 측정함으로써 평가될 수 있다.
- [0119] 또 다른 실시예로서, 치료 분자는 치료 분비 단백질을 암호화하는 핵산일 수 있다. 간은 혈액 속으로 분비되는 단백질(주요 혈장 단백질, 지혈 및 섬유소 용해 인자, 담체 단백질, 호르몬, 호르몬 전구체 및 아포지단백질을 포함함)을 생산하는 데 중요한 역할을 한다. 인간 ASGR1 단백질은 간에서 특정 당단백질의 세포 내 유입을 매개하는 막관통 단백질이다. 결과적으로, 인간 ASGR1과 함께 복합체화된 분자는 인간 ASGR1과 함께 내재화될 수 있다. 이와 같이, 치료 분비 단백질을 암호화하는 핵산은 인간 ASGR1을 표적으로 하는 복합체에 이것을 전달함으로써 간으로 표적화될 수 있다. 핵산은 ASGR1을 통해 내재화될 수 있고, 간은 치료 단백질을 제조하고 분비할 수 있다.
- [0120] 따라서, 본원에서 제공된 방법들 중 일부는 생체 내 인간 ASGR1을 통해 치료 분비 단백질을 암호화하는 핵산의 전달을 평가하는 방법이다. 예를 들어, 이러한 방법은 (a) 본원의 다른 곳에 기술된 비인간 동물에게 치료 분비 단백질을 암호화하는 핵산(예컨대, DNA)을 투여하는 단계, 여기서 핵산은 인간 ASGR1에 특이적으로 결합하는 치료 복합체 내로 전달되고, ASGR1은 복합체의 내재화를 매개함; (b) 비인간 동물에서 치료 분비 단백질의 분비 수준(예컨대, 혈청 수준) 또는 활성을 평가하는 단계를 포함할 수 있다.
- [0121] 치료 단백질의 생산 및 분비는 임의의 알려진 수단에 의해 평가될 수 있다. 예를 들어, 도입된 핵산의 발현은 알려진 분석법을 사용하여 비인간 동물의 간에 있는 비인간 동물 또는 암호화된 치료 단백질의 간에 있는 암호화된 mRNA의 수준을 측정함으로써 평가될 수 있다. 치료 단백질의 분비는 알려진 분석법을 사용하여 비인간 동물에서 암호화된 치료 단백질을 측정하거나 혈청 수준에 의해 측정 또는 혈청 수준을 측정함으로써 평가될 수 있다. 또한, 분비된 치료 단백질이 특정 세포 유형, 조직, 또는 기관에 작용한다면, 분비된 치료 단백질의 활성은 표적 세포 유형, 조직, 또는 기관에서 평가될 수 있다.
- [0122] B. 인간-ASGR1-매개 내재화를 통해 표적 간세포 표면 단백질 또는 간에서 표적 수용성 단백질을 내재화하기 위한 치료 복합체의 생체 내 효능을 평가하는 방법

[0123] 인간 ASGR1을 통해 간으로 표적 간 세포 표면 단백질을 내재화하기 위해 설계된 치료 복합체의 *생체 내* 효능을 평가하기 위해 본원에서 다른 곳에 기술된 바와 같이 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물을 사용하기 위한 다양한 방법이 제공된다. 치료제 처리는 종종 세포에 작용하거나 세포 근처에서 작용하는 하나 이상의 표적 분자의 불활성화 또는 차단을 필요로 한다. 예를 들어, 항체를 기반으로 한 치료제는 종종 세포 표면에 발현된 특정 항원과의 결합 또는 가용성 리간드와의 결합에 의해 이러한 항원의 정상적인 생물학적 활성을 방해함으로써 기능한다. 이러한 유형의 치료제는 전형적으로 세포 신호 전달을 약화시키거나 억제하기 위해 사이토카인과 그 수용체 사이의 상호 작용을 차단함으로써 기능한다. 그러나 일부 맥락에서는 표적 분자의 또 다른 성분과의 물리적 상호 작용을 차단하는 것을 반드시 수반하는 것만은 아닌 방식으로 표적 분자의 활성을 불활성화하거나 억제하는 것이 치료적으로 유익할 것이다. 표적 분자의 그러한 비 차단성 약화를 달성할 수 있는 한 가지 방법은 표적 분자의 세포 외 또는 세포 표면 농도를 감소시키는 것일 것이다. 예를 들어, 표적 분자는 표적 분자와 ASGR1과 같은 내재화 이펙터 단백질 사이의 물리적 연결과 관련하여 촉진하거나 가져옴으로써 약독화 또는 불활성화될 수 있다. 이는, 예를 들어, 제1 항원-결합 도메인 및 제2 항원-결합 도메인을 포함하는 다중특이적(예컨대, 이중특이적) 항원-결합 분자의 사용을 통해 달성될 수 있다. 각각의 항원-결합 도메인은 상이한 분자와 결합한다: 처음 표적 분자와 결합하고, 두 번째는 ASGR1에 특이적으로 결합한다. 이런 유형의 물리적인 분자간 연결을 통해, 이러한 표적 분자는 ASGR1과 함께 세포 안으로 내재화되도록 강제되어 세포 내 분해 기구에 의해 처리되거나, 그렇지 않으면 약화, 격리, 또는 불활성화될 수 있다. 예를 들어, WO 2013/138400 및 US 2013/0243775를 참조하며, 이들 각각은 그 전체가 모든 목적을 위해 참조로서 본원에 통합된다.

[0124] 인간 ASGR1 단백질은 간에서 노출된 말단 갈락토오스 또는 N-아세틸갈락토사민 잔기를 갖는 당단백질의 세포 내 유입 및 리소좀 분해 매개를 매개하는 막관통 단백질이기 때문에, 인간 ASGR1과 복합체화된 세포 표면 단백질은 인간 ASGR1과 함께 내재화되거나, 분해 구획에 표적 세포 표면 단백질 또는 표적 가용성 단백질의 경로를 변경하거나, 표적 세포 표면 단백질 또는 내부 구획 또는 엑소좀에서 표적 가용성 단백질을 격리시킬 수 있다.

[0125] 따라서, 인간 ASGR1을 통해 간으로 표적 간 세포 표면 단백질 또는 간에서 표적 수용성 단백질을 내재화하기 위해 설계된 치료 복합체의 *생체 내* 효능을 평가하기 위해 본원에서 다른 곳에 기술된 바와 같이 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물을 사용하기 위한 방법이 본원에서 제공된다. 예를 들어, 이러한 방법은 (a) 본원의 다른 곳에서 기술된 비인간 동물에 치료 복합체를 투여하는 단계, 여기서 치료 복합체는 표적 세포 표면 단백질 또는 표적 가용성 단백질에 특이적으로 결합하고 인간 ASGR1에 특이적으로 결합하며 인간 ASGR1에 특이적으로 결합하는 이중특이적 항원-결합 단백질을 포함하고, 여기서 ASGR1은 표적 세포 표면 단백질 또는 표적 가용성 단백질의 내재화를 매개함; 및 (b) 표적 간 세포 표면 단백질의 세포 표면 수준 또는 활성을 평가하거나, 비인간 동물의 간에서의 표적 가용성 단백질의 세포 표면 수준 또는 활성을 평가하는 단계를 포함할 수 있다. 치료 복합체의 투여는 본원의 다른 곳에서 기술된 바와 같이, 임의의 적절한 수단에 의한 것일 수 있다.

[0126] 표적 세포 표면 단백질은 간에서 발현된 임의의 세포 표면 단백질을 포함할 수 있다. 표적 가용성 단백질은 간에서 발현된 임의의 가용성 단백질을 포함할 수 있다. 비인간 동물의 간에서 표적 간 세포 표면 단백질의 세포 표면 수준 또는 활성은 수용체 또는 다른 단백질의 세포 표면 수준을 측정할 때 알려진 분석법에 의해 평가될 수 있다. 마찬가지로, 가용성 단백질의 수준은 알려진 분석법에 의해 평가될 수 있다.

[0127] **C. 비인간 동물에 분자를 투여하는 단계**

[0128] 본원에 개시된 방법은 핵산, 단백질, 또는 단백질 복합체를 포함하여, 하나의 다양한 분자를 비인간 동물 내에 도입하는 단계를 포함한다. 도입하는 단계는 임의의 수단에 의해 달성될 수 있다. 예를 들어, 이러한 분자는 비인간 동물 내에, 예컨대, 벡터 전달, 입자-매개 전달, 엑소좀-매개 전달, 지질-나노입자-매개 전달, 세포-관통-펩티드-매개 전달, 또는 이식형-장치-매개 전달에 의해 도입될 수 있다. 특정 실시예로서, 분자는 폴리(락트산)(PLA) 마이크로스피어, 폴리(D,L-락티드-코글리콜산)(PLGA) 마이크로스피어, 리포솜, 미셀, 역미셀, 지질 코클레이트(lipid cochleate), 또는 지질 미세관(lipid microtubule)과 같은 담체 안에서 세포 또는 비인간 동물 내에 도입될 수 있다. 비인간 동물로 전달의 일부 특정 실시예는 유체역학적 전달, 바이러스-매개 전달(예컨대, 아데노-연관 바이러스(AAV)-매개 전달), 및 지질-나노입자-매개 전달을 포함한다.

[0129] 핵산 및 단백질, 또는 다른 성분의 비인간 동물 내로 도입은 유체역학적 전달(HDD)에 의하여 달성될 수 있다. 유체 역학적 전달은 *생체 내*에서 세포 내 DNA 전달을 위한 완벽한 방법으로서 출현하였다. 실질 세포로의 유전자 전달을 위해, 필수 DNA 서열만이 선택된 혈관을 통해 주입될 필요가 있고, 이것으로 현재 바이러스 및 합성

벡터와 연관된 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 순환 중인 비압축성 혈액으로의 신속한 주입에 의해 생성된 힘을 사용한다. DNA의 전달에 더불어, 이 방법은 RNA, 단백질, 및 다른 작은 화합물의 생체 내에서의 효율적인 세포 내 전달에 유용하다. 예를 들어, Bonamassa 외 (2011) *Pharm. Res.* 28(4):694-701을 참조하고, 그 전체는 모든 목적을 위해 참조로서 본원에 통합된다.

[0130] 핵산의 도입은 또한 AAV-매개 전달 또는 렌티바이러스-매개 전달 같은, 바이러스-매개 전달에 의해 달성될 수 있다. 다른 예시적인 바이러스/바이러스벡터는 레트로바이러스, 아데노바이러스, 우두 바이러스, 폭스바이러스 및 단순 포진 바이러스를 포함한다. 바이러스는 분열세포, 비분열 세포, 또는 분할과 비분열 세포 모두를 감염시킬 수 있다. 바이러스는 숙주 게놈 내로 통합될 수 있거나, 대안적으로 숙주 게놈 내로 통합되지 않는다. 이러한 바이러스는 또한 감소된 면역성을 갖도록 설계될 수 있다. 바이러스는 복제 능력을 갖거나 복제-결합(예컨대, 비리온 복제 및/또는 패키징의 추가 횟수에 필요한 하나 이상의 유전자에서 결합)일 수 있다. 바이러스는 일시적인 발현, 장기간 발현(예컨대, 적어도 1주, 2주, 1개월, 2개월, 또는 3개월), 또는 영구적 발현(예컨대, Cas9 및/또는 gRNA)을 유발할 수 있다. 예시적인 바이러스 역가(예: AAV 역가)는  $10^{12}$ ,  $10^{13}$ ,  $10^{14}$ ,  $10^{15}$ , 및  $10^{16}$  벡터 게놈/mL를 포함한다.

[0131] ssDNA AAV 게놈은 상보적 DNA 가닥의 합성을 허용하는 두 역방향 말단 반복부(inverted terminal repeat)에 의해 측면에 위치한 두 오픈 리딩 프레임인, Rep 및 Cap으로 이루어진다. AAV 전달 플라스미드를 구성할 때, 두 ITR 사이, 및, Rep와 Cap 사이에 위치하는 이식유전자는 트랜스로 공급될 수 있다. Rep 및 Cap과 더불어, AAV는 아데노바이러스의 유전자를 함유하는 헬퍼 플라스미드를 필요로 할 수 있다. 이들 유전자(E4, E2a, 및 VA)는 AAV 복제를 매개한다. 예를 들어, 전달 플라스미드인, Rep/Cap, 및 헬퍼 플라스미드는 아데노바이러스 유전자 E1+를 함유하는 HEK293 세포로 형질감염되어서 감염성 AAV 입자들을 생성할 수 있다. 대안적으로, Rep, Cap, 및 아데노바이러스 헬퍼 유전자는 단일 플라스미드로 조합될 수 있다. 유사한 패키징 세포 및 방법들은 레트로바이러스와 같은 다른 바이러스에 사용될 수 있다.

[0132] AAV의 여러 혈청형을 확인하였다. 이들 혈청형은 그들이 감염시키는 세포의 유형(즉, 그들의 향성)이 상이하여, 특정 세포 유형의 우선적인 형질 도입을 허용한다. 간 조직을 위한 혈청형은 AAV7, AAV8, 및 AAV9, 그리고 구체적으로 AAV8을 포함한다.

[0133] 향성은 거짓형 분석을 통해 더 정제될 수 있는데, 이는 상이한 바이러스 혈청형으로부터의 캡시드 및 게놈의 혼합이다. 거짓형 바이러스의 사용은 형질 도입 효율을 개선할 뿐만 아니라, 향성을 바꿀 수 있다. 상이한 혈청형으로부터 유도된 하이브리드 캡시드는 또한 바이러스 향성을 바꾸기 위해 사용될 수 있다. 예를 들어, AAV-DJ는 8개의 혈청형으로부터의 하이브리드 캡시드를 함유하고 생체 내에서 광범위한 세포 유형에 걸쳐 높은 감염성을 나타낸다. AAV 혈청형은 또한 돌연변이를 통해 변형될 수도 있다. 다른 거짓유형/변형된 AAV 변이체는 AAV2/1, AAV2/6, AAV2/7, AAV2/8, AAV2/9, AAV2.5, AAV8.2, 및 AAV/SASTG를 포함한다.

[0134] 이식 유전자 발현을 가속화하기 위해, 자가-상보적 AAV(scAAV) 변이체를 사용할 수 있다. AAV는 AAV의 단일-가닥 DNA 게놈의 상보적 가닥을 합성하기 위한 세포의 DNA 복제 기전에 의존하기 때문에, 이식 유전자 발현은 지연될 수 있다. 이러한 지연을 해결하기 위해, 감염 시 자발적으로 어닐링할 수 있는 상보적인 서열을 함유하는 scAAV를 사용하여, 숙주 세포 DNA 합성에 대한 요건을 제거할 수 있다.

[0135] 패키징 용량을 증가시키기 위해, 더 긴 이식 유전자는 두 AAV 전달 플라스미드(첫 번째는 3' 스플라이스 공여자와 함께, 두 번째는 5' 스플라이스 수용자)사이에서 분할될 수 있다. 세포의 공동-감염 시, 이들 바이러스는 연쇄동일순서를 형성하고, 함께 스플라이싱되며, 전체 길이 이식 유전자가 발현될 수 있다. 이는 더 긴 이식 유전자 발현을 허용하지만, 발현은 덜 효율적이다. 용량을 증가시키기 위한 유사 방법은 상동성 재조합을 이용한다. 예를 들어, 이식 유전자는 두 개의 전사 플라스미드 사이에서 분할될 수 있지만, 공동-발현이 전체 길이 이식 유전자의 상동성 재조합 및 발현을 유도하도록 실질적인 서열 중첩을 가질 수 있다.

[0136] 핵산 및 단백질의 도입은 또한 지질 나노입자(LNP) 매개 전달에 의해 달성될 수 있다. 예를 들어, LNP-매개 전달은 RNA의 형태로 가이드 RNA를 전달하는 데 사용될 수 있다. 이러한 방법을 통한 전달은 가이드 RNA의 일시적인 존재를 초래하고, 생분해성 지질은 클리어런스를 개선하고, 내약성을 개선하고, 면역원성을 감소시킨다. 지질 제형은 그들의 세포 흡수를 개선하면서 분해로부터 생물학적 분자를 보호할 수 있다. 지질 나노입자는 분자간 힘에 의해 서로 물리적으로 연관된 복수의 지질 분자를 포함하는 입자이다. 이들은 마이크로스피어(유니라멜라(unilamellar) 및 다중라멜라(multilamellar) 소포, 예컨대, 리포솜을 포함함), 유화액에 분산된 상, 미셀,

또는 현탁액에서 분산상(internal phase)을 포함한다. 이러한 지질 나노입자는 전달을 위해 하나 이상의 핵산 또는 단백질을 캡슐화하는 데 사용될 수 있다. 양이온성 지질을 함유하는 제형은 핵산과 같은 폴리 음이온을 전달하는 데 유용하다. 포함될 수 있는 기타 지질은 중성 지질(즉, 하전되지 않거나 쌍성이온성 지질), 음이온성 지질, 형질감염을 증강시키는 헬퍼 지질, 및 나노 입자가 생체 내에 존재할 수 있는 시간 길이를 증가시키는 은폐 지질이다. 적절한 양이온성 지질, 중성 지질, 음이온성 지질, 헬퍼 지질, 및 은폐 지질의 실시예는 그 전문이 모든 목적을 위해 참조로 본원에 통합된, WO 2016/010840 A1에서 발견될 수 있다. 예시적인 지질 나노 입자는 양이온성 지질 및 하나 이상의 다른 성분을 포함할 수 있다. 일 실시예에서, 다른 성분은 콜레스테롤과 같은 헬퍼 지질을 포함할 수 있다. 또 다른 실시예에서, 다른 성분은 콜레스테롤 및 DSPC와 같은 중성 지질 같은 헬퍼 지질을 포함할 수 있다. 또 다른 실시예에서, 다른 성분은 콜레스테롤 및 DSPC와 같은 임의의 중성 지질, 및 S010, S024, S027, S031, 또는 S033과 같은 은폐 지질과 같은 헬퍼 지질을 포함할 수 있다.

[0137] 전달 모드는 면역원성을 감소시키도록 선택될 수 있다. 상이한 모드는 대상체 전달 분자에 상이한 약력학 또는 약동학적 특성을 부여할 수 있다. 예를 들어, 상이한 모드는 상이한 조직 분포, 상이한 반감기, 또는 상이한 일시적 분포를 초래할 수 있다. 일부 전달 모드(예컨대, 자율 증식 또는 게놈 통합에 의해 세포에서 지속되는 핵산 벡터의 전달)는 보다 지속적인 발현 및 분자의 존재를 초래하는 반면, 다른 전달 모드는 일시적이고 덜 지속적이다(예컨대, RNA 또는 단백질의 전달). 더 일시적인 방식으로 성분의 전달은, 예를 들어, RNA 또는 단백질로서, 성분이 단시간 동안만 존재하고 활성화되고 면역원성을 감소시킬 수 있다는 것을 보장할 수 있다.

[0138] 생체 내 투여는, 예를 들어, 비경구, 정맥 내, 피하, 동맥 내 또는 복강 내(intraperitoneal)를 포함하는 임의의 적절한 경로에 의할 수 있다. 전신 투여 모드는, 예를 들어, 정맥 내, 동맥 내, 피하, 및 복강 내 경로와 같은 비경구 경로를 포함한다. 특정 실시예는 정맥 내 주입이다. 국소 투여 모드는 또한 사용할 수 있다.

[0139] 핵산 또는 단백질을 포함하는 조성물은 하나 이상의 생리학 및 약학적으로 허용되는 담체, 희석제, 부형제 또는 보조제를 사용하여 제형화될 수 있다. 제형은 선택된 투여 경로에 따라 달라질 수 있다. 용어 "약학적으로 허용되는"은 담체, 희석제, 부형제 또는 보조제가 제제의 다른 성분과 양립하고 이의 수용자에게 실질적으로 유해하지 않다는 것을 의미한다.

[0140] 투여 빈도와 투여량은 투여 횟수는 도입된 성분의 반감기와 다른 인자들 중 투여 경로에 따라 달라질 수 있다. 핵산 또는 단백질을 비인간 동물 내로 도입하는 것은 한 번 수행되거나 일정 시간에 걸쳐 여러 번 수행될 수 있다. 예를 들어, 도입은 일정 기간에 걸쳐서 적어도 2회, 일정 기간에 걸쳐서 적어도 3회, 일정 기간에 걸쳐서 적어도 4회, 일정 기간에 걸쳐서 적어도 5회, 일정 기간에 걸쳐서 적어도 6회, 일정 기간에 걸쳐서 적어도 7회, 일정 기간에 걸쳐서 적어도 8회, 일정 기간에 걸쳐서 적어도 9회, 일정 기간에 걸쳐서 적어도 10회, 일정 기간에 걸쳐서 적어도 11회, 일정 기간에 걸쳐서 적어도 12회, 일정 기간에 걸쳐서 적어도 13회, 일정 기간에 걸쳐서 적어도 14회, 일정 기간에 걸쳐서 적어도 15회, 일정 기간에 걸쳐서 적어도 16회, 일정 기간에 걸쳐서 적어도 17회, 일정 기간에 걸쳐서 적어도 18회, 일정 기간에 걸쳐서 적어도 19회, 또는 일정 기간에 걸쳐서 적어도 20회 수행될 수 있다.

[0141] **IV. 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물을 만드는 방법**

[0142] 본원의 다른 곳에 개시된 바와 같이 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물을 만드는 다양한 방법이 제공된다. 이러한 유전적으로 변형된 비인간 동물을 생산하는 데는 유전적으로 변형된 유기체를 생산하기 위한 임의의 편리한 방법 또는 프로토콜이 적합하다. 예를 들어, Cho 외 (2009) *Current Protocols in Cell Biology* 42:19.11:19.11.1-19.11.22 및 Gama Sosa 외 (2010) *Brain Struct. Funct.* 214(2-3):91-109를 참조하고, 그 전체는 모든 목적을 위해 참조로서 본원에 통합된다. 이러한 유전적으로 변형된 비인간 동물은, 예를 들어, 표적화된 *Asgr1* 유전자좌에서 유전자 녹인(knock-in)을 통해 생성될 수 있다.

[0143] 예를 들어, 인간화 *Asgr1* 유전자좌를 포함하는 비인간 동물을 생산하는 방법은: (1) 인간화 *Asgr1* 유전자좌를 포함하도록 다능성 세포의 게놈을 변형시키는 단계; (2) 인간화 *Asgr1* 유전자좌를 포함하는 유전적으로 변형된 다능성 세포를 식별 또는 선택하는 단계; (3) 유전적으로 변형된 다능성 세포를 비인간 동물 숙주 배아에 도입하는 단계; (4) 대리모에 숙주 배아를 이식 및 임신시키는 단계를 포함할 수 있다. 임의로, 변형된 다능성 세포(예를 들어, 비인간 ES 세포)를 포함하는 숙주 배아는 대리모에 이식되고 임신되기 전에 배반포 단계까지 배양되어 F0 비인간 동물을 생산할 수 있다. 대리모는 그런 다음 인간화 *Asgr1* 유전자좌를 포함하는 F0 세대 비인간 동물을 생산할 수 있다.

[0144] 방법은 변형된 표적 게놈 유전자좌를 갖는 세포 또는 동물을 동정하는 단계를 추가로 포함할 수 있다. 표적화된

유전적 변형을 갖는 세포 및 동물을 동정하기 위해 다양한 방법이 사용될 수 있다.

- [0145] 스크리닝 단계는, 예를 들어, 부모 염색체의 대립유전자(MOA)의 변형을 평가하기 위한 정량 분석을 포함할 수 있다. 예를 들어, 정량 분석은 실시간 PCR과 같은 정량 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 이용할 수 있다. 프라이머 세트는 증폭된 서열을 인식하는 형광 프로브를 포함할 수 있다.
- [0146] 적절한 정량적 검정의 다른 예에는, 형광-매개 인시츄 혼성화(fluorescence-mediated in situ hybridization; FISH), 비교 게놈 혼성화(comparative genomic hybridization), 등온 DNA 증폭(isothermic DNA amplification), 고정 프로브(immobilized probe(s)), INVADER<sup>®</sup> 프로브, TAQMAN<sup>®</sup> 분자 비콘 프로브, 또는 ECLIPSE<sup>®</sup> 프로브 기술이 포함된다(예를 들어, US 2005/0144655를 참조하며, 그 전체는 모든 목적을 위해 참조로서 본원에 통합됨).
- [0147] 적절한 다능성 세포의 실시예는 배아 줄기(ES) 세포(예컨대, 마우스 ES 세포 또는 랫트 ES 세포)이다. 변형된 다능성 세포는, 예를 들어, (a) 5' 및 3' 표적 부위에 상응하는 5' 및 3' 상동성 아암(arm)에 의해 측면에 위치한 삽입 핵산을 포함하는 하나 이상의 표적화 벡터를 세포 내에 도입하는 단계, 여기서 삽입 핵산은 인간화 *Asgr1* 유전자좌를 포함함; 및 (b) 표적 게놈 유전자좌에 통합된 삽입 핵산을 그 게놈에서 포함하는 적어도 하나의 세포를 동정하는 단계에 의한 재조합을 통해, 생산될 수 있다. 대안적으로, 변형된 다능성 세포는 (a) 세포 내로 도입되는 단계: (i) 뉴클레아제 제제, 여기서 뉴클레아제 제제는 표적 게놈 유전자좌 내의 인식 부위에서 Nick(nick) 또는 이중 가닥 절단을 유도함; 및 (ii) 인식 부위에 충분히 근접하여 5' 및 3'; 표적 부위에 상응하는 5' 및 3' 상동성 아암에 의해 측면에 위치하는 삽입 핵산을 포함하는 하나 이상의 표적화 벡터, 여기서 삽입 핵산은 인간화 *Asgr1* 유전자좌를 포함함; 및 (c) 표적 게놈 유전자좌에서 변형(예컨대, 삽입 핵산의 통합)을 포함하는 적어도 하나의 세포를 식별하는 단계에 의해 생산될 수 있다. 원하는 인식 부위에 Nick 또는 이중 가닥 절단을 유도하는 임의의 뉴클레아제 제제를 사용할 수 있다. 적절한 뉴클레아제의 실시예는 전사 활성제-유사 이펙터 뉴클레아제(TALEN), 징크-핑거 뉴클레아제(ZFN), 메가뉴클레아제, 및 규칙적으로 반복되는 짧은 팔린드롬 조각(CRISPR)/CRISPR-연관(Cas) 시스템 또는 이러한 시스템의 성분(예컨대, CRISPR/Cas9)을 포함한다. 예를 들어, US 2013/0309670 및 US 2015/0159175를 참조하며, 이들 각각은 그 전체가 모든 목적을 위해 참조로서 본원에 통합된다.
- [0148] 공여자 세포는 배반포 단계 또는 상실배 전단계(즉, 4세포 단계 또는 8세포 단계)와 같은 임의 단계의 숙주 배아에 도입될 수 있다. 생식선을 통해 유전적 변형을 전달할 수 있는 자손이 생산된다. 예를 들어, 미국 특허 제 7,294,754호를 참조하며, 그 전체는 모든 목적을 위해 참조로서 본원에 통합된다.
- [0149] 대안적으로, 본원의 다른 곳에서 기술된 비인간 동물을 생산하는 방법은: (1) 다능성 세포를 변형하기 위해 상기 기술된 방법을 사용하여 인간화 *Asgr1* 유전자좌를 포함하도록 1세포 단계의 게놈을 변형시키는 단계; (2) 유전적으로 변형된 배아를 선택하는 단계; 및 (3) 대리모에 유전적으로 변형된 배아를 이식하고 임신시키는 단계를 포함할 수 있다. 생식선을 통해 유전적 변형을 전달할 수 있는 자손이 생산된다. 핵 전달 기술도 비인간 포유류 동물을 생성하는 데 사용될 수 있다.
- [0150] 간략하게, 핵 전달을 위한 방법은: (1) 난모 세포를 탈핵시키거나 탈핵된 난모 세포를 제공하는 단계; (2) 탈핵된 난모 세포와 결합될 공여자 세포 또는 핵을 단리하거나 제공하는 단계; (3) 공여자 세포 또는 핵을 탈핵된 난모 세포에 삽입하여 재구성 세포를 형성시키는 단계; (4) 재구성 세포를 동물의 자궁에 이식하여 배아를 형성시키는 단계; 및 (5) 배아를 성장시키는 단계를 포함할 수 있다. 이러한 방법에서, 난모 세포를 살아있는 동물의 난관(oviducts) 및/또는 난소(ovaries)에서 단리할 수도 있지만, 일반적으로는 죽은 동물로부터 수거한다. 탈핵에 앞서, 잘 알려진 다양한 배지에서 난모 세포를 성숙시킬 수 있다. 난모세포의 탈핵은 잘 알려진 다수의 방식으로 수행될 수 있다. 탈핵된 난모 세포 내에 공여자 세포 또는 핵을 삽입하여 재구성 세포를 형성시키는 것은, 융합 이전에 투명대(zona pellucida) 아래에 공여자 세포를 미세주입하는 것에 의한 것일 수 있다. 융합은, 접촉/융합면에 걸쳐 DC 전기 펄스를 적용(전기 융합)시키거나, 폴리에틸렌 글리콜과 같은 융합 촉진 화학물질에 세포를 노출시키거나, 센다이 바이러스(Sendai virus)와 같은 비활성화 바이러스를 통해 유도될 수 있다. 재구성 세포는 핵 공여자와 수여자 난모 세포의 융합 이전, 도중 및/또는 이후에 전기적 및/또는 비전기적 수단에 의해 활성화될 수 있다. 활성화 방법에는 전기 펄스, 화학적으로 유도된 충격, 정자에 의한 침투, 난모 세포에서 2가 양이온 레벨의 증가, 및 (키나아제 억제제를 통한) 난모 세포에서 세포 단백질의 인산화 감소가 포함된다. 활성화된 재구성 세포 또는 배아는 잘 알려진 배지에서 배양된 후, 동물의 자궁에 옮겨질 수 있다. 예를 들어, US 2008/0092249, WO 1999/005266, US 2004/0177390, WO 2008/017234, 및 미국 특허 제 7,612,250호를 참

조하며, 이들 각각은 그 전체가 모든 목적을 위해 참조로서 본원에 통합된다.

[0151] 본원에 제공된 다양한 방법을 사용해 유전적으로 변형된 비인간 F0 동물을 생산할 수 있으며, 유전적으로 변형된 F0 동물의 세포는 인간화 *Asgr1* 유전자좌를 포함한다. 인간화 *Asgr1* 유전자좌를 갖는 F0 동물 내의 세포의 수는 F0 동물을 생성하는 데 사용되는 방법에 따라 달라질 것임을 알 수 있을 것이다. 예를 들어, VELOCIMOUSE<sup>®</sup> 방법을 통해 상용 유기체(예: 8세포 단계 마우스 배아)의 상실 전단계 배아 내에 공여자 ES 세포를 도입하면, F0 동물의 세포 집단 중 더 많은 백분율이 표적화된 유전적 변형을 포함하여 관심있는 뉴클레오타이드 서열을 갖는 세포를 포함할 수 있다. 예를 들어, 비인간 F0 동물에 기여한 세포의 적어도 50%, 60%, 65%, 70%, 75%, 85%, 86%, 87%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100%는 표적화된 변형을 갖는 세포 집단을 포함할 수 있다.

[0152] 유전적으로 변형된 F0 동물의 세포는 인간화 *Asgr1* 유전자좌에 대한 이형접합체일 수 있거나, 인간화 *Asgr1* 유전자좌에 대한 동형접합체일 수 있다.

[0153] 상기 또는 하기 인용된 모든 특허 출원, 웹사이트, 다른 간행물, 수탁 번호 등은 각각의 개별 항목이 구체적으로 및 개별적으로 그렇게 참조로 통합되도록 지시되어 있는 것처럼 그와 동일한 정도로 모든 목적을 위해 전체적으로 참조로 통합된다. 상이한 버전의 서열이 다른 시점에서 수탁 번호와 연관되는 경우, 본 출원의 유효 출원일에서 수탁 번호와 관련된 버전을 의도된다. 유효 출원일은 실제 출원일 또는 적용가능한 경우 수탁 번호를 언급하는 우선 출원의 출원일 중 더 빠른 것을 의미한다. 마찬가지로 다양한 버전의 간행물, 웹사이트 등이 다양한 시점에서 공개된 경우, 달리 지시되지 않는 한 본 출원의 유효 출원일에서 가장 최근에 공개된 버전이 의도된다. 본 발명의 임의의 특징부, 단계, 요소, 실시 형태, 또는 태양은 달리 구체적으로 지시되지 않는 한 임의의 다른 것과 조합하여 사용될 수 있다. 비록 본 발명이 명료함 및 이해를 위해 예시와 실례에 의하여 일부 상세하게 설명되긴 하지만, 일정한 변화와 변형이 첨부된 청구항의 범위 내에서 실시될 수 있다는 것은 명백할 것이다.

[0154] **서열의 간단한 설명**

[0155] 첨부하는 서열 목록에서 열거된 뉴클레오타이드 및 아미노산 서열은 뉴클레오타이드 염기에 대하여 표준 문자 약어, 그리고 아미노산에 대하여 3-문자 코드를 사용하여 도시된다. 뉴클레오타이드 서열은 서열의 5' 말단에서 개시하고 3' 말단까지 정방향(즉, 각각의 라인내 좌부터 우까지) 진행되는 표준 협약을 따른다. 각각의 뉴클레오타이드 서열의 단 하나의 가닥은 보여지지만, 상보성 가닥은 표시된 가닥의 임의의 참조에 의해 포함되도록 이해된다. 아미노산 서열은 서열의 아미노 말단에서 개시하고 카복시 말단까지 정방향(즉, 각각의 라인내 좌부터 우까지) 진행되는 표준 협약을 따른다.

**표 1**

**서열의 설명**

[0156]

서열 번호	유형	설명
1	단백질	hASGR1 - P07306
2	단백질	mAsgr1 - P34927
3	단백질	7302 인간화 <i>Asgr1</i>
4	DNA	mAsgr1 암호화 서열 - NM_009714(뉴클레오타이드 154-1008)
5	DNA	hASGR1 암호화 서열 - NM_001671(뉴클레오타이드 401-1276)
6	DNA	7302mTU Fwd
7	DNA	7302mTU 프로브(BHQ)
8	DNA	7302mTU Rev
9	DNA	7302mTD Fwd
10	DNA	7302mTD 프로브(BHQ)
11	DNA	7302mTD Rev
12	DNA	7302hTU Fwd
13	DNA	7302hTU 프로브(BHQ)
14	DNA	7302hTU Rev
15	DNA	7302hTD Fwd
16	DNA	7302hTD 프로브(BHQ)
17	DNA	7302hTD Rev
18	DNA	7302 경계 A

19	DNA	7302 경계 B
20	DNA	7302 경계 C
21	DNA	7302 대립유전자
22	DNA	7303 경계 A
23	DNA	7303 경계 D
24	DNA	7303 대립유전자
25	단백질	c-Myc 에피토프
26	DNA	c-Myc 에피토프
27	단백질	hASGR1 꼬인 나선 도메인
28	단백질	hASGR1 C형 렉틴 도메인
29	단백질	mASGR1 세포질 도메인
30	단백질	mASGR1 막관통 도메인
31	단백질	Asgr1의 인간화 영역에 의해 암호화되는 ASGR1 분절
32	DNA	pAAVRCBs iWF 프라이머
33	DNA	N587mycR 프라이머
34	DNA	N587mycF 프라이머
35	DNA	pAAVRCPmeR 프라이머
36	DNA	AAV2 ITR-F 프라이머
37	DNA	AAV2 ITR-R 프라이머
38	DNA	AAV2 ITR 프로브
39	단백질	Macaca Asgr1 - XP_005582755.1
40	DNA	Macaca Asgr1 암호화 서열 - XM_005582698.2(뉴클레오티드 310-1185)
41	단백질	랫트 Asgr1- NP_036635.1
42	DNA	랫트 Asgr1 암호화 서열 - NM_012503.2(뉴클레오티드 66-920)

[0157] 실시예

[0158] 실시예 1. 인간화 *Asgr1* 유전자좌를 포함하는 마우스 생산

[0159] 아시알로당단백질 수용체(Ashwell Receptor, ASGR)는 렉틴 수용체의 C형 클래스에 속하는 주로 간 막 결합 탄수화물 결합 단백질 수용체(~33 kDa)이다. ASGR은 말단 갈락토오스 또는 N-아세틸갈락토사민(GalNac) 모티프를 갖는 당단백질과 결합한다. 이는 수용체 매개 세포 내 유입에 의한 순환으로부터 탈시알릴화 당단백질을 제거할 수 있다. 리간드는 리소좀 분해를 겪지만, ASGR은 세포 표면으로 다시 재활용된다. 수용체는 두 개의 서브유닛인, ASGR1 및 ASGR2(H1 및 H2)의 이종올리고머이다. 고효능 세포 내 유입 간 특이적 수용체로서, ASGR1은 항체, 작은 분자(항체-약물 접합체의 일부로서) 및 DNA와 같은 치료제의 간 특이적 전달을 위해 사용될 수 있다. 그러나, 인간 ASGR1 세포 외 도메인에 대해 생성된 항체 및 이종특이적 항체는 마우스 *Asgr1* 동원체(Biacore 데이터 미도시)에 결합하지 않는다. 따라서, 우리는 다수의 상이한 접근법을 이용하여 상이한 치료제의 간-특이적 전달의 검증에 사용하기 위해 인간화 *Asgr1* 마우스를 생산하였다.

[0160] bMQ-69B11로부터 24 kb를 포함하는 5' 상동성 아암 및 bMQ-69B11로부터 67 kb를 포함하는 3' 상동성 아암은 생성되어 *ASGR1*의 인간 서열에 상응하는 마우스 *Asgr1*의 암호화 엑손 3~8(종결 코돈 및 3' UTR을 통해, 3' UTR의 바로 하류 서열 내에)을 대체하였다. 암호화된 *Asgr1* 단백질은 마우스 *Asgr1* 막관통 도메인에 이어서 인간 꼬인 나선 및 C형 렉틴 도메인을 가질 것이다. **참조: 도 2a 및 3** 돌연변이 대립유전자를 생성하기 위해, CRISPR/Cas9 성분을 거대 표적화 벡터와 함께 F1H4 마우스 배아 줄기 세포에 도입하였다. 프라이머 및 **표 2**에 제시된 프로브를 사용하여 대립유전자 상실 분석을 수행하여 내인성 마우스 대립유전자 상실을 검출하였고, 프라이머 및 **표 3**에 제시된 프로브를 사용하여 대립 유전자 획득 분석을 수행하여 인간화 대립유전자 획득을 검출하였다. 대립유전자 상실 및 대립유전자 획득 검정은, 예를 들어 US 2014/0178879, US 2016/0145646, WO 2016/081923, 및 Friendewey 외 (2010) *Methods Enzymol.* 476:295-307에 기술되어 있고, 그 전체는 모든 목적을 위해 참조로서 본원에 통합된다.

**표 2**

[0161] 마우스 TAQMAN® 대립유전자 상실 분석

LOA 분석	프라이머/프로브	서열
7302mTU	Fwd	TCCCAACTCCGGGAAGATC(서열 번호 6)

	프로브(BHQ)	TGCTGGCTCTAAGGCAGAATTTCA(서열 번호 7)
	Rev	TCAGTGCTCACAGTGAGGTT(서열 번호 8)
7302mTD	Fwd	GGGTTGGCTCATGTTAGGAAGG(서열 번호 9)
	프로브(BHQ)	TCAGCAGCCGAGCTGTGAAA(서열 번호 10)
	Rev	CAGGCTGTGCTACCCAAAGTTC(서열 번호 11)

표 3

인간 TAQMAN® 대립유전자 획득 분석

GOA 분석	프라이머/프로브	서열
7302hTU	Fwd	GGAGGCAATGTGGGAAGAAAG(서열 번호 12)
	프로브(BHQ)	TGAAGTCGCTAGAGTCCCAGCTGG(서열 번호 13)
	Rev	TCAGGTCCTTCTGCTGTTTC(서열 번호 14)
7302hTD	Fwd	GATTGGGAATCCGCCATCT(서열 번호 15)
	프로브(BHQ)	CCTCTTCTGCTTCTCGGGAATTTTCATC(서열 번호 16)
	Rev	AAAGCGCCACGGGTTTCAAG(서열 번호 17)

[0162]

[0163]

FO 마우스를 그린 다음 VELOCIMOUSE® 방법을 사용하여 생산하였다. 예를 들어, US 7,576,259, US 7,659,442, US 7,294,754, US 2008/007800, 및 Poueymirou 외 (2007) *Nature Biotech.* 25(1):91-99를 참조하고, 그 전체는 모든 목적을 위해 참조로서 본원에 통합된다. 암호화 엑손 3의 인트론 상류의 일부가 또한 인간화되었다. 모든 경우, 1863 bp의 마우스 *Asgr1* 서열을 3907 bp의 인간 *ASGR1* 서열로 대체하였다. loxP-Hub1-em7-히그로마이신 저항성 유전자-폴리아데닐화 신호-mPrml1-Crei-폴리아데닐화 신호-loxP 카세트(5218 bp)를 또한 카세트 직전에 3' UTR의 대략 190 bp의 3' 인간 서열의 완충제과 함께, 인간 3' UTR의 하류에 삽입하였다. 히그로 마이신 저항성 자가-결실 카세트를 갖는, 생성된 부분적 인간화 마우스 *Asgr1* 대립유전자를 서열 번호 21에 제시하였다(7302 대립유전자를 참조함). 도 2a에서 서열 경계 영역 A, B, 및 C를 서열 번호 18, 19 및 20에 각각 제시하였다. 인간 ASGR1 단백질(서열 번호 1), 마우스 *Asgr1* 단백질(서열 번호 2), 및 부분적 인간화 마우스 *Asgr1* 단백질(서열 번호 3)의 비교를 도 3에 도시하였다.

[0164]

Cre 재조합효소로 자가-결실 카세트를 제거하면, loxP 및 클로닝 부위(77 bp)는 인간 3' UTR의 하류에 남아있고, 나머지 loxP 바로 직전에 3' UTR 이후 대략 190 bp의 3' 인간 서열의 완충제가 있다. 참조: 도 2b. 생성된 부분적 인간화 마우스, 카세트-결실 *Asgr1* 대립유전자를 서열 번호 24에 제시하였다(7303 대립유전자를 참조함). 도 2b에서 서열 경계 영역 A 및 C를 서열 번호 22 및 23에 각각 제시하였다.

[0165]

실시예 2. 인간화 *Asgr1* 유전자좌를 포함하는 마우스 검증

[0166]

검증 모델로서 인간화 *Asgr1* 마우스(*Asgr1*<sup>hu/hu</sup>)를 검증하기 위해, *Asgr1*<sup>hu/hu</sup> 마우스를 표현하고, 이들의 표현형을 야생형 한배 새끼의 표현형과 비교하였다. *Asgr1*<sup>hu/hu</sup> 마우스는 야생형 한배 새끼와 비교하여 혈장 지질 수준(총 콜레스테롤, 중성지방, HDL-C, LDL-C)에 차이가 없음을 보여주었다. 참조: 도 4. 마찬가지로, *Asgr1*<sup>hu/hu</sup> 마우스는 야생형 한배 새끼와 비교하여 체중 또는 혈당 수준에 차이가 없음을 보여주었다. 참조: 도 5. 인간 ASGR1 단백질을 마우스 *Asgr1*과 유사하게, 간 막에 공동 국소화하였다. 참조: 도 6. 결론적으로, *Asgr1*<sup>hu/hu</sup> 마우스는 간 막에 인간 ASGR1 단백질을 발현하며, 정상 플라즈마 지질 프로파일을 가진다.

[0167]

방법

[0168]

**Asgr1 인간화 마우스에서 순환 지질 수준 평가.** 수컷 *Asgr1* 인간화 마우스(*Asgr1*<sup>hum/hum</sup>) 및 그들의 야생형 한배 새끼(*Asgr1*<sup>+/+</sup>)로부터의 혈장을 비-공복 조건에서 수집하고 혈청 화학 분석기 ADVIA® 화학 XPT 시스템(지멘스(Siemens))을 사용하여 혈청 지질(중성 지질(TG), 총 콜레스테롤, 저밀도 리포단백질 콜레스테롤(LDL-C), 고밀도 리포단백질 콜레스테롤(HDL-C))을 분석하였다. N=8/그룹, 11주령. 데이터는 각각에 대한 평균 ± SEM으로 표현함.

[0169]

**혈당 평가.** 공복(16시간) 및 급식(비-공복) 상태에서 Accu-Chek 혈당계(로슈)를 사용하여 꼬리 팁으로부터 혈당

을 측정하였다.

[0170] **Asgr1의 웨스턴 블롯 평가.** 전체 간을 *Asgr1* 인간화 마우스 및 WT 마우스(n=8/유전자형)로부터 수확하고, 처리할 때까지 -80°C에서 동결 저장하였다. 각 시료에 대해, 동결된 간의 ~40 mg 조각을 전체 간으로부터 절제하고, 각 조각을 다운스(dounce) 균질기에 넣고 균일한 현탁액을 수득할 때까지 균질화시켰다. 각 간 시료의 세포질 분획과 막 분획을 세제-기반 상업용 키트(Thermo #89842)를 사용하여, 제조사 "연조직" 프로토콜에 따라 분리하였다. 일단 각 시료의 세포질 분획과 막 분획을 분리하고, 마이크로플레이트 절차를 위한 키트 프로토콜에 따라, BCA 단백질 정량 분석(Thermo #23225)을 각 시료의 두 분획에 대해 실행하였다. 웨스턴 블롯 시료를 5X 환원 염료를 사용하여 각 시료의 각 세포질 분획과 막 분획에 대해 제조하였다. 모든 시료를 0.8 µg/µL로 준비하고, 각 시료의 총 단백질 중 20 µg을 웨스턴 블롯 겔에 로딩하였다. 토끼 항-hASGR1 다클론 항체(Abgent #AP16133a, TBS-T에서 2.5% 차단유(blocking milk) 중 1:1,000 희석으로)를 사용하여 hASGR1을 검출하였다. 항체가 마우스 단백질에 대한 교차-반응성을 나타내기 때문에, 한배 새끼 마우스의 마우스 *Asgr1*을 동일한 항체로 검출하였다. hASGR1 검출을 위한 이차 항체는 당나귀 항-토끼 IgG-HRP(Jackson #711-035-152, TBS-T에서 2.5% 차단유 중 0.1 µg/mL로)이었다. 세포질 분획에 대한 로딩 대조군으로서, 토끼 항-GAPDH 단클론 항체(Cell Signaling #2118S, TBS-T에서 2.5% 차단유 중 1:10,000 희석으로)를 사용하여 모든 시료에서 GAPDH를 검출하였다. GAPDH 검출을 위한 이차 항체는 당나귀 항-토끼 IgG-HRP(Jackson #711-035-152, TBS-T에서 2.5% 차단유 중 0.1 µg/mL로)이었다. 막 분획에 대한 로딩 대조군으로서, 토끼 항-트랜스페린 수용체 다클론성 항체(R&D #AF2472, TBS-T에서 2.5% 차단유 중 0.25 µg/mL로)를 사용하여 모든 시료에서 트랜스페린 수용체를 검출하였다. 트랜스페린 수용체 검출을 위한 이차 항체는 당나귀 항-염소 IgG-HRP(Jackson #705-035-147, TBS-T에서 2.5% 차단유 중 0.2 µg/mL로)이었다.

[0171] **실시예 3. 이중 에피토프를 갖는 아데노-연관 바이러스 벡터의 생산**

[0172] 이어서, 이중 특이적 항-myc-ASGR1 항체가 scAAV-N587myc 바이러스 입자를 실시예 1에서 생산한 마우스에서 생체 내 인간화 ASGR1을 발현하는 간 세포로 재표적할 수 있는지의 여부를 결정하기 위해 실험을 수행하였다. 이를 시험하기 위해, 바이러스 입자를 먼저 생산하였다.

[0173] PCR을 사용해 AAV 캡시드 단백질을 여러 개의 이중 에피토프(FLAG, c-myc, 헥사히스티딘 등) 중 하나를 함유하도록 변형시켜, 재조합 캡시드 단백질을 암호화하는 플라스미드를 생성한다. 간략하게, FLAG, c-myc 또는 헥사히스티딘을 암호화하는 서열을 AAV2의 N587 또는 AAV6 VP1 캡시드 단백질의 Q585를 암호화하는 코돈 후에 프레임에 삽입한다.

[0174] 아데노-연관 바이러스 생산은 HEK293 세포를 사용해 삼중 형질감염 방법을 사용하여 수행한다(예를 들어, Erik Arden 및 Joseph M. Metzger의 *J Biol Methods*. 2016; 3(2) 참조). 세포를 플레이팅하고, 1일 후에 다음과 같은 적절한 벡터로 세포를 PEFpro(Polyplus transfection, 뉴욕주 뉴욕 소재)로 매개하여 형질감염시킨다:

[0175] - 헬퍼 플라스미드, pHelper(Agilent, Cat #240074);

[0176] - 야생형 또는 변형된 AAV rep/cap 유전자(pAAV RC2(Cell Biolabs, Cat# VPK-422), 예를 들어, pAAV RC2/6(Cell Biolabs, Cat# VPK-426), pAAV RC2-N587myc, pAAV RC2/6-Q585myc) 등을 암호화하는 플라스미드; 및

[0177] - 관심 뉴클레오티드 및 AAV ITR 서열, 예를 들어, pscAAV-CMV-eGFP, pAAV-CMVGFP(Agilent Cat# 240074), pAAV-EF1a-eGFP 또는 pAAV-CAGG-eGFP 등을 암호화하는 플라스미드.

[0178] 형질감염 후 72시간차에, 배지를 수집하고 완충액[50 mM 트리스-HCl, 150 mM NaCl, 및 0.5% 테옥시콜린산나트륨(Sigma, Cat# D6750-100G)] 중에서 세포를 용해시킨다. 다음으로, 벤조나아제(Sigma, 미주리주 세인트루이스 소재)를 0.5 U/µL의 최종 농도가 될 때까지 배지 및 세포 용해물 모두에 첨가한 후, 37°C에서 60분 동안 인큐베이션한다. 세포 용해물을 4000 rpm에서 30분 동안 회전시킨다. 세포 용해물 및 배지를 함께 조합하고, 최종 농도 8%의 PEG 8000(Teknova Cat# P4340)으로 침전시킨다. 펠릿을 400 mM NaCl에서 재현탁하고 10000 g에서 10분 동안 원심분리한다. 상등액 내 바이러스를 149,000 g에서 3시간 동안 초원심분리에 의해 펠렛화하고 qPCR에 의해 역가한다.

[0179] qPCR로 AAV 계수를 적정하기 위해, AAV 샘플을 DNaseI(ThermoFisher Scientific, Cat #EN0525)로 1시간 동안 37°C에서 처리하고, DNA extract All Reagents(ThermoFisher Scientific Cat# 4403319)를 사용해 용해시킨다. 캡시드화된 바이러스 계수를 QuantStudio 3 실시간 PCR 시스템을 사용하여, AAV2 ITR에 대해 유도된 프라이머를 사용해 정량화하였다. AAV2 ITR 프라이머의 서열은 5'-GGAACCCCTAGTGATGGAGTT-3'(정방향 ITR; 서열번호 36) 및 5'-CGGCCTCAGTGAGCGA-3'(역방향 ITR; 서열번호 37)(Aurnhammer 외, 2012)이며, 각각 AAV의 좌측 내부 역전 반

복(ITR) 서열 및 AAV의 우측 내부 반복(ITR) 서열로부터 유도된다. AAV2 ITR 프로브의 서열은 5'-6-FAM-CACTCCCTCTCTGCGCTCG-TAMRA-3'이다(서열 번호 38)(Aurnhammer 외 (2012) *Hum. Gene Ther. Methods* 23:18-28). 95°C에서 10분 동안의 활성화 단계 후, 95°C에서 15초 동안 및 60°C에서 30초 동안의 2단계 PCR 사이클을 40사이클 동안 수행한다. qPCR에는 TaqMan Universal PCR Master Mix(Thermofisher Scientific, Cat #4304437)를 사용하였다. DNA 플라스미드(Agilent, Cat #240074)를 표준 역가를 결정하기 위한 표준으로 사용한다.

[0180] c-myc 에피토프가 삽입된 캡시드를 포함하는 아데노-연관 바이러스 벡터를 생산하였다. 본 실시예에서, c-myc 에피토프(EQKLISEEDL; 서열 번호 25)를 AAV2 VP1 캡시드 단백질의 아미노산 N587과 R588 사이에 삽입하였는데, 다시 말해, cmyc 에피토프를 암호화하는 뉴클레오티드 서열(GAA CAA AAA CTC ATC TCA GAG GAT CTG; 서열번호 26)을 플라스미드 pAAV RC2(Cell Biolabs, Inc., 캘리포니아주 샌디에고 소재)에 삽입하였고, 변형된 pAAV RC2-N587Myc 플라스미드를 사용해, AAV 바이러스 벡터가 중화된 향성을 갖도록, 변형된 캡시드 단백질을 암호화하였다.

[0181] 구체적으로, (5'에서 3'까지) BsiW1 제한 부위, pAAV RC2의 위치 3050과 3773 사이의 뉴클레오티드 서열, 및 c-myc 에피토프 돌출 뉴클레오티드 서열을 포함하는 제1 중합효소 연쇄 반응(PCR) 산물 및 (5'에서 3'까지) c-myc 에피토프 돌출 뉴클레오티드 서열, pAAV RC2의 위치 3774와 4370 사이의 뉴클레오티드 서열, 및 Pme1 제한 부위를 포함하는 제2 PCR 산물을, 표 4에 제시된 프라이머를 사용해 생성하였다. pAAV RC2-N587Myc 플라스미드(즉, VP1 캡시드 단백질의 아미노산 N587과 R588 사이의 c-myc 에피토프를 암호화하도록 변형된 pAAV RC2 플라스미드)는 BsiW1(New England Biolabs, R0553L) 및 Pme1(New England Biolabs, R0560L)로 pAAV RC2를 분해하고, Li 외 (2012) *Methods Mol. Biol.* 852:51-59에 기술된 바와 같은 연결-독립적 클로닝을 통해 2개의 PCR 산물을 삽입함으로써 생성하였다.

표 4

PCR 산물	프라이머 명칭	5' -서열-3' 및 서열 번호
3050-3773	pAAVRCBsiWF	GGAGTACCAGCTCCCGTACG (BsiW1)(서열 번호 32)
pAAVRC2-cMyc	N587mycR	CTCTTCTGAGATGAGTTTTTGTTCGTTGCCTCTCTGGAGGTIG(서열 번호 33)
cMyc-3774-4370	N587mycF	AAACTCATCTCAGAAGAGGATCTGAGACAAGCAGCTACCGCAG(서열 번호 34)
pAAVRC2	pAAVRCPmeR	TCCGCCCGCTGTTAAAC (Pme1)(서열 번호 35)
밑줄 서열은 제한 효소 인식 부위를 나타낸다.		

[0183] 구체적으로, 3757과 3758 사이에 삽입된 c-myc-에피토프 서열을 갖는 pAAV RC2/6의 위치 3700과 3940을 포함하는 g블록 DNA 단편을 Integrated DNA Technologies (아이오와주 코랄빌 소재)에 주문하였다. pAAV RC2/6-Q585Myc 플라스미드는 Li 외 (2012) *Methods Mol. Biol.* 852:51-59에 기술된 바와 같은 연결 독립적 클로닝을 통해 MsCi(New England Biolabs, Cat# R0534L) 및 AfIII(New England Biolabs, Cat #R0520L)로 분해한 pAAV RC2/6 내에 g블록 단편을 삽입하여 생성하였다.

[0184] 구체적으로, BamHI 및 NotI 제한 부위를 사용하여 pscAAV MCS 벡터(Cell Biolabs, Cat# VPK-430) 내로 GFP 단편을 도입함으로써 pscAAV-CMV-eGFP를 생산하였다. pAAV-EF1a-eGFP 플라스미드 및 pAAV-CAGG-eGFP를 Thermofisher Scientific(메사추세츠주 월섬 소재)의 드 노보 합성(de novo synthesis)으로 만들었다.

[0185] **실시예 4. 생체 내 인간화 ASGR1 유전자좌가 있는 마우스에서 scAAV-N587Myc 입자의 이중특이적-항체-매개 내재화**

[0186] 이중특이적 항-myc-ASGR1 항체가 생체 내에서 scAAV2-N587myc-CMV-eGFP 바이러스 벡터를 hASGR1을 발현하는 간 세포로 재표적화할 수 있는지 여부를 결정하기 위해, C57BL/6 백그라운드에서 간 세포가 hASGR1을 발현하도록 유전자 변형된 마우스 및 대조군 야생형 C57BL/6 마우스에게 1x10<sup>11</sup>개의(qPCR에 의해 측정됨) 야생형 scAAV2-CMV-eGFP 한 가지 또는 이중특이적 항-myc-ASGR1 항체와 1:8의 바이러스 게놈 대 항체 분자의 비율로 조합된 scAAV2-N587myc-CMV-eGFP 바이러스 벡터를 혈관 내 주사하였다. 식염수[250 mM NaCl]를 주사한 마우스 또는 scAAV2-N587myc-CMV-eGFP 바이러스 벡터 한 가지만 주사한 마우스를 대조군에 포함시켰다. 주사 후 10일차에, 마우스를 희생시키고 4% PFA로 경심 관류시켰다. 간, 신장, 및 심장의 기관을 채취하여 15% 수크로스에 이어서 30% 수크로스로 탈수하였다. 그런 다음, 기관을 슬라이드 상에서 동결 절단하고 닭 항-EGFP 항체(Jackson ImmunoResearch Labs, Inc. 펜실베이니아주 웨스트그로브 소재) 및 알렉사-488 접합된 항-닭 2차 항체(Jackson

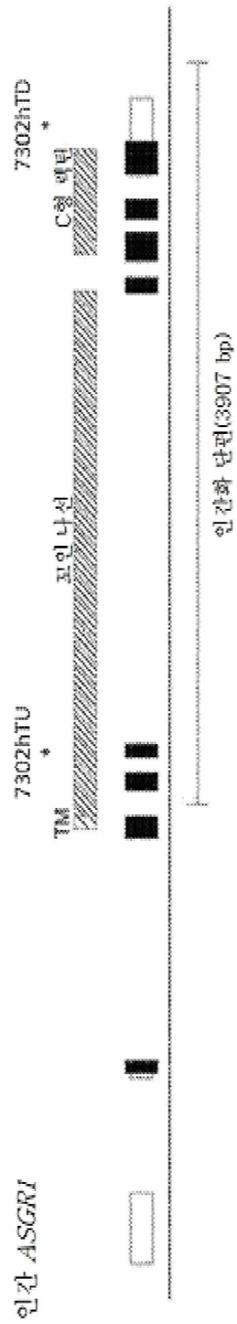
ImmunoResearch Labs, Inc. 상동)로 염색하였다(도 8a~8c). 간에서 ASGR1을 발현하도록 유전자 변형되고 야생형 scAAV2-CMV-eGFP가 주입되거나 scAAV2-N587myc-CMV-eGFP가 이중특이적 항-myc-ASGR1 항체와 함께 주입된 이들 유전자 이식 동물 유래의 간(도 8a(i) 및 8a(iv))과 야생형 scAAV2-CMV-eGFP가 주입된 야생형 C57BL/6 마우스의 간(도 8a(v))에서 GFP-양성 세포를 검출하였다. GFP는, 식염수가 주입되거나, scAAV2-N587myc-CMV-eGFP 한 가지가 주입된 동물의 비장이나 신장 샘플(도 8b 및 8c)에서 검출되지 않았고, 간, 비장, 또는 신장 샘플에서도 검출되지 않았으며(도 8a(ii, iii, vi, vii)), scAAV2-N587myc-CMV-eGFP가 이중특이적 항-myc-ASGR1 항체와 함께 주입된 야생형 C57BL/6 동물에서 채취한 간 샘플에서도 검출되지 않았다(도 8a(viii)). 요약하면, scAAV2-N587myc-CMV-eGFP 바이러스 벡터 및 이중특이적 항-myc-ASGR1 항체의 조합은 hASGR1을 발현하는 이들 (간) 세포만을 감염시켰는데, 이는 scAAV2-CMV-eGFP 바이러스 벡터가 캡시드 단백질의 변형에 의해 불활성화되었다는 것, 예를 들어, scAAV 바이러스 벡터의 천연 항성이 예컨대 c-myc 에피토프에 의해 중화될 수 있고, 이러한 바이러스 벡터가 예를 들어, 이중특이적 항-myc-ASGR1 항체에 의해 생체 내에서 특이적으로 재활성화될 수 있고, 예를 들어, 간 세포를 향해 특이적으로 재표적화될 수 있다는 것을 강하게 시사한다.

[0187] 유사하게, 이중특이적 항-myc-ASGR1 항체가 생체 내에서 ssAAV2-N587myc-CAGG-eGFP 바이러스 벡터를 hASGR1을 발현하는 간 세포로 재표적화할 수 있는지 여부를 결정하기 위해, C57BL/6 백그라운드에서 간 세포가 hASGR1을 발현하도록 유전자 변형된 마우스 및 대조군 야생형 C57BL/6 마우스에게  $2.18 \times 10^{11}$ 개의(qPCR에 의해 측정됨) 야생형 ssAAV2-CAGG-eGFP 한 가지 또는 이중특이적 항-myc-ASGR1 항체와 1:4의 바이러스 게놈 대 항체 분자의 비율로 조합된 ssAAV2-N587myc-CAGG-eGFP 바이러스 벡터를 혈관 내 주사하였다. PBS 또는 ssAAV2-N587myc-CAGG-eGFP 바이러스 벡터 한 가지만 주사한 마우스를 대조군에 포함시켰다. 주사 후 4주차에, 마우스를 희생시키고 4% PFA로 경심 관류시켰다. 간, 신장, 및 심장의 기관을 채취하여 15% 수크로스에 이어서 30% 수크로오스로 탈수하였다. 그런 다음, 기관을 슬라이드 상에서 동결 절단하고 닭 항-EGFP 항체(Jackson ImmunoResearch Labs, Inc. 펜실베이니아주 웨스트그로브 소재) 및 알렉사-488 접합된 항-닭 2차 항체(Jackson ImmunoResearch Labs, Inc. 상동)로 염색하였다(도 9a~9o). 간에서 ASGR1을 발현하도록 유전자 변형되고 야생형 ssAAV2-CAGG-eGFP가 주입되거나 ssAAV2-N587myc-CAGG-eGFP가 이중특이적 항-myc-ASGR1 항체와 함께 조합되어 주입된 이들 유전자 이식 동물 유래의 간(도 9e~9f, 9p~9r) 및 야생형 ssAAV2-CAGG-eGFP가 주입된 야생형 C57BL/6 마우스의 간(도 9b~9c)에서 GFP-양성 세포를 검출하였다. 놀랍게도, 이중특이적 항-myc-ASGR1 항체와 조합된 ssAAV2-N587myc-CAGG-eGFP의 감염 효율은 야생형 ssAAV2-CAGG-eGFP보다 훨씬 높다(도 9e~9f, 9p~9r). GFP는, 식염수가 주입되거나 ssAAV2-N587myc-CAGG-eGFP 바이러스 벡터 한 가지만 주입된 임의의 동물 유래의 간 샘플에서는 검출되지 않았거나 거의 검출되지 않았고(도 9a, 9d, 9g~9i), 이중특이적 항-myc-ASGR1 항체와 함께 ssAAV2-N587myc-CAGG-eGFP가 주입된 야생형 C57BL/6 동물에서 채취한 간 샘플에서도 검출되지 않았다(도 9m~9o). 요약하면, ssAAV2-N587myc-CAGG-eGFP 바이러스 벡터 및 이중특이적 항-myc-ASGR1 항체의 조합은 hASGR1을 발현하는 이들 (간) 세포만을 감염시켰는데, 이는 ssAAV2-N587myc-CAGG-eGFP 바이러스 벡터가 캡시드 단백질의 변형에 의해 불활성화되었다는 것, 예를 들어, scAAV 바이러스 벡터의 천연 항성이 예컨대 c-myc 에피토프에 의해 중화될 수 있고, 이러한 바이러스 벡터가 예를 들어, 이중특이적 항-myc-ASGR1 항체에 의해 생체 내에서 특이적으로 재활성화될 수 있고, 예를 들어, 간 세포를 향해 특이적으로 재표적화될 수 있다는 것을 강하게 시사한다.

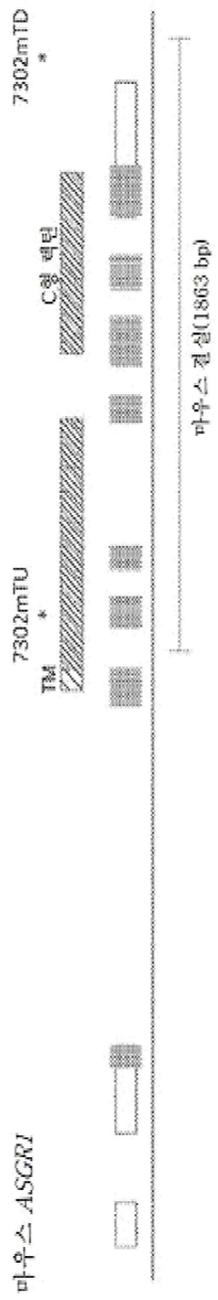
[0188] 본 실시예는 인간화 ASGR1 마우스가 간 세포로 표적화된 전달을 위해 인간 ASGR1에 특이적으로 결합하는 잠재적 치료제를 시험하는데 사용될 수 있음을 입증한다.

도면

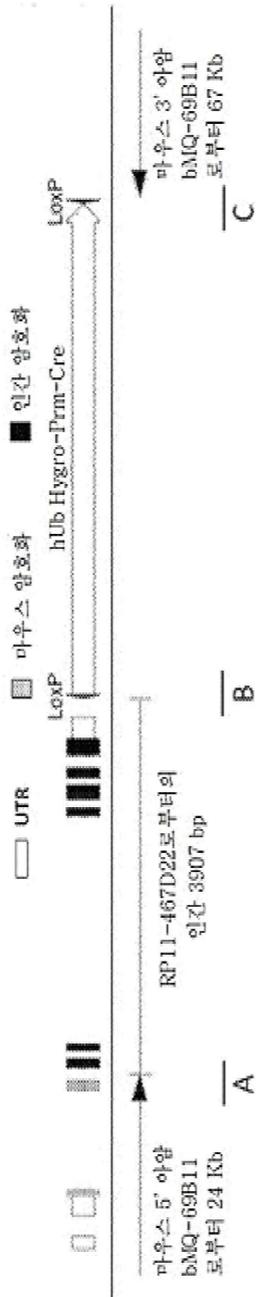
도면1a



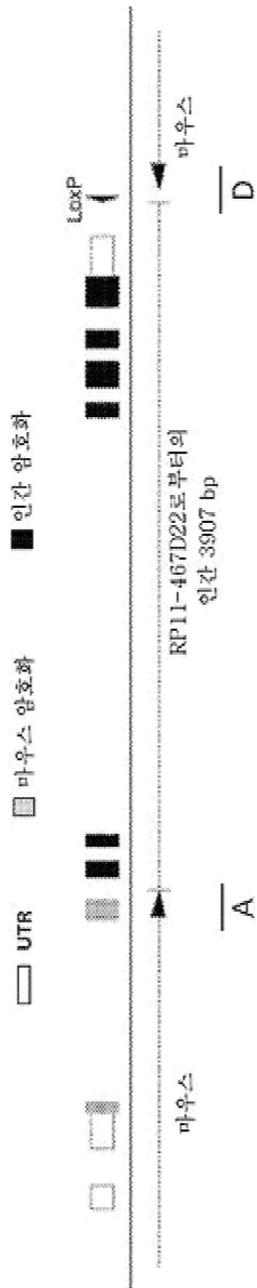
도면1b



도면2a

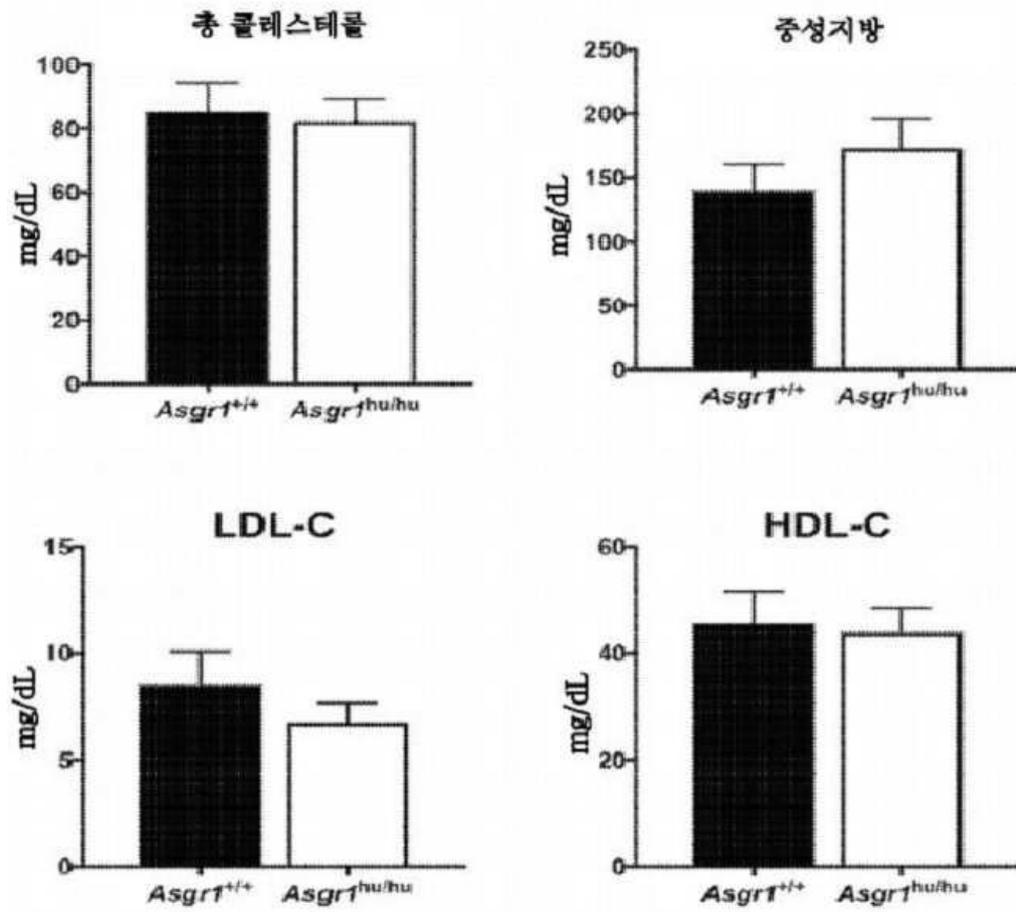


도면2b

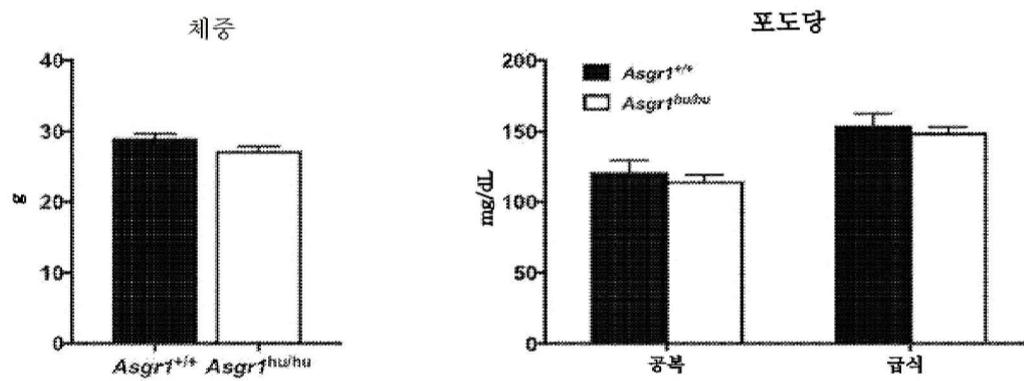




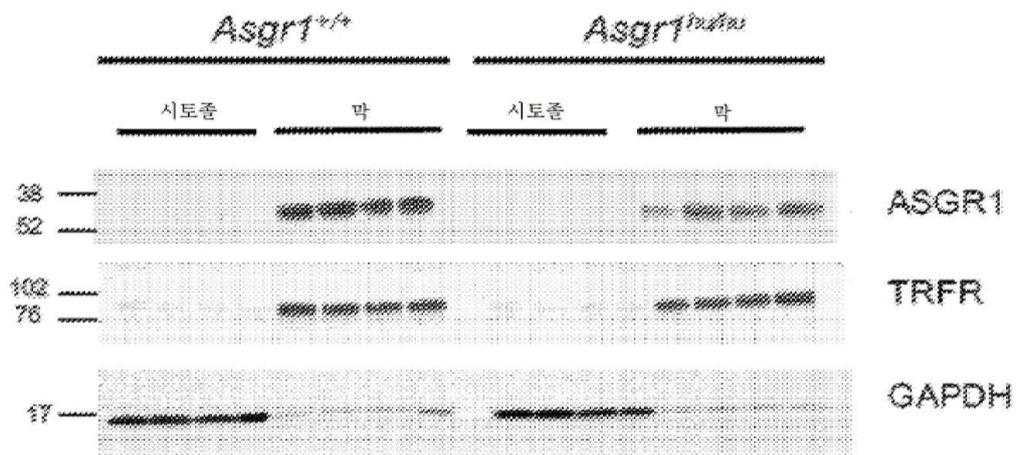
도면4



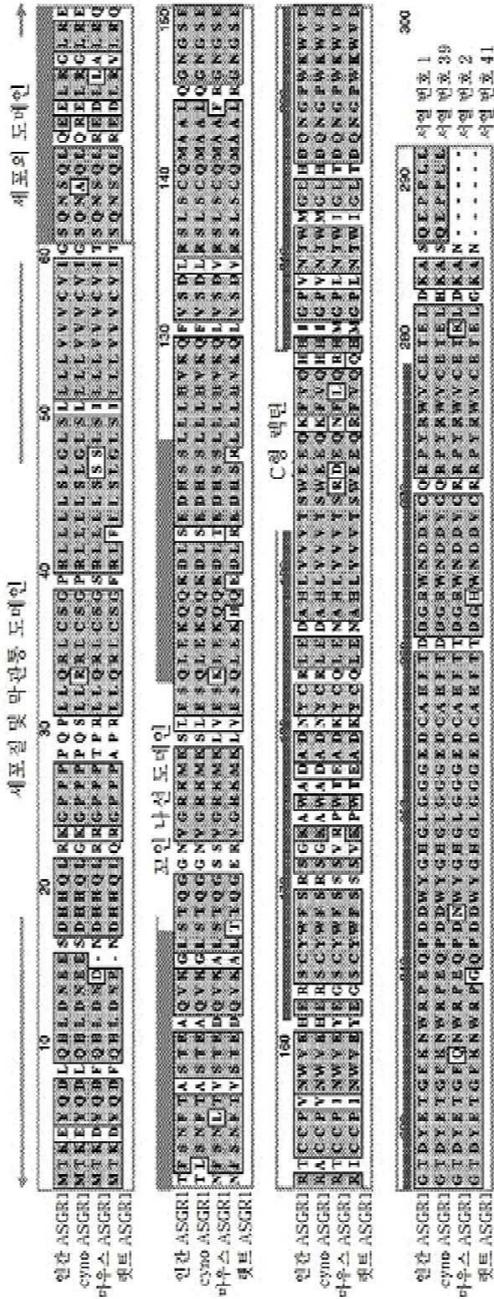
도면5



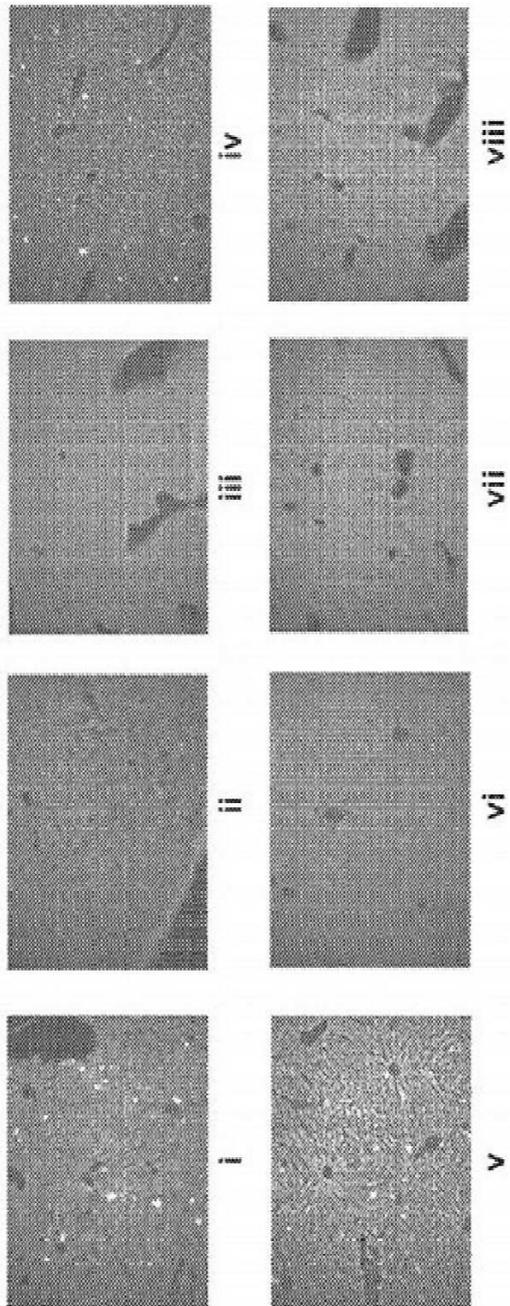
도면6



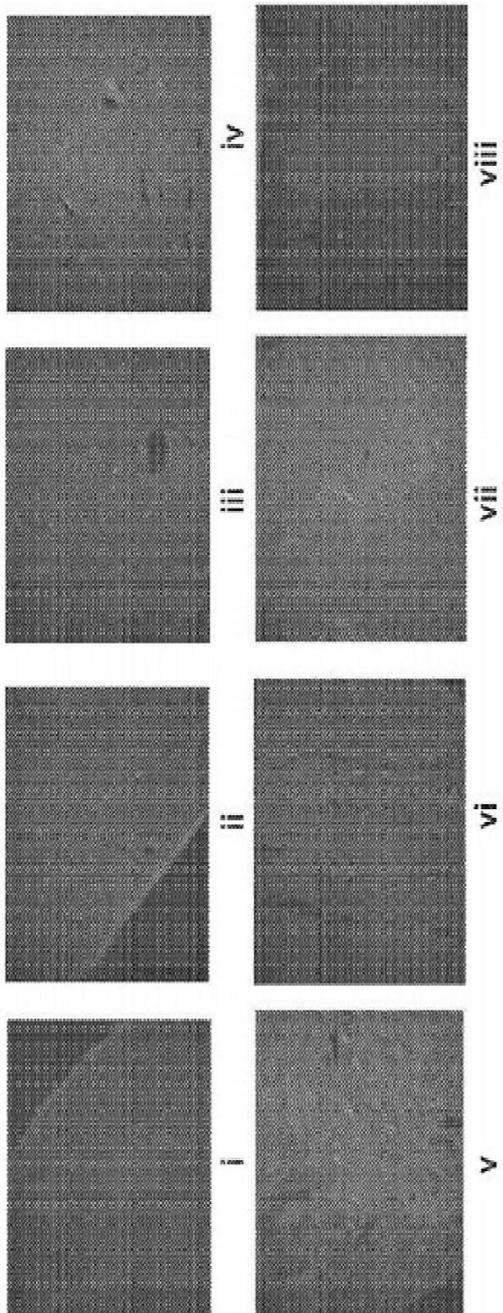
도면7



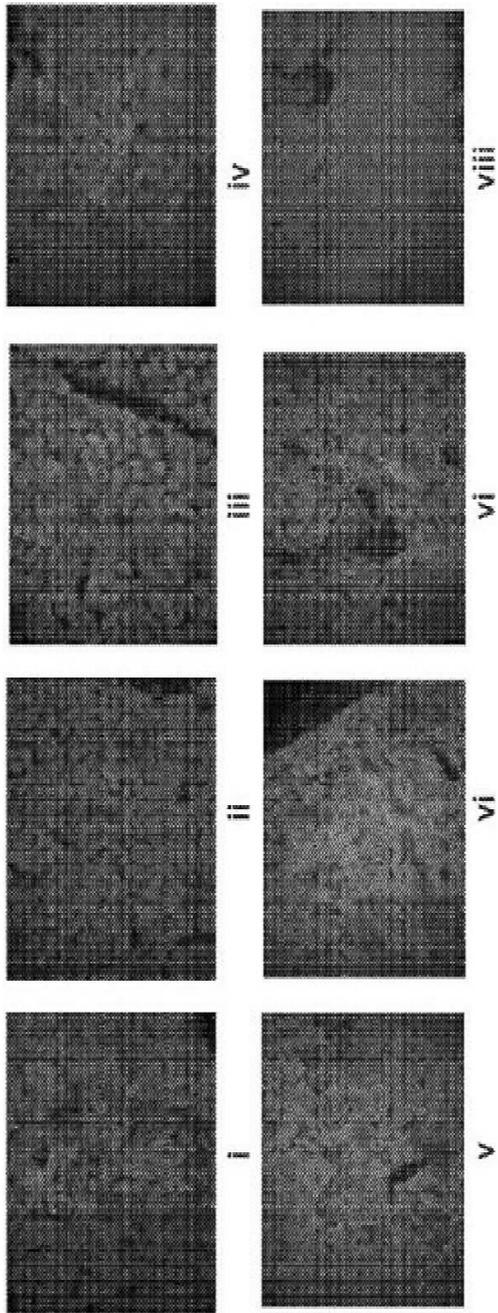
도면8a



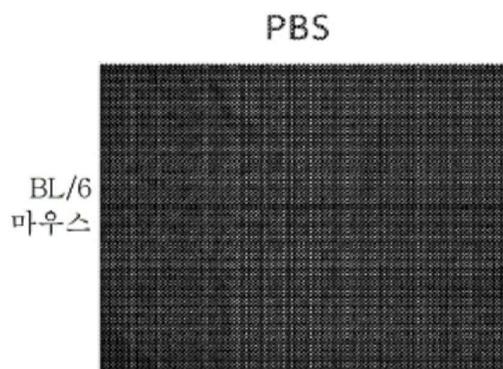
도면8b



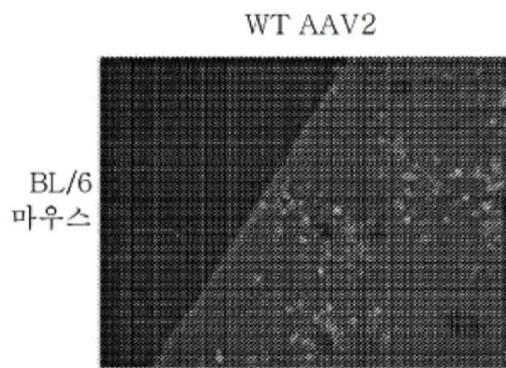
도면8c



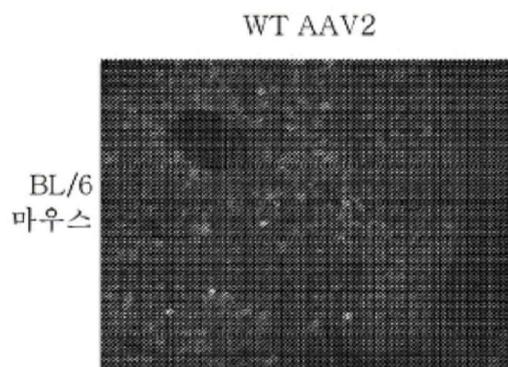
도면9a



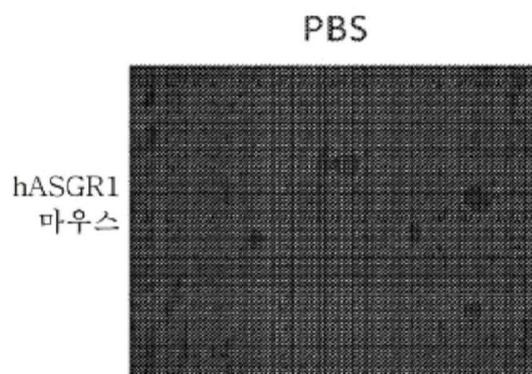
도면9b



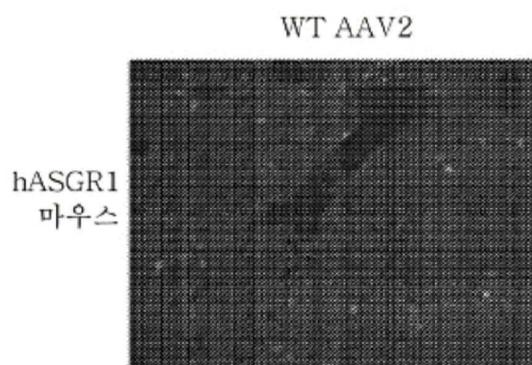
도면9c



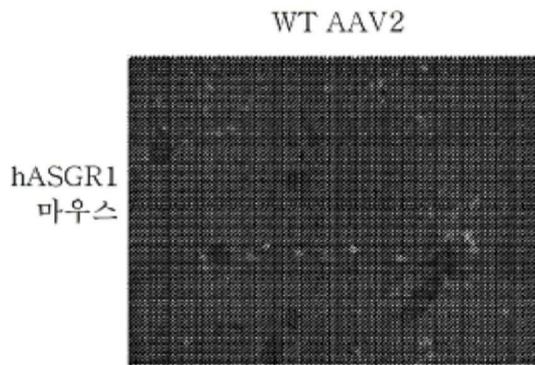
도면9d



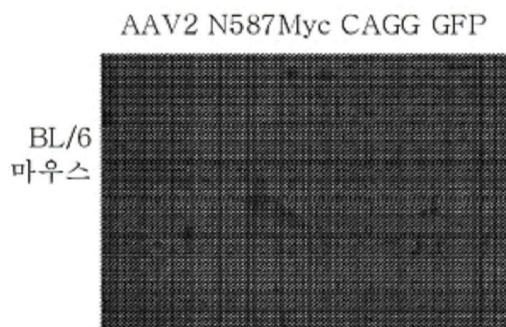
도면9e



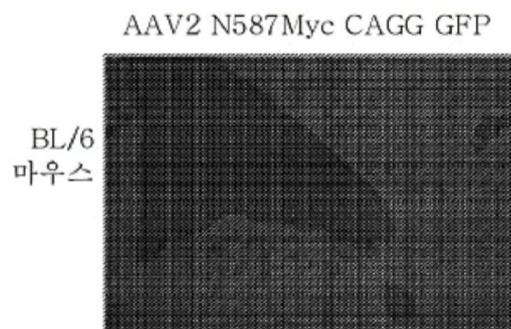
도면9f



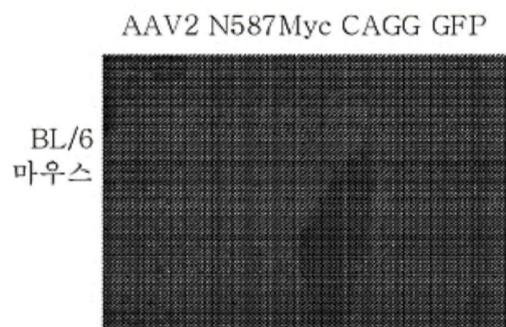
도면9g



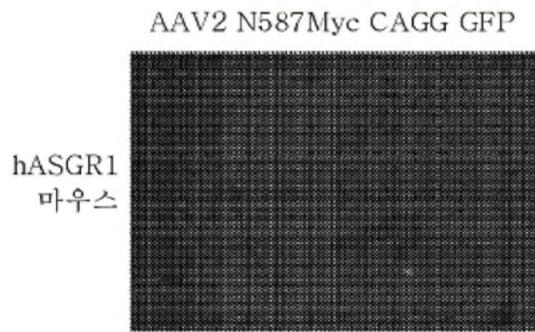
도면9h



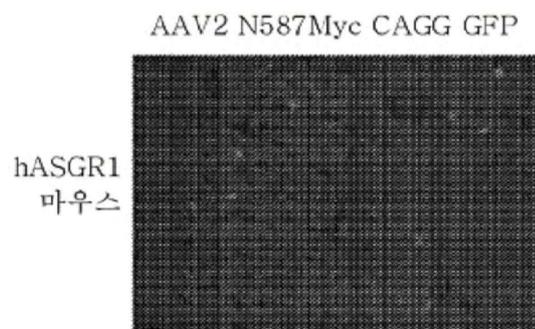
도면9i



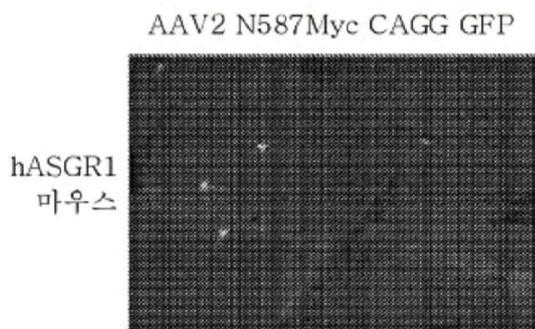
도면9j



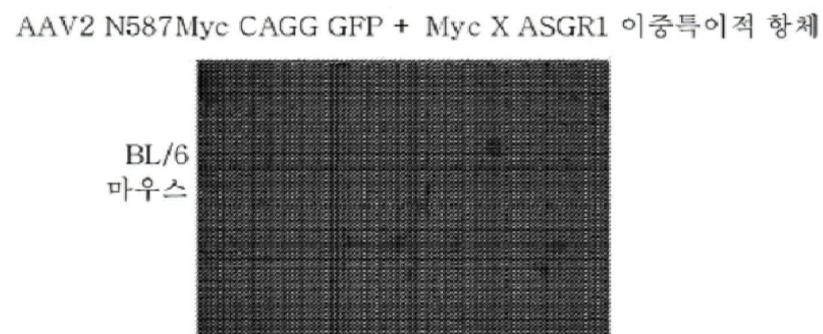
도면9k



도면9l



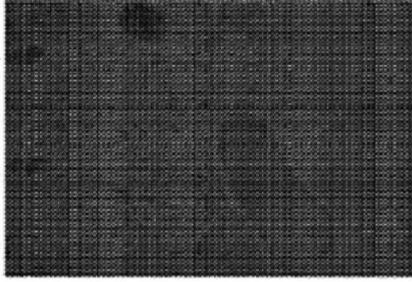
도면9m



도면9n

AAV2 N587Myc CAGG GFP + Myc X ASGR1 이중특이적 항체

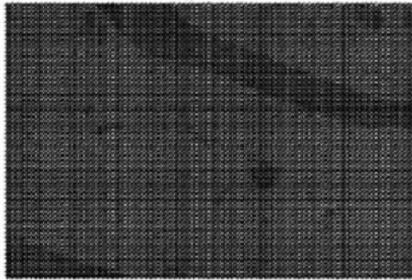
BL/6  
마우스



도면9o

AAV2 N587Myc CAGG GFP + Myc X ASGR1 이중특이적 항체

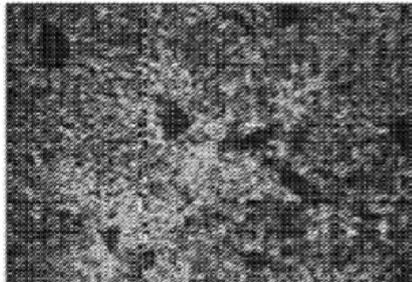
BL/6  
마우스



도면9p

AAV2 N587Myc CAGG GFP + Myc X ASGR1 이중특이적 항체

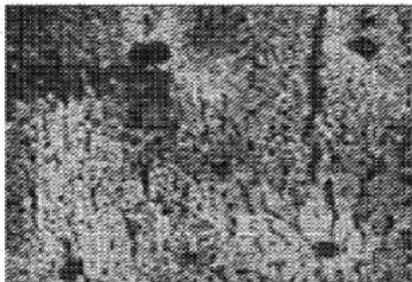
hASGR1  
마우스



도면9q

AAV2 N587Myc CAGG GFP + Myc X ASGR1 이중특이적 항체

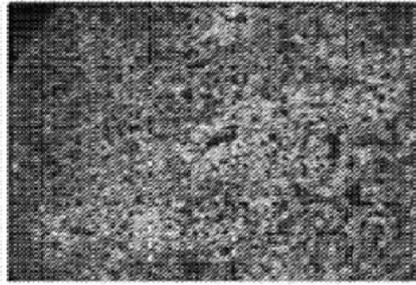
hASGR1  
마우스



도면9r

AAV2 N587Myc CAGG GFP + Myc X ASGR1 이중특이적 항체

hASGR1  
마우스



서열목록

SEQUENCE LISTING

<110> Regeneron Pharmaceuticals, Inc.

Mujica, Alejo

Gusarova, Viktoria

Wang, Cheng

Kyratsous, Christos

Potocky, Terra

Cygnar, Katherine

Martin, Joel

<120> NON-HUMAN ANIMALS COMPRISING A HUMANIZED ASGR1 LOCUS

<130> 009108.362W01

<150> 62/525,524

<151> 2017-06-27

<160> 42

<170> PatentIn version 3.5

<210> 1

<211> 291

<212> PRT

<213> Homo sapiens

<220><221> MISC\_FEATURE

<222> (1)..(40)

<223> Cytoplasmic Domain

<220><221> MISC\_FEATURE

<222> (41)..(61)

<223> Transmembrane Domain

<220><221> MISC\_FEATURE

<222> (61)..(123)

<223> Coiled-Coil Domain

<220><221> MISC\_FEATURE

<222> (161)..(278)

<223> C-Type Lectin Domain

<400> 1

Met Thr Lys Glu Tyr Gln Asp Leu Gln His Leu Asp Asn Glu Glu Ser

1                    5                    10                    15

Asp His His Gln Leu Arg Lys Gly Pro Pro Pro Pro Gln Pro Leu Leu

                  20                    25                    30

Gln Arg Leu Cys Ser Gly Pro Arg Leu Leu Leu Leu Ser Leu Gly Leu

                  35                    40                    45

Ser Leu Leu Leu Leu Val Val Val Cys Val Ile Gly Ser Gln Asn Ser

                  50                    55                    60

Gln Leu Gln Glu Glu Leu Arg Gly Leu Arg Glu Thr Phe Ser Asn Phe

65                    70                    75                    80

Thr Ala Ser Thr Glu Ala Gln Val Lys Gly Leu Ser Thr Gln Gly Gly

                  85                    90                    95

Asn Val Gly Arg Lys Met Lys Ser Leu Glu Ser Gln Leu Glu Lys Gln

                  100                    105                    110

Gln Lys Asp Leu Ser Glu Asp His Ser Ser Leu Leu Leu His Val Lys

                  115                    120                    125

Gln Phe Val Ser Asp Leu Arg Ser Leu Ser Cys Gln Met Ala Ala Leu

                  130                    135                    140

Gln Gly Asn Gly Ser Glu Arg Thr Cys Cys Pro Val Asn Trp Val Glu

145                    150                    155                    160

His Glu Arg Ser Cys Tyr Trp Phe Ser Arg Ser Gly Lys Ala Trp Ala

                  165                    170                    175

Asp Ala Asp Asn Tyr Cys Arg Leu Glu Asp Ala His Leu Val Val Val

                  180                    185                    190

Thr Ser Trp Glu Glu Gln Lys Phe Val Gln His His Ile Gly Pro Val  
 195 200 205  
 Asn Thr Trp Met Gly Leu His Asp Gln Asn Gly Pro Trp Lys Trp Val  
 210 215 220  
  
 Asp Gly Thr Asp Tyr Glu Thr Gly Phe Lys Asn Trp Arg Pro Glu Gln  
 225 230 235 240  
 Pro Asp Asp Trp Tyr Gly His Gly Leu Gly Gly Gly Glu Asp Cys Ala  
 245 250 255  
 His Phe Thr Asp Asp Gly Arg Trp Asn Asp Asp Val Cys Gln Arg Pro  
 260 265 270  
 Tyr Arg Trp Val Cys Glu Thr Glu Leu Asp Lys Ala Ser Gln Glu Pro  
 275 280 285

Pro Leu Leu

290

<210> 2

<211> 284

<212> PRT

<213> Mus musculus

<220><221> MISC\_FEATURE

<222> (1)..(39)

<223> Cytoplasmic Domain

<220><221> MISC\_FEATURE

<222> (40)..(60)

<223> Transmembrane Domain

<220><221> MISC\_FEATURE

<222> (60)..(122)

<223> Coiled-Coil Domain

<220><221> MISC\_FEATURE

<222> (160)..(277)

<223> C-Type Lectin Domain

<400> 2

Met Thr Lys Asp Tyr Gln Asp Phe Gln His Leu Asp Asn Asp Asn Asp

1                    5                    10                    15  
 His His Gln Leu Arg Arg Gly Pro Pro Pro Thr Pro Arg Leu Leu Gln  
                          20                    25                    30  
 Arg Leu Cys Ser Gly Ser Arg Leu Leu Leu Leu Ser Ser Ser Leu Ser  
                          35                    40                    45  
 Ile Leu Leu Leu Val Val Val Cys Val Ile Thr Ser Gln Asn Ser Gln  
                          50                    55                    60  
 Leu Arg Glu Asp Leu Leu Ala Leu Arg Gln Asn Phe Ser Asn Leu Thr  
 65                    70                    75                    80  
  
 Val Ser Thr Glu Asp Gln Val Lys Ala Leu Ser Thr Gln Gly Ser Ser  
                          85                    90                    95  
 Val Gly Arg Lys Met Lys Leu Val Glu Ser Lys Leu Glu Lys Gln Gln  
                          100                    105                    110  
 Lys Asp Leu Thr Glu Asp His Ser Ser Leu Leu Leu His Val Lys Gln  
                          115                    120                    125  
 Leu Val Ser Asp Val Arg Ser Leu Ser Cys Gln Met Ala Ala Phe Arg  
                          130                    135                    140  
  
 Gly Asn Gly Ser Glu Arg Thr Cys Cys Pro Ile Asn Trp Val Glu Tyr  
 145                    150                    155                    160  
 Glu Gly Ser Cys Tyr Trp Phe Ser Ser Ser Val Arg Pro Trp Thr Glu  
                          165                    170                    175  
 Ala Asp Lys Tyr Cys Gln Leu Glu Asn Ala His Leu Val Val Val Thr  
                          180                    185                    190  
 Ser Arg Asp Glu Gln Asn Phe Leu Gln Arg His Met Gly Pro Leu Asn  
                          195                    200                    205  
  
 Thr Trp Ile Gly Leu Thr Asp Gln Asn Gly Pro Trp Lys Trp Val Asp  
                          210                    215                    220  
 Gly Thr Asp Tyr Glu Thr Gly Phe Gln Asn Trp Arg Pro Glu Gln Pro  
 225                    230                    235                    240  
 Asp Asn Trp Tyr Gly His Gly Leu Gly Gly Gly Glu Asp Cys Ala His  
                          245                    250                    255

Phe Thr Thr Asp Gly Arg Trp Asn Asp Asp Val Cys Arg Arg Pro Tyr  
 260 265 270

Arg Trp Val Cys Glu Thr Lys Leu Asp Lys Ala Asn  
 275 280

<210> 3

<211> 290

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<220><221> MISC\_FEATURE

<222> (1)..(39)

<223> Cytoplasmic Domain

<220><221> MISC\_FEATURE

<222> (40)..(60)

<223> Transmembrane Domain

<220><221> MISC\_FEATURE

<222> (60)..(122)

<223> Coiled-Coil Domain

<220><221> MISC\_FEATURE

<222> (62)..(290)

<223> Residues Encoded by Introduced Human Exons

<220><221> MISC\_FEATURE

<222> (160)..(277)

<223> C-Type Lectin Domain

<400> 3

Met Thr Lys Asp Tyr Gln Asp Phe Gln His Leu Asp Asn Asp Asn Asp  
 1 5 10 15

His His Gln Leu Arg Arg Gly Pro Pro Pro Thr Pro Arg Leu Leu Gln  
 20 25 30

Arg Leu Cys Ser Gly Ser Arg Leu Leu Leu Leu Ser Ser Ser Leu Ser  
 35 40 45

Ile Leu Leu Leu Val Val Val Cys Val Ile Thr Ser Gln Asn Ser Gln

50                      55                      60  
 Leu Gln Glu Glu Leu Arg Gly Leu Arg Glu Thr Phe Ser Asn Phe Thr  
 65                      70                      75                      80  
 Ala Ser Thr Glu Ala Gln Val Lys Gly Leu Ser Thr Gln Gly Gly Asn  
                          85                      90                      95  
 Val Gly Arg Lys Met Lys Ser Leu Glu Ser Gln Leu Glu Lys Gln Gln  
                          100                      105                      110  
 Lys Asp Leu Ser Glu Asp His Ser Ser Leu Leu Leu His Val Lys Gln  
  
                          115                      120                      125  
 Phe Val Ser Asp Leu Arg Ser Leu Ser Cys Gln Met Ala Ala Leu Gln  
                          130                      135                      140  
 Gly Asn Gly Ser Glu Arg Thr Cys Cys Pro Val Asn Trp Val Glu His  
 145                      150                      155                      160  
 Glu Arg Ser Cys Tyr Trp Phe Ser Arg Ser Gly Lys Ala Trp Ala Asp  
                          165                      170                      175  
 Ala Asp Asn Tyr Cys Arg Leu Glu Asp Ala His Leu Val Val Val Thr  
  
                          180                      185                      190  
 Ser Trp Glu Glu Gln Lys Phe Val Gln His His Ile Gly Pro Val Asn  
                          195                      200                      205  
 Thr Trp Met Gly Leu His Asp Gln Asn Gly Pro Trp Lys Trp Val Asp  
                          210                      215                      220  
 Gly Thr Asp Tyr Glu Thr Gly Phe Lys Asn Trp Arg Pro Glu Gln Pro  
 225                      230                      235                      240  
 Asp Asp Trp Tyr Gly His Gly Leu Gly Gly Gly Glu Asp Cys Ala His  
  
                          245                      250                      255  
 Phe Thr Asp Asp Gly Arg Trp Asn Asp Asp Val Cys Gln Arg Pro Tyr  
                          260                      265                      270  
 Arg Trp Val Cys Glu Thr Glu Leu Asp Lys Ala Ser Gln Glu Pro Pro  
                          275                      280                      285  
 Leu Leu  
                          290  
 <210> 4

<211> 855

<212> DNA

<213> Mus musculus

<400> 4

```

atgacaaagg attatcaaga tttccagcac ctggacaatg ataatgacca tcatcaactc      60
cggagagggc cgcctcccac tccacggctc ttgcagcgac tctgctctgg atccccctc      120

ctcctgctct cctcagacct cagcattctg ttgctgggtg ttgtctgtgt gatcacatcc      180
caaaattccc aactccggga agatctgctg gctctaaggc agaatttcag caacctcact      240
gtgagcactg aggaccaggt caaggccctg agcaccaggc gaagttagtg gggaagaaag      300
atgaagttag tggagtcgaa gctggaaaaa cagcagaagg atctgactga agatcactcc      360
agtttgctac tgcacgtgaa gcagtttagt tctgacgtgc gaagcttgag ctgccagatg      420
gctgcatttc ggggcaatgg ctctgaaagg acctgctgcc ccatcaactg ggtggagtat      480
gaaggcagct gctactggtt ctccagctct gtgaggcctt ggactgaagc tgacaagtac      540

tgccagctgg aaaatgccca tctggtgggt gtgacctcca gggatgagca gaacttcctc      600
cagcgccaca tgggcccttt aaacacttgg attggcctaa ctgaccagaa cgggccctgg      660
aatgggtgg atggaacaga ctacagaca ggcttcagaa attggagacc agagcagcca      720
gataactggt acggacatgg gcttggagga ggcgaggact gtgcccactt cacgacggat      780
ggccgctgga atgacgactg ctgcaggagg ccctaccgct gggctctgtga gacaaagttag      840
gataaggcta attag                                             855

```

<210> 5

<211> 876

<212> DNA

<213> Homo sapiens

<400> 5

```

atgaccaagg agtatacaaga ccttcagcat ctggacaatg aggagagtga ccacatcag      60
ctcagaaaag ggccacctcc tcccagccc ctctgcagc gtctctgctc cggacctcgc      120
ctctctctgc tctccctggg cctcagctc ctgctgcttg tggttgtctg tgtgatcgga      180
tccccaaact cccagctgca ggaggagctg cggggcctga gagagacgtt cagcaacttc      240
acagcgagca cggaggccca ggtcaagggc ttgagcacc agggaggcaa tgtgggaaga      300
aagatgaagt cgctagagtc ccagctggag aaacagcaga aggacctgag tgaagatcac      360
tccagcctgc tgetccactg gaagcagttc gtgtctgacc tgcggagcct gagctgtcag      420

```

atggcggcgc tccagggcaa tggctcagaa aggacctgct gcccgggtcaa ctgggtggag 480  
 cacgagcgca gctgctactg gttctctcgc tccgggaagg cctgggctga cgccgacaac 540  
 tactgccggc tggaggacgc gcacctggtg gtggtcacgt cctgggagga gcagaaattt 600  
 gtccagcacc acataggccc tgtgaacacc tggatgggcc tccacgacca aaacgggccc 660  
 tggaagtggg tggacgggac ggactacgag acgggcttca agaactggag gccggagcag 720  
 ccggacgact ggtacggcca cgggctcgga ggaggcgagg actgtgcca cttaccgac 780  
 gacggccgct ggaacgacga cgtctgccag aggcctacc gctgggtctg cgagacagag 840

ctggacaagg ccagccagga gccacctctc ctttaa 876

<210> 6

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 6

tccaactcc gggaagatc 19

<210> 7

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 7

tgctggctct aaggcagaat ttca 24

<210> 8

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 8

tcagtgtca cagtgaggtt 20

<210> 9

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic  
 <400> 9  
 gggttgctc atgtaggaa gg 22  
 <210> 10  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 10  
 tcagcagccg agctgtgaaa 20  
 <210> 11  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 ><223> Synthetic  
 <400> 11  
 caggctgtgc tacccaaagt tc 22  
 <210> 12  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 12  
 ggaggcaatg tggaagaaa g 21  
 <210> 13  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 13  
 tgaagtcgct agagtcccag ctgg 24  
 <210> 14  
 <211> 20  
 <212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 14

tcaggtcctt ctgctgtttc 20

<210> 15

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 15

gattgggaat ccgcccattc 20

<210> 16

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 16

cctcttctgc tttctcggga attttcatc 29

<210> 17

<211> 20

<212> DNA

<213>

Artificial Sequence

<220><223> Synthetic

<400> 17

aaagcgccac gggtttcaag 20

<210> 18

<211> 120

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<220><221> misc\_feature

<222> (1)..(60)

<223> Mouse Sequence

<220><221> misc\_feature  
 <222> (61)..(120)  
 <223> Human Sequence  
 <400> 18  
 catcccaaag tgggtggcca gggctgggca gagaaagggg gcaacttcgg gtgtgtgtga 60  
 caagggagtg gtgggtgcag tgggtggcga cacagcgatc ccgttttctt ctctctgcac 120  
  
 <210> 19  
 <211> 160  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <220><221> misc\_feature  
 <222> (1)..(60)  
 <223> Human Sequence  
 <220><221> misc\_feature  
 <222> (61)..(66)  
 <223> XhoI  
 <220><221> misc\_feature  
 <222> (67)..(100)  
 <223> loxP  
 <220><221> misc\_feature  
 <222> (101)..(160)  
 <223> Cassette  
 <400> 19  
 tgttatttac agatacgtga gtttgggcaa attattgttc tctgtgtccc agctgtaaac 60  
 ctcgagataa cticgtataa tgtatgctat acgaagtat atgcatggcc tccgcgccgg 120  
  
 gttttggcgc ctcccgcggg cgccccctc ctcacggcga 160  
 <210> 20  
 <211> 191  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <220><221> misc\_feature

<222> (1)..(60)  
 <223> Cassette  
 <220><221> misc\_feature  
 <222> (61)..(94)  
 <223> loxP  
 <220><221> misc\_feature  
 <222> (100)..(125)  
 <223> I-CeuI  
 <220><221> misc\_feature  
 <222> (126)..(131)  
 <223> NheI  
 <220><221> misc\_feature  
 <222> (132)..(191)  
 <223> Mouse Sequence  
 <400> 20  
 ttctactgca ttctagttgt ggtttgtcca aactcatcaa tgtatcttat catgtctgga 60  
  
 ataacttcgt ataatgtatg ctatacgaag ttatgctagt aactataacg gtcctaaggt 120  
 agcgagctag ccgtggacag atacagcaac gtgagctagt tattctgtcc taaagtctca 180  
 gttggaagat g 191  
 <210> 21  
 <211> 9692  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <220><221> misc\_feature  
 <222> (1)..(69)  
 <223> Mouse Sequence  
 <220><221> misc\_feature  
 <222> (70)..(3976)  
 <223> Human Sequence  
 <220><221> misc\_feature  
 <222> (3977)..(3982)  
  
 <223> XhoI

<220><221> misc\_feature  
 <222> (3983)..(4016)  
 <223> loxP  
 <220><221> misc\_feature  
 <222> (4023)..(5235)  
 <223> Human Ubiquitin Promoter  
 <220><221> misc\_feature  
 <222> (5236)..(5302)  
 <223> EM7 Promoter  
 <220><221> misc\_feature  
 <222> (5303)..(6328)  
 <223> Hygromycin Resistance Gene  
 <220><221> misc\_feature  
 <222> (6329)..(6813)  
 <223> PGK polyA  
 <220><221> misc\_feature  
 <222> (6825)..(7506)  
 <223> Prm Promoter  
 <220><221> misc\_feature  
 <222> (7507)..(8646)  
  
 <223> CreI  
 <220><221> misc\_feature  
 <222> (8909)..(9119)  
 <223> SV40 polyA  
 <220><221> misc\_feature  
 <222> (9124)..(9157)  
 <223> loxP  
 <220><221> misc\_feature  
 <222> (9163)..(9188)  
 <223> I-CeuI  
 <220><221> misc\_feature  
 <222> (9189)..(9194)  
 <223> NheI  
 <220><221> misc\_feature

<222> (9195)..(9692)

<223> Mouse Sequence

<400> 21

```

gtgtgatcac atcccaaagt gggtggccag gctgggcag agaaaggggg caacttcggg      60
tgtgtgtgac aaggagatgg tgggtgcagt ggtggcggac acagcgatcc cgttttcttc      120

tctctgcacg ctgtcctggc cagactccca gctgcaggag gagctgcggg gcctgagaga      180
gacgttcagc aacttcacag cgagcacgga ggcccaggtc aagggttga gcaccaggg      240
tgaggcgct gggcggggc tgggctggg gctgggctg ggggctgtc gggaacgtg      300
agcgacctc tccgcagga ggcaatgtg gaagaaagat gaagtcgcta gagtccagc      360
tggagaaca gcagaaggac ctgagtgaag gtcagagagg gagtgtgtgt gtgtgtgtgt      420
gtgtgtgaaa gagagtgaga atgtgtggat gtgtgtgaga aagtgtgagt gtgtgtggat      480
gtgtgtgaga atgagagga gtgtgtgtgt gtgtgagtct gtgtgtgaga atgaggggga      540

gtgtgttttg ggtgtgtgta tgagagcctt gtgtggatgt gagaatgaga gggagtgtgt      600
atgtctgtga gtgtgagaat gagatggagt gtgtgtgagt ctgtgtgtga gaatgagtg      660
tgtgtgtgtg agaatgagat ggtgtgtgtg tgggaatgag aggggtgtg tgtctgagtg      720
tgagaatgag atagagtgtg tgtgagacag tctgtgggaa tgagaggag tgtgtgtgag      780
agtgtgagaa tgacggagtg tgtctgtgag tgtgataatg aggtgtgtgt gagtctgagt      840
gtaagaatga gatggggtgt gtgtgtctgt gagtgtgaga gtgtgagaat gaggggtgtt      900
tgtgtctgag tgtgagtctg tgagtgtgag aatgagatgg ggtgtgtgag tgagtgtgag      960

aatgagatgg ggtgtgtgtg tctgtgagtg tgtgtgtgtt tgtgagtgtg agaatgagat      1020
ggggtgtgtg tgagtgtgag aatgagatgg ggtgtgtgtc tgtgtgtgag aatgagatgg      1080
gtgtgtgtgt gacagagtct gagtgtgaga atgagagga gtgtgtgtga gtgtgagaat      1140
gagaaggagt ggatgggtgt gtgagtctgt gtgaatgagg gagggtgt gtgtacgagt      1200
tgagatctgt gtttatgtgt gagaatgtgt cagtgtatgt gtgtgagaac gtgtgtatgt      1260
gtgttagtgt gtgttgcgtg tgtgggggaa tgagaggat tgtgtctgtg agtgtgagaa      1320
tgagatggag tgtctgtgag actgtgtgtg aggagtggga gtgtgtgtga gaatgagatg      1380

gggtgtgtgt tctgagtgtg tgtctgtgag aatgagaggg agtgtgtgtg tgtgtgagag      1440
cctgtgtgaa aatgagaagg agtgtggatg ggtgtttgtg agtgggagag tctgtgtgtt      1500
tatgtgtgtg agaatgaggg agtgtgggtg tgtgtgcaa tgtgagtctg tgtttatgtg      1560
tgtgagaatg tgcagtgtg tgtgagaacg tgtgtgttag tgtgttgcgt gtgtgagaat      1620
gtaagtatat gtgtaagtc atgtgagtgt gtgtatgtgc gtgtgtgtg aatgtgcatt      1680

```

gtgtgtgcat gtgtgaaaga gtatatgtgt gttgtgggtg agtgtgtgtg gtgtgtgtag 1740  
 tgggtgaggg tgtgttgtat gtgtgggtgt gcgttgtgtg aatgtgtgta tgtgggtgag 1800  
  
 ggtgtgtgtg cctgtgtgag ggtgtgttgt ggtttttgtg tgtgtttggg tgagggtgtg 1860  
 ttgtgtgtgt gtgtgggtga aggtgtgttg tgtgtgtcgt gggtgaaggt gtgttgtgtg 1920  
 tgtagtact gtagattagg gtgtttccg tgtgtgtgtg tgagggtgta tgttgtgggt 1980  
 gttttgtgtg tgagtgggtg tgaaggggtg tgttgtgtgt atgtgggtta aggtgtgtta 2040  
 tgcgtgaggg tgtattgtgt gtgtgttttg tgtgtgttgt gtgtatgtgg gttagggtgt 2100  
 gttgtgtgtt tgtgtgtttt gtgtgttttc tgtgtatgtg ggttacgggt tgttgtgcgt 2160  
 gtgagggtgt gttgtgatg gtgtgtttgt tgtgttgtgt gagtgtgtat gtgagttagg 2220  
  
 gtgtgttgtg tctatgatg tgtgtgtaag ggtgtgttgt gtgtctgtgg gtgtgttttg 2280  
 tgtatgtggg ttagggtgtg ttgtgtgttc tgtattgtgt gttttatgtg ttgtctgtat 2340  
 gtgggttatg tgtgttgtgt gtgtttgtga tgtatgtggg ttagggtgtg ttgtgtgtct 2400  
 ctgtgtgttg tctgcgtttg tgtctgtggg ttagggtgtg ttgtatgtgt tgtgttttgt 2460  
 gtgttgtccg tgtgtgtgta tgtgggttag gttgtgtgtg tgtgtgttgt atattgtctg 2520  
 tgtgtgtgtg ttaggatgtg ttgtgtgtct gtgtgagtgt gtgtgtaagg gtgtgttgtg 2580  
 tgtgtaggag tgtgtgtgtg tgtgtgatg ggggtctctc aggccaactc cgctgctgtt 2640  
  
 tgtggcaatg cgacgggtgt tccgggtcca gcaggaggat gtagggctga cctcgtttcc 2700  
 cgtttccctc cccgtggttt ccgcatctcc tcccgtccc ctccgcccg tctcccaga 2760  
 tcaactccage ctgctgtccc acgtgaagca gttcgtgtct gacctgcgga gcctgagctg 2820  
 tcagatggcg gcgtccagg gcaatggtaa ggaggccagc ccggcccgt ctctgcctcc 2880  
 ccccttctct gggcagcgt tagccctgc gccctttc tcccgtcag gctcagaaag 2940  
 gacctgctgc ccgtcaact ggggtggagca cgagcgagc tgctactggt tctctcgctc 3000  
 cgggaaggcc tgggctgacg ccgacaacta ctcccgtg gaggacgcgc acctggtggt 3060  
  
 ggtcacgtcc tgggaggagc aggtgaggac ccggagggtc tgggaggctg gctggcctcg 3120  
 gagagatcac caccgcctt ctctctctc agaaattgt ccagcaccac ataggcctg 3180  
 tgaacacctg gatgggcctc cagacaaa acgggccctg gaagtgggtg gacgggacgg 3240  
 actacagagac gggcttcaag tgagtgcgcg cctcctctg gcctgggtcc ggccgcttc 3300  
 gcgcctggg gcctgggct gaggagtctg gagcgaccg cctgcggatc cgacctctg 3360  
 gggcccacag ctggctctgt cccaggaac tggaggccgg agcagccgga cgactgttac 3420

ggccacgggc tcggaggagg cgaggactgt gcccaattca ccgacgacgg ccgctggaac 3480  
  
gacgacgtct gccagaggcc ctaccgctgg gtctgcgaga cagagctgga caaggccagc 3540  
caggagccac ctctccttta atttatttct tcaatgectc gacctgccgc aggggtccgg 3600  
gattgggaat ccgccatct gggggcctct tctgcttct cggaatttt catctaggat 3660  
tttaaggga ggggaaggat aggggatgt tccgaaggat aggagcttga aaccctggc 3720  
gctttctgca gtttgcagg taccattgtg aactttttt ttttaagagt aaaaagaaat 3780  
atacctaaac ctctgttag ttgtctggtt attggggatt cggaagcagg agtgggctgg 3840  
ttggcattac gaagccttag cgggtgctgt ggcatcatga gaactgtgtg ggctttgggc 3900  
  
cagaatggcc agactttgtt atttacagat acgtgagttt gggcaaatta ttgttctctg 3960  
tgtccagct gtaaacctcg agataacttc gtataatgta tgctatacga agttatatgc 4020  
atggcctccg cgcgggtttt tggcgcctcc cgcgggcgcc cccctctca cggcgagcgc 4080  
tgccacgtca gacaagggc gcagcgagcg tctgatcct tccgcccgga cgctcaggac 4140  
agcggcccgc tgctcataag actcggcctt agaaccccag taccagcaga aggacatttt 4200  
aggacgggac ttgggtgact ctagggcaact ggttttttt ccagagagcg gaacaggcga 4260  
ggaaaagtag tcccttctcg gcgattctgc ggagggatct ccgtggggcg gtgaacgccg 4320  
  
atgattatat aaggacgcgc cgggtgtggc acagctagtt ccgtcgagc cgggatttgg 4380  
gtcgcggttc ttgttttgg atcgtgtga tctcacttg gtgagtagcg ggctgctggg 4440  
ctggccgggg ctctctggc gcgccggccg ctcggtggga cggaagcgtg tggagagacc 4500  
gccaaaggct gtagtctggg tccgcgagca aggttgcct gaactggggg ttggggggag 4560  
cgcagcaaaa tggcggctgt tcccagctt tgaatggaag acgcttgtga ggcgggctgt 4620  
gaggtcgttg aaacaagggt gggggcatgg tggcgcgcaa gaaccaagg tctttaggcc 4680  
ttcgtaatg cgggaaagct cttattcggg tgagatgggc tggggcacca tetggggacc 4740  
  
ctgacgtgaa gtttgtcact gactggagaa ctcggtttgt cgtctgttc gggggcggca 4800  
gttatggcgg tccggttggg cagtgcacce gtaccttgg gagcgcgcgc cctcgtctg 4860  
tctgacgtc accgttctg ttggcttata atgcagggtg gggccacctg ccgtaggtg 4920  
tgcgtaggc tttctccgt cgcaggacgc agggttcggg cctagggtag gctctcctga 4980  
atcgacagc gccggacctc tggtaggggg agggataagt gaggcgtcag tttctttggt 5040  
cggttttatg tacctatctt cttaagtagc tgaagctccg gttttgaact atgcgctcgg 5100  
ggttggcgag tgtgtttgt gaagttttt aggcacctt tgaatgtaa tcatttgggt 5160

caatatgtaa ttttcagtgt tagactagta aattgtccgc taaattctgg cgtttttgg 5220  
cttttttgtt agacgtgttg acaattaatc atcggcatag tatacggca tagtataata 5280  
cgacaagggtg aggaactaaa ccatgaaaaa gcctgaactc accgcgacgt ctgtcgagaa 5340  
gtttctgac gaaaagtctg acagcgtgtc cgacctgatg cagctctcgg agggcgaaga 5400  
atctcgtgct ttcagcttcg atgtaggagg gcgtggatat gtcctcggg taaatagctg 5460  
cgccgatggt ttctacaaag atcgttatgt ttatcggcac tttgcatcgg ccgcctccc 5520  
gattccgga gtccttgaca ttggggaatt cagcgagagc ctgacctatt gcatctccc 5580  
  
ccgtgcacag ggtgtcacgt tgcaagacct gcctgaaacc gaactgcccg ctgttctgca 5640  
gccggctcgg gaggccatgg atgcgattgc tgcggccgat cttagccaga cgagcgggtt 5700  
cggccattc ggaccgcaag gaatcggca atactaca tggcgtgatt tcatatcgc 5760  
gattgctgat ccccatgtgt atcactggca aactgtgatg gacgacaccg tcagtgcctc 5820  
cgtcgcgag gctctgatg agctgatgct ttgggccgag gactgcccc aagtccgca 5880  
cctcgtgac gcggatttcg gctccaaca tgcctgacg gacaatggc gcataacagc 5940  
ggtcattgac tggagcgagg cgatgttcgg ggattccaa tacgaggtcg ccaacatctt 6000  
  
cttctggagg ccgtggttgg cttgtatgga gcagcagacg cgctacttcg agcggaggca 6060  
tccggagctt gcaggatcgc cgcgctccg ggcgtatatg ctccgattg gtcttgacca 6120  
actctatcag agcttgggtg acggcaattt cgatgatgca gcttggcgc agggctgatg 6180  
cgacgcaatc gtccgatccg gagccgggac tgcgggcgt acacaaatcg cccgcagaag 6240  
cgccgcccgc tggaccgatg gctgtgtaga agtactgcc gatagtggaa accgacgcc 6300  
cagcactcgt ccgagggcaa aggaataggg ggatccctg taagtctgca gaaattgatg 6360  
atctattaa caataaagat gtccactaaa atggaagttt ttctgtcat actttgttaa 6420  
  
gaagggtgag aacagagtac ctacattttg aatggaagga ttggagctac gggggtgggg 6480  
gtggggtggg attagataaa tgctgtctt ttactgaagg ctcttacta ttgctttatg 6540  
ataatgtttc atagttggat atcataattt aaacaagcaa aaccaatta agggccagct 6600  
cattctccc actcatgac tatagatcta tagatctctc gtgggatcat tgttttctc 6660  
ttgattccca cttgtggtt ctaagtactg tggtttcaa atgtgtcagt tcatagcct 6720  
gaagaacgag atcagcagc tctgttccac atacactca ttctcagtat tgttttgca 6780  
agttctaatt ccatcagacc tcgaactgca gccctagcc cgggcgccag tagcagcacc 6840  
  
cacgtccacc ttctgtctag taatgtccaa cacctcctc agtccaaaca ctgctctgca 6900  
tccatgtgge tccatttat acctgaagca ctgatgggg cctcaatgtt ttactagagc 6960  
ccacccccct gcaactctga gaccctctgg atttgtctgt cagtgcctca ctgggctgtt 7020

ggataatttc ttaaaaggtc aagttccctc agcagcattc tctgagcagt ctgaagatgt 7080  
 gtgcttttca cagttcaaat ccatgtggct gtttcacca cctgcctggc cttgggttat 7140  
 ctatcaggac cttagcctaga agcaggtgtg tggcacttaa cacctaagct gagtgactaa 7200  
 ctgaacactc aagtggatgc catctttgtc acttcttgac tgtgacacaa gcaactcctg 7260  
  
 atgccaaagc cctgcccacc cctctcatgc ccatatttgg acatggtaca ggtcctcact 7320  
 ggccatggte tgtgaggctc tggctctctt tgacttcata attcctaggg gccactagta 7380  
 tctataagag gaagaggggtg ctggctccca ggccacagcc cacaaaattc cacctgctca 7440  
 caggttggct ggctcgacc aggtgggtgc ccttgcctcg agccagctcc cggccaagcc 7500  
 agcaccatgg gtacccccaa gaagaagagg aaggtgcgta ccgatttaaa ttccaattta 7560  
 ctgaccgtac accaaaattt gcctgcatta ccggctgatg caacgagtga tgaggttcgc 7620  
 aagaacctga tggacatgtt cagggatcgc caggcgtttt ctgagcatac ctggaaaatg 7680  
  
 cttctgtccg ttgcccgtc gtgggcggca tggtgcaagt tgaataaccg gaaatggttt 7740  
 cccgcagaac ctgaagatgt tcgcgattat cttctatata ttcaggcgcg cggctctggca 7800  
 gtaaaaacta tcagcaaca ttggggccag ctaaacatgc ttcactctcg gtccgggctg 7860  
 ccacgaccaa gtgacagcaa tgctgtttca ctggttatgc ggcgatccg aaaagaaaac 7920  
 gttgatgccg gtgaactgc aaaacaggct ctagecttcg aacgcactga tttcgaccag 7980  
 gttctttcac teatggaaaa tagtgatcgc tgccaggata tacgtaatct ggcatttctg 8040  
 gggattgctt ataacacct gttactata gccgaaatg ccaggatcag ggttaaagat 8100  
  
 atctcacgta ctgacggagg gagaatgtta atccatattg gcagaacgaa aacgctggtt 8160  
 agcaccgcag gttagagaaa ggcacttagc ctgggggtaa ctaactggt cgagcgatgg 8220  
 atttccgtct ctggtgtagc tgatgatccg aataactacc tgttttgccg ggtcagaaaa 8280  
 aatggtgttg ccgcgccatc tgccaccagc cagctatcaa ctgcgcctt ggaagggatt 8340  
 tttgaagcaa ctcatcgatt gatttacggc gctaaggtaa atataaaatt tttaatgtga 8400  
 taatgtgtta aactactgat tctaattgtt tgtgtatfff aggatgactc tggtcagaga 8460  
 tacctggcct ggtctggaca cagtgccctg gtcggagccg cgcgagatat ggccccgcct 8520  
  
 ggagtttcaa taccggagat catgcaagct ggtggctgga ccaatgtaa tattgtcatg 8580  
 aactatatec gtaacctgga tagtgaaaca ggggcaatgg tgcgcctgct ggaagatggc 8640  
 gattgatcta gataagtaat gatcataatc agccatatca catctgtaga ggttttactt 8700  
 gctttaaaaa acctcccaca cctccccctg aacctgaaac ataaaatgaa tgcaattgtt 8760  
 gttgttaaac ctgccctagt tgcggccaat tccagctgag cgtgcctccg caccattacc 8820  
 agttggtctg gtgtcaaaaa taataataac cgggcagggg ggatctaagc tctagataag 8880

taatgatcat aatcagccat atcacatctg tagaggtttt acttgcttta aaaaacctcc 8940  
  
 cacacctccc cctgaacctg aaacataaaa tgaatgcaat tgttggttgt aacttgttta 9000  
 ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 9060  
 ttttttcaact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 9120  
 ggaataaactt cgtataatgt atgctatacg aagttatgct agtaactata acggctctaa 9180  
 ggtagcgcgagc tagccgtgga cagatacagc aacgtgagct agttattctg tectaaagtc 9240  
 tcagttggaa gatgggagga tttttgacct ctgtctgctg ggggcaggac caaccaccag 9300  
 ggaactgcag cccccctgtg ctgagtgcac cagagacttg gaatggaaca cactggcctg 9360  
  
 cgacactcat cacaacgaac agaaactgct ttgtacactg aataaacgca gtgaatacce 9420  
 agctcaggat cacagacaca tgaatgcaaa gttatattag tataaccaag ggtgggaatg 9480  
 agggcaatta cagataactt atagacatga attactaaca aaacagggca aatgtttgc 9540  
 tcataaataa catgaaaata caatatatag tcatatgtat atatacatgt atatatataa 9600  
 atgacataat atgtatatat ttttacaat aactgcagg aaaataatat ttttcctcta 9660  
 cagaagagat tggcaaatct gacatctaaa at 9692  
  
 <210> 22  
 <211> 160  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><223> Synthetic  
 <220><221> misc\_feature  
 <222> (1)..(90)  
 <223> Mouse Sequence  
 <220><221> misc\_feature  
 <222> (91)..(160)  
 <223> Human Sequence  
 <400> 22  
 catgatgttt ctttcttagg aaagccaggg catttctcta ttctccaatc tcttggtca 60  
 atgccttgg ccctctttt gttccactag tgaagcctct ccagccaggg gctgaggtcc 120  
 cgggtggtgtg ggcccaggag ggggctcctg cccagctccc 160  
  
 <210> 23  
 <211> 197  
 <212> DNA

<213> Artificial Sequence  
 <220><223> Synthetic  
 <220><221> misc\_feature

<222> (1)..(60)  
 <223> Human Sequence  
 <220><221> misc\_feature

<222> (61)..(66)  
 <223> XhoI  
 <220><221> misc\_feature

<222> (67)..(100)  
 <223> loxP  
 <220><221> misc\_feature

<222> (106)..(131)  
 <223> I-CeuI  
 <220><221> misc\_feature

<222> (132)..(137)  
 <223> NheI  
 <220><221> misc\_feature

<222> (138)..(197)  
 <223> Mouse Sequence

<400> 23

tgttatttac agatacgtga gtttgggcaa attattgttc tctgtgtccc agctgtaaac 60  
 ctcgagataa ctctgtataa tgatgctat acgaagtat gctagtaact ataacggtec 120

taaggtagcg agctagccgt ggacagatac agcaacgtga gctagttatt ctgtcctaaa 180  
 gtctcagttg gaagatg 197

<210> 24

<211> 4551

<212> DNA

<213> Artificial Sequence  
 <220><223> Synthetic  
 <220><221> misc\_feature

<222> (1)..(69)  
 <223> Mouse Sequence

<220><221> misc\_feature

<222> (70)..(3976)

<223> Human Sequence

<220><221> misc\_feature

<222> (3977)..(3982)

<223> XhoI

<220><221> misc\_feature

<222> (3983)..(4016)

<223> LoxP

<220><221> misc\_feature

<222> (4022)..(4047)

<223> I-CeuI

<220><221> misc\_feature

<222> (4048)..(4053)

<223> NheI

<220><221> misc\_feature

<222> (4054)..(4551)

<223> Mouse Sequence

<400> 24

gtgtgatcac atcccaaagt gggtggccag gctgggcag agaaaggggg caacttcggg 60

tgtgtgtgac aaggagtggt tgggtgcagt gctggcggac acagcgatcc cgttttcttc 120

tctctgcacg ctgtcctggc cagactcca gctgcaggag gagctgcggg gcctgagaga 180

gacgttcagc aacttcacag cgagcacgga ggcccaggtc aagggttga gcaccaggg 240

tgaggcgct gggcggggc tgggctggg gctgggctg ggggctgtc gggaacgctg 300

agcgagcctc tcccgcagga ggcaatgtgg gaagaaagat gaagtcgcta ggtcccagc 360

tggagaaca gcagaaggac ctgagtgaag gtgagagagg gagtgtgtgt gtgtgtgtgt 420

gtgtgtgaaa gagagtgaga atgtgtggat gtgtgtgaga aagtgtgagt gtgtgtggat 480

gtgtgtgaga atgagagga gtgtgtgtgt gtgtgagtct gtgtgtgaga atgaggggga 540

gtgtgttttg ggtgtgtgta tgagagcctt gtgtggatgt gagaatgaga gggagtgtgt 600

atgtctgtga gtgtgagaat gagatggagt gtgtgtgagt ctgtgtgtga gaatgaggtg 660

tgtgtgtgtg agaatgagat ggtgtgtgtg tgggaatgag agggggtgtg tgtctgagtg 720

tgagaatgag atagagtgtg tgtgagacag tctgtgggaa tgagaggag tgtgtgtgag 780

agtgtgagaa tgacggagtg tgtctgtgag tigtataatg aggtgtgtgt gagtctgagt 840  
 gtaagaatga gatggggtgt gtgtgtctgt gagtgtgaga gtgtgagaat gaggggtgtt 900  
 tgtgtctgag tgtgagtctg tgagtgtgag aatgagatgg ggtgtgtgag tgagtgtgag 960  
 aatgagatgg ggtgtgtgtg tctgtgagtg tigtgtgtgtt tigtgagtgtg agaatgagat 1020  
 ggggtgtgtg tgagtgtgag aatgagatgg ggtgtgtgtc tigtgtgtgag aatgagatgg 1080  
  
 gtgtgtgtgt gacagagtct gagtgtgaga atgagaggga gtgtgtgtga gtgtgagaat 1140  
 gagaaggagt ggatgggtgt gtgagtctgt gtgaatgagg gagtgggtgt gtgtacgagt 1200  
 gtgagtctgt gtttatgtg gagaatgtgt cagtgtatgt gtgtgagaac gtgtgtatgt 1260  
 gtgttagtgt gtgttgcgtg tgtgggggaa tgagagggat tigtctgtg agtgtgagaa 1320  
 tgagatggag tigtctgtgag actgtgtgtg aggagtggga gtgtgtgtga gaatgagatg 1380  
 ggtgtgtgtg tctgagtgtg tgtctgtgag aatgagaggg agtgtgtgtg tigtgtgagag 1440  
 cctgtgtgaa aatgagaagg agtgtggatg ggtgtttgtg agtgggagag tctgtgtgtt 1500  
  
 tatgtgtgtg agaatgaggg agtgtgggtg tigtgtcgaa tigtgagtctg gtttatgtg 1560  
 tigtgagaat tigtgagtga tigtgagaacg tigtgttag tigtgtgctg gtgtgagaat 1620  
 gtaagtatat gtgtaagtc atgtgagtgt gtgtatgtgc gtgtgtgtg aatgtgcatt 1680  
 gtgtgtgcat gtgtgaaaga gtatatgtgt gttgtgggtg agtgtgtgtg gtgtgtgtag 1740  
 tgggtgaggg tigtttgat gtgtgggtgt gcgttgtgtg aatgtgtgta tgtgggtgag 1800  
 ggtgtgtgtg cctgtgtgag ggtgtgtgtt ggtttttgtg tgtgtttggg tgagggtgtg 1860  
 ttgtgtgtgt gtgtgggtga aggtgtgttg tigtgtcgt gggtgaaggt gtgtgtgtg 1920  
  
 tgtagtact gtagattagg gtgtgttccg tigtgtgtgtg tgagggtgta tgttgtgggt 1980  
 gttttgtgtg tgagtgggtg tgtaaggggtg tgtgtgtgt atgtgggtta aggtgtgtta 2040  
 tgcgtgaggg tglatttgtgt gtgtgttttg tigtgttgtgt gtgtatgtgg gttagggtgt 2100  
 gttgtgtgtt tigtgttttt gtgtgttct tigtgtatgtg gggtacgggtg tgttgtcgt 2160  
 gtgagggtgt gttgtgtatg gtgtgttgt tigtgtgtgt gagtgtgtat gtgagttagg 2220  
 gtgtgtgtg tctatgtatg tigtgtgtaag ggtgtgtgt gtgtctgtgg gtgtgtttg 2280  
 tgtatgtggg ttagggtgtg ttgtgtgttc tgtatttgtgt gtttatgtg ttgtctgtat 2340  
  
 gtgggttatg tigtgtgtgt gtgtttgtgga tgtatgtggg ttagggtgtg ttgtgtgtct 2400  
 ctgtgtgttg tctcgttttg tgtctgtggg ttagggtgtg ttgtatgtgt tgtgtttgt 2460  
 gtgttgtccg tigtgtgtga tgtgggttag gttgtgtgtg tigtgttgt atattgtctg 2520  
 tigtgtgtgt ttaggatgtg ttgtgtgtct gtgtgagtgt gtgtgtaagg gtgtgtgtg 2580  
 tigttaggag tigtgtgtgt tigtgtatg ggggtctctc aggccaactc cgctgtgtt 2640

tgtggcaatg cgacgggtgt tccgggtccca gcaggaggat gtagggctga cctcgtttcc 2700  
 cgtttccctc cccgtggttt ccgcatctcc tcccgtccc ctccgcccg tctccccaga 2760  
  
 tcactccage ctgctgctcc acgtgaagca gttcgtgtct gacctgcgga gectgagctg 2820  
 tcagatggcg gcgctccagg gcaatggtaa ggaggccagc ccggcccgt ctctgctcc 2880  
 ccccttctct gggcagcgt tagcccctgc gccccgttc tcccgtcag gctcagaaag 2940  
 gacctgctgc cgggtcaact ggggtggagca cgagcgagc tgctactggt tetctcgctc 3000  
 cgggaagcc tgggctgacg ccgacaacta ctcccggctg gaggacgcgc acctgggtgt 3060  
 ggtcacgtcc tgggaggagc aggtgaggac ccggagggtc tgggaggctg gctggcctcg 3120  
 gagagatcac caccgcctt ctctctctc agaaattgt ccagcaccac ataggcctg 3180  
  
 tgaacacctg gatgggcctc cagacaaaa acgggccctg gaagtgggtg gacgggacgg 3240  
 actacagagc gggcttcaag tgagtgcgcg cctccctcg gcctgggtcc ggccccttc 3300  
 gcgccctggg gcctgggct gaggagtctg gagcgaccg cctgcggatc cgacctctg 3360  
 gggcccacag ctggctctgt ccccaggaac tggaggccgg agcagccgga cgactgttac 3420  
 ggccacgggc tgggaggagg cgaggactgt gccacttca ccgacgacgg ccgctggaac 3480  
 gacgagctct gccagagcc ctaccgctgg gtctgcgaga cagagctgga caaggccagc 3540  
 caggagccac ctctcttta atttattct tcaatgctc gacctgccgc aggggtccgg 3600  
  
 gattgggaat ccgccatct gggggcctct tctgcttct cgggaatctt catctaggat 3660  
 ttaaggaa ggggaaggat aggggatgt tccgaagggt aggagcttga aaccctggc 3720  
 gctttctgca gttgcaggt tatcattgtg aactttttt ttttaagagt aaaaagaaat 3780  
 atacctaac ctctgttag ttgtctggtt attggggatt cggaagcagg agtgggctgg 3840  
 ttggcattac gaagccttag cgggtgctgt ggcatcatga gaactgtgtg ggctttgggc 3900  
 cagaatggcc agactttgtt atttacagat acgtgagttt gggcaaatta ttgttctctg 3960  
 tgtcccagct gtaaacctcg agataactc gtataatgta tgctatacga agttatgcta 4020  
  
 gtaactataa cggctctaag gtagcgagct agccgtggac agatacagca acgtgagcta 4080  
 gtattctgt cctaaagtct cagttggaag atgggaggat tttgacctc tgtctgctgg 4140  
 gggcaggacc aaccaccagg gaactgcagc cccctgtgc tgagtgcac agagacttg 4200  
 aatggaacac actggcctgc gacctcacc acaacgaaca gaaactgctt tgtacctga 4260  
 ataaacgag tgaataccca gctcaggatc acagacacat gaatgcaaag ttatattagt 4320  
 ataaccaagg gtgggaatga gggcaattac agataacta tagacatgaa ttactaacia 4380

aacagggcaa aatgtttgct cataaataac atgaaaatac aatatatagt catatgtata 4440

tatacatgta tatatataaa tgacataata tgtatatatt ttacaaata cactgcagga 4500

aaataatatt tttcctctac agaagagatt ggcaaatctg acatctaaaa t 4551

<210> 25

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 25

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu

1 5 10

<210> 26

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 26

gaacaaaaac tcatctcaga agaggatctg 30

<210> 27

<211> 63

<212> PRT

<213> Homo sapiens

<400> 27

Ser Gln Asn Ser Gln Leu Gln Glu Glu Leu Arg Gly Leu Arg Glu Thr

1 5 10 15

Phe Ser Asn Phe Thr Ala Ser Thr Glu Ala Gln Val Lys Gly Leu Ser

20 25 30

Thr Gln Gly Gly Asn Val Gly Arg Lys Met Lys Ser Leu Glu Ser Gln

35 40 45

Leu Glu Lys Gln Gln Lys Asp Leu Ser Glu Asp His Ser Ser Leu

50 55 60

<210> 28

<211> 118

<212> PRT

<213> Homo sapiens

<400> 28

His Glu Arg Ser Cys Tyr Trp Phe Ser Arg Ser Gly Lys Ala Trp Ala

1                    5                    10                    15

Asp Ala Asp Asn Tyr Cys Arg Leu Glu Asp Ala His Leu Val Val Val

                  20                    25                    30

Thr Ser Trp Glu Glu Gln Lys Phe Val Gln His His Ile Gly Pro Val

                  35                    40                    45

Asn Thr Trp Met Gly Leu His Asp Gln Asn Gly Pro Trp Lys Trp Val

                  50                    55                    60

Asp Gly Thr Asp Tyr Glu Thr Gly Phe Lys Asn Trp Arg Pro Glu Gln

65                    70                    75                    80

Pro Asp Asp Trp Tyr Gly His Gly Leu Gly Gly Gly Glu Asp Cys Ala

                  85                    90                    95

His Phe Thr Asp Asp Gly Arg Trp Asn Asp Asp Val Cys Gln Arg Pro

                  100                    105                    110

Tyr Arg Trp Val Cys Glu

                  115

<210> 29

<211> 39

<212> PRT

<213> Mus musculus

<400> 29

Met Thr Lys Asp Tyr Gln Asp Phe Gln His Leu Asp Asn Asp Asn Asp

1                    5                    10                    15

His His Gln Leu Arg Arg Gly Pro Pro Pro Thr Pro Arg Leu Leu Gln

                  20                    25                    30

Arg Leu Cys Ser Gly Ser Arg

                  35

<210> 30

<211> 21

<212> PRT

<213> Mus musculus

<400

> 30

Leu Leu Leu Leu Ser Ser Ser Leu Ser Ile Leu Leu Leu Val Val Val

1 5 10 15

Cys Val Ile Thr Ser

20

<210> 31

<211> 229

<212> PRT

<213> Homo sapiens

<400> 31

Asn Ser Gln Leu Gln Glu Glu Leu Arg Gly Leu Arg Glu Thr Phe Ser

1 5 10 15

Asn Phe Thr Ala Ser Thr Glu Ala Gln Val Lys Gly Leu Ser Thr Gln

20 25 30

Gly Gly Asn Val Gly Arg Lys Met Lys Ser Leu Glu Ser Gln Leu Glu

35 40 45

Lys Gln Gln Lys Asp Leu Ser Glu Asp His Ser Ser Leu Leu Leu His

50 55 60

Val Lys Gln Phe Val Ser Asp Leu Arg Ser Leu Ser Cys Gln Met Ala

65 70 75 80

Ala Leu Gln Gly Asn Gly Ser Glu Arg Thr Cys Cys Pro Val Asn Trp

85 90 95

Val Glu His Glu Arg Ser Cys Tyr Trp Phe Ser Arg Ser Gly Lys Ala

100 105 110

Trp Ala Asp Ala Asp Asn Tyr Cys Arg Leu Glu Asp Ala His Leu Val

115 120 125

Val Val Thr Ser Trp Glu Glu Gln Lys Phe Val Gln His His Ile Gly

130 135 140

Pro Val Asn Thr Trp Met Gly Leu His Asp Gln Asn Gly Pro Trp Lys



<210> 35  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 35  
 tccgcccgct gttaaac 18  
 <210> 36  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 36  
 ggaacccta gtgatggagt t 21  
 <210> 37  
 <211> 16  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 37  
 cggcctcagt gagcga 16  
 <210> 38  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 38  
 cactccctct ctgcgcgctc g 21  
 <210> 39  
 <211> 291  
 <212> PRT  
 <213> Macaca fascicularis  
 <400> 39  
 Met Thr Lys Glu Tyr Gln Asp Leu Gln His Leu Asp Asn Glu Glu Ser

1                    5                    10                    15  
 Asp His His Gln Leu Gly Lys Gly Pro Pro Pro Pro Gln Ser Leu Leu  
                          20                    25                    30  
  
 Arg Arg Leu Cys Ser Gly Pro Arg Leu Leu Leu Leu Ser Leu Gly Leu  
                          35                    40                    45  
 Ser Leu Leu Leu Leu Val Val Val Cys Val Ile Gly Ser Gln Asn Ala  
                          50                    55                    60  
 Gln Leu Gln Arg Glu Leu Arg Gly Leu Arg Glu Thr Leu Ser Asn Phe  
 65                    70                    75                    80  
 Thr Ala Ser Thr Glu Ala Gln Val Lys Gly Leu Ser Thr Gln Gly Gly  
                          85                    90                    95  
  
 Asn Val Gly Arg Lys Met Lys Ser Leu Glu Ser Gln Leu Glu Lys Gln  
                          100                    105                    110  
 Gln Lys Asp Leu Ser Glu Asp His Ser Ser Leu Leu Leu His Val Lys  
                          115                    120                    125  
 Gln Phe Val Ser Asp Leu Arg Ser Leu Ser Cys Gln Met Ala Ala Leu  
                          130                    135                    140  
 Gln Gly Asn Gly Ser Glu Arg Ala Cys Cys Pro Val Asn Trp Val Glu  
 145                    150                    155                    160  
  
 His Glu Arg Ser Cys Tyr Trp Phe Ser Arg Ser Gly Lys Ala Trp Ala  
                          165                    170                    175  
 Asp Ala Asp Asn Tyr Cys Arg Leu Glu Asp Ala His Leu Val Val Val  
                          180                    185                    190  
 Thr Ser Trp Glu Glu Gln Lys Phe Val Gln His His Ile Gly Pro Val  
                          195                    200                    205  
 Asn Thr Trp Met Gly Leu His Asp Gln Asn Gly Pro Trp Lys Trp Val  
                          210                    215                    220  
  
 Asp Gly Thr Asp Tyr Glu Thr Gly Phe Lys Asn Trp Arg Pro Glu Gln  
 225                    230                    235                    240  
 Pro Asp Asp Trp Tyr Gly His Gly Leu Gly Gly Gly Glu Asp Cys Ala  
                          245                    250                    255

His Phe Thr Asp Asp Gly Arg Trp Asn Asp Asp Val Cys Gln Arg Pro  
 260 265 270

Tyr Arg Trp Val Cys Glu Thr Glu Leu Asp Lys Ala Ser Gln Glu Pro  
 275 280 285

Pro Leu Leu  
 290

<210> 40

<211> 876

<212> DNA

<213> *Macaca fascicularis*

<400> 40

atgaccaagg agiatcagga cctgcagcat ctggacaatg aggagagtga ccaccatcag 60  
 ctcgaaaaag ggccacctcc tccgcagtcc ctctctcggc gtctctgctc eggccctcgc 120  
 ctctctctgc tctccttggg cctcagcctc ctgctgctgg tggttgtctg tgtgatcgga 180  
 tcccaaacg cccagctgca gcgggagctg cggggcctga gagagacgct cagcaacttc 240  
 acagcgagca ccgaggccca ggtcaagggc ttgagcacc agggaggcaa tgtgggaaga 300  
 aagatgaagt cgttggagtc ccagctggag aaacagcaga aggacttgag tgaagatcac 360  
  
 tccagcctgc tgctccactg gaagcagttc gtgtctgacc tgcggagcct gagctgtcag 420  
 atggcggcgc tccagggcaa tggctcgga agggcctgct gccagtcaa ctgggtggag 480  
 cacgagcgca gctgctactg gttctctcgc tccgggaagg cctgggccga cgccgacaac 540  
 tactgccgc tggaggacgc gcacctggtg gtggtcacgt cctgggagga gcagaaattt 600  
 gtccagcacc acataggtcc tgtgaacacc tggatgggcc tccacgacca aaacgggccc 660  
 tggaaagtg tggacgggac ggactacgag acgggcttca agaactggag accggagcag 720  
 ccggacgact ggtacggcca cgggctcggg ggaggggagg actgtgccca cttaccgac 780  
  
 gacggccgct ggaacgacga cgtctgccag aggcctacc gctgggtctg cgagacagag 840  
 ctggacaagg ccagtcagga gccacctctc ctttaa 876

<210> 41

<211> 284

<212> PRT

<213> *Rattus norvegicus*

<400> 41

Met Thr Lys Asp Tyr Gln Asp Phe Gln His Leu Asp Asn Glu Asn Asp

1                    5                    10                    15  
 His His Gln Leu Gln Arg Gly Pro Pro Pro Ala Pro Arg Leu Leu Gln  
                           20                    25                    30  
 Arg Leu Cys Ser Gly Phe Arg Leu Phe Leu Leu Ser Leu Gly Leu Ser  
                           35                    40                    45  
 Ile Leu Leu Leu Val Val Val Cys Val Ile Thr Ser Gln Asn Ser Gln  
                           50                    55                    60  
 Leu Arg Glu Asp Leu Arg Val Leu Arg Gln Asn Phe Ser Asn Phe Thr  
 65                    70                    75                    80  
 Val Ser Thr Glu Asp Gln Val Lys Ala Leu Thr Thr Gln Gly Glu Arg  
                           85                    90                    95  
 Val Gly Arg Lys Met Lys Leu Val Glu Ser Gln Leu Glu Lys His Gln  
                           100                    105                    110  
 Glu Asp Leu Arg Glu Asp His Ser Arg Leu Leu Leu His Val Lys Gln  
                           115                    120                    125  
 Leu Val Ser Asp Val Arg Ser Leu Ser Cys Gln Met Ala Ala Leu Arg  
                           130                    135                    140  
 Gly Asn Gly Ser Glu Arg Ile Cys Cys Pro Ile Asn Trp Val Glu Tyr  
 145                    150                    155                    160  
 Glu Gly Ser Cys Tyr Trp Phe Ser Ser Ser Val Lys Pro Trp Thr Glu  
                           165                    170                    175  
 Ala Asp Lys Tyr Cys Gln Leu Glu Asn Ala His Leu Val Val Val Thr  
                           180                    185                    190  
 Ser Trp Glu Glu Gln Arg Phe Val Gln Gln His Met Gly Pro Leu Asn  
                           195                    200                    205  
 Thr Trp Ile Gly Leu Thr Asp Gln Asn Gly Pro Trp Lys Trp Val Asp  
                           210                    215                    220  
 Gly Thr Asp Tyr Glu Thr Gly Phe Lys Asn Trp Arg Pro Gly Gln Pro  
                           225                    230                    235                    240  
 Asp Asp Trp Tyr Gly His Gly Leu Gly Gly Gly Glu Asp Cys Ala His  
                           245                    250                    255

Phe Thr Thr Asp Gly His Trp Asn Asp Asp Val Cys Arg Arg Pro Tyr  
 260 265 270

Arg Trp Val Cys Glu Thr Glu Leu Gly Lys Ala Asn  
 275 280

<210> 42  
 <211> 855  
 <212> DNA  
 <213> Rattus norvegicus  
 <400> 42

atgacaaagg attatcaaga tttccagcac ttggacaatg agaacgacca ccatcaactc 60

cagagagggc cacctcccgc tccaaggctc ttgcagcgac tctgctctgg attccgtctc 120  
 ttcttgcttt ccctgggcct cagcatcctg ctgctggtgg ttgtctgtgt gatcacatcc 180  
 caaaattccc aactccggga agatctgcgg gttctaaggc agaatttcag caactttacc 240  
 gtgagcactg aggaccaggt caaggccctg accaccagg gagagagagt gggaagaaag 300  
 atgaagttag tcgagtcaca gctggaaaaa catcaggagg atctgagga agaccactct 360  
 agattgctac tgcatgtaaa gcagttagt tctgacgtgc gaagcttgag ctgccagatg 420  
 gccgcacttc ggggcaatgg ctctgaaagg atctgctgcc ccatcaactg ggtggagtat 480

gaaggcagct gctactggtt ctccagctct gtgaagcctt ggacggaagc tgacaagtac 540  
 tgccagctgg agaacgcccc cctggtggtg gtgacttctt gggaggagca gagattcgtc 600  
 cagcaacaca tgggccctt aaatacttgg attggcctaa ctgaccagaa cggaccctgg 660  
 aatgggtgg atgggacaga ctatgagaca ggcttcaaga actggagacc agggcagcca 720  
 gatgactggt acggacatgg gcttggaggg ggtgaagact gtgcccactt caccaccgat 780  
 ggccactgga atgatgact ctgcaggagg ccctaccgct gggctctgtga gacagagttg 840  
 ggcaaggcca attag 855